scientific reports

Check for updates

OPEN

Collective patterns and stable misunderstandings in networks striving for consensus without a common value system

Johannes Falk[1✉], Edwin Eichler[2,3], Katja Windt[3,4] & Marc-Thorsten Hütt[1]

Collective phenomena in systems of interacting agents have helped us understand diverse social, ecological and biological observations. The corresponding explanations are challenged by incorrect information processing. In particular, the models typically assume a shared understanding of signals or a common truth or value system, i.e., an agreement of whether the measurement or perception of information is 'right' or 'wrong'. It is an open question whether a collective consensus can emerge without these conditions. Here we introduce a model of interacting agents that strive for consensus, however, each with only a subjective perception of the world. Our communication model does not presuppose a definition of right or wrong and the actors can hence not distinguish between correct and incorrect observations. Depending on a single parameter that governs how responsive the agents are to changing their world-view we observe a transition between an unordered phase of individuals that are not able to communicate with each other and a phase of an emerging shared signalling framework. We find that there are two types of convention-aligned clusters: one, where all social actors in the cluster have the same set of conventions, and one, where neighbouring actors have different but compatible conventions ('stable misunderstandings').

Self-organisation in networks of communicating agents is a fundamental process for the emergence of order in social systems [1–3]. Such spontaneous order occurs in the form of consensus [4], social segregation and opinion formation [5,6] or as the establishment of norms and conventions, e.g. a language or set of signals [7–11]. Beyond the social sciences, also Biology [12,13] and Computer Science [14–18] are interested in how interacting agents do and could organise.

Consequently, there exist a variety of models that analyse or explain different aspects of self-organisation within communities: In the Voter Model [19–21] or the Axelrod Model [22], spatially distributed agents copy or invert opinions or attributes from their neighbours according to certain rules, globally leading to a transition between an ordered homogeneous and a disordered state [23–26]. The Lewis Signalling Game [27] explains how a sender and a receiver agree on common signals and the naming game [28,29] describes how agents agree on a common vocabulary for objects in their environment. Another approach is the analysis of learning in social networks [30]. As an example, DeGroot [31] showed how agents that start with individual (subjective) knowledge about a parameter can reach a consensus by repeatedly assimilating information about the parameter that they observe in the rest of the group. Besides these purely theoretical studies, there are also findings with empirical support. Building a dynamical model around given data of scientific collaboration it was e.g. possible to find signatures of self-organisation in real social processes [32].

Less investigated is the emergence of order (or synchronisation) under erroneous or subjective perception and in the absence of a common value system, an objective instance or a universal truth. Exceptions to the first point (erroneous perception) include the effect of information processing noise on vote dynamics [33,34] or naming games [35]. The lack of a common value system has partly been the motivation behind variations of the naming game [28,29,36], the signalling game [10,29,37–42] and several other models [43–45], even though all these models still rely on the global notion of 'true' and 'false' or a shared interpretation of signals/actions.

Here, we present a model that addresses both challenges—erroneous perceptions and lack of a universal truth—simultaneously. Our model comprises a network of agents that each only has a subjective perception of

[1]Department of Life Sciences and Chemistry, Jacobs University, Bremen, Germany. [2]EICHLER Consulting AG, Weggis, Switzerland. [3]SMS Group GmbH, Düsseldorf, Germany. [4]Global Production Logistics, Jacobs University Bremen, Bremen, Germany. ✉email: j.falk@jacobs-university.de
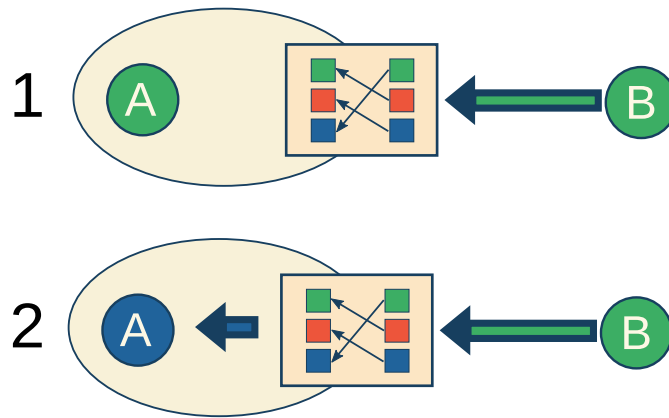
Figure 1. (top) *A* observes the colour of agent *B*. From an objective view, *B*'s colour is green. (bottom) However, due to *A*'s dictionary, *A* observes blue and changes its colour accordingly.

the world. Thereby, they are faced with a cognitive dissonance between their own cognition and what they perceive as their neighbours' cognition [46]. In order to reach conformity [47], each agent strives to minimise its cognitive dissonance, based only on its own subjective observations. We show that under these conditions collective behaviour still can emerge. We also show that stable misunderstandings can form, i.e. an emerging pattern where many nodes are without perceived conflict, despite heterogeneity in subjective perception. Often this emerges as an alternating arrangement of compatible but distinct subjective perceptions on suitable network topologies. To elaborate on the interplay between our model's dynamics and the topology of the underlying network, we analyse the dynamics on regular lattices and random regular graphs. The results help to better understand the emergence of order within a connected community of agents without an objective instance.

