

## Article

# Epistemic communities under active inference

Mahault Albarracín<sup>1</sup>, Daphne Demekas<sup>2</sup>, Maxwell Ramstead<sup>3,4</sup> and Conor Heins<sup>5,6,7,\*</sup>

<sup>1</sup> Department of Cognitive Computing, Université du Québec à Montréal, Montréal, QC H2K4M1

<sup>2</sup> Department of Computing, Imperial College London

<sup>3</sup> Wellcome Centre for Human Neuroimaging, University College London, London, UK.

<sup>4</sup> VERSES Labs/Spatial Web Foundation, Los Angeles, California, USA.

<sup>5</sup> Department of Collective Behaviour, Max Planck Institute of Animal Behavior, Konstanz, Germany

<sup>6</sup> Centre for the Advanced Study of Collective Behaviour, and

<sup>7</sup> Department of Biology, University of Konstanz, Konstanz, Germany

\* Correspondence: cheins@ab.mpg.de

**Abstract:** The spread of ideas is a fundamental concern of today's news ecology. Understanding the dynamics of the spread of information and its co-option by interested parties is of critical importance. Research on this topic has shown that individuals tend to cluster in echo-chambers and are driven by confirmation bias. In this paper, we leverage the active inference framework to provide an *in silico* model of confirmation bias and its effect on echo-chamber formation. We build a model based on active inference, where agents tend to sample information in order to justify their own view of reality, which eventually leads to them to have a high degree of certainty about their own beliefs. We show that, once agents have reached a certain level of certainty about their beliefs, it becomes very difficult to get them to change their views. This system of self-Confirming beliefs is upheld and reinforced by the evolving relationship between agent's beliefs and its observations, which over time will continue to provide evidence for their ingrained ideas about the world. The epistemic communities that are consolidated by these shared beliefs, in turn, tend to produce perceptions of reality that reinforce those shared beliefs. We provide an active inference account of this community formation mechanism. We postulate that agents are driven by the epistemic value that they obtain from sampling or observing the behaviors of other agents. Inspired by digital social networks like Twitter, we build a generative model in which agents generate observable social claims or posts (e.g. 'tweets') while reading the socially-observable claims of other agents, that lend support towards one of two mutually-exclusive abstract topics. Agents can choose which other agent they pay attention to at each timestep, and crucially who they attend to and what they choose to read influences their beliefs about the world. Agents also assess their local network's perspective, influencing which kinds of posts they expect to see other agents making. The model was built and simulated using the freely-available Python package pymdp. The proposed active inference model can reproduce the formation of echo-chambers over social networks, and gives us insight into the cognitive processes that lead to this phenomenon.

**Keywords:** epistemic community, social media, active inference, opinion dynamics

---

## 1. Introduction

### 1.1. Confirmation bias and conformity

The practice of exchanging ideas, sharing concepts and values between different minds, is a fundamental process that allows humans and other living agents to coordinate and operate socially. By sharing of ideas, individuals and communities can better pursue their pragmatic goals and improve their understanding of the world and each other. Humans are compulsory cooperators [1]: human survival itself is predicated on the ability to access and leverage bodies of accumulated cultural knowledge. Over the course of evolutionary history, humans have developed an exquisitely sensitive capacity to discriminate reliable sources of information from unreliable ones, and to learn from other relevant human agents to improve their understanding or model of their world [2,3].

This epistemic process is not, however, without its flaws. There is evidence that humans process information by reasoning heuristically, which is hypothesized to limit

the consumption of energy and facilitate rapid decision-making [4–6]. One such heuristic is confirmation bias, which implies that, all other things being equal, individuals prefer sticking to their own beliefs over changing their minds [7]. There is an extensive literature documenting the phenomenon of *confirmation bias* and its relation to cognitive dissonance. Individuals faced with information that conflicts with their core beliefs may be prone to cognitive dissonance, which is experienced as undesirable [8–10]. Tolerance for cognitive dissonance varies across individuals, but in general, the phenomenon significantly influences decision-making [9,11]. To avoid such dissonance, individuals tend to selectively seek information from 'others like me', others whom they expect will share similar ideas, concepts, and values [9]. Confirmation bias has a social influence; in particular, individuals prefer sampling data from their in-group, and will seek to confirm their own ideas by foraging for confirmatory information from their in-group [12,13]. To make sure that they have access to other like-minded allies, agents are more likely to choose to belong to communities where their deeply held beliefs are promoted and shared, which limits the cognitive effort that is already expanded in the foraging of information [11]. In-group delivery of information influences how strongly this information is integrated, especially if group membership is important for the individual [14]. This sampling extends beyond other agents, to choice of media and environment. For instance, individuals generally choose news sources that fit their expectations [9].

This phenomenon of confirmation bias is echoed in another heuristic: conformity, the need to cohere with the beliefs of one's in-group [15,16]. It is adaptive for agents to conform to the behaviours of others in their niche, in part for the very reasons highlighted above [17]. Conformity limits how much information any one agent has to gather to act appropriately, and the sources sampled from their ingroup are generally trusted [18]. This is partly due to the fact that members of an in-group can be most precisely predicted: their behaviours are normed, and expected by the members of the group, in ways that generally benefit its members [19,20]. But conformity has other benefits as well. Being able to sample from the group entails a continued relationship to other members. This will also enable members to acquire pragmatic resources beyond information (e.g., food and shelter), as the group generally provides for its members [21,22]. Being cut off from the group can lead to existential difficulties [23,24]. Group members can be sanctioned if they fail to conform to the norms, including epistemic norms [25,26].

### 1.2. The spread of ideas

These two heuristics, confirmation bias and conformity, mutually reinforce each other. Specifically, to save energy, confirmation bias leads to agents' being drawn to groups that validate their opinion, and thus increases the probability of behavioural and epistemic conformity [27]. Importantly, these two heuristics form the basis for information spread. Agents spread information through media and through connections to one another, given a network structure [28]. The spread of ideas and behaviors from one agent to another serves both local and larger-scale coordination [28,29].

The spread of ideas more straightforward when agents are already attuned to them. Individuals are more likely to adopt ideas that they believe will have a positive effect on them, especially if the outcome of sharing that information will be positive [30]. According to Falk and Scholz, this entails that sharing among group members of news that dovetails with group norms is likely to lead to the adoption of these ideas among the other group members, following the conformity heuristic mentioned earlier. One way to predict whether information will be coherent with the group norms is to assess it with one's respect to own value system. Naturally, similar individuals within a given group, who share values, will be more likely to spread ideas [31,32].

This notion of attunement or synchronisation is fundamental. Synchronisation across network nodes lowers the cost of information flow [33], and increases the certainty of the message being spread, as well as the quality of its reception, even if the message itself may be prone to errors [34,35]. Specifically, a message will be more intelligible to group

members who share a common set of codes, and agents are more likely to integrate new information if it fits with their understanding of the world [36–38].

Hashtags have been shown to be heavy carriers of information in echo-chambers. They tend to be used in partisan ways, to reach people of similar mindsets, as well as to signal one's own partisanship affiliation [39]. The spread of information is optimised through hashtags as pseudo-meta-linguistic categorisation makers [40].

### 1.3. Communities forming around ideas

Thus, the beliefs and epistemic communities of agents develop together, synchronously. We label communities formed in this process of belief sharing as 'epistemic communities'. Such communities share and spread a worldview, or a paradigm, and normalize sampling behaviors (i.e., manners of observing and engaging with the shared social world) that reinforce this view of the world [41]. Individuals in the community are tied together by these epistemic practices, further reinforcing the social signals which act as evidence for the shared model of the world [20].

One such example of these communities is the *echo-chamber*, a phenomenon that has been studied significantly in social media [39,42–48]. Echo-chambers are an extreme example of epistemic communities, and they have components that enforce their formation and maintenance [49–51]. Echo-chambers tie people with similar views together, and tend to actively work against the engagement with, and assessment and evaluation of, external sources (e.g., information provided by members of the outgroup) [43,49]. Echo-chambers can become epistemically vulnerable when members can no longer assess whether an information is true or not [49,52]. Similarly, only having access to a few sources limits how much information can be gathered, and relevant sources of evidence may fall through the gaps [50,53]. According to [53], error will be propagated, and it will be difficult to check errors against anything, as most minds in the echo-chamber are synchronized, and poised to make the same mistakes.

### 1.4. Volatility and habit formation

Studies on the perception of environmental volatility range from economics to psycho-education for the autism spectrum [54–56]. Optimal inference in a changing world requires integrating incoming sensory data with beliefs about the intrinsic volatility of the environment. Intuitively, environments with higher volatility change more quickly and thus have a shorter intrinsic timescale—and conversely for environments with lower volatility. For example, autistic individuals tend to pay more attention to small changes in the environment, giving them a better ability to track potentially important fluctuations in information [54]. On the other hand, this increased attention to environmental fluctuations may also lead to increased sensitivity to random, non-informative changes in the environment, a phenomenon that might be called (from a signal-detection perspective) a higher 'false-positive' rate [54].

When this type of precision dynamics [57] is applied to the social field at large, emergent epistemic phenomena can be explained. For instance, during the COVID-19 pandemic, the certainty around knowledge was very low, as information about the pandemic and the biology of the virus was limited [58,59]. In addition, alternative sources of information (e.g., anti-vaccine conspiracies) had become more prevalent and more influential in some social networks [60]. The gravity of the affliction, and the strength of the governmental response, also made any information on the topic vitally important, and worth one's attention [61]. This prompted an intensive use of information technology in order for individuals to find answers ("doing one's own research"). This excessive use points to the awareness by laypeople of the high volatility of the topic. [62] measured emotional volatility on social media in China during the pandemic, and explored the social dynamics underlying the emotional volatility.

Individuals can deal with volatility by using various coping mechanisms. One such mechanism is to constrain the uncertainty related to their own behaviors via *habit formation*

[63–72]. In this paper, we model habit formation as a form of behavioral reinforcement, where behaviors become more probable as a function of how often they are engaged [73,74]. If behavior is initially goal- or information-driven, habit-learning can then ‘zero-in’ and isolate the invariant features of such (initially) goal-directed behavior [75], mirroring the so-called transition from ‘model-based’ to ‘model-free’ decision-making in reinforcement learning [76,77]. After an agent has engaged in a given behavior enough, even if that behavior is initially pursued in a goal-driven manner, a habit can then be formed and become hard to ‘unlearn’ [78]. This view also supports the idea that, initially, habit-formation can be goal-driven. In the model we introduce here, behavior is driven by information-seeking drives that, due to confirmation bias, leads agents to preferentially sample information from other agents with beliefs that (they believe) are similar to their own. In this sense, confirmation bias serves as the original ‘motivation’ that later underwrites preferential sampling behavior. In combination with habit learning, this peer-specific sampling can then become impossible to stop enacting, even in the face of changing information.

### 1.5. An active inference model of epistemic communities

This paper introduces a computational model of epistemic communities, wherein individual agents share information with one another and come to form beliefs not only about their local environment, but also about the beliefs of other agents in their community. To understand this phenomenon, we leverage the active inference framework, a first principles theory of cognition, which explains the manner in which agents select actions based on their causal model or understanding of the world. Active inference says that organisms act to minimize a quantity called variational free energy, which quantifies the divergence between expected and sensed data. From this point of view, to select an action is to infer ‘what I must be doing, given what I believe and what I sense’. Extensive work has been done in the field of active inference to study social systems and the way in which the minimisation of free energy could give rise to (eventually large-scale) behavioural coordination [3,16,79–84]. However, much of this work is still theoretical.

At first glance, it might appear difficult to model a phenomenon like confirmation bias using an active inference formulation, because action selection in active inference is guided by the principle of maximizing Bayesian surprise or salience, which requires constantly seeking out information that is expected to ‘challenge’ one’s world model [85–87].

However, the key notion that allows ‘confirmation bias’ to nonetheless emerge under active inference, is ultimately the *subjective* nature of information gain, also known as ‘epistemic value’. Crucially, this Bayesian surprise or information gain term is always an *expected* surprise—that is, what counts as an ‘information-maximizing’ observation is always defined in relation to agent’s set of beliefs or generative model. Due to this inherent subjectivity, the true informativeness or epistemic value of an action can be arbitrarily far from the agent’s expectation thereof. Taking advantage of this, in the model presented here, we endow agents with what we refer to as *epistemic confirmation bias*. This is implemented by building a prior belief into the generative model, namely that agents are more likely to sample informative observations from agents with whom they agree *a priori*. Therefore, agents will sample agents with whom they agree under the (not necessarily true) beliefs that such agents are more likely to provide higher quality information.

We can make two important distinctions between the kind of polarisation that we observe in traditional opinion dynamics and the kind achieved through multi-agent active inference modelling. First, in traditional approaches, the implementation of bounded confidence to motivate polarisation is essentially a hard-coded restriction on the agents’ ability to perceive and therefore update their beliefs [88–91]. In contrast, in the active inference approach, polarisation is instead motivated by the positive effect of confirmation bias, which is integrated directly in the agents’ (likelihood) model of the world, which allows agents to get more evidence about their environment if the information comes from another agent that shares the same worldview. This means that agents are motivated implicitly in their generative models to gain more evidence about the world if this evidence

confirms their preexisting beliefs. Second, in the traditional approaches, agents can directly perceive the 'belief state' of other agents, and are thus the opinion of one agent directly influences that of another [88,90]. This is an unrealistic assumption, since human agents have to infer the belief states of others by interpreting their behaviour. This aspect of belief inference is a cornerstone of the active inference approach: the belief of another agent is modelled as a hidden state of the world — thus agents do not have direct access to each others' belief states. Instead, through inference, they come to hold beliefs about each others' beliefs, in addition to a belief about some agent-independent 'world states' [20,92].

More recently, researchers have begun to build Bayesian models of opinion dynamics, motivated by the Bayesian brain hypothesis and the notion that decision-making is inherently probabilistic [93–97]. Generally, the active inference approach falls within the theoretical umbrella of Bayesian agent-based modelling, because there is a deep assumption that environmental states are inherently hidden (in our case, the belief states of other agents) and need to be inferred on the basis of prior beliefs and sensory observations (i.e., observing the behaviour of other agents). However, as sketched above, a crucial point that distinguishes approaches like active inference and planning as inference from the general Bayesian approach is the notion that *actions themselves are inferred* [98,99]. While there have been models that use Bayesian inference for inference of opinions (i.e., Bayesian belief states about some particular idea), the process of action selection within these works is still often added on after the fact using an arbitrary decision rule (e.g., a softmax function of an arbitrary value vector). Action selection is often cast as a noisy signal of the true belief state, such as in [97], which is then used to update neighbouring agents' beliefs through Bayesian inference. Crucially, in active inference, behavior itself is cast as the result of inference, specifically by sampling actions from a posterior distribution over actions. The posterior over actions is obtained by minimizing the expected free energy of future beliefs, conditioned on actions. In other words, actions are selected in order to achieve goals and minimize future uncertainty, i.e., to maximize a lower bound on Bayesian model evidence.

Importantly for our purposes, one can supplement this goal-directed aspect of policy inference, driven by the expected free energy, with an inflexible 'prior preferences over actions', i.e., habits. If this prior preference over actions is learned over time, then in the context of the opinion dynamics model presented here, this can lead to a propensity to continue sampling agents that have been sampled previously. The idea of choosing actions through inference in accordance with the minimisation of uncertainty is powerful as a modelling technique, because through the choice of policy preferences, one can encode various social behaviors, such as conformity, habit formation, hostility, or indifference. While in this report, only habit formation, conformity, and polarisation are explored, we emphasize the potential of augmenting the current model to capture a wider range of features observed in human social behaviour.

### 1.6. Hypotheses

In this paper, we present a multi-agent model of opinion dynamics based on the active inference formulation. Our simulated agents are situated in a social network where they observe the behavior of other agents and update their beliefs about a pair of abstract, mutually-exclusive "Ideas" (e.g., the truth values of two competing claims), as well as the beliefs of their neighbours in the social network. Agents themselves have a prior preference to announce their beliefs via an action that is observable by other agents (e.g., posting/tweeting a "hashtag"). We show that the proposed active inference model can replicate confirmation bias, exposure effects, the formation of echo-chambers, and exacerbation of these phenomena via habit-learning. These effects can be modelled by changing the parameters of individual generative models, i.e., the cognitive features of the individuals comprising the group. We also uncover interesting interactions between individual-level cognitive features and the network architecture that constrains their social interactions. The large-scale behaviour of the model can be used to test three hypotheses, which are motivated by the existing literature. We formulate and test three hypotheses as follows:

**Hypothesis 1:** We cast confirmation bias in active inference as a form of ‘biased curiosity,’ in which agents selectively gather information from other agents with whom (they believe) they agree, under the assumption that like-minded agents provide higher-quality, more reliable information. We hypothesize that this ‘epistemic confirmation bias’ can mediate the formation of echo-chambers and polarisation in social networks of active inference agents. However, we further hypothesize that epistemic confirmation bias and network connectivity will bidirectionally modulate the formation of polarised epistemic communities, tuning the collective tradeoff between deadlock (polarisation) and agreement (consensus).

**Hypothesis 2:** We also consider the effect of agents’ beliefs about the volatility of their social environments. In particular, we examine how beliefs about social volatility impact exploratory sampling of other agents’ perspectives, which itself may interact with epistemic confirmation bias to determine the formation of echo-chambers. In particular, we hypothesize that beliefs about less-quickly-changing social environment (a belief in lower social volatility) will increase the likelihood of polarisation, as opposed to consensus.

**Hypothesis 3:** Finally, we also hypothesize that we can model selective exposure effects and conformity through habit formation, which naturally emerges through Bayes-optimal learning of a prior distribution over policies. To do so, we show that agents will begin to sample only those who belong to a particular epistemic community. We hypothesize that a greater learning rate for habit formation will lead to clusters within the network, thus amplifying and quickening the formation of echo-chambers.

Using the multi-agent active inference model of opinion dynamics, we achieve simulation outcomes that replicate common phenomena observed in the opinion dynamics literature, such as polarisation and consensus. In the sections to follow, we first describe the generative model that each agent uses to engage in active inference, and then discuss how we couple the agents together in an opinion dynamics network. We conclude by presenting numerical results that investigate each of the three hypotheses laid out above.

## 2. An active inference model of opinion dynamics

### 2.1. Overview

We present an multi-agent active inference model of opinion dynamics on an idealized social network. In the model, a group of agents simultaneously updates their beliefs about an abstract, binary hidden state (that represents two conflicting “Ideas”) and the opinion states about these ideas, held by a limited set of neighbouring agents. Each agent also generates an action that is observable to other agents. In the context of digital social networks like Twitter, these observable actions could be analogized to ‘posts’, ‘tweets’ or ‘hashtags’, i.e. some abstract expression carrying information about the belief state of the agent generating that expression. Hereafter we refer to these actions as ‘tweeting a Hashtag’ and describe agents’ behaviour as the decision to ‘tweet Hashtag 1 vs. Hashtag 2’, etc. Over time, each agent updates a posterior distribution (or belief) about which of the two Ideas is true, as well as a belief about what a connected set of other agents in the network believe (namely, those agents who they ‘follow’ or are ‘followed by’ in the social network). Both of these inferences are achieved by observing the behaviour of other agents, where crucially, this behaviour depends on each agent’s beliefs (notably about other agents). In our formulation, agents can only observe the behaviour of other agents to which they are specifically connected.

It is worth emphasizing that in this formulation, there is no *true* hidden state that corresponds to the competing truth status of the two “Ideas.” Rather, this abstract binary hidden state is only contained in the *generative model* or internal representation of each agent. The only ‘real’ states of the system are the social agents who comprise the network and their observable behaviour.

In the sections to follow, we will first briefly summarize the previous literature on computational approaches to the study of opinion dynamics. We then review the formalism of active inference, from the specification of the generative models that will each agent

will use to represent their external world, to the update equations for state estimation and decision-making. Finally, we describe the simulations of multi-agent dynamics by linking an ensemble of such active inference agents into a network.

## 2.2. Opinion dynamics models

In previous models of opinion dynamics, individual agents are often characterized by one or a few variables that encode the current belief or opinion held by that agent [100–102]. Collections of agents then update their respective opinion variables by ‘observing’ other variables that (either deterministically or stochastically) depend on the opinions of other agents in the ensemble. The nature of the inter-agent interactions varies across different models, ranging from homogeneous, ‘mean-field’-like global potentials [103,104] to structured, heterogeneous networks with fixed or dynamic weights between agents [105,106]. The opinion variables can take scalar or vector-values [107,108], and have either discrete or continuous support [109–112].

Bayesian variants of opinion dynamics models explicitly take into account the *uncertainty* associated with the observations and decisions of agents, where now, the updates to opinion variables become (exact or approximate) Bayesian updates [96,97,113,114]. The active inference model we present here is an example of such a Bayesian approach, with a few crucial distinctions, such as the approximate (as opposed to exact) nature of the Bayesian belief updating, and the fact that actions, in addition to opinions, are the result of inference. We will detail these distinctions further in the sections below on active inference.

## 2.3. Active inference

Active inference is a biologically motivated framework that rests on first principles of self-organization in complex, adaptive systems [87,98,115]. Particularly, it is premised on the notion that the internal states of any biological system are statistically insulated from the environment that generates sensory observations, and thus must engage in inference (about the causes of its sensory states) to behave optimally [116]. Active inference finesse this fundamental uncertainty by adding a Bayesian twist, proposing that biological systems entertain or entail a generative model of the latent environmental causes of their sensory inputs. Therefore, unlike classic reinforcement learning or reflexive behavioral algorithms (e.g., state-action policy mapping [73,117]), actions taken under active inference are guided by internal beliefs, which themselves are optimized with respect to an internal ‘world model,’ or representation of the world’s causal and data-generating structure.

Crucially, active inference agents represent their own actions (and their typical sensory consequences) in their generative model. By performing inference with respect to both hidden environment states of the world and the consequences of their own actions, active inference agents can evince behavior that both 1) achieves their goals or fulfills preferences and 2) actively reduces uncertainty in the agent’s world-model [87,98,116]. An active inference agent’s only imperative is to increase model evidence, or equivalently, to *reduce surprise*. Processes like learning, perception, planning, and goal-directed behavior emerge from this single drive to increase evidence for the agent’s generative model of the world.

In active inference, the agents never act directly on sensory data, but rather, change their beliefs about what causes that data. Thus, the core step in active inference consists in optimizing these beliefs using a generative model. This process is also known as Bayesian inference or Bayesian model inversion. Inference answers the question: “what is my best guess about the state of the world, given my sensory data and prior beliefs”? This can be formalized using Bayes’ rule:

$$P(\vartheta|y) = \frac{P(y|\vartheta)P(\vartheta)}{\sum_{\vartheta} P(y|\vartheta)P(\vartheta)} \quad (1)$$

where the optimal belief about ‘hidden’ or latent variables  $\vartheta$ , given some sensory data  $y$ , is called the *posterior distribution*  $P(\vartheta|y)$ . Bayes’ rule yields an analytic relationship

between the generative model  $P(y, \theta)$  and the posterior. Bayesian inference consists in calculating (either analytically or approximately)  $P(\theta|y)$ . Active inference is no different: perception (the generation of a best guess about the current hidden states of the world) is formalized as the computation of a posterior distribution over hidden states  $s$ , and action (the *active* part of active inference) is formalized as the computation of a posterior distribution over policies  $\pi$ . In active inference, however, this problem is turned into one of *approximate* Bayesian inference, where instead of finding the optimal posterior  $P(s|o)$ , active inference agents instead approximate this optimal posterior with a variational posterior  $Q(s; \phi)$ , i.e., a belief over hidden states that is parameterised by variational parameters  $\phi$ . The reason for this is that the exact inference is often computationally intractable. The marginalization problem involved in exact Bayesian inference (expressed in Equation (1)) is often intractable for many realistic generative models. Variational inference turns this intractable calculation of the marginal into an optimization problem, where a variational upper bound on *surprise* known as variational free energy (aka negative model evidence in statistics) is minimized:

$$Q^*(s; \phi) = \underset{\phi}{\operatorname{argmin}} \underbrace{D_{KL}(Q(s; \phi) \parallel P(o, s))}_{\text{surprise bound}}$$

$$D_{KL}(Q(s; \phi) \parallel P(o, s)) = D_{KL}(Q(s; \phi) \parallel P(s|o)) \underbrace{- \log P(o)}_{\text{surprise}} \quad (2)$$

where  $D_{KL}(q \parallel p)$  is the Kullback-Leibler divergence, a non-negative measure of difference between probability distributions, where  $D_{KL}(q \parallel p) = 0$  when  $q = p$ . Variational inference thus consists in optimizing the variational parameters  $\phi$  in order to minimize the free energy, which itself renders the variational posterior a better approximation to the true posterior. When variational inference is exact, the bound becomes exact and the free energy reduces to the surprise or negative log evidence. The remaining (negative) surprise can be itself used as a score for model averaging and model selection [118,119].