Our model is motivated by Gotthard Günther's polycontextural logic [48,49], where two subjects observing the same situation can come to different conclusions, even when each of the subjects adheres to binary (but distinct, contexture-dependent) logic. We hence denote our model *polycontextural networks*.

Our findings not only have an impact on sociological questions but also on the ongoing problem of how distributed machine learning systems can negotiate a common signalling system. It is also an ongoing debate in philosophy how a consensus can emerge out of observer-dependent facts [50].

The remainder of the paper is organised as follows: In the next section, we introduce our model and illustrate its static properties with two simple network motifs. In Sect. 3 we then analyse the dynamics of the model on random regular graphs, before we focus on triangular and square lattices to investigate the observed self-similarity and develop a mechanistic understanding of the observed dynamics. Subsequently, in Sect. 4 we discuss the implications of our model and draw some conclusions in Sect. 5.

Model

The polycontextural network is a simple model where N agents interact over a network. Each agent A_n with $n \in N$ is equipped with a characteristic c_n whose expression is taken from a pool of size C. To simplify the notation, the characteristic of each agent is given as a standard basis vector $\boldsymbol{e}_i$ of length C with 1 in the ith position and 0 in every other position.

To incorporate subjectivity, each agent has an individual dictionary that bijectively maps the 'outside world' of the agent to its personal cognition. Formally, this dictionary is a bijective function $\sigma : C \rightarrow C$ and can be written as a $C \times C$ permutation matrix T_n. If one agent A_n observes the characteristic c_m of another agent, the observing agent sees $T_n c_m$ instead of the 'true' (objective) c_m (as depicted in Fig. 1). To give it an intuitive meaning, in the following we will assume that the characteristics c_n are colours. Due to this definition, our model does not have objective truth values—a predefined understanding of colour—but $C!$ different and equally correct world-views (here: colour mappings). In the following, we understand the term *world-view* to mean a set of truth values that determine how an agent perceives the environment. Each world-view hence refers to a specific choice of a value system.

The different agents are spatially distributed and partially connected, whereby they form a network structure where each agent is one node. During this investigation, we will analyse our model with different network topologies of different sizes.

The only interaction in our model is a simple version of social influence where all agents strive for consensus. A single update step of our model proceeds similarly to the standard voter model with C different opinions and asynchronous dynamics, which means that a randomly chosen agent A_n adopts the opinion (the colour) of one of its neighbours A_m. However, and in difference to the voter model, A_n can not observe the 'true and objective' colour c_m but sees the characteristic $T_n c_m$. A sequence of N updates forms a time step, which means that on average at every time step every node is selected once.

While formally similar to the voter model, due to the different world-views (T_x) and hence the different perception of colours, the model's dynamic would in general not converge to a uniform colouring. To illustrate this, let us imagine two nodes A and B, which can be either red or blue. The nodes shall have two different, but
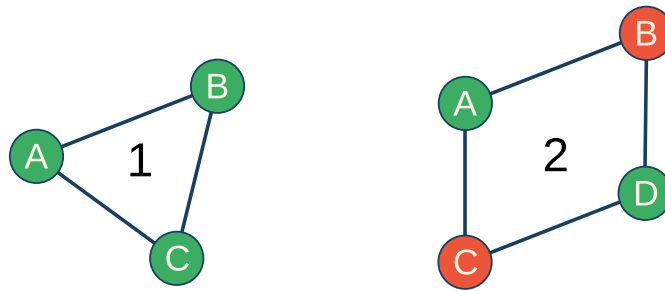
Figure 2. Two different minimal models to illustrate possible solved states of our model.

fixed world-views: A recognises colours as they are and B recognises the colours reversed (blue to red, red to blue). If A is red and is observed by B, B will turn blue. In some subsequent time step, A will observe B and also turn blue. However, as soon as B observes A again, B will turn red and the process starts again.

What is missing is the single agent's ability to sense that its own world-view is not aligned with that of its neighbours. According to relational epistemology [51] and the philosophy of world-views [52,53], world-views are shaped by and changed according to lived experiences and determine how one understands the world and responds to it.

Following this concept, within our model each agent is equipped with two internal counters: O_n and K_n, and every update step proceeds as follows:

- One agent A_n and one of its neighbours A_m are selected randomly following a uniform distribution.
- Agent A_n subjectively observes the feature of A_m, which means A_n sees the characteristic $T_n c_m$.
- If agent A_n's own characteristic is already equal to the observed one, A_n only increments its internal O_n counter by 1.
- Otherwise, A_n changes its own characteristic to the subjectively observed one and increments both its internal O_n and K_n counter by one.
- If the fraction K_n/O_n is larger than the parameter q (which means that in more than q percent of the observations the observed characteristic was not equal to the own), the agent changes its own world-view T_n to a random selection out of the $C!$ possibilities and resets both counters to zero. Note that O_n gets always incremented at least once before this last step. The fraction is hence always defined.

The individuals in our model hence share a predefined response once a threshold number of conflicts is detected and, following the definitions given in Ref [54], our model would have to be considered to belong to the group of quorum sensing models, although the term 'subjective quorum-sensing' would probably fit best. In terms of everyday experience, it may seem strange that opinions in our model are changed immediately, regardless of past observations. However, what our model reflects are the different time scales for a change of opinion vs. a change of world-view.

The illustrative idea behind the dynamics of the model is that agents react—according to their subjective interpretation—to the states of other agents. Since each agent is also an object of other observations, the state of the observed neighbours is sometimes already the reaction to the observation of the own state. This enables each agent to perform a self-reflection and a repeated observation of the neighbour's state can hence indicate whether the neighbourhood confirms the own world-view (for a more detailed philosophical interpretation see Sect. 4). The parameter q (threshold parameter) therefore controls how sensitive an agent decides that its own belief system does not conform to the neighbourhood, subsequently changing it.

Before we proceed, we define two special terms to simplify the notation: Two connected nodes i, j are called *compatible*, if their dictionaries mutually agree in all colours, which means $T_i \times T_j = \mathbb{I}$. A network is considered *solved* if all connected nodes are compatible. Note that our definition of compatible and solved networks only depends on the dictionary of the nodes and not on the current colour. As we will show in the following examples, a solved network does not imply that no more colour changes occur.

Two illustrative examples. Let us refer to two minimal network motifs as illustrative examples. Figure 2(1) shows three connected nodes (a triangular structure). Following the above-given definition of a solved system, node A has to be compatible with respect to both nodes B and C, which means:

$$T_A \times T_B = \mathbb{I} \tag{1}$$

$$T_A \times T_C = \mathbb{I}. \tag{2}$$

At the same time, the nodes B and C need to be compatible as well, hence:

$$T_B \times T_C = \mathbb{I}. \tag{3}$$

This is only possible if

$$T_A = T_B = T_C = T_A^{-1}, \tag{4}$$

which is true for all 2-cycle permutation matrices. Note that the definition of a solved system is a local definition and does explicitly not imply that no more colour changes occur. Let us hereto assume that $C = 3$ and for all dictionaries it holds:

$$T_{\{A,B,C\}} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5}$$

Let us further assume that node A is in state $c_A = (1,0,0)$. If both node B and C observe A, both change their states to $c_{A/B} = (0,1,0)$. However, if now B observes C it needs to change its state again to $c_B = (1,0,0)$ which would create a further state change after an observation by A, and so on. For the system to reach a state without any more colour changes, it needs to hold:

$$T_A T_B T_C = T_A^3 = \mathbb{I}. \tag{6}$$

This would require T_A to be a 3-cycle permutation matrix. However, since we already know that T_A needs to be a 2-cycle permutation matrix this is only possible if T_A is the identity.

Things are different for the four-node system shown in Fig. 2(2). For A it again holds:

$$T_A \times T_B = \mathbb{I} \tag{7}$$

$$T_A \times T_C = \mathbb{I}. \tag{8}$$

However, B and C are not connected directly, but only via D. To reach a solved system there are hence two more equations that need to be satisfied:

$$T_D \times T_B = \mathbb{I} \tag{9}$$

$$T_D \times T_C = \mathbb{I}, \tag{10}$$

which means that:

$$T_A = T_D = T_B^{-1} \tag{11}$$

$$T_B = T_C = T_A^{-1}. \tag{12}$$

In contrast to the triangular structure, these constraints cause a system where no more changes of colour will happen. To see how this comes, let us start with node A and cycle over the other nodes. It follows:

$$T_B T_D T_C T_A = T_B T_A T_B T_A = T_B T_B^{-1} T_B T_B^{-1} = \mathbb{I}. \tag{13}$$

Since in this configuration a system can be stable although the agents have different world-views, we call this novel effect a *stable misunderstanding*. The important point here is that—due to their subjectivity—the nodes involved can not detect this misunderstanding. To our knowledge, our polycontextural network model is the first that allows for and demonstrates the impact of such stable misunderstandings.