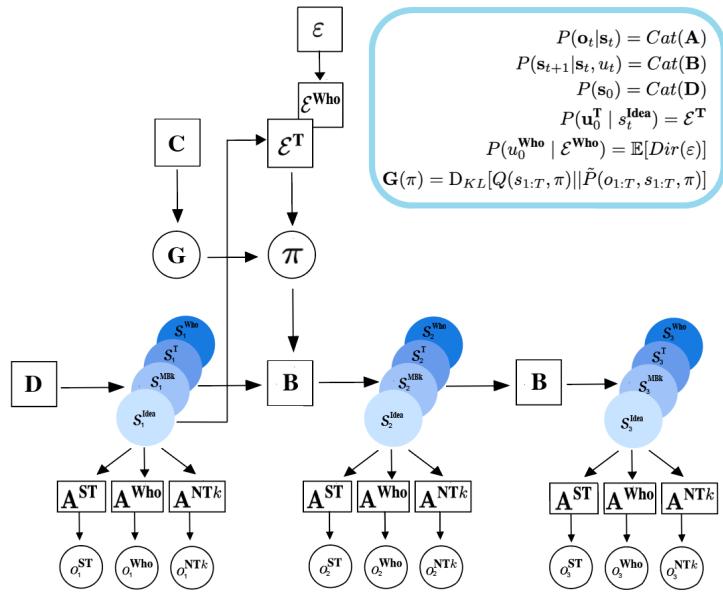
Active inference agents achieve perception and action by minimizing the surprise bound in Equation (2) with respect to variational beliefs about particular variables of their generative model. Optimizing beliefs about variables that represent latent environmental states (often denoted  $s$ ) is proposed as a formal model of perception, while optimizing beliefs about variables that correspond to policies or control of the environment (often denoted  $u$  or  $\pi$ ) is the formal analogue of planning and action. Therefore, active inference agents both infer the hidden states (perception) and policies (action) through a process of variational inference. The update equations used for perception and planning under active inference are detailed in sections [State estimation](#), [Policy inference](#) and [Action Selection](#).

Specifying a generative model  $P(o, s)$  is critical to determining the behavior of active inference agents. In the following sections we introduce the discrete state space model, a partially observed Markov Decision Process or POMDP, with which we equip agents in the multi-agent opinion dynamics setting.

#### 2.4. Generative model

Formally, the generative model is a joint probability distribution  $P(o, \varphi)$  over observations  $o$  and latent variables  $\varphi$ . Intuitively, one can think of the generative model as the agent's 'representation' of its environment, and specifically how that environment elicits observations [120]. In the discrete generative model described below, this generative model comprises assumptions about how hidden states  $s$  and actions  $u$  are probabilistically related to one another and to observations  $o$ .

In the current study, agents entertain *partially-observed Markov Decision Process* generative models, or POMDPs [121,122]. POMDPs are a class of decision-making models commonly used to simulate planning and decision-making in environments where agents must at each timestep select one of a discrete set of mutually-exclusive options. This is



**Figure 1.** Bayesian network representation of the POMDP generative model. Squares represent priors, likelihoods, or ‘factors’ that relate random variables to one another, and circles represent random variables (stochastic nodes). Different hidden state factors are represented as state variables and the different modality-specific  $A^{(m)}$  arrays of the observation model shown are side by side, since they lead independently to the observations generated in that modality, but dependent conjunctively on hidden state factors. Note that the  $\mathbf{B}$  array can be similarly decomposed into different sub-arrays, one per hidden state factor, but is shown as a single square here for simplicity. The prior over policies is parameterised by  $\mathcal{E}$ , which has separate prior over control states ( $\mathcal{E}^{\text{Who}}$  and  $\mathcal{E}^T$ ) for each control state factor. The box at the top right contains mathematical descriptions of each component in the generative mode. Note that while included in the graphical model, we leave out the  $\mathbf{C}$  vector since it is not relevant for the current model.

Variable Name	Notation	Meaning
<b>Observations</b>	$\mathbf{o} = \{o^{(1)}, \dots, o^{(M)}\}$	The focal agent's tweets $o^{\text{ST}} \in \mathbb{Z}^{1 \times H}$ Neighbour $k$ 's tweets $o^{\text{NT}k} \in \mathbb{Z}^{1 \times (H+1)}$ The sampled agent $o^{\text{Who}} \in \mathbb{Z}^{1 \times K}$
<b>Hidden States</b>	$\mathbf{s} = \{s^{(1)}, \dots, s^{(F)}\}$	The focal agent's beliefs $s^{\text{Idea}} \in \mathbb{Z}^{1 \times 2}$ Neighbour $k$ 's beliefs $s^{\text{MB}k} \in \mathbb{Z}^{1 \times 2}$ The Hashtag tweeted by focal agent $s^{\text{T}} \in \mathbb{Z}^{1 \times H}$ The neighbour sampled focal agent $s^{\text{Who}} \in \mathbb{Z}^{1 \times n}$
<b>Actions</b>	$\mathbf{u} = \{u^{(1)}, \dots, u^{(F)}\}$	The Hashtag control state $u^{\text{T}} \in \mathbb{Z}^{1 \times H}$ The neighbour attendance control state $u^{\text{Who}} \in \mathbb{Z}^{1 \times n}$
<b>Observation model</b>	$P(o_t^{(m)} = i   s_t^{(1)} = j, s_t^{(2)} = k, \dots) = [A^{(m)}]_{ijk\dots}$	Self tweet likelihood $\mathbf{A}^{\text{ST}} \in (\mathbb{R}_{>0})^{2 \times 2 \times 2K \times H \times K}$ neighbour tweet likelihood $\mathbf{A}^{\text{NT}k} \in (\mathbb{R}_{>0})^{2 \times 2 \times 2K \times H \times K}$ Neighbour Attend Likelihood $\mathbf{A}^{\text{Who}} \in (\mathbb{R}_{>0})^{K \times 2 \times 2K \times H \times K}$
<b>Transition model</b>	$P(s_{t+1}^{(f)} = i   s_t^{(f)} = j, u_t^{(f)} = k) = [B^{(f)}]_{ijk}$	Environmental dynamics and volatility $\mathbf{B}^{\text{Idea}} \in \mathbb{R}_{>0}^{2 \times 2}$ Meta-belief dynamics and volatility $\mathbf{B}^{\text{MB}k} \in (\mathbb{R}_{>0})^{2 \times 2}$ Tweet control $\mathbf{B}^{\text{T}} \in (\mathbb{R}_{>0})^{H \times H \times H}$ Neighbour attendance control $\mathbf{B}^{\text{Who}} \in (\mathbb{R}_{>0})^{K \times K \times K}$
<b>Initial State</b>	$p(s_0^{(f)} = i) = [D^{(f)}]_i$	Initial state distribution $\mathbf{D} \in (\mathbb{R}_{>0})^{2 \times 2K \times H \times K}$
<b>Control State Prior</b>	$P(u_0^{\text{T}}   s^{\text{Idea}}) = \mathcal{E}^{\text{T}}$ $P(u_0^{\text{Who}}   \mathcal{E}^{\text{Who}}) = \mathbb{E}[\text{Dir}(\varepsilon)]$	Empirical prior over Hashtag control state $\mathcal{E}^{\text{T}} \in (\mathbb{R}_{>0})^{H \times 2}$ Dirichlet hyperparameters over neighbour attendance control state $\varepsilon \in (\mathbb{R}_{>0})^{1 \times K}$

**Table 1.** Variables of the POMDP generative model of single agent opinion formation. The abstract name of each variable is written in the left column, its mathematical notation is in the middle column, and the right column shows how these variables correspond to different components of the opinion formation generative model.  $M$  is the total number of observation modalities and  $F$  is the number of hidden state / control factors. The observation model is a categorical likelihood distribution encoded by  $\mathbf{A}$ , which comprises a collection of modality-specific  $A^{(m)}$  arrays. The transition model is also a likelihood, mapping each state to its successor in time, encoded by the  $B^{(f)}$  arrays. The initial distribution over hidden states is encoded by the  $\mathbf{D}$  vector, and the prior distribution over control factors is encoded by the  $\mathcal{E}$  and  $\varepsilon$  distributions.

often represented using several random variables: a discrete set of actions  $u$  (also known as control states); hidden states  $s$ , which evolve according to (action-dependent) Markovian dynamics; and observations  $o$ , which probabilistically depend upon current hidden states. In most active inference models using POMDP generative models, hidden states, observations, and actions are discrete random variables—namely, they can take one of a finite set of values at a given time.

We include an additional latent variable, policies  $\pi$ , in the generative model. Policies are simply sequences of control states  $u$ . Using the terminology above, our generative model can be written down as  $P(\tilde{o}, \tilde{s})$  where  $\tilde{\varphi} = \{\tilde{s}, \tilde{u}, \pi\}$ . The tilde notation  $\tilde{x}$  denotes a sequence of random variables over time, e.g.  $\tilde{s} = s_{1,\dots,T}$ .

We can now write down the Markovian generative model as follows:

$$P(\tilde{o}, \tilde{s}, \tilde{u}, \pi) = P(s_1)P(\pi) \prod_{\tau=2}^T P(s_\tau | s_{\tau-1}, u_\tau)P(u_\tau | \pi) \prod_{\tau=1}^T P(o_\tau | s_\tau) \quad (3)$$

The observation likelihood  $P(o_\tau | s_\tau)$  represents the agent's probabilistic understanding of the relationship between hidden states  $s_\tau$  and concurrent observations  $o_\tau$ . Because both

observations  $\tilde{o}$  and states  $\tilde{s}$  are discrete, this likelihood distribution will be represented as a multidimensional array, which we hereafter denote **A**. Similarly, the transition distributions  $P(s_\tau|s_{\tau-1}, u_\tau)$ , which are denoted **B**, encode the agent's beliefs about how hidden states and control states determine subsequent hidden states. It is by changing actions  $u_\tau$  that the agent can exert control on its environment, since the evolution of hidden states depends both on the past state  $s_{\tau-1}$  as well as the concurrent action  $u_\tau$ . Finally, the distribution  $P(u_\tau|\pi_\tau)$  represents the mapping between policies and actions.

In many POMDP models, we segregate observations  $\tilde{o}$  and hidden states (and controls)  $\tilde{s}$  (resp.  $\tilde{u}$ ) into distinct *modalities* (for observations) and *factors* (for hidden states / control states):

$$\tilde{\mathbf{o}} = \left\{ \tilde{o}^{(1)}, \tilde{o}^{(2)}, \dots, \tilde{o}^{(M)} \right\} \quad \tilde{\mathbf{s}} = \left\{ \tilde{s}^{(1)}, \tilde{s}^{(2)}, \dots, \tilde{s}^{(F)} \right\} \quad \tilde{\mathbf{u}} = \left\{ \tilde{u}^{(1)}, \tilde{u}^{(2)}, \dots, \tilde{u}^{(F)} \right\} \quad (4)$$

where the superscripts refer to the index of the modality or factor index, respectively.

Observation modalities can be thought of as sensory 'channels' that provide distinct sorts of information. For example, in the context of human cognition, observation modalities might correspond to the information originating in different sense organs, e.g., the ears, eyes, skin.

Hidden state factors may be thought of as the generative model's latent representation of different features of the external world. Each of these factors has its own dynamics and can be thought of as statistically independent from other factors. For instance, an object might be described by both its spatial location and its color—'location' and 'color' would thus be candidates for distinct hidden state factors in a generative model of an object. This factorization is motivated by our intuition that something like an object's color and location are independent. An additional, minor note is that control states (the agent's representation of its own actions or ability to intervene on hidden states) are also divided into a set of control factors, with one control factor for every hidden state factor.

Given this factorization, at any given time a single observation will thus comprise a *set* of modality-specific observations, one from each sensory channel, and a hidden state will comprise of a *set* of hidden states, one from each distinct hidden state factor.

Now that we've introduced the class of discrete generative models with which our active inference agents will be endowed, we are now in a position to articulate the particular structure of the generative model for a single agent. From here, using active inference to perform inference and action with respect to each single agent's generative model, we can then 'link together' ensembles of these agents to form a complete opinion dynamics simulation.

## 2.5. An individual model of opinion formation

We describe a generative model of opinion formation for a single agent. Note that each active inference agent in the multi-agent simulations described below will be equipped with this same basic generative model. A single agent (hereafter: the 'focal agent') observes the actions of other agents, forms beliefs about an abstract binary environmental state, and chooses actions, which themselves are observable to other agents. The focal agent's action consists of two simultaneous choices: an 'expression' action (choosing which observable expression to make) and an 'observation' action (choosing which other agent to attend to). As mentioned above, we analogize the 'expression' actions to posts made by users on online social networks (e.g. 'tweets', 're-tweets', 'shares', 'likes'), and the contents of these actions we refer to as 'Hashtags.' Crucially, an agent can only observe one neighbouring agent at a time. Therefore, at each timestep, a focal agent both tweets its own Hashtag, and also chooses to read the Hashtag tweeted by another single agent. See Figure 1 and Table 1 for a summary of the distributions and random variables that comprise a single agent's generative model of opinion formation.

### Hidden states

Each agent's generative model comprises hidden states that fall into four categories—however, the actual number of hidden state factors per agent depends on their local network connectivity, so a particular agent will usually have *more than* four hidden state factors, but we nevertheless classify each hidden state factor into one of these four categories:

1.  $s^{\text{Idea}}$ : A binary random variable that encodes the agent's beliefs about an abstract environmental state that represents the truth value of two mutually-exclusive Ideas or claims. This binary variable can thus take a value of either 0 or 1, which we assign arbitrary labels of **Idea 1** and **Idea 2**. If **Idea 1** is true, then necessarily **Idea 2** is false, and vice-versa.
2.  $s^{\text{MetaBelief}}$  (shortened to:  $s^{\text{MB}}$ ): a set of binary random variables, each of which corresponds to a *particular neighbour's belief* about which of the two Ideas is true. As a representation of another agent's belief, we hereafter refer to this class of hidden state factor (and corresponding posteriors) as 'meta-beliefs'. The values of this variable we label **Believe Idea 1** and **Believe Idea 2**. Each agent will have one hidden state factor belonging to this category for each of its  $K$  neighbours, e.g.  $s^{\text{MB1}}, s^{\text{MB2}}, \dots, s^{\text{MBK}}$ .
3.  $s^{\text{SelfTweet}}$  (shortened to:  $s^{\text{T}}$ ): A binary random variable corresponding to what the focal agent is currently doing. In analogy to Twitter and other digital social media platforms, we refer to this action as 'tweeting' or 'posting', and the variable can take a value of either 0 or 1, representing one of two possible contents ('Hashtags'). These two actions are thus labeled **Tweet Hashtag 1** ( $s^{\text{SelfTweet}} = 0$ ) and **Tweet Hashtag 2** ( $s^{\text{SelfTweet}} = 1$ ).
4.  $s^{\text{WhoAttend}}$  (shortened to:  $s^{\text{Who}}$ ): A multinomial random variable with as many discrete levels as the focal agent has neighbours, representing which of their neighbours' actions the focal agent is currently attending to. For example, for an agent with three neighbours, this variable could take three values:  $[0, 1, 2]$  which we label **Attend Neighbour 1**, **Attend Neighbour 2**, **Attend Neighbour 3**.

For a single agent's generative model, the precise *number* of 'meta-belief' hidden state factors (those belonging to the  $s^{\text{MB}}$  class of factors) depends on how many neighbours the focal agent has. For instance, if a given agent  $i$  has three neighbours, then that agent's generative model will have three meta-belief hidden state factors:  $s^{\text{MB1}}, s^{\text{MB2}}, s^{\text{MB3}}$ , each representing the belief state of one of agent  $i$ 's three neighbours. Each agent has only *one* hidden state factor belonging to the other categories:  $s^{\text{Idea}}, s^{\text{T}}, s^{\text{Who}}$ . However, the *cardinality* (i.e. number of levels) for the  $s^{\text{Who}}$  hidden state factor will be equal to the focal agent's number of neighbours. So in the case of our agent  $i$  with 3 neighbours, then the possible values of  $s^{\text{Who}}$  will be  $[0, 1, 2]$ , corresponding to the action of attending to one of the three neighbours.

### Control states

Each agent is also equipped with two *control state* factors. These state factors are the agent's representation of its own actions in the environment. Control factors interact with hidden state factors to determine the next hidden state—thus, certain hidden state factors are deemed 'controllable' if they are paired with a control factor. In the current model, these two control state factors are paired with hidden state factors in Categories 3 and 4 above:

1.  $u^{\text{T}}$ : A binary random variable corresponding to which 'tweet action' to take, i.e., **Tweet Hashtag 1** vs. **Tweet Hashtag 2**. This control factor interacts with the  $s^{\text{T}}$  hidden state factor.
2.  $u^{\text{Who}}$ : A multinomial random variable corresponding to which neighbour to attend to, e.g., **Attend Neighbour 1**, **Attend Neighbour 2**, **Attend Neighbour 3**,... This control factor interacts with the  $s^{\text{Who}}$  hidden state factor.

### Observation modalities

Just as we did for the hidden states, now we describe three categories of observation modalities for a single agent's generative model:

1.  $o^{\text{SelfTweet}}$  or  $o^{\text{ST}}$ : A binary random variable representing the focal agent's observation of its own tweet actions—these 'self-observations' take the values of **Hashtag 1** and **Hashtag 2**.
2.  $o^{\text{NeighbourTweet}}$  or  $o^{\text{NT}}$ : A ternary random variable representing the observation of a neighbour agent's actions — these take the values of **Null**, **Hashtag 1**, **Hashtag 2**. Each agent has one 'tweet observation' modality for each of its  $K$  neighbours:  $o^{\text{NT1}}, o^{\text{NT2}}, o^{\text{NT3}}, \dots, o^{\text{NTK}}$ , in the same way that the number of  $s^{\text{MB}}$  factors depends on the number of neighbours. The purpose of the **Null** observation level will be clarified later on.
3.  $o^{\text{WhoAttend}}$  or  $o^{\text{Who}}$ : A multinomial random variable representing the observation of which neighbour the focal agent is attending to. This random variable has as many discrete levels as the focal agent has neighbours. For example, for an agent with three neighbours, this variable could take three values:  $[0, 1, 2]$  which we label **Attend Neighbour 1**, **Attend Neighbour 2**, **Attend Neighbour 3**.

A focal agent receives a full multi-modality observation per timestep, i.e.

$$\mathbf{o}_t = \{o_t^{\text{ST}}, o_t^{\text{NT1}}, o_t^{\text{NT2}}, \dots, o_t^{\text{NTK}}, o_t^{\text{Who}}\} \quad (5)$$

Each single observation is thus a collection of observations, one from each modality. Because one observation is collected from each modality at every timestep, the cardinality of some modalities is increased by 1, creating an additional observation level which we can call the "**Null**" observation level. The **Null** observation is included to effectively 'block' the focal agent from seeing the Hashtags of neighbours they are not actively attending to. This observation level is designed to have maximal ambiguity with respect to hidden states—in other words, seeing a **Null** observation affords no information about hidden states and thus has no effect on inference. This will become more clear when the observation and transition likelihoods of the generative model are described.

### Likelihoods

Having specified the random variables that form the support of a single agent's POMDP generative model, we can now move onto describing the likelihoods that determine how hidden states relate to observations, and how hidden states relate to each other over time. The construction of these likelihoods is indispensable for understanding both the belief updating and choice behavior of active inference agents.

We begin with the observation likelihood model  $P(\mathbf{o}_t | \mathbf{s}_t)$ . This is also known as the 'sensory likelihood' or observation model, and is parameterised by a series of categorical distributions whose parameters we collectively encode as the columns of a multidimensional array called **A**. In other words:

$$P(\mathbf{o}_t | \mathbf{s}_t) = \text{Cat}(\mathbf{A})$$

The entire **A** array is actually a set of tensors, with one sub-tensor per observation modality:

$$\mathbf{A} = \{A^{\text{ST}}, A^{\text{NT1}}, A^{\text{NT2}}, \dots, A^{\text{NTK}}, A^{\text{Who}}\}$$

Each modality-specific likelihood tensor  $A^m$  is a potentially multidimensional array, that encodes the conditional dependencies between each combination of hidden states

$s_t = \{s_t^1, s_t^2, \dots, s_t^F\}$  and observations  $o_t^m$  for that modality. For example, in a likelihood array with two hidden state factors, entry  $[A^m]_{ijk}$  encodes the conditional probability  $P(o_t^m = i | s_t^1 = j, s_t^2 = k)$ , i.e., the probability of observing outcome  $i$  within observation modality  $m$ , under hidden state factor 1 being level  $j$  and hidden state factor 2 being level  $k$ . In the case of the generative model for opinion formation, these likelihood arrays will be much higher dimensional than 3-D tensors, so we will generally refer to the elements of a modality-specific  $A^m$  array with the notation  $[A^m]_{ijk\dots}$  where the ellipses refer to an indefinite number of index-able lagging dimensions.

Each agent in the opinion dynamic model will have one  $A^m$  array per observation modality. We will now step through them to describe their role in the generative model.

#### Self Tweet Likelihood

The array  $A^{ST}$  represents the agent's beliefs about how hidden states relate to  $o^{ST}$  (which content the agent is tweeting, either **Hashtag 1** or **Hashtag 2**). By construction,  $A^{ST}$  encodes an assumption that  $o^{ST}$  only depends on  $s^T$ , the controllable hidden state factor corresponding to the tweet action. This is an unambiguous or isomorphic mapping, which we can express as follows:

$$A^{ST} = P(o_t^{ST} | s_t^T) = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (6)$$

In other words, the agent believes that the  $s^T$  factor unambiguously signals its true value via the  $o^{ST}$  observation modality. Each column of the matrix in Equation (6) represents a (conditioning) value of  $s^T$ , and each row represents a (conditioned) value of  $o^{ST}$ . The value of  $o^{ST}$  does not depend on any of the other hidden state factors, which means that this identity matrix is uniformly 'tiled' across the other dimensions of the  $A^{ST}$  array that represent the mapping between the remaining hidden state factors  $\{s^{(1)}, s^{(2)}, \dots\} \notin s^T$  and  $o^{ST}$ .