One might argue that this observed effect of stable misunderstandings is just an artefact of whether the considered cycle has an even or odd number of nodes. While this is true from a mathematical point of view, we argue that the artefact mainly arises because of our simplified world-view representation. We assume that for each node perception of colours takes place deterministically (no single colour is perceived in multiple ways without a change in world-view) and without information loss (two distinct colours are never perceived as the same colour; no 'colour blindness'). This set of requirements leads to the restriction to bijective translation tables. If we would e.g. allow for non-bijective associations within the translation tables, the effect would vanish. However, to keep the model simple and comprehensible, we will stick to our definition.

Over time, we observe that connected nodes synchronise their world-views and form clusters of nodes with the same understanding of the world. We are hence mainly interested in the dynamics and organisation of the world-views (the tables). The (fluctuating) colours of the nodes are just signs (their language) to communicate with their surroundings and are—in our investigation—only of limited interest. In the following section we show that, depending on the value of the threshold value, the sizes of these table-clusters stay small or expand over all scales, indicating a critical state and a phase transition. Besides the change of the threshold parameter, we also demonstrate how the topology of the network affects the type and size of clusters.

In what follows, we present the results of the simulations for four different network topologies. To avoid any grid artefacts, we first analyse random 3-regular and 4-regular graphs. To better understand and visualise the dynamics of our model we then focus on regular triangular and square lattices. For all models, we set $C = 3$ and—to avoid boundary effects—use periodic boundary conditions if applicable.

Results

Phenomenology.
We will first illustrate the time-dependent behaviour of our model before we proceed to analyse the dependence on the q parameter. Figure 3 shows the time evolution of colour changes per time step, table changes per time step and the relative mean size of clusters of the same translation table T_X within a
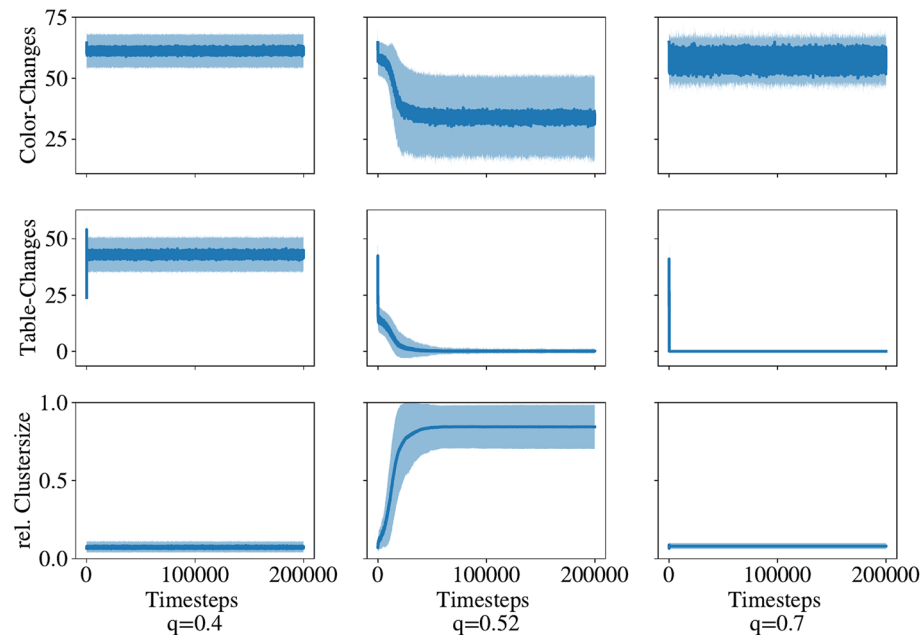
Figure 3. Time series of colour changes per time step, table changes per time step and relative mean size of the clusters for different values of the parameter q (average over 100 runs). The light blue area indicates the standard deviation. The analysed network is a random 4-regular graph with 200 nodes.

random 4-regular lattice with 200 nodes for three different values for q. If the threshold value is small ($q < q_c$) the nodes constantly change their world-views, corresponding to a constant number of table changes. These frequent changes prevent a build-up of clusters with equal world-views. Since no clusters of compatible world-views emerge, there are also frequent colour changes. If the threshold value is too large $q > q_c$, there are only a few table changes before the system reaches a frozen state (the agents are satisfied with their world-view). The dynamics freeze before large clusters can emerge, hence neighbouring world-views are incompatible and frequent colour changes are observed. For an intermediate value $q \approx q_c$ we observe a slow decrease in the change of the world-view tables that finally leads to the build-up of a large cluster, similar to the behaviour observed in naming games [55] or in threshold voter models [7].

To gain insight into how the growth of clusters is taking place, Fig. 4 shows a snapshot of a random 4-regular graph. The colours of the nodes indicate the time span the respective node has not changed. The more yellowish the node, the longer no change. For illustrative purposes, the six nodes that did not change for the longest have been shifted out of the bulk. As can be seen from the figure, three of these nodes each form a separate triangular motif. These triangular motifs were the nuclei for the growth of a large cluster of shared world-views. The evolution of the cluster sizes can thus be understood as a nucleation and coarsening process. If the cluster does not span the full system we can still see colour fluctuations, however, the colours change less often than in the other two cases. There is hence an intermediate regime between continuous table changes and an immediately frozen state, where world-views can synchronise and cluster.

q-Dependence. To gain more insight into the critical behaviour, in the following we analyse in detail how the model behaves under a change of the parameter q. Figure 5 shows exemplary behaviours of the relative size of the largest cluster over the value of q for both a random 3-regular and 4-regular graph of size $N = 200$. For a small window of q the largest cluster spans the full system, indicating system-wide correlations between the tables. Additionally, the lower plots show the fraction of the six different translation tables, where T_A denotes the identity, $T_{\{B,C,D\}}$ denote the three possible 2-cycle matrices and $T_{\{E,F\}}$ denote the two 3-cycle matrices, with $T_E \times T_F = \mathbb{I}$.

In terms of the motivation of our model the results so far already prove the emergence of a consensus. However, to better understand the characteristics of our model as well as its properties at criticality, we will now turn to two simplified network topologies and analyse their behaviour.

Regular lattices. To gain more insight into the critical behaviour of our polycontextural network model, in the following we focus on regular triangular and square lattices. Figure 6 shows a snapshot of two critical systems after the evolution of $t = 20,000$ steps. Here, the colours of the nodes do not indicate their current colour but are illustrations for their respective translation table. Depending on the topology of the network, the clusters are only formed by equal translation tables or also by patterns of two alternating tables, indicating the occurrence of stable misunderstandings. This is in line with the analysis of the motifs in Sect. 2.

Figure 7a, b shows the mean cluster size over the value of q for different system sizes for both the triangular and square lattices. As expected, for small values of q the mean cluster size is close to one: the fluctuations in the
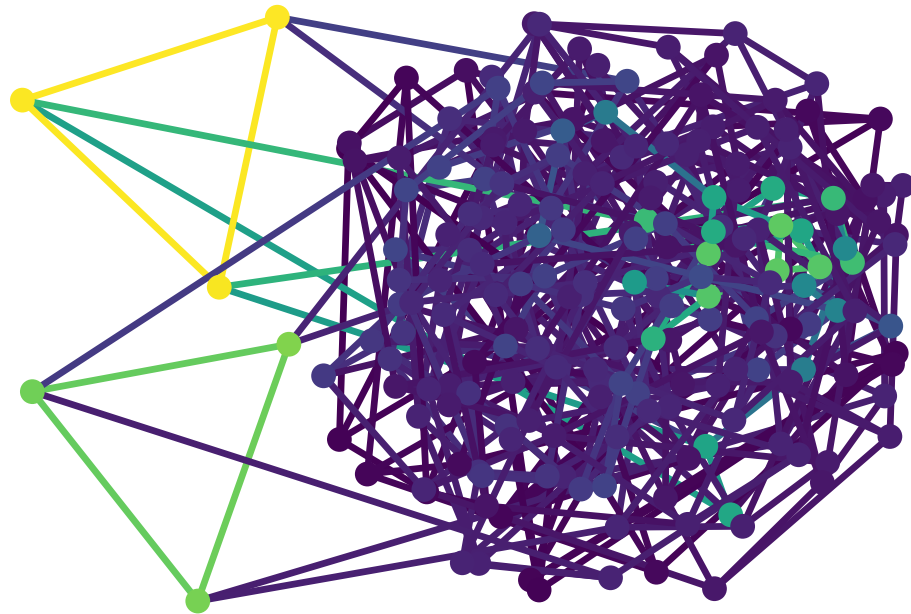
Figure 4. Illustration of a random 4-regular graph. The colours indicated the time since the last change of the respective node. The more yellowish the node, the longer no change of the translation table. The six nodes with the longest time since the last change have been shifted out of the bulk for illustrative purposes.
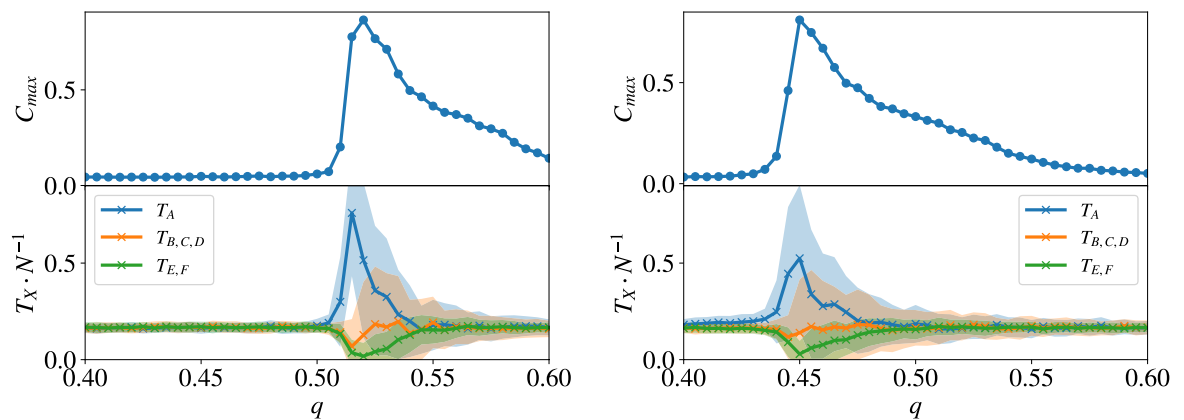