#### Neighbour Tweet Likelihood

The array  $A^{NTk}$  represents the focal agent's beliefs about how hidden states relate to  $o^{NTk}$ , the focal agent's observation of neighbour  $k$ 's tweet content.  $A^{NTk}$  encodes an assumption that  $o^{NTk}$  probabilistically depends neighbour  $k$ 's belief about the two Ideas, i.e., that  $o^{NTk}$  depends on  $s_t^{MBk}$ . This can be expressed as:

$$A^{NTk} = P(o_t^{NTk} | s_t^{MBk}, s_t^{Who} = k) = \begin{bmatrix} \mathbf{0} \\ \mathbf{h} \end{bmatrix} \quad (7)$$

where  $\mathbf{0}$  represents a  $1 \times 2$  vector of 0s, and  $\mathbf{h}$  is a  $2 \times 2$  matrix that represents the 'Hashtag semantics,' i.e. the assumed relationship between neighbour  $k$ 's beliefs and what Hashtag they are expected to tweet. Importantly, the first row of the likelihood matrix in Equation (7) represents the probability of encountering the **Null** observation, for the various settings of hidden states. This observation always has probability 0 when the focal agent is sampling neighbour  $k$ , as represented by the condition  $s_t^{Who} = k$ . Otherwise, when  $s_t^{Who} \neq k$ , the **Null** value will be expected with certainty. This can be expressed as:

$$A^{NTk} = P(o_t^{NTk} | s_t^{MBk}, s_t^{Who} \neq k) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (8)$$

This inclusion of the **Null** is necessary to ensure that a focal agent only expects to read one of Neighbour  $k$ 's tweet, if they are actively attending to Neighbour  $k$ —otherwise, they receive a 'blank' observation that affords no information about hidden states (as

represented by a maximally ambiguous likelihood over hidden states, i.e. a row of 1's). The lower two rows of the likelihood matrix in Equation (7) are occupied by the Hashtag semantics  $\mathbf{h}$ , which we stipulatively define with a 'Hashtag reliability' parameter  $p_h$ :

$$\mathbf{h} = \begin{bmatrix} p_h & 1 - p_h \\ 1 - p_h & p_h \end{bmatrix} \quad (9)$$

Here,  $p_h$  parameterizes two Bernoulli distributions, that respectively map between the two levels of  $s^{Idea}$  and the two levels of  $o^{NTk}$ . In the limiting case of  $p_h = 1$ , this means that the focal agent believes that neighbour  $i$ 's tweet content is unambiguous evidence for what Idea Neighbour  $k$  believes in. On the other hand, as  $p_h \rightarrow 0$ ,  $\mathbf{h}$  comes to resemble a maximum entropy distribution — in this case, according to the focal agent's generative model, Neighbour  $k$ 's tweet activity provides no information about its beliefs.

This basic conditional relationship outlined in Equations (7) - (9) enables agents to update their beliefs about the beliefs of their neighbours  $s^{MB}$  according what they observe their neighbours tweeting. Intuitively, this mapping captures the focal agent's beliefs that what their neighbours tweet is representative of what they believe. The accuracy of this mapping (the value of  $p_h$ ) determines how strongly Hashtags reflect opinions or the strength of beliefs. However, in order to allow agents to update their beliefs about the truth-values of the Ideas per se (i.e., update a posterior distribution over  $s^{Idea}$ ), we also construct  $A^{NTk}$  such that agents believe that the validity or truth-values of the Ideas themselves  $s^{Idea}$  probabilistically relates to  $o^{NTk}$ . Importantly, we make this conditional relationship 'biased' in the sense that, according to  $A^{NTk}$ , tweet observations are more precisely related to a particular setting of the  $s^{MBk}$  factor, if any only if  $s^{Idea}$  is aligned with that belief, i.e. when  $s_t^{MBk} = s^{Idea}$ . This can be formalized as an increased precision  $\gamma$  for subsets of those conditional distributions encoded by  $P(o_t^{NTk} | s_t^{MBk}, s^{Idea})$ , importantly those subsets when  $s_t^{MBk} = s^{Idea}$ . As we will describe later, in the context of action, this leads to an 'epistemic' drive for the focal agent to attend to neighbours who (are believed to) share their opinions, leading to a *confirmation bias* effect. We therefore refer to this 'biased precision'  $\gamma$  as the *epistemic confirmation bias* (ECB).

$$P(o_t^{NTk} = i | s_t^{MBk} = j, s^{Idea} = j, s_t^{Who} = k, \gamma) = \frac{e^{\gamma \mathbf{h}_{ij}}}{\sum_l e^{\gamma \mathbf{h}_{lj}}} \quad (10)$$

Note that this additional precision term  $\gamma$  exponentiates the Hashtag semantics matrix  $\mathbf{h}$ , which is already parameterised by the 'Hashtag reliability' parameter  $p_h$ . In the context of inference, an increasing value of  $\gamma$  means that the focal agent believes that tweet observations  $o_t^{NTk}$  will provide more information about hidden states, only in the case that the neighbour  $k$  generating that tweet has 'correct' beliefs, i.e., their beliefs are aligned with the true Idea. In the context of decision-making, this means that agents believe that most informative observations come from those neighbours that have the 'correct' beliefs. Under active inference, actions that evince informative observations (i.e., observations that resolve the most uncertainty) are preferred. This drive is known as the 'epistemic value' or 'salience' [87]. Therefore, higher levels of  $\gamma$  will lead to increased epistemic value associated with sampling only those neighbours that the focal agent believes have veridical beliefs, according to its own beliefs about  $s^{Idea}$ .

### Neighbour Attend Likelihood

The array  $A^{Who}$  represents the agent's beliefs about hidden states relate to  $o^{Who}$ . This observation model is constructed such that  $o^{Who}$  only depends on  $s^{Who}$ , and specifically that agents can always unambiguously infer who they are currently attending to, based on  $o^{Who}$ . This can be expressed succinctly as a  $K$ -dimensional identity matrix:

$$A^{\text{Who}} = P(o_t^{\text{Who}} | s_t^{\text{Who}}) = I_K \quad (11)$$

where  $K$  is the number of the focal agent's neighbours. Since the value of  $o^{\text{Who}}$  does not depend on any hidden state factors besides  $s^{\text{Who}}$ ,  $I_K$  is 'tiled' across the remaining dimensions of the  $A^{\text{Who}}$  array.

#### Transition Model

Now we move onto the transition likelihood model  $P(\mathbf{s}_t | \mathbf{s}_{t-1}, \mathbf{u}_{t-1})$ . This is also known as the 'dynamical likelihood' and is parameterized by a series of categorical distributions whose parameters stored in a tensor  $B$ :

$$P(\mathbf{s}_t | \mathbf{s}_{t-1}) = \text{Cat}(\mathbf{B})$$

As there are multiple hidden state factors in our generative model, the full  $\mathbf{B}$  array is actually split into a collection of sub-arrays, one for each hidden state factor:

$$\mathbf{B} = \{B^{\text{Idea}}, B^{\text{MB}1}, B^{\text{MB}2}, \dots, B^{\text{MB}K}, B^{\text{T}}, B^{\text{Who}}\}$$

Each sub-array  $B^f$  contains the categorical parameters of the factor-specific transition likelihood  $P(s_t^f | s_{t-1}^f | u_{t-1}^f)$ . Note that this construction means that hidden state factors are assumed to be independent by the generative model. In the context of the opinion dynamics model, this mean that a single agent assumes that the hidden state  $s^{\text{Idea}}$  both does not affect, and is not affected by, the belief states of neighbouring agents  $s^{\text{MB}k}$ , and furthermore that the belief states of neighbours do not affect one another. In the following sections, we summarize the transition models for each hidden state factor.

#### Environmental dynamics and volatility

The dynamics of  $s^{\text{Idea}}$  according a focal agent's generative model, are described by  $B^{\text{Idea}}$ . Since this is an uncontrollable hidden state factor, this can be expressed as a simple  $2 \times 2$  matrix, which expresses the focal agent's beliefs about the probability that  $s^{\text{Idea}}$  (which Idea is "true") switches over time. We parameterise this matrix with a precision parameter that we call 'inverse environmental volatility'  $\omega^{\text{Idea}}$ :

$$B^{\text{Idea}} = P(s_t^{\text{Idea}} = i | s_{t-1}^{\text{Idea}} = j, \omega^{\text{Idea}}) = \frac{e^{\omega^{\text{Idea}} I_{ij}}}{\sum_l e^{\omega^{\text{Idea}} I_{lj}}} \quad (12)$$

where  $I$  is the  $2 \times 2$  identity matrix. The higher the value of  $\omega^{\text{Idea}}$ , the more the focal agent believes that the same Idea remains valid over time (e.g. **Idea 1** is likely to remain the 'valid' idea from one timestep to the next). Consequently, a lower value of  $\omega^{\text{Idea}}$  (and thus a higher value of 'environmental volatility') means that the focal agent believes that the truth value of the two Ideas changes less predictably over time (the hidden state is likely to oscillate between **Idea 1** and **Idea 2**).

#### Meta-belief dynamics and volatility

The dynamics of  $s^{\text{MB}k}$ , or the meta-belief associated with neighbour  $k$  according to a focal agent's generative model, is described by  $B^{\text{MB}k}$ . Like  $s^{\text{Idea}}$ ,  $s^{\text{MB}k}$  is an uncontrollable hidden state factor, so the  $B^{\text{MB}k}$  array can be expressed as a  $2 \times 2$  matrix. Like  $B^{\text{Idea}}$ , we parameterize  $B^{\text{MB}k}$  with a precision parameter that we term 'inverse social volatility'  $\omega^{\text{Soc}}$ :

$$B^{\text{MB}k} = P(s_t^{\text{MB}k} = i \mid s_{t-1}^{\text{MB}k} = j, \omega^{\text{Soc}}) = \frac{e^{\omega^{\text{Soc}} I_{ij}}}{\sum_l e^{\omega^{\text{Soc}} I_{lj}}} \quad (13)$$

The interpretation of  $\omega^{\text{Soc}}$  is similar to that of  $\omega^{\text{Idea}}$ : a higher value of  $\omega^{\text{Soc}}$  implies that the focal agent assumes its neighbours have 'stubborn' opinions and aren't likely to change over time. A lower value means that the focal agent assumes its neighbours' opinions can easily change over time, or that its neighbours are 'fickle'.

#### Tweet control

Now we discuss the controllable dynamics of the hidden state factor corresponding to the Hashtag that the focal agent is tweeting:  $s^{\text{SelfTweet}}$  or  $s^{\text{T}}$ . Under the focal agent's generative model, this factor only depends on the control state factor  $u^{\text{T}}$ , and the corresponding  $B^{\text{T}}$  array can thus be expressed as an identity matrix that maps from the action (**Tweet Hashtag 1** vs. **Tweet Hashtag 1**) at timestep  $t-1$  to the next tweet value at timestep  $t$ :

$$B^{\text{T}} = P(s_t^{\text{T}} \mid u_{t-1}^{\text{T}}) = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (14)$$

This means that the agent can unambiguously determine what it tweets next (the value of  $s_{t+1}^{\text{T}}$ ) by means of actions  $u_t^{\text{T}}$ .

#### Neighbour Attendance control

Similarly for the dynamics of  $s^{\text{Who}}$ , under the focal agent's generative model, this factor only depends on the control state factor  $u^{\text{Who}}$ , and the corresponding  $B^{\text{Who}}$  array can thus be expressed as an identity matrix that maps from the action of which of  $K$  neighbours to attend to at timestep  $t-1$ , to the next value of  $s^{\text{Who}}$  at timestep  $t$ , namely which neighbour is being attended to:

$$B^{\text{Who}} = P(s_t^{\text{Who}} \mid u_{t-1}^{\text{Who}}) = I_K \quad (15)$$

Just like the dynamics of  $s^{\text{T}}$ ,  $s^{\text{Who}}$  is thus fully controllable by the agent, i.e. determined by the value of  $u^{\text{Who}}$ .

#### Priors

The next component of the generative model are the *priors* over both observations  $P(\mathbf{o})$ , hidden states  $P(\mathbf{s}_0)$ , and actions  $P(\mathbf{u})$ . In discrete active inference models, we represent these as vectors  $\mathbf{C}$ ,  $\mathbf{D}$ , and  $\mathbf{E}$ , respectively.

#### Observation prior $\mathbf{C}$

In active inference, goal-directed action is often motivated by appealing to a baseline prior over observations  $P(o \mid \mathbf{C})$  that specifies the agent's preferences to encounter particular outcomes over others. This caches out value in terms of log probabilities or information, rather than classical constructs like 'reward.' Interestingly, this prior over observations does not come into play when performing inference about hidden states (i.e., it is not part of the generative model in Equation (3)), but only during decision-making and action. Under active inference, actions are selected to minimize a quantity called the *expected free energy*, a quasi-value function that scores policies by their ability to bring expected observations in alignment with preferred observations, while also maximizing information gain (see the section on [Policy inference](#) for more details). In the current model, we do not rely on this  $\mathbf{C}$  vector to encode goals, but rather motivate action through a conditional action prior (see

the section on the **E** vector below). For this reason, in our model the **C** is a flat distribution over observations and doesn't contribute to decision-making in this context.

### State prior **D**

The prior over hidden states at the initial timestep is encoded by the so-called **D** vector,  $P(\mathbf{s}_0 \mid \mathbf{D})$ . The **D** vector encodes the agent's beliefs about the initial state of the world, prior to having made any observations. In the context of the opinion formation generative model, it encodes baseline beliefs about which Idea is true, the metab-beliefs of the focal agent's neighbours, as well as the initial tweet that the focal agent is making and the initial neighbour to which the focal agent is attending.

### Empirical prior over Hashtag control state: $\mathcal{E}^T$

We furnish the generative model with a special conditional prior over Hashtag control states  $P(u_0^T \mid s_t^{\text{Idea}})$ , parameterised by a mapping denoted by  $\mathcal{E}^T$ . This quasi-likelihood or link function renders the prior over Hashtag control states  $u_0^T$  an *empirical prior*, because of an explicit dependence on  $s_t^{\text{Idea}}$ . Under active inference, the final posterior over control states  $Q(\mathbf{u}_t)$  becomes a Bayesian average of the 'value' of each control state, as determined by the (negative) expected free energy (see the corresponding section on **Policy inference** below), as well as the prior probability of each control state as encoded by  $P(\mathbf{u}_0)$ . In the current model, we make the prior over control states an *empirical prior* parameterised by a 'link function' denoted  $\mathcal{E}^T$  vector. This makes the prior over the Hashtag control state  $u_0^T$  conditionally dependent on the  $s_t^{\text{Idea}}$  hidden state factor of the generative model. In practice, this implies that the prior over those control states corresponding to tweet actions  $P(u_0^T)$  depends on the posterior over  $s_t^{\text{Idea}}$ , the hidden state corresponding to which Idea is true. This can be expressed as follows:

$$P(u_0^T \mid s_t^{\text{Idea}}) = \mathbf{Cat}(\mathcal{E}^T)$$

where the mapping encoded by the entries of  $\mathbf{Cat}(\mathcal{E}^T)$  is an identity matrix, that maps each value of  $s_t^{\text{Idea}}$  to a single Hashtag control state (value of  $u_0^T$ ). At each timestep we approximate the prior at timestep  $t$  over  $s_t^{\text{Idea}}$  with the agent's current posterior belief  $Q(s_t^{\text{Idea}})$ . See the following sections on belief updating explain how one optimizes the variational posterior over hidden states  $Q(\mathbf{s}_t)$  using observations. Once approximated this way, we can re-express the empirical prior over Hashtag control states  $P(u_0^T)$  as:

$$P(u_0^T) = \mathbb{E}_{Q(\mathbf{s}_t)}[P(u_0^T \mid s_t^{\text{Idea}})]$$

Agents are therefore more likely to take the action  $u^T = \text{Tweet Hashtag 1}$  if they believe more in **Idea 1** than **Idea 2** (as reflected in the value of  $Q(s_t^{\text{Idea}})$ ), and likewise more likely to take the action  $u^T = \text{Tweet Hashtag 2}$  if they believe more in **Idea 1** than **Idea 2**. This empirical prior formulation thus renders the probability of taking a particular **Tweet Hashtag** action directly proportional to the agent's belief in one of the two Ideas, as encoded in the variational posterior  $Q(s_t^{\text{Idea}})$ .

### Prior over Neighbour Attendance control state: $\mathcal{E}^{\text{Who}}$

In addition to the prior over Hashtag control states  $P(u_0^T)$ , the generative model also contains a prior over the **Neighbour Attendance** control state  $u_0^{\text{Who}}$ . We parameterize this prior over control states using a categorical distribution  $\mathcal{E}^{\text{Who}}$ , whose probability itself is given by a Dirichlet distribution with parameters  $\varepsilon$ :

$$P(u_0^{\text{Who}} \mid \mathcal{E}^{\text{Who}}) = \mathbb{E}[\text{Dir}(\varepsilon)]$$

The Dirichlet parameters  $\varepsilon$ , unlike the parameters of categorical distributions, are positive but not constrained to integrate to 1.0. As hyperparameters of a conjugate prior distribution, they are often analogized to ‘pseudo-counts’ that score the prior number of times a given action has been taken (in this case, sampling a particular neighbour via the control state  $u_0^{\text{Who}}$ ). For instance, if the  $\varepsilon$  vector for an agent with three neighbours is initialized to have the values [5, 2, 1], this means that the focal agent has a built-in propensity to take the action **Attend Neighbour 1** rather than the actions **Attend Neighbour 2** or **Attend Neighbour 3**. And in turn, taking the action **Attend Neighbour 2** is twice as probable as taking the action **Attend Neighbour 3**. As we will see in the following sections, this ‘habit vector’  $\varepsilon$  can be learned over time by optimizing a variational beliefs over  $\mathcal{E}^{\text{Who}}$ , which involves incrementing a Dirichlet  $\varepsilon$  vector that parameterises the posterior  $Q(\mathcal{E}^{\text{Who}})$ .

### Summary

This concludes the specification of a single agent’s generative model for opinion formation. Now that we have specified this generative model, we move on to define the family of the approximate posteriors (the agent’s beliefs) over hidden states and policies  $Q(\mathbf{s}, \pi; \phi)$  as well as the variational free energy. In conjunction with the generative model, these can be used to derive the update equations used to perform active inference.

### 2.6. Approximate posteriors and free energy

Under active inference, both perception and decision-making are cast as approximate inference problems, wherein the variational free energy (or bound on surprise) is minimized in order to optimize beliefs about hidden states (perception) and beliefs about policies (decision-making/action). In order to derive the equations that perform this optimization, we therefore have to define the variational free energy. This free energy, equivalent to the bound defined in Equation (2), requires both an approximate posterior and a generative model. We defined a POMDP generative model for our active inference agents in the previous section; the remaining step before writing out the free energy is then to define an approximate posterior distribution. For compatibility with the categorical prior and likelihood distributions of the generative model defined in Equation (3), we will also define the approximate posterior as categorical distributions. Additionally, we will invoke a particular factorization of the approximate posterior, also known as a *mean-field approximation*, that allows us to factorize the approximate posterior over hidden states across timesteps. We define the approximate posterior over hidden states and policies as follows:

$$\begin{aligned} Q(s_\tau | \pi) &= \text{Cat}(\mathbf{s}_{\pi\tau}) \\ Q(\pi) &= \text{Cat}(\boldsymbol{\pi}) \\ Q(s_{1:T}, \pi) &= Q(\pi) \prod_{\tau=1}^T Q(s_\tau | \pi) \end{aligned} \tag{16}$$

where the notation  $P(x) = \text{Cat}(\phi)$  denotes a categorical distribution over some random variable  $x$  with parameters  $\phi$ . While this simplification assumes that posterior beliefs at subsequent timesteps are statistically independent, as we will see below, the Markovian temporal structure of the generative model means that, in practice, beliefs about hidden states at one timestep are contextualized by empirical priors from past timesteps (posterior beliefs from earlier timesteps).

The full free energy for the POMDP generative model and the approximate posterior specified in (16) can be written as follows:

$$\mathcal{F}_{1:T} = \mathbb{E}_{Q(s_{1:T}, \pi)} [\ln Q(s_{1:T}, \pi) - \ln P(o_{1:T}, s_{1:T}, \pi)]$$

Equipped with the free energy, we can now derive update equations for hidden state estimation and policy inference that involve minimizing  $F_{1:T}$ .

### 2.7. State estimation

Under active inference, hidden state estimation is analogized to perception — this is achieved by optimizing the variational posterior  $Q(s_{1:T}|\pi)$  over hidden states, given policies. Because our approximate posterior and generative model are defined using categorical distributions, the problem of state estimation becomes minimizing free energy gradients of the form  $\frac{\partial \mathcal{F}}{\partial \mathbf{s}}$ , where  $\mathbf{s}$  are the parameters of the approximate posterior distribution over hidden states,  $Q(\mathbf{s}) = \text{Cat}(\mathbf{s})$ .

At each timestep, the agent can take advantage of the mean-field factorization of the posterior and the Markovian structure of the generative model to update only its beliefs about the current state of the world:  $Q(s_t)$ . The optimal posterior at timestep  $t$  is then found by finding the solution to  $Q(s_t)$  that minimizes the timestep-specific free energy  $\mathcal{F}_t$ :

$$\begin{aligned} \mathcal{F}_t &= \mathbb{E}_{Q(s_t)Q(\pi)}[\ln Q(s_t) - \ln P(o_t, s_t | s_{t-1}, \pi)] \\ \implies \frac{\partial \mathcal{F}_t}{\partial Q(s_t)} &= 0 \iff Q^*(s_t) = \sigma(\ln P(o_t | s_t) + \ln(P(s_t | s_{t-1}, u_{t-1})P(s_{t-1}))) \end{aligned} \quad (17)$$

This furnishes a simple belief update scheme for perception, where the optimal posterior  $Q^*(s_t)$  is a Bayesian integration of a likelihood term  $P(o_t | s_t)$  and a prior term  $P(s_t | s_{t-1}, u_{t-1})P(s_{t-1})$ .

Further details on the form of the approximate posterior and the derivation of the time-dependent free energy can be found in Appendix A.

### 2.8. Policy inference

Under active inference, policies  $\pi$  are also a latent variable of the generative model and thus must be inferred. Accordingly, planning and action also emerge as results of (approximate) Bayesian inference, where now the inference is achieved by optimizing a variational posterior over policies  $Q(\pi)$ .

The optimal posterior that minimizes the full variational free energy  $\mathcal{F}_{1:T}$  is found by taking the derivative of  $\mathcal{F}_{1:T}$  with respect to  $Q(\pi)$  and setting this gradient to 0, yielding the following free-energy-minimizing solution for  $Q(\pi)$ :

$$Q^*(\pi) = \underset{Q(\pi)}{\operatorname{argmin}} \mathcal{F} = \sigma(\ln P(\pi) - F(\pi)) \quad (18)$$

Therefore in the same way that state estimation or optimization of  $Q(s)$  in Equation (17) resembles a Bayesian average of a likelihood and a prior term, policy inference also becomes an average of the policy prior  $P(\pi)$  and the ‘evidence’ afforded to each policy, scored by  $F(\pi)$ . See Appendix A for a more detailed derivation of the optimal policy posterior  $Q^*(\pi)$ .

The crucial component in understanding the behavior of active inference agents lies in the specification of the policy prior,  $P(\pi)$ . Under the standard construct of active inference,<sup>1</sup> the probability of a policy is defined *a priori* to be proportional to the negative *expected free energy* of that policy:

$$P(\pi) = \sigma(-\mathbf{G}(\pi)) \quad (19)$$

<sup>1</sup> But see alternative derivations as in [123] and [124]

The expected free energy or EFE is denoted  $\mathbf{G}(\pi)$ , and measures the free energy *expected* under pursuit of a policy. This expected or predictive nature of the EFE is the crucial: although the standard free energy is typically a direct function of observations (and functional of beliefs), when evaluating the consequences of a policy in the future, observations are not known—therefore, the expected free energy must deal with predicted observations or predictive densities over observations. As we will see below, this counterfactual nature of the expected free energy is what endows action selection with inherently both goal-directed and information-seeking components.