Figure 5. (upper) Relative cluster size of the largest cluster C_{max} vs. the threshold parameter q on an (left) random 3-regular graph and (right) random 4-regular graph ($N = 200$, after $t = 80{,}000$ simulated steps, averaged over 100 runs). (lower) The fraction of the different tables depending on q. Here, T_X denotes the different types of possible translation tables as described in the full text. The light area around the curves indicates the standard deviation.

system do not permit the buildup of correlations. At a rather sharp value $q = q_c$ we observe a sudden jump in the cluster size, indicating a phase transition. Then, for values $q > q_c$ the cluster size slowly decreases and converges to the initial mean cluster size of one. It is important to note that for $q \gg q_c$ the typical size of clusters is much smaller than the system size and hence the mean cluster size is not limited by the system size.

As already for the random regular graphs, a key quantity for our system is the size of the largest cluster. Figure 7c, d shows exemplary behaviours of the relative size of the largest cluster. For a small window of q the largest cluster of the system spans the full system, indicating system-wide correlations between the tables. Additionally, the lower plots show the fraction of the six different translation tables, where T_A denotes the identity, $T_{\{B,C,D\}}$ denote the three possible 2-cycle matrices and $T_{\{E,F\}}$ denote the two 3-cycle matrices, with $T_E \times T_F = 1$. As already observed in Fig. 6 the topology of the underlying network determines which translation tables cluster: At criticality, the triangular network predominantly consists of nodes that hold the identity matrix T_A. Contrarily, in the grid network, all translation tables have the same probability. However, as the diverging variance at the critical point already indicates, this is just an averaging effect. In a single system, only one type of cluster configuration wins, but the probability to win is equal for all configurations.

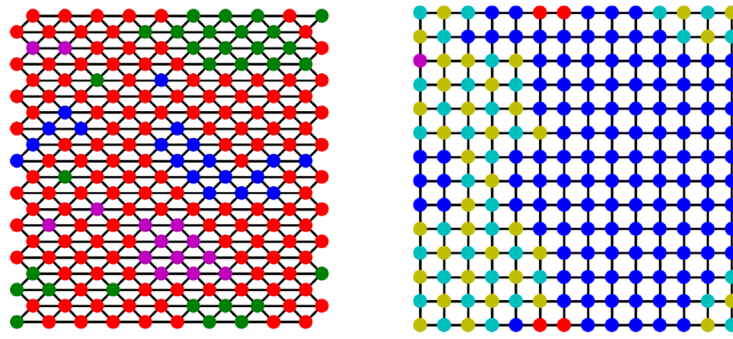We will now turn to an analysis of the scale-freeness of the cluster size distribution.