The expected free energy is defined mathematically as:

$$\mathbf{G}(\pi) = D_{KL}[Q(s_{1:T}, \pi) \parallel \tilde{P}(o_{1:T}, s_{1:T}, \pi)] \quad (20)$$

where  $\tilde{P}$  represents a generative model ‘biased’ towards the preferences of the agent. We can write this biased generative model at a single timestep as  $\tilde{P}(o_\tau, s_\tau, \pi) = P(s_\tau | o_\tau) \tilde{P}(o_\tau)$ , where  $\tilde{P}(o_\tau)$  represents a ‘biased prior’ over observations. Given the factorization of the approximate posterior  $Q(s, \pi)$  over time as defined in (16), the EFE for a single timestep can also be defined as follows:

$$\begin{aligned} \mathbf{G}(\pi)_\tau &= D_{KL}[Q(s_\tau | \pi) \parallel \tilde{P}(o_\tau, s_\tau)] \\ &\approx -\underbrace{\mathbb{E}_{Q(o_\tau | \pi)}[D_{KL}[Q(s_\tau | o_\tau, \pi) \parallel Q(s_\tau | \pi)]]}_{\text{Epistemic Value}} - \underbrace{\mathbb{E}_{Q(o_\tau | \pi)}[\ln \tilde{P}(o_\tau)]}_{\text{Utility}} \end{aligned} \quad (21)$$

where the first term, the epistemic value, scores policies according to how much information observations  $o_\tau$  expected under that policy provide about hidden states. This term is expressed here as the divergence between the states predicted under a policy, with and without conditioning on observations. The second term represents the degree to which expected outcomes under a policy will align with the biased prior over observations in the generative model. Since the prior over policies *minimize* expected free energy, policies with thus favor states that resolve uncertainty (maximize epistemic value) and satisfy prior preferences (maximize utility).

Having specified the prior over policies in terms of the (negative) expected free energy, we can now rewrite Equation (18) by expanding the prior in terms of  $\mathbf{G}(\pi)$ :

$$Q^*(\pi) = \sigma(-\mathbf{G}(\pi) - F(\pi)) \quad (22)$$

Additionally, in extensions introduced in [73], one has the option of augmenting the prior over policies with a ‘baseline policy’ or ‘habit vector’  $P(\pi_0)$ , also referred to as the  $\mathbf{E}$  distribution. This means the full expression for the optimal posterior can be written as (expanding  $\ln P(\pi)$  as  $\ln P(\pi_0) - \mathbf{G}$ ):

$$Q^*(\pi) = \sigma(-\mathbf{G}(\pi) + \ln P(\pi_0) - F(\pi)) \quad (23)$$

We introduce this ‘habit vector’  $P(\pi_0)$  explicitly here, because it will be one of the parameters we explore in the multi-agent model. Note that in the sections on [Habit Learning](#) below, we reformulate the prior over policies in terms of two separate priors over *control states*, to disentangle the prior over policies that include particular Hashtag control states  $u^T$  from the prior over policies that are specific to neighbour-attendance control states  $u^{\text{Who}}$ .

## 2.9. Action Selection

Action selection results from sampling from the marginal posterior over actions, or ‘control states’. The marginal posterior over actions can be computed by marginalizing out the posterior probability of policies using the policy-to-control mapping  $P(u_t | \pi)$ :

$$Q(u_t) = \sum_{\pi} P(u_t|\pi)Q(\pi) \quad (24)$$

This marginalization is necessary because the mapping between policies and actions is not necessarily one-to-one: in the case of multi-timestep policies or multi-factor generative models, a particular control state  $u_t$  might be entailed by more than one policy. Therefore, this marginalization effectively computes the value of each action by summing together the posterior probabilities of all policies that include it. This entailment relation is encoded in the likelihood  $P(u_t|\pi)$ .

Once the posterior over control states  $Q(u_t)$  has been computed, an action  $a$  is simply sampled from this posterior marginal—this is then the action that the agent takes at timestep  $t$ :

$$a_t \sim Q(u_t) \quad (25)$$

### 2.10. Habit Learning

Under active inference, *learning* also emerges as a form of variational inference. However, this inference is not over hidden states, but rather over model parameters [73]. Such parameter inference is referred to as ‘learning’ because it is often assumed to occur on a fundamentally slower timescale than hidden state and policy inference. However, the update equations for model parameters follow the exact same principles as hidden state inference—namely, we optimize a variational posterior over model parameters  $Q(\phi)$  by minimizing the variational free energy  $\mathcal{F}$ .

In the current model, we use ‘habit learning’ as originally described in [73] to model the development of so-called ‘epistemic habits,’ or the tendency for an originally epistemically-motivated behavior to become habitually driven, mimicking the transfer from model-based to model-free learning in the context of behavioral conditioning [76,77]. Technically, habit-learning reduces to updating a variational posterior over the categorical vector  $\mathcal{E}^{\text{Who}}$ , which parameterises the prior over the neighbour-attendance control state  $u^{\text{Who}}$ .

Recall from the final section on that  $\mathcal{E}^{\text{Who}}$  is vector of categorical parameters whose prior probability is given as a Dirichlet distribution:

$$P(u_0^{\text{Who}} | \mathcal{E}^{\text{Who}}) = \mathbb{E}[\text{Dir}(\varepsilon)] \quad (26)$$

The Dirichlet distribution is a conjugate prior for categorical distributions, meaning that the resulting posterior will also be Dirichlet distributed. Motivated by this conjugacy, we can define a variational posterior over the ‘habits’  $Q(\mathcal{E}^{\text{Who}})$  parameterised by variational Dirichlet parameters  $\varepsilon$ . One then simply augments the generative model from Equation (3) with the prior over the categorical  $\mathcal{E}^{\text{Who}}$  parameters, which then allows one to define a new variational free energy, supplemented with the approximate posterior over  $\mathcal{E}^{\text{Who}}$ . Solving for the free-energy minimizing solution with respect to the variational Dirichlet parameters  $\varepsilon$  leads to the following fixed-point solution for  $Q(\mathcal{E}^{\text{Who}})$  [125]:

$$\begin{aligned} Q(\mathcal{E}^{\text{Who}}) &= \text{Dir}(\varepsilon) \\ \varepsilon^* &= \varepsilon + \eta \cdot Q(u_t^{\text{Who}}) \end{aligned} \quad (27)$$

where  $\eta$  is a so-called ‘learning rate’ and  $Q(u_t^{\text{Who}})$  are current posterior beliefs about  $u^{\text{Who}}$  controls states. In other words, agents will update their posterior over actions or ‘habit vector’ according to how often they attend to a particular neighbour, as measured by the probability of each  $u^{\text{Who}}$  action. In the current work, we eschew the usual ‘separation

of timescales' assumption used in learning simulations (e.g. in [73,126]) and update the posterior habit vector  $\epsilon$  at every timestep, i.e. after every action. This means that agents in this context simultaneously infer which neighbour to attend to, based on the prerogative to minimize expected free energy, while also incorporating a continuously-learned 'habit' based on the frequencies with which they attend to different neighbours.

### 2.11. Multi-agent simulations

Now that we've introduced the generative model used by single agents and the ensuing inference, action, and learning rules that each agent will use to update its beliefs over time, we proceed to describe the multi-agent simulation itself.

A single multi-agent opinion dynamics simulation consists of a group of  $N$  active inference agents, where in the current work  $N$  ranged from 12 – 30 agents. Each agent is equipped with the single generative model of opinion formation, as described in the previous sections.

At each timestep, all agents simultaneously 1) update their beliefs as a function of observations and then 2) take an action (i.e., selecting which Hashtag to tweet and which neighbour to attend to). Crucially, each agent's observations are a function of its own actions at the previous timestep, as well as the actions of a select set of neighbours at the previous timestep. Each agent has a fixed set of neighbours, where the particular neighbours are determined by a randomly-chosen network topology. In the current study, we set the neighbour-to-neighbour connectivity for all simulations using Erdős-Rényi (ER) networks with some connection parameter  $p$ , meaning that agents are connected with fixed probability  $p$  [127]. For the current purposes, we make these networks undirected or symmetric, so that any agents that share an edge can both observe each other's tweet actions and choose to read each other's tweets. The components of each agent's generative model (i.e., the number of observation modalities, number of hidden state factors) is a function of its local connectivity and the number of neighbours that it has. For example, a random agent in the network that was initialized to have 3 other neighbours will have 3 hidden state factors corresponding to the 'meta-beliefs' of these three neighbours:  $s^{\text{MB1}}, s^{\text{MB2}}, s^{\text{MB3}}$  as well as three observation modalities that it will use to read each of those neighbours' tweets:  $o^{\text{NT1}}, o^{\text{NT2}}, o^{\text{NT3}}$ . Each of those neighbouring active inference agents' actions (which Hashtag they tweet) will thus feed into the focal agent's various **Neighbour Tweet** modalities at every timestep. Because edges are bidirectional, each of the neighbouring agents themselves will have a hidden state factor and observation modality in their respective generative models, that represent the beliefs and **Tweet Hashtag X** actions of the focal agent.

In the results section to follow, we investigate the opinion dynamics under active inference by testing the hypotheses stated in the **Hypotheses** section. We do this by systematically varying both the network connectivity  $p$  as well as the parameters of individual generative models, in an effort to investigate the extent to which 'epistemic communities' depend on both network properties as well as the cognitive features of individuals.

## 3. Model parameterisation

### 3.1. Fixed parameters

It is worth mentioning the vast parameter space one encounters when simulating multi-agent active inference models. In the current work, each active inference agent is equipped with an entire POMDP generative model that contains hundreds of individual parameters (consider, for example, all the categorical parameters that comprise the observation model  $P(\mathbf{o}_\tau | \mathbf{s}_\tau)$ ). Importantly, this parameter explosion is exacerbated in the multi-agent setting, since not only does the number of total parameters scale simply in the size of the network  $N$ , but connections between agents render this scaling supra-linear in  $N$ , since each agent is equipped with  $K_i + 2$  hidden state factors and observation modalities, where  $K_i$  is the number of neighbours that agent  $i$  is connected to.

This means that the possible parameter space that one must explore in order to understand the behavior of the model is combinatorially explosive. To enable transparency and efficient parameter exploration, we employ several simplifications and low-dimensional parameterizations of every agent's generative model, which render the resulting space easier to explore.

First of all, we assume that every agent's observation model relating the tweet content of others to their beliefs has the same basic form. Recall from Equation (7) the 'Hashtag semantics' matrix  $\mathbf{h}$  that comprises observation model for the observation of neighbour  $k$ 's tweet content:  $P(o^{NTk} | s)$ , parameterised with a 'Hashtag reliability' parameter  $p_h$ . We fix this matrix to have the same parameter  $p_h = 0.73$  for all agents:

$$\mathbf{h} = \begin{bmatrix} 0.73 & 0.27 \\ 0.27 & 0.73 \end{bmatrix} \quad (28)$$

A focal agent believes that if it sees some neighbour  $k$  tweeting **Hashtag 1**, then the likelihood that neighbour  $k$  believes in **Idea 1** is 73%, and the likelihood that they believe in **Idea 2** is 27%. The relationship is inverted in case the focal agent sees neighbour  $k$  tweeting **Hashtag 2**. In the current study we assume this basic Hashtag semantics matrix in Equation (28) is common to all agents, and for all neighbours (relative to some focal agent). This enables us to selectively explore the effect of *epistemic confirmation bias*, a single (scalar) precision  $\gamma$  that can be used to up- or down-weight columns of the Hashtag semantics matrix, according to whether a given neighbouring agent's belief aligns with (the focal agents belief about) the environmental hidden state factor  $s^{Idea}$  (see the section on **Neighbour Tweet Likelihood** for a more detailed explanation).

Another restriction is in space of network architectures we explore; for the present study, we constrain the connectivity to be defined by random graphs (also known as Erdős-Rényi or ER networks), that are characterized by two parameters: the network size  $N$  and the connectivity  $p$ . We render the simulations computationally tractable by exploring small networks (in the range of  $N = 12 - 30$  agents) while systematically varying the connection probability  $p$ . We also assume that all agents' transition models (those for both the environmental hidden state factor  $s^{Idea}$  and meta-belief factors  $s^{MB}$ ) are a scaled version of the  $2 \times 2$  identity matrix  $I_2$ . This further enables their systematic exploration in terms of single scalar (the precision), rather than exploring all possible parameterizations of  $2 \times 2$  transition matrices. In addition, while we systematically explore the inverse volatility parameter  $\omega^{Soc}$  and epistemic confirmation bias precision  $\gamma$ , we fix the value of  $\omega^{Idea}$  to be 9.0 for all simulations. We leave the full combinatorial exploration of all parameters, including  $\omega^{Idea}$ , to future work.

Finally, while parametrically exploring the dependence of collective outcomes on individual parameters, we usually restricted parameter sweeps to vary at most two parameters at a time. We did this in order to simulate a sufficient number of trials for each condition while also investigating each parameter with as fine a resolution as possible. Under both these constraints, the computation time would explode when varying more than just 2 parameters simultaneously, so we fix the values of the non-varied parameters to limit computational burden (e.g. fix  $\omega^{Soc}$  while varying  $\gamma$  and  $\eta$ ). In practice, we clamped the value of the fixed parameters to 'insensitive' regions of parameter space where we know that the collective measure of interest (e.g. polarisation) didn't depend on small changes in that parameter.

### 3.2. Parameters of interest

In the following results section we describe four sets of parameters that we systematically varied to investigate their role in determining emergent phenomena in the multi-agent simulations. Below we briefly step through each parameter and rehearse its interpretation, and our motivation for investigating it.

### Epistemic Confirmation Bias

Recall from the section **Neighbour Tweet Likelihood** that *epistemic confirmation bias* or ECB is a precision parameter  $\gamma$  that selectively scales the Hashtag semantics matrix of the agent's observation model, linking  $s^{\text{Idea}}$  and  $s^{\text{MB}k}$  to  $o^{\text{NT}k}$ . The ECB precision  $\gamma$  scales the Hashtag semantics matrix in such a way that some focal agent  $i$  receives evidence for the  $s^{\text{Idea}}$  hidden state factor's value (**Idea 1** vs. **Idea 2**) from the tweet output of some neighbour  $k$ , in proportion to how much neighbour  $k$  agrees with agent  $i$ .

This means that a focal agent with a higher  $\gamma$  believes that tweets more reliable if they come from neighbouring agents that are believed to share the opinion of the focal agent. The consequence of this is an ironically-named 'epistemic' sort of confirmation bias, where agents believe more reliable information about  $s^{\text{Idea}}$  comes from neighbours who are believed to be 'like-minded' to themselves. This can be revealed by recalling the expected free energy, the key determinant in action selection under active inference. As decomposed in Equation (21), this comprises an information gain term and a utility term. By means of the ECB parameter, the epistemic value term is preferentially higher for those actions that entail attending to a neighbour who the focal believes is like-minded. This can be analyzed more quantitatively by inspecting the 'negative ambiguity' term of the epistemic value,  $\mathcal{H}$ , which we show to be directly proportional to epistemic confirmation bias:

$$\mathcal{H} = \mathbb{E}_{Q(s^{\text{Idea}}, s^{\text{MB}k})} \left[ \frac{1}{C} [p_h^\gamma \log p_h^\gamma + (1 - p_h)^\gamma \log(1 - p_h)^\gamma - 2 \log C] + [p_h \log p_h + (1 - p_h) \log(1 - p_h)] \right]$$

See Appendix B for a complete derivation of the relationship between  $\gamma$  and epistemic value.

Given this relationship, we expect that higher epistemic confirmation bias will drive agents to preferentially attend to the actions of agents that share their beliefs. On a collective level, we hypothesize that ECB will increase the probability of both polarisation (two clusters of oppositely-minded agents) and consensus (all agents have the same or similar beliefs about the **Idea**).

### Inverse Social Volatility

Recall the inverse temperature parameter introduced in Section (2.4), where we parameterise a focal agent's beliefs about the stochasticity of the social dynamics using precision parameters  $\omega^{\text{Soc}}$  (following the notation used in [85]). The inverse social volatility scales the transition model that describes the dynamics of  $s^{\text{MB}k}$ , such that a higher  $\omega^{\text{Soc}}$  induces an assumption of less stochasticity in the belief evolution of neighbours' 'meta-beliefs.' This relationship also implies that the inverse social volatility is related to the epistemic value of actions that involve attending to particular neighbours. In particular, higher volatility (i.e. more entropy in the columns of the  $\mathbf{B}^{\text{MB}k}$  matrices) leads to higher overall uncertainty in beliefs about hidden states. In other words, for lower values of  $\omega^{\text{Soc}}$  the uncertainty of the posterior marginal  $Q(s^{\text{MB}k})$  will accumulate faster, as long as the focal agent isn't attending to neighbour  $k$ . Actions that entail attending to these unattended neighbours will therefore grow in epistemic value, the more time elapses that those neighbours remain unattended. Importantly, the growth in epistemic value will scale inversely with  $\omega^{\text{Soc}}$  (see Appendix B for details). This means that the particular value of  $\omega^{\text{Soc}}$  sets an effective 'refresh rate' for how often a neighbour should be re-attended to, in order to resolve uncertainty about their beliefs.

Given this relationship, we hypothesize that high 'meta-belief' volatility (low  $\omega^{\text{Soc}}$ ) will lead agents to re-read their neighbours' tweet content with a higher rate - whether or not they (believe they) agree with them, in order to resolve uncertainty about their beliefs. We expect that this continuous, epistemically-driven 're-sampling' will counteract the tendency of the group to polarise and thus favor collective agreement or consensus. An interesting question will be whether the inverse social volatility parameters directly

'reverses' the effect of  $\gamma$ , where the two jointly determine a collective trade-off between consensus and polarisation.

### Learning Rate

The learning rate  $\eta$  associated with updating the habit vector over neighbour-attendance control states  $u^{\text{Who}}$  represents the degree to which agents will preferentially sample those neighbours that they've attended in the past. In the presence of a higher learning rate, the Dirichlet hyperparameters over the habit vector  $\mathcal{E}^{\text{Who}}$  will be "bumped up" by a larger amount after choosing to attend to any particular agent, such that a focal agent will form preferences to attend to those agents whose Hashtags they habitually read. We expect therefore that a higher value of  $\eta$  will lead to increasingly-preferential neighbour-attendance patterns among agents, and eventually to a change in the overall collective belief distribution of the group. Specifically, we hypothesize that 'echo-chamber' like dynamics will be exacerbated by a higher value of  $\eta$ , such that it will be harder to 'escape' from polarised dynamics in the presence of a large habit-learning rate  $\eta$ .

### Network connectivity

In addition to individual generative model parameters like  $\gamma$ ,  $\omega^{\text{Soc}}$  and  $\eta$ , we also quantitatively investigate whether and how the topology of agent-to-agent communication determines emergent behaviour. To quantitatively investigate this using a simple, 1-dimensional parameterization, we initialized the agent-to-agent communication network (i.e. which agents can read which other agents' Hashtags) using a fixed random graph with connection probability  $p$ . For random graphs,  $p$  encodes the probability that any two agents have an edge between them. In the current context, an edge between any two agents determines whether they can view each other's Hashtags, and thus form beliefs about one another's beliefs. We hypothesize that denser communication topologies, represented by random graphs with increasing connection probability  $p$ , will obviate the risk of polarization and lead to consensus with higher probability. In investigate this network effect, we also hope to reveal interactions between  $\gamma$  (which we hypothesize will *induce* polarization) and connection probability  $p$ .

In the following sections, we describe the results of numerical experiments wherein we systematically vary the parameters discussed above, and reveal how they modulate the collective formation of 'epistemic communities' (e.g. echo-chambers, polarization, consensus).

## 4. Results

In the following sections we summarize the results of numerical experiments that validate the basic dynamics of the opinion formation generative model and then systematically investigate each of our three hypotheses. The results sections are organized as follows:

First, we demonstrate the basic dynamics of an active inference agent engaged in opinion formation. Over time, we show how a single focal agent updates its beliefs about the world in the face of conflicting Hashtag observations from two neighbours. In this process, the agent simultaneously forms beliefs about the abstract, environmental hidden state (**Idea 1** vs. **Idea 2**) as well as beliefs about the meta-beliefs of two neighbouring agents, whose Hashtags it is exposed to. We examine the dependence of a single agent's belief-updating dynamics depend on different settings of the epistemic confirmation bias  $\gamma$  and the inverse social volatility  $\omega^{\text{Soc}}$ , under a fixed value of  $\omega^{\text{Idea}} = 9.0$ .

Next, we demonstrate the emergent formation of epistemic communities and the diverse dynamics that can observed under the current active inference model. These are meant as proof-of-principle validation of the opinion dynamics model and the rich sorts of collective behaviours it can give rise to.

Finally, in order to test the three hypotheses that frame our study of epistemic communities under active inference, we systematic vary parameters like  $\gamma$ ,  $\omega^{\text{Soc}}$ ,  $\eta$ , and  $p$  to investigate how they determine collective dynamics. In these collective dynamics experi-

ments, we link groups of active inference agents together and simulate their multi-agent dynamics for up to  $T = 100$  timesteps. We then study collective outcomes by averaging the results of hundreds of independent realizations.

When systematically varying parameter configurations, we define a single condition as a combination of the parameters of interest. This includes the network connectivity  $p$  and a vector of generative model parameters, e.g.  $\gamma = 3.5, \omega^{\text{Soc}} = 0.5, \eta = 1.5$ . For each condition, we ran 100 independent multi-agent simulations with a network size  $N = 15$  agents. We chose relatively small networks in order to limit the computational burden of each simulation.

#### 4.1. Opinion formation in a single agent

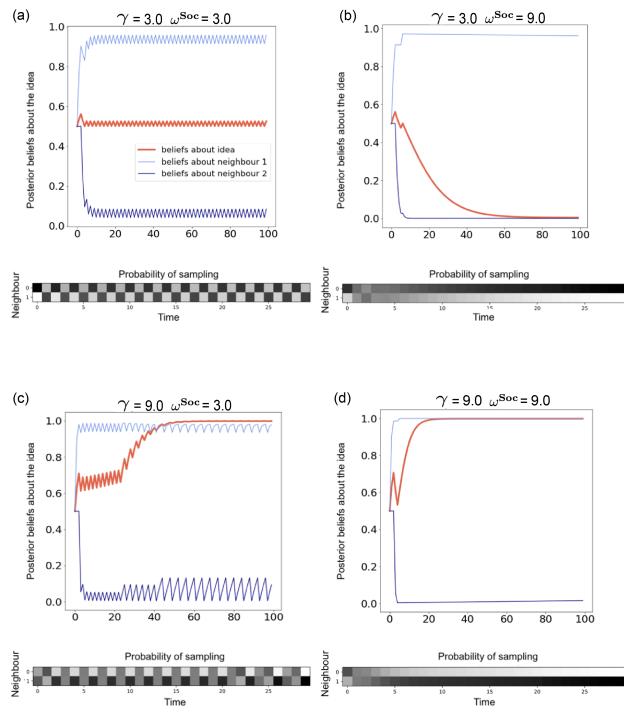
Figure 2 visualises opinion formation in a single active inference agent, and sheds light on the relationship between  $\omega^{\text{Soc}}$  and  $\gamma$  in determining the rate of belief updating and action selection. We investigate this using a simplified 3-agent set-up, where one focal agent is exposed to a sequence of conflicting information from two neighbours. At each timestep, the focal agent chooses to read a Hashtag from one of its two neighbours, and the two neighbours are not actually active inference agents, but are simply sources of a sequence of discrete Hashtag observations (**Hashtag 1** issue from neighbour 1, **Hashtag 2** issue from neighbour 2). We can see anecdotally how belief updating and sampling behaviour are bidirectionally modulated by different combinations of  $\omega^{\text{Soc}}$  and  $\gamma$ . In general, Figure 2 shows that beliefs in more meta-belief volatility (lower  $\omega^{\text{Soc}}$ ) lead to higher posterior uncertainty about the  $s^{\text{Idea}}$  hidden state, as is shown by the red lines in subplots (a) and (c). Higher epistemic confirmation bias  $\gamma$ , on the other hand, induces a positive feedback effect, wherein the focal agent comes to agree with one of its two neighbours with high certainty, most likely whichever neighbour it happens to attend to at the first timestep.

With high enough  $\gamma$  or high enough  $\omega^{\text{Soc}}$ , the focal agent's beliefs, faced with these two conflicting sources of information, converge to one Idea. This choice is consistently reinforced by the focal agent continuing to sample the agent it agrees with (lower insets in each subplot of Figure 2). There is also an interesting interaction between  $\gamma$  and  $\omega^{\text{Soc}}$ , such that  $\omega^{\text{Soc}}$  drives down posterior uncertainty in the focal agent's beliefs about its neighbour  $Q(s^{\text{MB}k})$ . This in turn decreases the information gain term in the expected free energy, such that the agent has stronger prior beliefs about its neighbour's beliefs and there is less information gain afforded to attending to that neighbour. On other hand, higher  $\gamma$  drives up epistemic value, even in the face of precise beliefs about the neighbour's belief state, making the agent expect to artificially resolve more uncertainty from its observations.

It is clear that for configurations with high inverse social volatility, as the focal agent's beliefs converge toward the beliefs of neighbour 1, it also begins to attend to neighbour 1 more often than neighbour 2 (subplot (d)). However, with low inverse social volatility, the focal agent is driven to periodically attend to both neighbours, due to the increasing epistemic value associated with un-attended neighbours. Interestingly, when  $\omega^{\text{Soc}}$  is low and  $\gamma$  is high (Figure 2(c)), the focal agent continues to periodically re-attend to the neighbour it disagrees with, due to increasing uncertainty about that neighbour's belief, induced by high volatility associated to it. Note however that the total probability of attending to the like-minded neighbour is still higher due to the presence of high epistemic confirmation bias. In the presence of both low epistemic confirmation bias and low inverse social volatility, posterior uncertainty is high all-around and the focal agent is 'ambivalent' between both **Idea 1** and **Idea 2**. Nonetheless, the focal agent succeeds in inferring the belief-states of its two neighbours as it repeatedly alternates between sampling them.

#### 4.2. Epistemic community dynamics

Figure 3 shows examples of the collective opinion dynamics (i.e. 'epistemic communities') that emerge when simulating networks of active inference. Unlike in Figure 2, in these simulations the observations for every agent are generated by the actions of other active inference agents, who are all collectively reading the Hashtag actions of other agents



**Figure 2.** Belief dynamics and actions of a single agent in response to a sequence of Hashtag observations from two fictive neighbours. Shown are the history of Bernoulli parameters defining three marginal posterior beliefs of the focal agent: the belief about the truth value of **Idea 1** ( $Q(s_t^{\text{Idea 1}} = \text{Idea 1})$ , in red), and its beliefs about the beliefs of its two neighbours regarding **Idea 1** ( $Q(s_t^{\text{MB1}} = \text{Idea 1})$  and  $Q(s_t^{\text{MB2}} = \text{Idea 1})$ , shown in two shades of blue). Through its generative model, the focal agent believes its Hashtag observations are caused by two neighbour ‘meta-belief’ states. The focal agent is exposed to a sequence of Hashtag observations for 100 timesteps, where in case of attending to the first neighbour ( $u_t^{\text{Who}} = 0$ ), the agent receives observation  $o_t^{\text{NT1}} = \text{Hashtag 1}$ ,  $o_t^{\text{NT2}} = \text{Null}$ , and in case of sampling the other neighbour ( $u_t^{\text{Who}} = 1$ ), the agent receives observation  $o_t^{\text{NT1}} = \text{Null}$ ,  $o_t^{\text{NT2}} = \text{Hashtag 2}$ . Due to the ‘Hashtag semantics’ matrix in its generative model, these two Hashtags respectively lend evidence for the two levels of  $s^{\text{Idea}}$ . At each timestep the focal agent performs inference with respect to hidden states  $Q(s_t)$  as well as policies (control states)  $Q(u_t)$ , and then samples a **Neighbour Attendance** action from the posterior over control states  $Q(u^{\text{Who}} = 0, u^{\text{Who}} = 1)$ . Below each subplot is a heatmap showing the temporal evolution of the probability of sampling neighbour 1 vs neighbour 2 over time.

while and generating their own. We include this to showcase the rich phenomenology displayed by collectives of active inference agents, validating our model alongside known opinion dynamics models that can capture phenomena like consensus and polarisation. In the following sections we investigate the dependence of these dynamics on the parameters of generative models and network density quantitatively.

#### 4.3. The dependence of epistemic communities on $\gamma$ and $p$

We first investigated Hypothesis 1, or how epistemic confirmation bias  $\gamma$  and network connectivity  $p$  determine collective formation of epistemic communities. We systematically varied both epistemic confirmation bias (15 values of  $\gamma$  tiling the range [3, 9]) and network connectivity (15 values of  $p$  tiling the range [0.2, 0.8]) in networks of  $N = 15$  agents, and simulated  $S = 100$  independent realisations of each condition for  $T = 100$  timesteps. Other parameters were fixed to constant values ( $\omega^{\text{Soc}} = 6.0$ ,  $\omega^{\text{Idea}} = 9.0$ ,  $\eta = 0.0$ ). Note that here, habit-learning was intentionally disabled ( $\eta = 0.0$ ) to selectively investigate the effect of  $\gamma$  while excluding the effect of habit learning on epistemic community formation. Within

each parameter configuration, every independent realisation and every agent had the same average value of epistemic confirmation bias  $\gamma$ , but for each agent, we sampled a vector of epistemic confirmation bias values from a normal distribution centered at the parameter setting with variance 1.0. Note that there are  $k$  different ECB parameters per agent because each agent has a collection of  $A^{NT^k}$  arrays, each corresponding to the observation model from a particular neighbour. Each of these  $k$  likelihood arrays is parameterised by a single  $\gamma$ . By sampling  $\gamma$  across  $A^{NT^k}$  arrays within each agent's generative model, we implicitly gave each agents a particular bias to believe that certain neighbours were more 'reliable' than others — some neighbours contribute more or less to the focal agent's confirmation bias tendency.

The aim was to investigate how higher epistemic confirmation bias, particularly in a sparse network, might drive the emergence of epistemic communities through the formation of belief clusters that are both dense and far apart in 'belief-space.' In general, it is known in the literature that clusters are more easily formed in sparsely connected networks, but less so in densely connected networks where all agents communicate with each other [95]. Therefore, one interesting hypothesis for this experiment was that increasing the value of  $\gamma$  could achieve the opposite effect: namely, a high degree of polarisation or belief-clustering behaviour in a densely connected network.

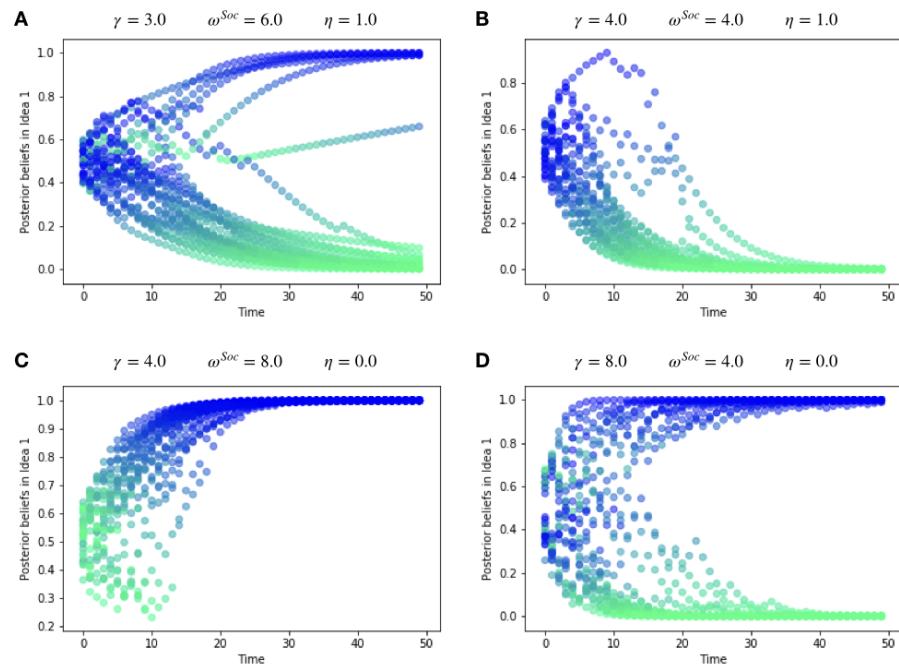
To assess the emergence of epistemic communities or clusters of like-minded individuals, we defined the polarisation index  $\rho$ , which measures the degree of 'epistemic spread' in a system. It is defined as the difference between the highest and the lowest values of the Bernoulli parameter defining  $Q(s_i^{\text{Idea}} = \text{Idea 1})$  across all agents at the final timestep of the simulation (where the choice of one 'side' of the belief  $Q(s_i^{\text{Idea}} = \text{Idea 1})$  is arbitrary). This final difference is then averaged across  $S$  independent realizations or trials to give the average value  $\langle \rho \rangle$  for a particular condition. This is directly proportional to the ratio of the number of trials in any configuration in which the simulation ends with two opposing clusters, as opposed to consensus, where consensus is defined at the final timestep when all agents' posterior beliefs about  $s_i^{\text{Idea}}$  are on the same side of 0.5.

$$\rho_s = \max_i [Q(s_i^{\text{Idea}} = \text{Idea 1}) - \min_i Q(s_i^{\text{Idea}} = \text{Idea 1})] |_{T=100} \in [0, 1]$$

$$\langle \rho \rangle = \frac{1}{S} \sum_{s=1}^S \rho_s \quad (29)$$

where  $S$  indicates the number of total trials (here,  $S = 100$ ).

A high value of  $\rho$  (close to 1) indicates more spread out beliefs and implies clustering, i.e. echo-chamber formation, whereas a low  $\rho$  implies that the network of agents have similar beliefs about  $s_i^{\text{Idea}}$  (i.e. consensus).

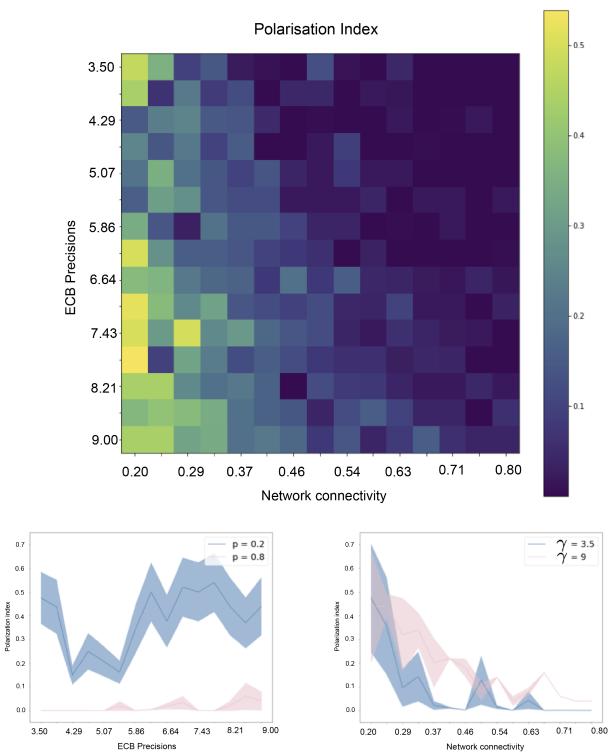


**Figure 3.** Collective belief dynamics of multi-agent active inference simulations under different generative model parameterizations. Above each panel are listed the parameter values of  $\gamma$ ,  $\omega^{\text{Soc}}$ , and  $\eta$  used in the simulation. Other parameters were fixed with  $T = 50$  timesteps,  $N = 30$ , network connectivity  $p = 0.2$ , and inverse environmental volatility  $\omega^{\text{Idea}} = 9.0$ . At the beginning of each simulation, every agent's beliefs about **Idea 1** were sampled from a uniform distribution over the interval  $Q(s^{\text{Idea}} = \text{Idea 1}) \in [0.4, 0.6]$ . Each panel displays the evolving beliefs of all agents about **Idea 1** (the Bernoulli parameter of each agent's respective posterior over  $s^{\text{Idea}}$ ), with proximity of the belief to 1.0 indicated by coloring along the green-to-blue spectrum (blue beliefs indicate  $Q(s^{\text{Idea}} = \text{Idea 1}) > 0.5$ ). Panels A and D demonstrate polarisation, where two subsets of agents end up believing in two different levels of the **Idea** hidden state with high certainty. Panels B and C on the other hand show examples of consensus, where the whole network converges to the same opinion by the end of the simulation.

Figure 4 shows the effects of varying  $\gamma$  and  $p$  on polarisation as measured by  $\langle \rho \rangle$ . It is clear from the first column of the heatmap that highly spread out beliefs can occur at all values of the epistemic confirmation bias in the presence of sparse connectivity. Denser networks in general reduce the risk of polarisation as seen by a drop-off in  $\langle \rho \rangle$  as  $p$  increases. However, epistemic confirmation bias can 'counteract' this effect to some extent by marginally bumping up the risk of polarisation, even in the presence of denser networks (high  $\gamma$  and high  $p$ ). The lower subplots of Figure 4 demonstrate this counteractive effect, where even at high connectivities (e.g.  $p = 0.8$ ) the epistemic confirmation bias can lead to a fraction of trials where polarisation is more likely.

Why, one might wonder, does polarisation still occur with some probability even when  $\gamma$  is small? When network connections are sparse, polarisation can still occur by virtue of the agents lacking access to a variety of neighbours — this forces them to attend to one of a limited set of neighbours that they start out connected to. Since all agents are initialised with flat prior beliefs about  $s^{\text{Idea}}$ , this leads to the formation of two clusters, since there is nothing correlating the beliefs of agents who are disconnected. Because there are two beliefs (**Idea 1** and **Idea 2**), this means that on average this fragmentation leads to distinct sub-clusters of connected agents that will believe in one of the two Ideas with approximately 50% probability.

As  $\gamma$  increases, even in the presence of increasing connectivity, agents are driven by epistemic value to preferentially attend to the neighbours that (they believe) share their



**Figure 4.** The interaction between epistemic confirmation bias and network connectivity in determining collective outcomes. Top: a heatmap of the mean polarisation index across  $S = 100$  independent realisations of the multi-agent opinion dynamics simulations, for unique combinations of of network connectivity  $p$  and epistemic confirmation bias precision  $\gamma$ . Bottom: Selected line plots show extreme settings of  $p$  (bottom left,  $p = 0.2$  and  $p = 0.8$ ) and  $\gamma$  ( $\gamma = 3.5$  and  $\gamma = 9.0$ ). Shaded areas around each line represent the standard deviation of the polarisation index across independent realisations.

beliefs. This accounts for the slower decrease in polarisation with increasing connectivity  $p$  at higher levels of  $\gamma$  shown in Figure 4. This can be compared to the faster decrease in polarization induced by  $p$  when  $\gamma$  is low (compare the first few rows of the heatmap in Figure 4 to the last few rows).

However, network connectivity seems to be a stronger effect than  $\gamma$  on enforcing consensus or at least the lack of polarization. This is because the exploration entailed by  $\gamma$  encourages agents to attend to a larger group of neighbours, leading to a higher average spread of beliefs and the ability for agents to serendipitously encounter other agents they agree with. However, because of the density of the network, it is much more difficult for agents to become polarised as they will more frequently be exposed to new information, despite their propensity towards confirmation bias.

#### 4.4. Effect of inverse social volatility on neighbour attendance and polarisation

Next, we explored Hypothesis 2, modelling behaviour under different values of inverse social volatility  $\omega^{\text{Soc}}$  to see how it would interact with  $\gamma$ . We swept over  $\gamma$  (15 values tiling the range [3, 9]) and  $\omega^{\text{Soc}}$  (15 values tiling the range [0.0, 0.8]) in networks of  $N = 15$  agents with  $p = 0.4$  connection probability. As before, each agent was equipped with a vector of  $k$  distinct  $\gamma$  and  $\omega^{\text{Soc}}$  parameters, which were sampled from a normal distribution centered around the parameter value characterising the condition. In this case, each sampled value parameterised the different neighbour-specific observation ( $A^{\text{NT}k}$ ) and transition models ( $B^{\text{MB}k}$ ) for a particular focal agent.

To assess the extent to which social attendance changes as a function of  $\gamma$  and  $\omega^{\text{Soc}}$ , we defined the re-attendance rate  $r$ . It scores the maximum number of times an agent samples the same neighbour throughout a parameter configuration, averaged over trials.

$$r_s = \max_i \sum_t \mathbf{1}_i(u_t^{\text{Who}})$$

$$\langle r \rangle = \frac{1}{S} \sum_{s=1}^S r_s$$

where  $\mathbf{1}$  is the indicator function.

We measured the re-attendance rate and polarisation index for each configuration, averaged across trials. Figure 5 portrays a complex picture on the relationship between  $\gamma$  and  $\omega^{\text{Soc}}$ . In the case of high volatility over meta-beliefs (low inverse social volatility), agents are driven to periodically re-attend to neighbours in order to resolve growing uncertainty about their beliefs. This is indicated by a higher average re-attendance rate  $\langle r \rangle$  (top right heatmap). Interestingly, there seems to be an interaction between re-attendance rate and epistemic confirmation bias, such that in the presence of both high volatility and low epistemic confirmation bias, the re-attendance rate is maximized. We speculate that the absence of ECB ( $\gamma = 0$ ) makes the epistemic value of attending to every neighbour equally high, and driven purely by  $\omega^{\text{Soc}}$ . In this case, agents will continually revisit neighbours sequentially, with the attendance-preference for any given neighbour solely dependent on the time elapsed since the last reading their Hashtag observation. In the absence of confirmation bias (which normally accelerates the the focal agent's beliefs not only about  $s^{\text{Idea}}$ , but also about  $s^{\text{MB}k}$ , c.f. Figure 2), this means that uncertainty about neighbours' beliefs will on average be higher. This will lead to diverse social attendance patterns, such that agents will prefer to constantly sample new neighbours, with no particular neighbour left out of this uncertainty-driven re-sampling.

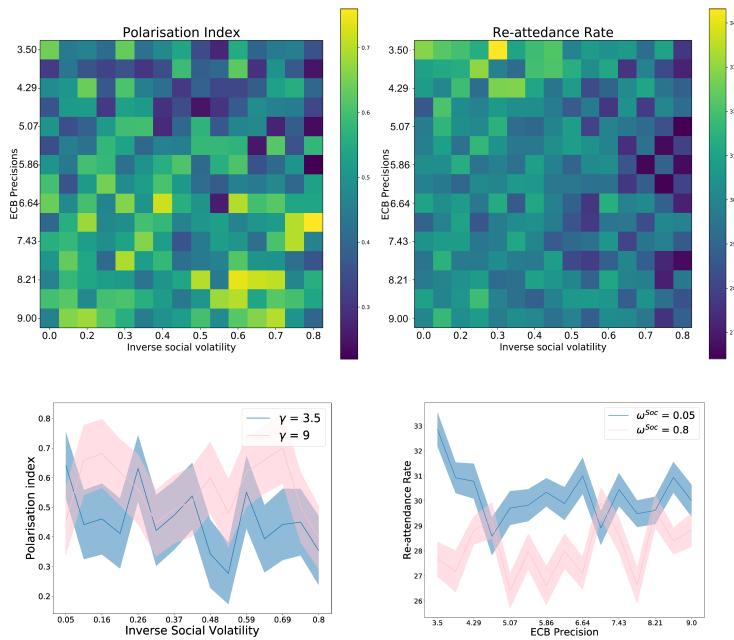
In terms of polarisation, it is not clear from the results that more volatility creates less polarised networks. However, in the presence of high ECB, the heatmap shown in the top left of Figure 5 suggests that increased volatility (low  $\omega^{\text{Soc}}$ ) may 'protect' the network from increased polarisation, which is maximised in the case of both high ECB and low volatility (high  $\omega^{\text{Soc}}$ ). We originally hypothesised that if agents are uncertain about the beliefs of their neighbours (low  $\omega^{\text{Soc}}$ ), it will become *more* difficult to induce polarisation and purposefully sample those who are thought to agree, due to the competing epistemic value induced by high volatility. So we seem to mildly observe this effect, but interestingly only in the case of high  $\gamma$ , i.e. low  $\omega^{\text{Soc}}$  induces a 'protective' effect in the case that polarisation is a likely outcome due to increased  $\gamma$ . A more robust effect is how social volatility induces the the tendency to re-attend neighbours (higher  $\langle r \rangle$ ).

#### 4.5. Habit formation and network initialization

For the final experiment, we explored Hypothesis 3, regarding the polarisation of networks via habit formation. We swept over  $\gamma$  (15 values tiling the range  $[3, 9]$ ) and  $\eta$  (15 values tiling the range  $[0.0, 0.9]$ ) in networks of  $N = 15$  agents with  $p = 0.4$  connection probability, where  $\omega^{\text{Soc}} = 6.0$  and as before  $\omega^{\text{Idea}} = 9.0$ . Here,  $\gamma$  was again normally distributed with a fixed mean (which varied by condition) and variance 1.0 across the  $k$  neighbours of each focal agent, but the learning rate  $\eta$  was fixed to the condition-dependent value across all trials and agents.

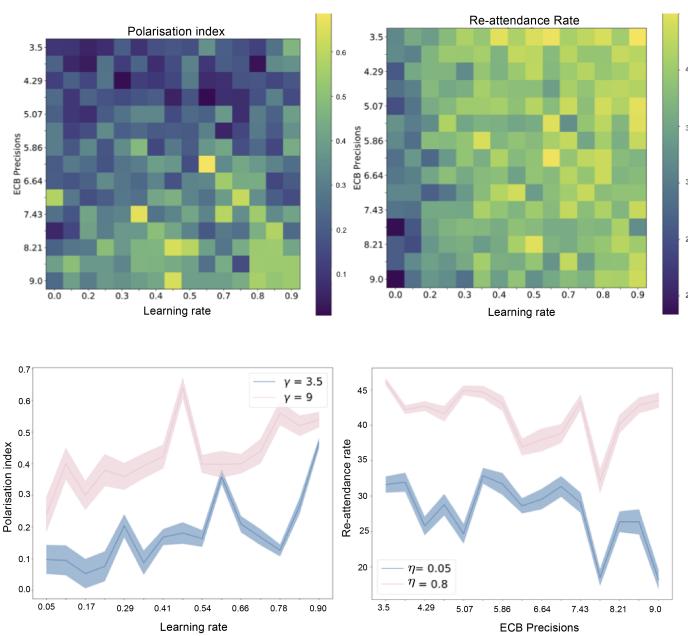
The learning rate  $\eta$  incentivises agents to re-attend to the same neighbour by forming a habit, which competes with the epistemic value of attending a new neighbour with unknown beliefs. This experiment tested the hypothesis that a higher learning rate, i.e. stronger habit-formation, will increase polarisation.