Figure 6. Snapshot of a (left) triangular and (right) square lattice for $q \approx q_c$. The colours encode the six possible translation-tables. The repeating patterns on the square lattice are *stable misunderstandings*.
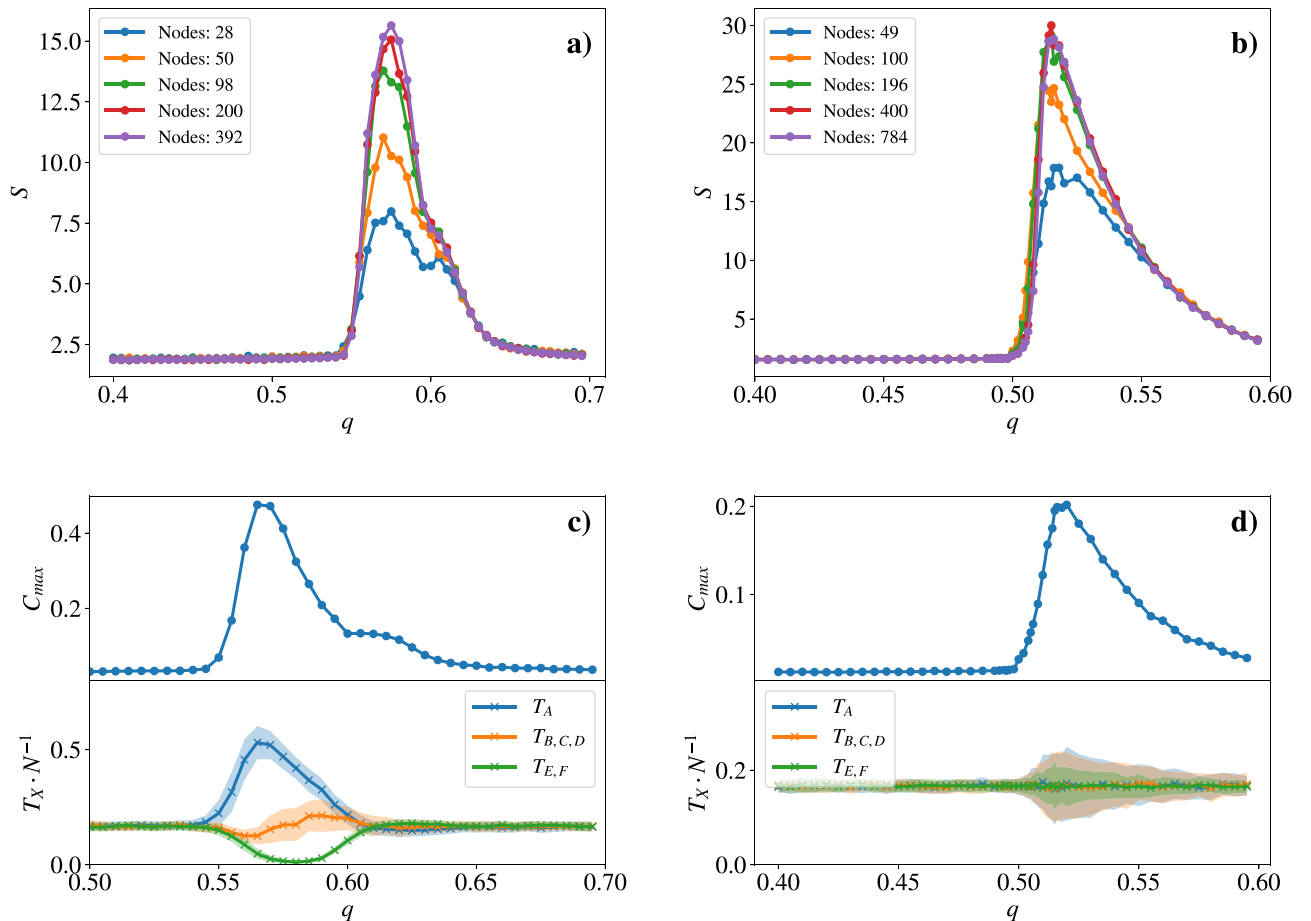