Figure 6 demonstrates how learning rate  $\eta$  and epistemic confirmation bias  $\gamma$  interact to influence outcomes at the collective level. Indeed, a higher learning rate induces more



**Figure 5.** Above left: a heatmap of the polarisation index for all 225 combinations of inverse belief volatility and epistemic confirmation bias precision. Above right: a heatmap of the re-attendance rate for all 225 combinations of inverse belief volatility and epistemic confirmation bias precision. Below left: a line plot of the most extreme rows of the polarisation heatmap. Below right: a line plot of the most extreme columns of the re-attendance rate heatmap.

polarisation, implying the formation of more ‘stubborn’ epistemic communities in the network. This effect appears at both low and high levels of epistemic confirmation bias, with on average a higher  $\langle \rho \rangle$  observed with increasing learning rate, even at low levels of  $\gamma$ . However, it seems the effect is most pronounced at the highest levels of  $\gamma$  and  $\eta$ . Examining the average re-attendance  $\langle r \rangle$  (right column of Figure 6) reveals a clear effect of  $\eta$  on neighbour re-attendance, with the rate seemingly maximized when the learning rate surpasses a value of  $\eta \approx 0.3$ . Interestingly, the effect of ECB on re-attendance is not very strong here, although it seems to have a mild negative effect. Namely, as ECB increases, the re-attendance rate tends to decrease. One counterintuitive explanation for this effect (which is similar to the effect observed in Figure 5) is the general increase in epistemic value of attending neighbours with unknown beliefs that is caused by increasing  $\gamma$ . Although by design  $\gamma$  is intended to ‘boost’ the epistemic value of only those actions that involve attending to neighbours that the focal agent believes it agrees with, there is still an overall ‘exploration bonus’ that scales with  $\gamma$ , *even for actions that entail attending to neighbours with whom the focal agent disagrees*. This is because in addition to the ambiguity term of the epistemic value, which captures the ‘confirmation bias’ effect encoded by  $\gamma$ , there is also a maximum-entropy component  $H[Q(\mathbf{o}_\tau | \pi)]$  (see Appendix B for details). This term is maximized when the posterior uncertainty over meta-beliefs  $Q(s^{\text{MB}})$  is high (maximal when  $(s^{\text{MB}}) = [0.5, 0.5]$ ). So although ECB ‘bends’ the epistemic value landscape towards sampling like-minded neighbours (see Figure A7 in Appendix B for a visualisation of this effect), when compared to neighbours with differing beliefs, the inherently uncertainty-resolving nature of the epistemic value as a whole means that higher  $\gamma$  still increases the value of actions that involve attending to *any* neighbours whose beliefs the focal agent is uncertain about. This may in fact may counteract the polarising effects we originally intended to capture by including the ECB parameter. This across-the-board ‘exploration bonus’ conferred by ECB may explain the mild effect we observe here, where increasing  $\gamma$  ends up decreasing average re-attendance  $\langle r \rangle$ . This may indeed explain the decrease observed in both in Figures 5 and 6.



**Figure 6.** Above left: a heatmap of the polarisation index for all 225 combinations of learning rate and epistemic confirmation bias precision. Above right: a heatmap of the re-attendance rate for all 225 combinations of learning rate and epistemic confirmation bias precision. The parameters represent the centres of the Normal distribution sampled from across trials for each configuration. Below left: the most extreme row of the polarisation index heatmap. Below right: the most extreme column of the re-attendance rate heatmap.

## 5. Discussion

In this paper, we focused on the way communities form around shared beliefs about abstract entities or meanings, symbolized by an abstract discrete hidden state: an ‘Idea.’ Shared belief around a particular ‘Idea’ emerges through coordination, which itself is individually driven by the desire to form accurate (Bayesian beliefs) about the world and the beliefs of one’s community. In particular, we modeled confirmation bias as a ‘epistemic’ phenomenon, wherein agents have a biased belief that agents with whom they believe they agree are more likely to provide uncertainty-resolving (information-availing) data — hence the proposed terminology of *epistemic confirmation bias*.

Twitter provides fertile ground for the academic study of the spread of ideas. The platform is extremely popular, easy to access, and has an API that enables researchers to collect and analyze data. It has also been one of the major vectors for misinformation, leading to large scale events, like the tensions around the 2016 elections results [128], or the vaccine for SARS-CoV-19 [129]. With its effective network structure in terms of follower-, like-, and retweet-networks, Twitter provides an ideal environment for the empirical study of the spread of ideas.

The formation of echo-chambers has been well studied on Twitter and Facebook. Echo-chambers tend to reinforce like-mindedness in users, and tend as well to enable the crafting of a shared narrative [42]. [42] analysed the different ways in which different social media platforms’ algorithms influence the mechanisms of formation. They defined the echo-chambers based on the distributions of leanings towards polar attitudes. These attitude distributions were found to range from monomodal to bimodal or more complex. Regardless, polarisation is rarely neutral, and tends to favor opposition between extreme opinions. According to their results, Twitter and Facebook showed the most striking echo-chambers. Using virality models, they also measured information spread. In Twitter and Facebook, information was most likely to be spread to other users sharing similar leanings. Similar findings were shown by [130], by following the online debates surrounding vac-

cination hesitancy in Italy. Despite the formation of distinct echo-chambers, they found within echo-chamber community structure also differed between vaccine advocates and skeptics and influenced information flow. Findings like these and others on polarised social network dynamics inspired us to analyse the model explored in the current work to online digital social media like Twitter, as well as to study how network structures influence echo-chamber formation. Alongside this, we chose to embrace an underlying active inference model as a cognitively-inspired, Bayesian model for single agents' belief formation.

To formalize confirmation bias as a fundamentally Bayesian phenomenon, we constructed our generative model to include a precision parameter that we named *epistemic conformation bias* or ECB. Specifically, ECB confers a higher weight to information that comes from peers that the reader (focal agent) believes are like-minded. This in turns leads an agent with higher ECB to selectively sample information that justifies what they already believe. We were able to replicate the formation of epistemic communities *in silico*, e.g. echo-chambers, on social networks such as Twitter. This unique formulation of confirmation bias as an epistemic phenomenon helps explain how individuals continuously forage their environment for information, but may become stuck in a so-called 'bad bootstrap' that simply reinforces existing beliefs about the world, which in the face of new information may lead to sub-optimal behaviour [131].

In agreement with previous work studying the relationship between synchronisation and network structure, we found that opinion dynamics depends heavily on network density. Our formalism allowed us to systematically vary the parameters of individual agents (e.g. cognitive biases or beliefs) as well as collective properties like network structure. We found the density of inter-agent connections, parameterised by connection probability of random graphs, determined the transition between echo-chamber formation (polarisation) and consensus. However, we found that in the presence of high ECB, one could observe polarisation even in the presence of dense connectivity (c.f. Figure 4). This result seems counter-intuitive, as we might think that network clustering is a necessary condition for more polarisation. However, clearly defined clusters and group boundaries can sometimes act as buffers [132–136]. Sub-clusters exchanging information are likely to average towards their local center [137–139], which entails a form of opinion stability within the group. They are generally sheltered from other opinions since they cut ties to other agents which are not part of their group, and have been selected out [140]. However, in networks without clusters, opinions can have a high degree of volatility and reach very polar tendencies even without being entirely clustered. By means of epistemic confirmation bias, agents were likely to give more weight to information that was similar to their own, even in the presence of network neighbours with different opinions.

The clustering phenomenon is exacerbated by adding the capacity to form habits. Specifically, we allowed agents to increase their likelihood of resampling the same agents based on how often they attended to them in the past. Since neighbour-attendance is driven by epistemic value (resolving uncertainty about the  $s^{\text{Idea}}$  and  $s^{\text{MB}}$  hidden state factors), this tendency to revisit previously-sampled neighbours is a form of 'epistemic habit formation,' where actions that are initially undertaken based on information gain become solidified over time due to a Pavlovian, model-free mechanism that simply reinforces past behaviour. We found that in addition to ECB, the presence of habit formation exacerbated polarisation, presumably due to the formation of echo-chambers or tight communities of agents that only read only the Hashtag content of their like-minded peers. On the other hand, we found that beliefs about social volatility (represented by our  $\omega^{\text{Soc}}$  parameter) pushed the agents to sample their social environments more frequently and diversely, counteracting the effect of confirmation bias and habit formation in driving polarisation. We speculate that increased social volatility increases each agent's incentive to sample a diverse array of network neighbours, which in turns lessens their susceptibility to believing in one Idea with high certainty. In other words, increased social volatility (low  $\omega^{\text{Soc}}$ ) makes agents more 'curious' about the beliefs of (potentially non-like-minded) neighbours, which in

turns increases their exposure to conflicting information and ‘protects’ them from falling into one or another echo-chamber.

The contributing influence of beliefs about social volatility to exploratory social sampling leads us to consider the role of norms in social settings. If an agent is incentivised (via e.g. epistemic value or curiosity) to pay attention to neighbours whom they are uncertain about, their social group could be a source of constant surprise, as long as their beliefs about their neighbours are constantly fickle (“I’m not sure what members of my social group believe from one time to the next”). In other words, even in the presence of a group of like-minded peers, we would expect that increased beliefs about social volatility leads to repeated attendance of peers among one another, even if those peers all agree (and believe as much about each other).

## 6. Conclusion

Our simulation showcased a novel opinion dynamics model based on multi-agent active inference, and highlights many interesting possibilities for future research. We introduced a new parameter, the epistemic confirmation bias, which can modulate the formation of epistemic communities by changing epistemic value in a biased way, namely towards attending preferentially to like-minded agents. In addition to the ECB, we also showed the importance of other features like network structure and habit formation in contributing to polarised dynamics. However, there are several limitations to this work which warrant further discussion. While we systematised our study design to explore several parameters simultaneously, this search was not exhaustive and vast regions of parameter space remain unexplored. Particular parameters like size of the network remained unexplored (we mainly explored networks with size  $N = 30$ ), and for computational efficiency we restricted both the resolution and the combinatorics of the parameter combinations explored. In future work, we could leverage distributed computing or GPU-accelerated operations to explore both larger network sizes and parameter combinations. However, in model spaces with high enough dimension, computational acceleration alone will not suffice, so one could also reduce the sampled region of parameter space by leveraging efficient search techniques (e.g. optimal experimental design [141]) or higher-order learning methods such as Bayesian hyperparameter optimisation [142].

The generative model used by the single agents was also limited, in the sense that we only modelled beliefs in one of two mutually exclusive Ideas. Previous research into opinion or collective dynamics has shown that such binarity may strongly determine the dynamics of the system [105,143]. From a construct-validity standpoint, such binarity also vastly simplifies the semantic complexity found in real epistemic communities. For example, the semantic expression of a particular idea or claim heavily depends on the community in which it circulates. In future designs, we should strive to make the ideas more complex and more porous. By porosity we mean ‘semantic cross-over’, in the sense that multiple ideas may entail more or less similar behavioural consequences, or indeed entail the truth value of one another. This porosity may give rise to groups who believe in the same idea from an inference standpoint, but have a different interpretation of it. Starting from there, we can begin to envisage a specific semantic embedding which leads us to social scripts [144]. These conceptual embeddings would lead to two different conceptions with distinct causal relations to the environment. The weak conception of the script corresponds to an embedding, linking the observation of an event to the belief in a particular idea. The strong conception of the script leads to a sequencing of the beliefs, such as an entailment relation (e.g. ‘if I believe X, this entails a belief in Y’). This type of conceptual entailment possible under a strong conception of social scripts, combined with the ability to express one’s beliefs, could engender a capacity to act and coordinate through language with other actors.

Future work could explicitly model these entailment relationships among semantic entities by violating the typical independence assumption used to factorize the generative model’s hidden state factors — for instance, instead of having each hidden state factor

$s(f)$  being conditionally dependent on only other states/control states *within that factor*, we could 'mix' hidden state factors to make states of factor  $i$  depend on states of factor  $j$ .

Another notable feature to include is the variation of prior beliefs about different ideas or claims. In the current model, agents were often initialised to have uniformly distributed beliefs about  $s^{\text{Idea}}$  around the 'ambivalence' line of [0.5, 0.5]. Future studies could quantitatively investigate the dependence of epistemic community formation on the initial distribution of prior beliefs and how that distribution intersects with structural features such as network position (e.g. 'is a very confident agent more influential in determining information spread, when it's a peripheral vs. central node in the network?'). In this way, we could study 'historical effects' like whether pre-existing echo-chambers or belief distributions influence the susceptibility of the network to incoming information or environmental fluctuations.

In future studies, we hope to investigate individual cognitive differences more quantitatively using the active inference framework. Under active inference, 'individual differences' can be formalized as variance among the parameters of generative models across agents — e.g. different settings of the inverse volatility parameters for different agents. Another interesting possibility that is accommodated within the active inference framework is the idea that agents may *learn* the parameters of their generative models, as opposed to keeping them fixed over time. For example, one could imagine that the epistemic confirmation bias associated to a particular neighbour  $k$  could change over time as a function of the reliability of Hashtags observed by the focal agent. This is easily cast as another form of inference under the Bayesian framework. All one would need to do is define appropriate priors and approximate posteriors over  $\gamma$ , from which an additional free energy term and appropriate belief updating scheme could be derived. Learning the parameters may add ecological validity to the model as well; for example, agents might become accustomed to their social environment and seek out an epistemic community in order to increase the predictability of their sensory information, thus requiring them to sample their social environment less frequently. This is the kind of phenomenon that could be modelled by letting the inverse social volatilities  $\omega^{\text{Soc}}$  become free, learn-able parameters. With larger networks, we may be able to simulate the emergence of similar but distant sub-communities, which become epistemically similar without coming into direct contact, or only through very distant contact with one another. This leads us to the possibility of simulating the way epistemic and pragmatic practices become cemented, giving way to social meaning semantics and scripts, which seem to separate cultures. Simulating the emergence of similar semantics and scripts across different communities may help us further understand their common underlying processes. Finally, in future studies, we could model an explicit state of conformity, by modeling the agent's assumptions about the groups they can identify around themselves, and be driven to model their behavior after the group they feel most kinship to.

## 7. Additional Information

### 8. Funding Statement

CH is supported by the U.S. Office of Naval Research (N00014-19-1-2556).

### 9. Competing Interests

We have no competing interests.

### 10. Authors' Contributions

MA and DD generated the concept for the paper. CH, DD, and MA formulated the generative model for single agent opinion formation. RH and DD wrote the code for performing the multi-agent simulations. DD and RH performed the numerical simulations and generated the figures. MA, DD, and CH wrote the paper. MJD provided valuable feedback and edits on drafts of the paper.

## References

1. Henrich, J. Culture and social behavior. *Current opinion in behavioral sciences* **2015**, *3*, 84–89.
2. Heerwagen, J.H.; Orians, G.H. Humans, habitats. *Biophilia hypothesis* **1995**, *138*, 138–172.
3. Vasil, J.; Badcock, P.B.; Constant, A.; Friston, K.; Ramstead, M.J. A world unto itself: Human communication as active inference. *Frontiers in psychology* **2020**, *11*, 417.
4. Goldstein, D.G.; Gigerenzer, G. Models of ecological rationality: the recognition heuristic. *Psychological review* **2002**, *109*, 75.
5. Lau, R.R.; Redlawsk, D.P. Advantages and disadvantages of cognitive heuristics in political decision making. *American Journal of Political Science* **2001**, pp. 951–971.
6. Sherman, S.J.; Corty, E. Cognitive heuristics. **1984**.
7. Oswald, M.E.; Grosjean, S. Confirmation bias. *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory* **2004**, *79*.
8. Donsbach, W. Cognitive dissonance theory—A roller coaster career: How communication research adapted the theory of cognitive dissonance. In *Media choice*; Routledge, 2009; pp. 142–162.
9. Knobloch-Westerwick, S.; Mothes, C.; Polavin, N. Confirmation bias, ingroup bias, and negativity bias in selective exposure to political information. *Communication Research* **2020**, *47*, 104–124.
10. Jonas, E.; Graupmann, V.; Frey, D. The influence of mood on the search for supporting versus conflicting information: Dissonance reduction as a means of mood regulation? *Personality and Social Psychology Bulletin* **2006**, *32*, 3–15.
11. Taber, C.S.; Lodge, M. Motivated skepticism in the evaluation of political beliefs. *American journal of political science* **2006**, *50*, 755–769.
12. Cookson, J.A.; Engelberg, J.; Mullins, W. Echo chambers. *Available at SSRN 3603107* **2021**.
13. Stanovich, K.E.; West, R.F. Natural myside bias is independent of cognitive ability. *Thinking & Reasoning* **2007**, *13*, 225–247.
14. Mallinson, D.J.; Hatemi, P.K. The effects of information and social conformity on opinion change. *PLoS one* **2018**, *13*, e0196600.
15. Pryor, C.; Perfors, A.; Howe, P.D. Conformity to the descriptive norms of people with opposing political or social beliefs. *PLoS one* **2019**, *14*, e0219464.
16. Constant, A.; Ramstead, M.J.; Veissière, S.P.; Friston, K. Regimes of expectations: An active inference model of social conformity and human decision making. *Frontiers in psychology* **2019**, *10*, 679.
17. Kameda, T.; Tindale, R.S. Groups as adaptive devices: Human docility and group aggregation mechanisms in evolutionary context. *Evolution and social psychology* **2006**, pp. 317–341.
18. Fu, J.H.y.; Morris, M.W.; Lee, S.I.; Chao, M.; Chiu, C.y.; Hong, Y.y. Epistemic motives and cultural conformity: need for closure, culture, and context as determinants of conflict judgments. *Journal of personality and social psychology* **2007**, *92*, 191.
19. Cialdini, R.B.; Goldstein, N.J. Social influence: Compliance and conformity. *Annu. Rev. Psychol.* **2004**, *55*, 591–621.
20. Veissière, S.P.; Constant, A.; Ramstead, M.J.; Friston, K.J.; Kirmayer, L.J. Thinking through other minds: A variational approach to cognition and culture. *Behavioral and brain sciences* **2020**, *43*.
21. Ladegard, G. Accessing Social Capital through Conformity: Investigating Network Cultures. *ADVANCES IN ORGANIZATION STUDIES* **2006**, *20*, 153.
22. Lieberman, M.D.; Eisenberger, N.I. Pains and pleasures of social life. *Science* **2009**, *323*, 890–891.
23. Pharo, H.H. The Behavioural and Psychological Effects of Ostracism In Adolescence and Emerging-Adulthood. PhD thesis, University of Otago, 2012.
24. Tang, H.H.Y. The eye of the ostracised: Investigating changes in attention and affect following and during single-and multiple-episodes of ostracism and inclusion **2011**.
25. Ellickson, R. Order without Law Cambridge, 1991.
26. Irwin, K.; Horne, C. A normative explanation of antisocial punishment. *Social science research* **2013**, *42*, 562–570.
27. Allen, M. Theory-led confirmation bias and experimental persona. *Research in Science & Technological Education* **2011**, *29*, 107–127.
28. Falk, E.B.; Bassett, D.S. Brain and social networks: fundamental building blocks of human experience. *Trends in cognitive sciences* **2017**, *21*, 674–690.
29. Baumeister, R.F.; Leary, M.R. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin* **1995**, *117*, 497.
30. Falk, E.; Scholz, C. Persuasion, influence, and value: Perspectives from communication and social neuroscience. *Annual review of psychology* **2018**, *69*, 329–356.

31. Baek, E.C.; Scholz, C.; O'Donnell, M.B.; Falk, E.B. The value of sharing information: a neural account of information transmission. *Psychological science* **2017**, *28*, 851–861.
32. Scholz, C.; Baek, E.C.; O'Donnell, M.B.; Kim, H.S.; Cappella, J.N.; Falk, E.B. A neural model of valuation and information virality. *Proceedings of the National Academy of Sciences* **2017**, *114*, 2881–2886.
33. Hong, H.; Choi, M.Y.; Kim, B.J. Synchronization on small-world networks. *Physical Review E* **2002**, *65*, 026139.
34. Etzlinger, B.; Wymeersch, H.; Springer, A. Cooperative synchronization in wireless networks. *IEEE Transactions on Signal Processing* **2014**, *62*, 2837–2849.
35. Masuda, N.; Aihara, K. Global and local synchrony of coupled neurons in small-world networks. *Biological cybernetics* **2004**, *90*, 302–309.
36. Schmälzle, R.; Häcker, F.E.; Honey, C.J.; Hasson, U. Engaged listeners: shared neural processing of powerful political speeches. *Social Cognitive and Affective Neuroscience* **2015**, *10*, 1137–1143.
37. Gumperz, J.J. The speech community. *Linguistic anthropology: A reader* **2009**, *1*, 66–73.
38. Wood, W. Attitude change: Persuasion and social influence. *Annual review of psychology* **2000**, *51*, 539–570.
39. Bodrunova, S.S.; Blekanov, I.S.; Kukarkin, M. Multi-dimensional echo chambers: language and sentiment structure of Twitter discussions on the Charlie Hebdo Case. International Conference on Human-Computer Interaction. Springer, 2018, pp. 393–400.
40. Garimella, V.R.K.; Weber, I. A long-term analysis of polarization on Twitter. Proceedings of the International AAAI Conference on Web and Social Media, 2017, Vol. 11.
41. Dunlop, C.A. Epistemic communities; Routledge, 2012.
42. Cinelli, M.; Morales, G.D.F.; Galeazzi, A.; Quattrociocchi, W.; Starnini, M. Echo chambers on social media: A comparative analysis. *arXiv preprint arXiv:2004.09603* **2020**.
43. Garimella, K.; De Francisci Morales, G.; Gionis, A.; Mathioudakis, M. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. Proceedings of the 2018 World Wide Web Conference, 2018, pp. 913–922.
44. Duseja, N.; Jhamtani, H. A sociolinguistic study of online echo chambers on twitter. Proceedings of the third workshop on natural language processing and computational social science, 2019, pp. 78–83.
45. Kozitsin, I.V.; Chkhartishvili, A.G. Users' Activity in Online Social Networks and the Formation of Echo Chambers. 2020 13th International Conference" Management of large-scale system development"(MLSD). IEEE, 2020, pp. 1–5.
46. Starnini, M.; Frasca, M.; Baronchelli, A. Emergence of metapopulations and echo chambers in mobile agents. *Scientific reports* **2016**, *6*, 1–8.
47. Wang, X.; Song, Y. Viral misinformation and echo chambers: The diffusion of rumors about genetically modified organisms on social media. *Internet Research* **2020**.
48. Baumann, F.; Lorenz-Spreen, P.; Sokolov, I.M.; Starnini, M. Modeling echo chambers and polarization dynamics in social networks. *Physical Review Letters* **2020**, *124*, 048301.
49. Nguyen, C.T. Echo chambers and epistemic bubbles. *Episteme* **2020**, *17*, 141–161.
50. Santos, B.R. Echo chambers, ignorance and domination. *Social epistemology* **2021**, *35*, 109–119.
51. Elzinga, B. Echo Chambers and Audio Signal Processing. *Episteme* **2020**, pp. 1–21.
52. Nguyen, C.T. Cognitive islands and runaway echo chambers: problems for epistemic dependence on experts. *Synthese* **2020**, *197*, 2803–2821.
53. Sullivan, E.; Sondag, M.; Rutter, I.; Meulemans, W.; Cunningham, S.; Speckmann, B.; Alfano, M. Vulnerability in social epistemic networks. *International Journal of Philosophical Studies* **2020**, *28*, 731–753.
54. Palmer, C.J.; Lawson, R.P.; Hohwy, J. Bayesian approaches to autism: Towards volatility, action, and behavior. *Psychological bulletin* **2017**, *143*, 521.
55. Engle, R.F.; Ghysels, E.; Sohn, B. Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics* **2013**, *95*, 776–797.
56. Fristad, M.A.; Gavazzi, S.M.; Mackinaw-Koons, B. Family psychoeducation: an adjunctive intervention for children with bipolar disorder. *Biological psychiatry* **2003**, *53*, 1000–1008.
57. Lawson, R.P.; Rees, G.; Friston, K.J. An aberrant precision account of autism. *Frontiers in human neuroscience* **2014**, *8*, 302.
58. Martin, G.P.; Hanna, E.; McCartney, M.; Dingwall, R. Science, society, and policy in the face of uncertainty: reflections on the debate around face coverings for the public during COVID-19, 2020.

59. Desforges, M.; Gurdasani, D.; Hamdy, A.; Leonardi, A.J. Uncertainty around the Long-Term Implications of COVID-19. *Pathogens* **2021**, *10*, 1267. [1582](#)  
[1583](#)

60. Tandoc Jr, E.C.; Lee, J.C.B. When viruses and misinformation spread: How young Singaporeans navigated uncertainty in the early stages of the COVID-19 outbreak. *New Media & Society* **2020**, *1584*, p. 1461444820968212. [1585](#)  
[1586](#)

61. Rettie, H.; Daniels, J. Coping and tolerance of uncertainty: Predictors and mediators of mental health during the COVID-19 pandemic. *American Psychologist* **2021**, *76*, 427. [1587](#)  
[1588](#)

62. Jiang, S.; Zhang, H.; Qi, J.; Fang, B.; Xu, T. Perceiving Social-Emotional Volatility and Triggered Causes of COVID-19. *International Journal of Environmental Research and Public Health* **2021**, *18*, 4591. [1589](#)  
[1590](#)  
[1591](#)

63. Berger, J.G.; Johnston, K. *Simple habits for complex times*; Stanford University Press, 2020. [1592](#)

64. Gillan, C.M.; Sahakian, B.J. Which is the driver, the obsessions or the compulsions, in OCD? *Neuropsychopharmacology* **2015**, *40*, 247. [1593](#)  
[1594](#)

65. Brewer, J.A.; Roy, A. Can Approaching Anxiety Like a Habit Lead to Novel Treatments? *American Journal of Lifestyle Medicine* **2021**, p. 15598276211008144. [1595](#)  
[1596](#)

66. Misztal, B.A. Trust in Habit: A Way of Coping in Unsettled Times. In *Trust in Contemporary Society*; Brill, 2019; pp. 41–59. [1597](#)  
[1598](#)

67. Pietarinen, A.V.; Beni, M.D. Active Inference and Abduction. *Biosemiotics* **2021**, pp. 1–19. [1599](#)

68. Hodgson, G.M. The ubiquity of habits and rules. *Cambridge journal of economics* **1997**, *21*, 663–684. [1600](#)

69. Dynes, R. Social capital: Dealing with community emergencies. *Homeland Security Affairs* **2006**, *2*. [1601](#)  
[1602](#)

70. Gardner, B.; Lally, P. Modelling habit formation and its determinants. *The psychology of habit* **2018**, pp. 207–229. [1603](#)  
[1604](#)

71. Schwabe, L.; Wolf, O.T. Stress prompts habit behavior in humans. *Journal of Neuroscience* **2009**, *29*, 7191–7198. [1605](#)  
[1606](#)

72. Wood, W.; Rünger, D. Psychology of habit. *Annual review of psychology* **2016**, *67*, 289–314. [1607](#)

73. Friston, K.; FitzGerald, T.; Rigoli, F.; Schwartenbeck, P.; Pezzulo, G.; et al. Active inference and learning. *Neuroscience & Biobehavioral Reviews* **2016**, *68*, 862–879. [1608](#)  
[1609](#)

74. Dezfouli, A.; Balleine, B.W. Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. *PLoS computational biology* **2013**, *9*, e1003364. [1610](#)  
[1611](#)  
[1612](#)

75. Pezzulo, G.; Rigoli, F.; Chersi, F. The mixed instrumental controller: using value of information to combine habitual choice and mental simulation. *Frontiers in psychology* **2013**, *4*, 92. [1613](#)  
[1614](#)

76. Pezzulo, G.; Rigoli, F.; Friston, K. Active inference, homeostatic regulation and adaptive behavioural control. *Progress in neurobiology* **2015**, *134*, 17–35. [1615](#)  
[1616](#)

77. Lee, S.W.; Shimojo, S.; O'Doherty, J.P. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* **2014**, *81*, 687–699. [1617](#)  
[1618](#)

78. Lally, P.; Van Jaarsveld, C.H.; Potts, H.W.; Wardle, J. How are habits formed: Modelling habit formation in the real world. *European journal of social psychology* **2010**, *40*, 998–1009. [1619](#)  
[1620](#)

79. Story, G.W.; Smith, R.; Moutoussis, M.; Berwian, I.M.; Nolte, T.; Bilek, E.; Dolan, R.J. Story et al. 2021 A Social Inference Model of Idealization and Devaluation **2021**. [1621](#)  
[1622](#)

80. Kaufmann, R.; Gupta, P.; Taylor, J. An active inference model of collective intelligence. *arXiv preprint arXiv:2104.01066* **2021**. [1623](#)  
[1624](#)

81. Ohata, W.; Tani, J. Investigation of the sense of agency in social cognition, based on frameworks of predictive coding and active inference: a simulation study on multimodal imitative interaction. *Frontiers in Neurorobotics* **2020**, *14*, 61. [1625](#)  
[1626](#)  
[1627](#)

82. Bouizegarene, N.; Ramstead, M.; Constant, A.; Friston, K.; Kirmayer, L. Narrative as active inference **2020**. [1628](#)  
[1629](#)

83. Tison, R.; Poirier, P. Communication as socially extended active inference: An ecological approach to communicative behavior. *Ecological Psychology* **2021**, pp. 1–39. [1630](#)  
[1631](#)

84. Kahl, S.; Kopp, S. How active inference can facilitate belief coordination in multi-agent interaction **2019**. [1632](#)  
[1633](#)

85. Parr, T.; Friston, K.J. Uncertainty, epistemics and active inference. *Journal of the Royal Society Interface* **2017**, *14*, 20170376. [1634](#)  
[1635](#)

86. Heins, R.C.; Mirza, M.B.; Parr, T.; Friston, K.; Kagan, I.; Pooreesmaeli, A. Deep active inference and scene construction. *Frontiers in Artificial Intelligence* **2020**, *3*, 81. [1636](#)  
[1637](#)

87. Friston, K.; Rigoli, F.; Ognibene, D.; Mathys, C.; Fitzgerald, T.; Pezzulo, G. Active inference and epistemic value. *Cognitive neuroscience* **2015**, *6*, 187–214. [1638](#)  
[1639](#)

88. Anderson, Y. Recent Advances in the Modelling and Analysis of Opinion Dynamics on Influence Networks. *SpringerLink* **2019**. [1641](#)

89. Parravano, Andina-Diaz, M.J. Bounded Confidence under Preferential Flip: A Coupled Dynamics of Structural Balance and Opinions. *PloS one*. [1642](#)

90. Blondel, Hendrickx, T. On Krause's Multi-Agent Consensus Model With State-Dependent Connectivity. *IEEE TRANSACTIONS ON AUTOMATIC CONTROL* **2009**, *54*. [1645](#)

91. Hegselmann, R.; Krause, U.; et al. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of artificial societies and social simulation* **2002**, *5*. [1646](#)

92. Ramstead, M.J.; Veissière, S.P.; Kirmayer, L.J. Cultural affordances: Scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in psychology* **2016**, *7*, 1090. [1649](#)

93. Knill, D.C.; Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences* **2004**, *27*, 712–719. [1650](#)

94. Friston, K. The history of the future of the Bayesian brain. *NeuroImage* **2012**, *62*, 1230–1233. [1652](#)

95. Madsen, J.K.; Bailey, R.M.; Pilditch, T.D. Large networks of rational agents form persistent echo chambers. *Scientific reports* **2018**, *8*, 1–8. [1653](#)

96. Acemoglu, D.; Ozdaglar, A. Opinion dynamics and learning in social networks. *Dynamic Games and Applications* **2011**, *1*, 3–49. [1655](#)

97. Fang, A.; Yuan, K.; Geng, J.; Wei, X. Opinion dynamics with Bayesian learning. *Complexity* **2020**, *2020*. [1657](#)

98. Friston, K.J.; Daunizeau, J.; Kiebel, S.J. Reinforcement learning or active inference? *PloS one* **2009**, *4*, e6421. [1659](#)

99. Botvinick, M.; Toussaint, M. Planning as inference. *Trends in cognitive sciences* **2012**, *16*, 485–488. [1661](#)

100. DeGroot, M.H. Reaching a consensus. *Journal of the American Statistical Association* **1974**, *69*, 118–121. [1663](#)

101. Holley, R.A.; Liggett, T.M. Ergodic theorems for weakly interacting infinite systems and the voter model. *The annals of probability* **1975**, pp. 643–663. [1664](#)

102. Deffuant, G.; Neau, D.; Amblard, F.; Weisbuch, G. Mixing beliefs among interacting agents. *Advances in Complex Systems* **2001**, p. 11. [1666](#)

103. Pais, D.; Hogan, P.M.; Schlegel, T.; Franks, N.R.; Leonard, N.E.; Marshall, J.A. A mechanism for value-sensitive decision-making. *PloS one* **2013**, *8*, e73216. [1669](#)

104. Pinkoviezky, I.; Couzin, I.D.; Gov, N.S. Collective conflict resolution in groups on the move. *Physical Review E* **2018**, *97*, 032304. [1671](#)

105. Bizyaeva, A.; Franci, A.; Leonard, N.E. A general model of opinion dynamics with tunable sensitivity. *arXiv preprint arXiv:2009.04332* **2020**. [1672](#)

106. Ye, M.; Liu, J.; Anderson, B.D.; Yu, C.; Başar, T. Evolution of social power in social networks with dynamic topology. *IEEE transactions on automatic control* **2018**, *63*, 3793–3808. [1674](#)

107. Friedkin, N.E.; Proskurnikov, A.V.; Tempo, R.; Parsegov, S.E. Network science on belief system dynamics under logic constraints. *Science* **2016**, *354*, 321–326. [1677](#)

108. Parsegov, S.E.; Proskurnikov, A.V.; Tempo, R.; Friedkin, N.E. Novel multidimensional models of opinion dynamics in social networks. *IEEE Transactions on Automatic Control* **2016**, *62*, 2270–2285. [1678](#)

109. Hołyst, J.A.; Kacperski, K.; Schweitzer, F. Social impact models of opinion dynamics. *Annual Reviews Of Computational PhysicsIX* **2001**, pp. 253–273. [1680](#)

110. Hegselmann, R.; Flache, A. Understanding complex social dynamics: A plea for cellular automata based modelling. *Journal of Artificial Societies and Social Simulation* **1998**, *1*, 1. [1682](#)

111. Latané, B.; Nowak, A. Self-organizing social systems: Necessary and sufficient conditions for the emergence of clustering, consolidation, and continuing diversity. *Progress in communication sciences* **1997**, pp. 43–74. [1684](#)

112. Weisbuch, G.; Deffuant, G.; Amblard, F.; Nadal, J.P. Interacting agents and continuous opinions dynamics. In *Heterogenous agents, interactions and economic performance*; Springer, 2003; pp. 225–242. [1688](#)

113. Smith, L.; Sørensen, P. Pathological outcomes of observational learning. *Econometrica* **2000**, *68*, 371–398. [1691](#)

114. Mann, R.P. Collective decision making by rational individuals. *Proceedings of the National Academy of Sciences* **2018**, *115*, E10387–E10396. [1692](#)

115. Da Costa, L.; Friston, K.; Heins, C.; Pavliotis, G.A. Bayesian mechanics for stationary processes. *arXiv preprint arXiv:2106.13830* **2021**. [1694](#)

116. Friston, K.; Fitzgerald, T.; Rigoli, F.; Schwartenbeck, P.; Pezzulo, G. Active inference: a process theory. *Neural computation* **2017**, *29*, 1–49. [1696](#)

117. Sutton, R.S.; Barto, A.G. *Reinforcement learning: An introduction*; MIT press, 2018. [1698](#)

118. MacKay, D.J. Bayesian interpolation. *Neural computation* **1992**, *4*, 415–447.
119. Beal, M.J. *Variational algorithms for approximate Bayesian inference*; University of London, University College London (United Kingdom), 2003.
120. Friston, K.J.; Wiese, W.; Hobson, J.A. Sentience and the origins of consciousness: From Cartesian duality to Markovian monism. *Entropy* **2020**, *22*, 516.
121. Puterman, M.L. Markov decision processes: Discrete stochastic dynamic programming. *Journal of the Operational Research Society* **1995**, *46*, 792–792.
122. Kaelbling, L.P.; Littman, M.L.; Cassandra, A.R. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* **1998**, *101*, 99–134. doi:10.1016/S0004-3702(98)00023-X.
123. Parr, T.; Friston, K.J. Generalised free energy and active inference. *Biological cybernetics* **2019**, *113*, 495–513.
124. Millidge, B.; Tschantz, A.; Buckley, C.L. Whence the expected free energy? *Neural Computation* **2021**, *33*, 447–482.
125. Da Costa, L.; Parr, T.; Sajid, N.; Veselic, S.; Neacsu, V.; Friston, K. Active inference on discrete state-spaces: a synthesis. *Journal of Mathematical Psychology* **2020**, *99*, 102447.
126. Schwartenbeck, P.; Passecker, J.; Hauser, T.U.; FitzGerald, T.H.; Kronbichler, M.; Friston, K.J. Computational mechanisms of curiosity and goal-directed exploration. *Elife* **2019**, *8*, e41703.
127. Erdős, P.; Rényi, A. On the evolution of random graphs. In *The structure and dynamics of networks*; Princeton University Press, 2011; pp. 38–82.
128. Grinberg, N.; Joseph, K.; Friedland, L.; Swire-Thompson, B.; Lazer, D. Fake news on Twitter during the 2016 US presidential election. *Science* **2019**, *363*, 374–378.
129. Rosenberg, H.; Syed, S.; Rezaie, S. The Twitter pandemic: The critical role of Twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. *Canadian journal of emergency medicine* **2020**, *22*, 418–421.
130. Cossard, A.; Morales, G.D.F.; Kalimeri, K.; Mejova, Y.; Paolotti, D.; Starnini, M. Falling into the echo chamber: the Italian vaccination debate on Twitter. Proceedings of the International AAAI conference on web and social media, 2020, Vol. 14, pp. 130–140.
131. Tschantz, A.; Seth, A.K.; Buckley, C.L. Learning action-oriented models through active inference. *PLoS computational biology* **2020**, *16*, e1007805.
132. Al-Qarni, B.H.; Almogren, A.; Hassan, M.M. An efficient networking protocol for internet of things to handle multimedia big data. *Multimedia Tools and Applications* **2019**, *78*, 30039–30056.
133. Le Goix, R. Gated communities: Sprawl and social segregation in Southern California. *Housing studies* **2005**, *20*, 323–343.
134. Aldrich, H.; Herker, D. Boundary spanning roles and organization structure. *Academy of management review* **1977**, *2*, 217–230.
135. Jaemthaworn, T.; Kalapanulak, S.; Saithong, T. Topological clustering of regulatory genes confers pathogenic tolerance to cassava brown streak virus (CBSV) in cassava. *Scientific reports* **2021**, *11*, 1–14.
136. Richter, A.W.; West, M.A.; Van Dick, R.; Dawson, J.F. Boundary spanners' identification, intergroup contact, and effective intergroup relations. *Academy of Management Journal* **2006**, *49*, 1252–1269.
137. Guan, J.; Li, S.; He, X.; Zhu, J.; Chen, J. Fast hierarchical clustering of local density peaks via an association degree transfer method. *Neurocomputing* **2021**, *455*, 401–418.
138. Canright, G.; Engø-Monsen, K. Roles in networks. *Science of Computer Programming* **2004**, *53*, 195–214.
139. Bell, V.; O'Driscoll, C. The network structure of paranoia in the general population. *Social psychiatry and psychiatric epidemiology* **2018**, *53*, 737–744.
140. Sasahara, K.; Chen, W.; Peng, H.; Ciampaglia, G.L.; Flammini, A.; Menczer, F. Social influence and unfollowing accelerate the emergence of echo chambers. *Journal of Computational Social Science* **2021**, *4*, 381–402.
141. Balietti, S.; Klein, B.; Riedl, C. Optimal design of experiments to identify latent behavioral types. *Experimental Economics* **2020**, pp. 1–28.
142. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; De Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE* **2015**, *104*, 148–175.
143. Sridhar, V.H.; Li, L.; Gorbonos, D.; Nagy, M.; Schell, B.R.; Sorochkin, T.; Gov, N.S.; Couzin, I.D. The geometry of decision-making in individuals and collectives. *Proceedings of the National Academy of Sciences* **2021**, *118*.
144. Albarracín, M.; Constant, A.; Friston, K.J.; Ramstead, M.J.D. A variational approach to scripts. *Frontiers in Psychology* **2021**, p. 3035.

145. Heins, C.; Millidge, B.; Demekas, D.; Klein, B.; Friston, K.; Couzin, I.; Tschantz, A. pymdp: A Python library for active inference in discrete state spaces. *Forthcoming 2022*.

## Appendix A

In this appendix we provide additional mathematical details on the update equations for perception and action (policy inference) under active inference.

We begin by recalling the POMDP generative model and the approximate posterior over hidden states and policies:

$$\begin{aligned}
 P(\tilde{o}, \tilde{s}, \tilde{u}, \pi) &= P(s_1)P(\pi) \prod_{\tau=2}^T P(s_\tau \mid s_{\tau-1}, u_\tau)P(u_\tau \mid \pi) \prod_{\tau=1}^T P(o_\tau \mid s_\tau) \\
 Q(s_\tau \mid \pi) &= \text{Cat}(\mathbf{s}_{\pi\tau}) \\
 Q(\pi) &= \text{Cat}(\boldsymbol{\pi}) \\
 Q(s_{1:T}, \pi) &= Q(\pi) \prod_{\tau=1}^T Q(s_\tau \mid \pi)
 \end{aligned} \tag{A30}$$

Given this generative model and approximate posterior, we can now write down the variational free energy over time:

$$\mathcal{F}_{1:T} = \mathbb{E}_{Q(s_{1:T}, \pi)} [\ln Q(s_{1:T}, \pi) - \ln P(o_{1:T}, s_{1:T}, \pi)]$$

Due to the Markovian nature of the POMDP generative model and the factorized form of the posterior, the free energy over trajectories can be factorized into a per-timestep free energy  $\mathcal{F}_\tau$ , which has the following simple form:

$$\mathcal{F}_\tau = \mathbb{E}_{Q(s_\tau \mid \pi)Q(\pi)} [\ln Q(s_\tau \mid \pi) - \ln P(o_\tau, s_\tau \mid s_{\tau-1}, \pi)] \tag{A31}$$

In the following subsections we show how state estimation (perception) and policy inference (decision-making/planning) are derived by minimizing the variational free energy functionals with respect to the parameters of the posterior  $Q(s_{1:T}, \pi)$ .

### State estimation

State estimation consists in optimizing the  $Q(s_{1:T} \mid \pi)$  over hidden states under different policies. Because our approximate posterior and generative model are defined using categorical distributions, the problem of state estimation becomes minimizing free energy gradients of the form  $\frac{\partial \mathcal{F}}{\partial \mathbf{s}}$ , where  $\mathbf{s}$  are the parameters of the approximate posterior distribution over hidden states,  $Q(s) = \text{Cat}(\mathbf{s})$ , where the notation  $P(x) = \text{Cat}(\boldsymbol{\phi})$  denotes a categorical distribution over some random variable  $x$  with parameters  $\boldsymbol{\phi}$ .

In the ‘full construct’ version of active inference (see e.g. [87]), the full joint posterior  $Q(s_{1:T}, \pi)$  is optimized simultaneously, meaning that the posterior over hidden states is conditioned on policies. This means that the full posterior beliefs at any timestep  $t$  include a separate  $Q(s_t \mid \pi)$  under each policy, where the beliefs about a given timestep under a given policy are often denoted by the sufficient statistics  $\mathbf{s}_{\pi\tau}$ . For the current model, we have simplified posterior inference to rely on an approximate posterior where hidden states are independent of policies. This move is justified because the practical differences between the ‘full construct’ and simplified versions are negligible, in the limit of small policy spaces and short time horizons (such as in the current work). Therefore in the current study we amended the variational posterior to have the following form:

$$Q(s_{1:T}, \pi) = Q(\pi) \prod_{\tau=1}^T Q(s_\tau)$$

Given this simplified posterior, state estimation proceeds by optimizing the posterior belief about hidden states at the current timestep  $Q^*(s_t)$  using the current observation  $o_t$ . This can be found using the gradients of the variational free energy from Equation (A31), now using the simplified form of the posterior:

$$\begin{aligned} \mathcal{F}_t &= \mathbb{E}_{Q(s_t)Q(\pi)}[\ln Q(s_t) - \ln P(o_t, s_t | s_{t-1}, \pi)] \\ \implies \frac{\partial \mathcal{F}_t}{\partial Q(s_t)} &= 0 \iff Q^*(s_t) = \sigma(\ln P(o_t | s_t) + \ln(P(s_t | s_{t-1}, u_{t-1})P(s_{t-1}))) \end{aligned} \quad (\text{A32})$$

where  $\sigma(x) = \frac{e^x}{\sum_x e^x}$  is the normalized exponential or softmax function. Equation (A32) is a type of ‘fixed-point solution’ for the posterior, where the beliefs about hidden states are directly set to the solution of minimal free energy (where  $\frac{\partial \mathcal{F}_t}{\partial Q(s_t)} = 0$ ). Note that this differs with the classic ‘gradient descent’ scheme used to optimize the variational posterior with marginal message passing or variational message passing, as proposed in [87], which was originally invoked as a biologically-plausible update scheme that could be implemented by neuronal population dynamics. Since we are not interested in simulating neurophysiological responses and the belief updating is simpler, for the simulations presented in this paper, we used this simpler update rule.

The functional form of (A32) invites a straightforward Bayesian interpretation: the ‘best’ posterior belief  $Q^*(s_t)$  is proportional to the product of a likelihood term  $P(o_t | s_t)$  and a prior term  $P(s_t | s_{t-1}, u_{t-1})P(s_{t-1})$ —the definition of Bayes rule. In practice, we use a ‘moving empirical prior’ rule, where the posterior from last timestep’s optimization  $Q^*(s_{t-1})$  becomes the prior  $P(s_{t-1})$ . This means the update rule can be re-written as follows:

$$\begin{aligned} P(s_{t-1}) &\approx Q(s_{t-1}) \\ \implies Q^*(s_t) &= \sigma(\ln P(o_t | s_t) + \ln \mathbb{E}_{Q(s_{t-1})}[P(s_t | s_{t-1}, u_{t-1})]) \end{aligned} \quad (\text{A33})$$

This means that at each timestep, the current posterior is a Bayesian average between the likelihood term and the previous timestep’s posterior belief, passed through the action-conditioned transition dynamics  $P(s_t | s_{t-1}, u_{t-1})$  of the generative model. Note that this update rule can be extended to generative models with factorized observations  $\mathbf{o}$  and hidden states  $\mathbf{s}$  by rewriting Equation (A33) for a particular marginal  $Q^*(s_t^f)$  as follows:

$$Q^*(s_t^f) = \sigma \left( \mathbb{E}_{Q^{i \setminus f}} [\ln P(\mathbf{o}_t | \mathbf{s}_t) + \ln \mathbb{E}_{Q(s_{t-1}^f)} [P(s_t^f | s_{t-1}^f, u_{t-1})]] \right) \quad (\text{A34})$$

where the expectation  $\mathbb{E}_{Q^{i \setminus f}}$  denotes an expectation with respect to all posterior marginals  $Q(s_t^i)$  besides the marginal  $Q(s_t^f)$  currently being optimized. To find the full, multi-factor posterior  $Q^*(\mathbf{s}_t)$ , this equation is iterated across marginals, holding the existing solutions for all other marginals fixed while a particular one is updated [145].

### Policy inference

Under active inference, policies  $\pi$  inferred, i.e. the agent optimizes a variational posterior over policies denoted by  $Q(\pi)$ . The optimal posterior over policies  $Q^*(\pi)$  is obtained by minimizing the free energy with respect to the categorical parameters  $\pi$  (c.f.

Equation (A30)). This can be shown by rewriting the full variational free energy over time  $\mathcal{F}_{1:T}$  as the sum of a complexity term that measures the divergence between the posterior and prior over policies, and an expected ‘accuracy’-like term:

$$\begin{aligned}\mathcal{F}_{1:T} &= \mathbb{E}_{Q(s_{1:T}, \pi)} [\ln Q(s_{1:T}, \pi) - \ln P(o_{1:T}, s_{1:T}, \pi)] \\ &= \mathbb{E}_{Q(s_{1:T}, \pi)} [\ln Q(\pi) + \sum_{\tau=1}^T \ln Q(s_{\tau} | \pi) - \ln P(\pi) - \ln P(o_{1:T}, s_{1:T} | \pi)] \\ &= D_{KL}[Q(\pi) \parallel P(\pi)] + \mathbb{E}_{Q(\pi)} [F(\pi)]\end{aligned}\quad (\text{A35})$$

where the second term is the expected variational free energy of policies  $F(\pi)$ , which is defined as follows:

$$F(\pi) = -\mathbb{E}_{Q(s_{1:T} | \pi)} [\ln P(o_{1:T}, s_{1:T} | \pi) - \mathbf{H}[Q(s_{1:T} | \pi)]] \quad (\text{A36})$$

The optimal posterior that minimizes the full variational free energy  $\mathcal{F}$  is found by taking the derivative of  $\mathcal{F}$  with respect to  $Q(\pi)$  and setting this gradient to 0, yielding the following free-energy-minimizing solution for  $Q(\pi)$ :

$$Q^*(\pi) = \underset{\pi}{\operatorname{argmin}} \mathcal{F} = \sigma(\ln P(\pi) - F(\pi)) \quad (\text{A37})$$

Therefore in the same way that state estimation or optimization of  $Q(s)$  in Equation (A33) resembles a Bayesian average of a likelihood and a prior term, policy inference also becomes an average of the policy prior  $P(\pi)$  and the ‘evidence’ afforded to each policy, scored by  $F(\pi)$ . Recall here that the policy prior  $P(\pi)$  is itself decomposed as a combination of the expected free energy prior and the ‘habit vector’:  $P(\pi) = P(\pi_0) - \mathbf{G}$ .

## Appendix B

In this appendix derive a quantitative relationship between the epistemic confirmation bias  $\gamma$  and the negative ambiguity term of the epistemic value. Recall the definition of the epistemic value:

$$\mathbb{E}_{Q(\mathbf{o}_{\tau} | \pi)} [D_{KL}(Q(\mathbf{s}_{\tau} | \mathbf{o}_{\tau}, \pi) \parallel Q(\mathbf{s}_{\tau} | \pi))] = -\mathbb{E}_{Q(\mathbf{s}_{\tau} | \pi)} [\mathbf{H}[P(\mathbf{o}_{\tau} | \mathbf{s}_{\tau})]] + \mathbf{H}[Q(\mathbf{o}_{\tau} | \pi)]$$

We define the first term on the RHS of the decomposition as the negative ambiguity  $\mathcal{H} = -\mathbb{E}_{Q(\mathbf{s}_{\tau} | \pi)} [\mathbf{H}[P(\mathbf{o}_{\tau} | \mathbf{s}_{\tau})]]$ . We drop the  $\tau$  subscript hereafter for simplicity, and restrict ourselves only to the computation of this term for the neighbour  $k$  tweet observation modality  $o^{\text{NT}k}$  and the hidden states that it depends on,  $s^{\text{Idea}}$  and  $s^{\text{MB}k}$ . We further condition our analysis only on those policies that entail sampling neighbour  $k$ , i.e. those policies where  $s^{\text{Who}} = u^{\text{Who}} = k$ . Therefore we redefine  $Q(\mathbf{s})$  here as  $Q(s^{\text{Idea}}, s^{\text{MB}k})$ .

**Theorem A1.** *The negative ambiguity is proportional to the epistemic confirmation bias parameter  $\gamma$ .*

$$\mathcal{H} \propto k^{\gamma} \log k^{\gamma}$$

**Proof.** WLOG, we simplify the state space to only consider the states  $s^{\text{Idea}}$  and  $s^{\text{MB}k}$  and the observation of neighbour  $k$ ’s Hashtag  $o^{\text{NT}k}$ . Recall that observation likelihood for  $o^{\text{NT}k}$ , in the case that  $s^{\text{Idea}} = s^{\text{MB}k}$  is as a softmax transformation of (certain columns of) the base likelihood  $P(o^{\text{NT}k} | s^{\text{Idea}}, s^{\text{MB}k})$  and the epistemic confirmation bias parameter  $\gamma$ :

$$P(o^{\text{NT}k} | s^{\text{Idea}} = s^{\text{MB}k}, \gamma) = \frac{e^{\gamma P(o^{\text{NT}k} | s^{\text{Idea}} = s^{\text{MB}k})}}{\sum e^{\gamma P(o^{\text{NT}k} | s^{\text{Idea}} = s^{\text{MB}k})}} \quad (\text{A38})$$

In this case, the likelihood  $P(o^{\text{NT}k} | s^{\text{MB}k})$  is comprised of two Bernoulli distributions, where

$$P(o^{\text{NT}k} | s^{\text{MB}k}) = \{\text{Bern}(p_h), \text{Bern}(1 - p_h)\}$$

where  $p_h$  is the 'Hashtag reliability' parameter of the matrix  $\mathbf{h}$ :

$$\mathbf{h} = \begin{bmatrix} p_h & 1 - p_h \\ 1 - p_h & p_h \end{bmatrix}$$

where this matrix is 'copied' across the dimension of the likelihood corresponding to the two settings of  $s^{\text{Idea}}$ .

Note also that the posterior over hidden states is factorized into two independent marginal posteriors.

$$Q(s^{\text{Idea}}, s^{\text{MB}k}) = Q(s^{\text{Idea}})Q(s^{\text{MB}k}) \quad (\text{A39})$$

The definition of negative ambiguity is

$$\mathcal{H} = -\mathbb{E}_{Q(\mathbf{s}|\pi)}[\mathbf{H}[P(o^{\text{NT}k} | \mathbf{s})]] \quad (\text{A40})$$

$$(\text{A41})$$

And we can write the negative entropy of the given likelihood as

$$-\mathbf{H}[P(o^{\text{NT}k} | s^{\text{MB}k})] = p_h \log p_h + (1 - p_h) \log(1 - p_h)$$

Using Equation (A38) we have

$$-\mathbf{H}[P(o^{\text{NT}k} | s^{\text{MB}k} = s^{\text{Idea}})] = \frac{p_h^\gamma}{C} \log \frac{p_h^\gamma}{C} + \frac{(1 - p_h)^\gamma}{C} \log \frac{(1 - p_h)^\gamma}{C}$$

where  $C = p^\gamma + (1 - p)^\gamma$

The negative entropy can be decomposed into a sum of negative entropies:

$$-\mathbf{H}[P(o^{\text{NT}k} | s^{\text{MB}k})] = -\mathbf{H}[P(o^{\text{NT}k} | s^{\text{MB}k} = s^{\text{Idea}})] - \mathbf{H}[P(o^{\text{NT}k} | s^{\text{MB}k} \neq s^{\text{Idea}})]$$

which then means the total negative ambiguity can be written as follows, expanding the expressions for the entropies in terms of the Bernoulli parameter  $p_h$ :

$$\begin{aligned} \mathcal{H} = \mathbb{E}_{Q(s^{\text{Idea}}, s^{\text{MB}k})} & \left[ \frac{1}{C} [p_h^\gamma \log p_h^\gamma + (1 - p_h)^\gamma \log(1 - p_h)^\gamma - 2 \log C] \right. \\ & \left. + [p_h \log p_h + (1 - p_h) \log(1 - p_h)] \right] \quad (\text{A42}) \end{aligned}$$

Since all terms in Equation (A42) are increasing in  $\gamma$  insofar as  $p_h \geq 0$ , then negative ambiguity is directly proportional to  $\gamma$ .  $\square$

**Theorem A2.** *For any  $\gamma > 1$ , and for any realization of the posterior  $Q(s^{\text{Idea}}, s^{\text{MB}k})$  the negative ambiguity will be maximized in the case that the posterior beliefs about  $s^{\text{Idea}}$  and the posterior beliefs about  $s^{\text{MB}k}$  are either both greater than 0.5 or both less than 0.5.*

$$\forall \gamma > 1$$

$$\max \mathcal{H} \in \{\mathcal{H} : \text{sgn}(Q((s^{\text{Idea}}) - 0.5) = \text{sgn}(Q(s^{\text{MB}k}) - 0.5)\}$$

**Proof.** We now use the fact that  $Q(s^{\text{Idea}})$  and  $Q(s^{\text{MB}k})$  are also Bernoulli probability distributions, such that

$$\begin{aligned} Q(s^{\text{Idea}}) &= \text{Bern}(p) \\ Q(s^{\text{MB}k}) &= \text{Bern}(q) \end{aligned}$$

We also note that since the  $\gamma$  parameter only scaled the likelihood in the case that  $s^{\text{Idea}} = s^{\text{MB}k}$ , the values of the posterior that these correspond to are only those along the diagonal of the joint probability distribution of  $Q(s^{\text{Idea}}, s^{\text{MB}k})$ , namely the joint line of solutions through  $p$  and  $q$  connected by the points  $qp$  and  $(1-q)(1-p)$ .

Expanding the expectation in Equation (A42), we can write the negative ambiguity as

$$\begin{aligned} -qp(\mathbf{H}[P(o^{\text{NT}k}|s^{\text{MB}k} = s^{\text{Idea}})]) - (1-q)(1-p)(\mathbf{H}[P(o^{\text{NT}k}|s^{\text{MB}k} = s^{\text{Idea}})]) \\ -q(1-p)(\mathbf{H}[P(o^{\text{NT}k}|s^{\text{MB}k} \neq s^{\text{Idea}})]) - p(1-q)(\mathbf{H}[P(o^{\text{NT}k}|s^{\text{MB}k} \neq s^{\text{Idea}})]) \end{aligned} \quad (\text{A43})$$

Suppose that  $\gamma > 1$ . This means that for nonzero  $p_h$ , the entropy terms in Equation (A43) will be exponentiated by a power greater than 1, which implies that

$$\mathbf{H}[P(o^{\text{NT}k}|s^{\text{MB}k} = s^{\text{Idea}})] < \mathbf{H}[P(o^{\text{NT}k}|s^{\text{MB}k} \neq s^{\text{Idea}})] \quad (\text{A44})$$

Now take the case that  $Q(s^{\text{Idea}}) > 0.5$  and  $Q(s^{\text{MB}k}) > 0.5$ . This means that the largest coefficient scaling the entropy of the likelihood will necessarily be  $qp$ , which scales  $-\mathbf{H}[P(o^{\text{NT}k}|s^{\text{MB}k} = s^{\text{Idea}})]$ .

Similarly, if  $Q(s^{\text{Idea}}) < 0.5$  and  $Q(s^{\text{MB}k}) < 0.5$ , the largest coefficient scaling the entropy of the likelihood will necessarily be  $(1-p)(1-q)$ , which also scales  $-\mathbf{H}[P(o^{\text{NT}k}|s^{\text{MB}k} = s^{\text{Idea}})]$ .

However, if  $\text{sgn}(Q((s^{\text{Idea}}) - 0.5) \neq \text{sgn}(Q(s^{\text{MB}k}) - 0.5)$ , the largest coefficients will be either  $(1-p)q$  or  $p(1-q)$  which will be scaling  $-\mathbf{H}[P(o^{\text{NT}k}|s^{\text{MB}k} \neq s^{\text{Idea}})]$ .

Therefore, because of Equations (A43) and (A44), the maximum negative ambiguity for any value of  $\gamma > 1$  will always be reached when  $\text{sgn}(Q((s^{\text{Idea}}) - 0.5) = \text{sgn}(Q(s^{\text{MB}k}) - 0.5)$

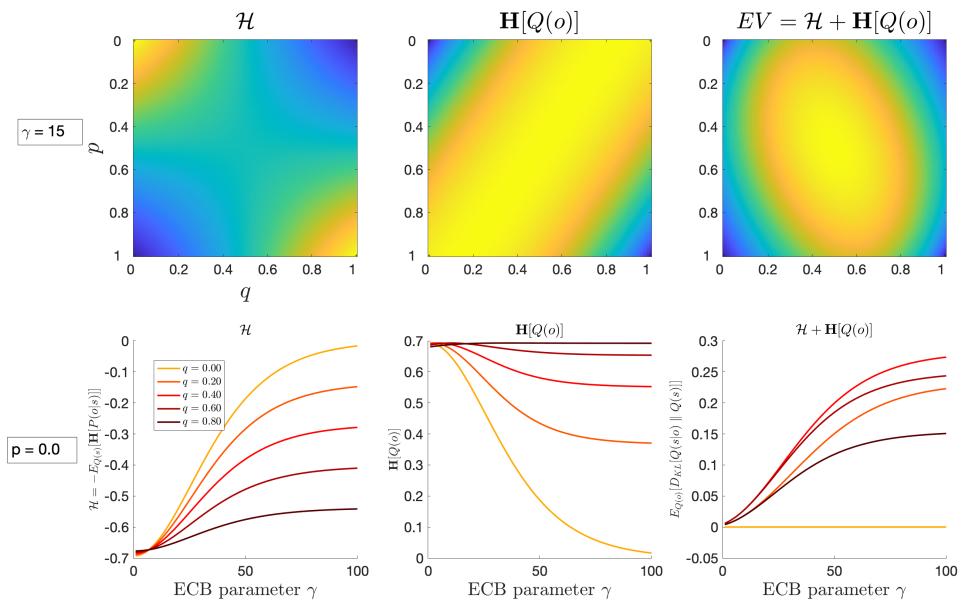
□

Figure A7 provides a visual intuition for the relationship between the two marginal posteriors (defined by Bernoulli parameters  $p$  and  $q$ ), the epistemic confirmation bias  $\gamma$  and the components of the epistemic value, decomposed here as the negative ambiguity  $\mathcal{H} = \mathbb{E}_{Q(\mathbf{s}|\pi)}[\mathbf{H}[P(\mathbf{o}|\mathbf{s})]]$  and the entropy of the predictive distribution over observations  $Q(\mathbf{o}|\pi)$ .

**Theorem A3.** *The inverse social volatility  $\omega^{\text{Soc}}$  is inversely related to the epistemic value of policies that entail sampling a particular neighbour - in other words  $EV \propto \frac{1}{\omega^{\text{Soc}}}$ .*

**Proof.** Recall the decomposition of the expected free energy in the section on [Policy inference](#) for a policy  $\pi$  into the negative instrumental value and the negative salience or epistemic value. For convenience, we define a pseudo-'value' function for policies as the negative of the expected free energy  $\mathbf{V}(\pi) \equiv -\mathbf{G}(\pi)$ :

$$\mathbf{V}(\pi) = \underbrace{E_{Q(\mathbf{o}_\pi|\pi)}[\ln \tilde{P}(\mathbf{o}_\pi)]}_{\text{Instrumental value}} + \underbrace{E_{Q(\mathbf{o}_\pi|\pi)}[D_{KL}[Q(\mathbf{s}_\pi|\mathbf{o}_\pi, \pi) \| Q(\mathbf{s}_\pi|\pi)]]}_{\text{Epistemic value}} \quad (\text{A45})$$



**Figure A7.** Exploration of the relationship between ‘agreement’ between an agent and one of its neighbours, the epistemic confirmation bias parameter  $\gamma$ , and the epistemic value of reading that neighbour’s tweet content. Here, the two marginal posteriors  $Q(s^{\text{Idea}})$  and  $Q(s^{\text{MB}k})$  are expressed as two Bernoulli distributions with respective parameters  $p$  and  $q$ , where ‘agreement’ is the case when  $p = q$  and hence  $(1 - p) = (1 - q)$ . The top row shows heatmaps of the negative ambiguity  $\mathcal{H}$ , entropy  $\mathbf{H}[Q(o)]$ , and the full epistemic value  $EV = \mathcal{H} + \mathbf{H}[Q(o)]$  for a fixed value of  $\gamma = 15.0$ , under all possible values of  $p$  and  $q$ . The ‘epistemic confirmation bias’ effect is seen in the negative ambiguity surface  $\mathcal{H}$  (upper left plot), which is maximized when posterior beliefs about the validity of **Idea 1**, measured by  $p$ , are aligned with posterior beliefs about a neighbour’s meta-belief about **Idea 1**,  $q$ . The bottom row of plots shows a complementary perspective, demonstrating the effect of increasing  $\gamma$  on the epistemic value and its components, for different settings of  $q$  when  $p = 0.0$ . The subplot on furthest to the right of the bottom row shows that increasing  $\gamma$  increases epistemic value most when  $q$  is on the same side of 0.5 as  $p$  ( $q = 0.2, q = 0.4$ ), and the effect of  $\gamma$  on epistemic value deceases once  $q$  passes 0.5. Note that the epistemic value is 0 when  $p = q = 0$ , because although the negative ambiguity is maximized in this case, it is counteracted by the entropy term which is 0 since both posteriors are certain.

It is straightforward to show the positive relationship between the entropy of the policy-conditioned beliefs  $H[Q(s_\tau|\pi)]$  and the epistemic value. We begin by isolating further analysis only to the meta-belief hidden state factor  $s^{\text{MB}k}$  for a particular neighbour  $k$  and the corresponding observation modality:  $o^{\text{NT}k}$ . For notational convenience, we let  $s = s^{\text{MB}k}$  and  $o = o^{\text{NT}k}$ . Using this notation, we can then rewrite the epistemic value as the predictive mutual information between states and observations, using the predictive distributions  $Q(o_\tau|\pi)$  and  $Q(s_\tau|\pi)$ :

$$\begin{aligned} E_{Q(o_\tau|\pi)}[D_{KL}[Q(s_\tau|o_\tau, \pi) \| Q(s_\tau|\pi)]] &= D_{KL}[Q(o_\tau, s_\tau|\pi) \| Q(s_\tau|\pi)Q(o_\tau|\pi)] \equiv I_{pred}(O; S) \\ &= \sum_{o_\tau, s_\tau} Q(o_\tau, s_\tau|\pi) \ln \frac{Q(o_\tau, s_\tau|\pi)}{Q(s_\tau|\pi)Q(o_\tau|\pi)} \\ &= -\mathbf{H}[Q(o_\tau, s_\tau|\pi)] + \mathbf{H}_{Q(o_\tau, s_\tau|\pi)}[Q(o_\tau|\pi)] - \sum_{o_\tau, s_\tau} Q(o_\tau, s_\tau|\pi) \ln Q(o_\tau, s_\tau|\pi) \end{aligned} \quad (\text{A46})$$

Using the factorization of the joint posterior predictive density  $Q(o_\tau, s_\tau|\pi) = P(o_\tau|s_\tau)Q(s_\tau|\pi)$ , the final term on the RHS of Equation (A46) can be rewritten:

$$-\sum_{o_\tau, s_\tau} Q(o_\tau, s_\tau|\pi) \ln Q(s_\tau|\pi) = -\sum_{o_\tau, s_\tau} P(o_\tau|s_\tau)Q(s_\tau|\pi) \ln Q(s_\tau|\pi) \quad (\text{A47})$$

$$\geq \mathbf{H}[Q(s_\tau|\pi)] + E_{Q(s_\tau|\pi)}[\ln P(o_\tau|s_\tau)] \quad (\text{A48})$$

where the inequality going from (A47) to (A48) follows from Jensen's inequality. Equation (A48) demonstrates that uncertainty about hidden states (as quantified by  $\mathbf{H}[Q(s_\tau|\pi)]$ ) is directly proportional to the drive to reduce that uncertainty, subject to the log probability of observations expected under hidden states  $\mathbb{E}_{Q(s_\tau|\pi)}[\ln P(o_\tau|s_\tau)]$ .

We can then use the dependence of  $P(s_\tau|s_{\tau-1}, \pi)$  on  $\omega^{\text{Soc}}$  to relate the inverse social volatility to the posterior entropy  $\mathbf{H}[Q(s_\tau|\pi)]$ .

$$Q(s_\tau|\pi) = P(s_\tau|s_{\tau-1}, \pi, \omega^{\text{Soc}})Q(s_{\tau-1}|\pi) = \frac{e^{\omega^{\text{Soc}}P(s_\tau|s_{\tau-1}, \pi)}}{\sum_{s_\tau} e^{\omega^{\text{Soc}}P(s_\tau|s_{\tau-1}, \pi)}}Q(s_{\tau-1}|\pi) \quad (\text{A49})$$

$$\ln Q(s_\tau|\pi) = \omega^{\text{Soc}}P(s_\tau|s_{\tau-1}, \pi) - \ln \sum_s e^{\omega^{\text{Soc}}P(s_\tau|s_{\tau-1}, \pi)} + \ln Q(s_{\tau-1}|\pi) \quad (\text{A50})$$

$$\ln Q(s_\tau|\pi) = \omega^{\text{Soc}}P(s_\tau|s_{\tau-1}, \pi) - C + \ln Q(s_{\tau-1}) \quad (\text{A51})$$

$$\mathbb{E}_{Q(s_\tau|\pi)}[\ln Q(s_\tau|\pi)] = \omega^{\text{Soc}}\mathbb{E}_{Q(s_\tau|\pi)}[P(s_\tau|s_{\tau-1}, \pi)] + \mathbb{E}_{Q(s_\tau|\pi)}[\ln Q(s_{\tau-1}) - C] \quad (\text{A52})$$

$$\implies \mathbf{H}[Q(s_\tau|\pi)] \propto -\omega^{\text{Soc}}\mathbb{E}_{Q(s_\tau|\pi)}[P(s_\tau|s_{\tau-1}, \pi)] \quad (\text{A53})$$

The final line demonstrates that the entropy of the predictive posterior is inversely proportional to the inverse social volatility  $\omega^{\text{Soc}}$ , and thus controls the rate at which the focal agent's uncertainty about their neighbours' belief-states increases, and therefore also determines the epistemic value of policies that entail reading that neighbour's tweets. Intuitively, if an agent believes their social world is volatile (the beliefs of neighbouring agents quickly grow uncertain), the agent will become incentivised to sample those neighbours more frequently, in order to resolve rapidly-growing uncertainty about  $s^{\text{MB}}$ . Importantly, this epistemic value will grow over time for a particular neighbour as long as that neighbour's tweets are not read, so a particular value of  $\omega^{\text{Soc}}$  entails a characteristic 'social re-attendance rate' for each neighbour.  $\square$