Figure 7. Variation of the mean cluster size (S) vs. the threshold parameter q for different numbers of nodes in a (**a**) triangular and (**b**) square lattice (Simulated steps $t = 20,000$, averaged over 100 runs). For $q \gg q_c$ the typical sizes are much smaller than the system size and the cluster size does not depend on N. For $q \approx q_c$ finite-size scaling appears and the cluster size depends on the system size. The similar behaviour of the two largest networks indicates that these systems did not finish growth after the 20,000 time steps. (**c**)/(**d**) (upper) Relative cluster size of the largest cluster C_{max} vs. the threshold parameter q on an triangular ($N = 392$) and square ($N = 784$) lattice (After $t = 20,000$ simulated steps, averaged over 100 runs). (lower) The fraction of the different tables depending on q. Here, T_X denotes the different type of possible translation tables as described in the full-text. The light area around the curves indicates the standard deviation. (After $t = 10,000$ simulated steps, averaged over 100 runs).
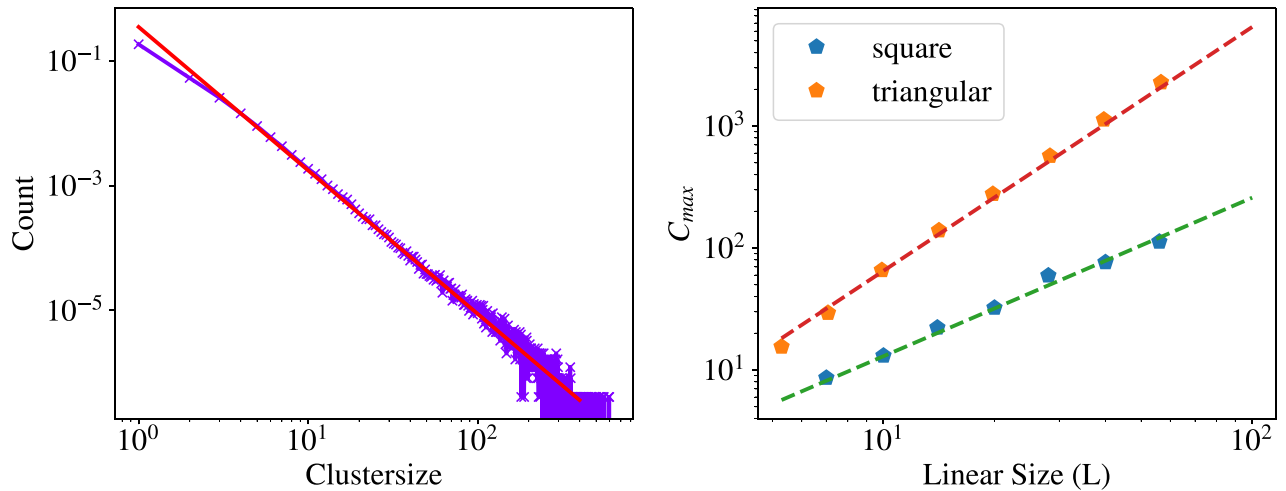
Figure 8. (left) Cluster size distribution of the triangular network for $q = 0.57$ at $t = 5000$ ($N = 2450$ nodes, averaged over 1000 runs). The red line indicates a power law with exponent $\alpha = -2.3$. (right) Size of the largest cluster C_{max} over the linear size $L = N^{0.5}$ of the system. Both systems were simulated with a $q \approx q_c$ for $t = 20,000$ time steps. The dashed lines indicate the respective power-law fits with (green) $d_f = 1.3$ and (red) $d_f = 2$.

Self-similarity. In some regard, the dynamics of our model are similar to grain growth and coarsening processes as observed in crystallisation. The clusters with compatible tables in our model (local consensus) play the role of domains with similar orientations in the coarsening process. A successful model to simulate such crystallisation processes is the Monte Carlo Potts Model [56]. The Potts model is defined on a spatial grid where each grid point (often called Monte Carlo Unit) can be in one of Q possible states that describe the respective orientation. At each time step, a random grid point is selected. Then, similar to the selection of a new translation table in our model, a new orientation is assigned to the selected grid point. However, in the Potts model, this assignment is on probation. The new orientation is selected with a probability that depends on the energy difference between the old and new state (with regard to the interactions with the nearest neighbours) as well as on an external parameter, the temperature T. It is well known that the cluster sizes in grain growth and particle coarsening show self-similarity [56,57]. This means that during coarsening different system configurations reveal similar behaviour when scaled to the same scale: they are scale-invariant. Whether such self-similarity is also present in social dynamics is an ongoing debate [58]. Due to the similarities between our model and the Potts model it is hence natural and interesting to ask if we also observe a scale-invariant behaviour in our model. A well-known characteristic of self-similarity is a power-law distribution of observable quantities. In the case of our model, this could e.g. be the cluster size distribution. In Fig. 8 (left) we show this distribution for the triangular network at $q = 0.57$, slightly larger than the critical value $q_c = 0.56$. In a log-log plot, this distribution shows a linear behaviour with slope $m = -2.3$, indicating a power law with an exponent of $\alpha = -2.3$. This cluster size distribution is hence scale-invariant. In terms of social systems, this would mean that an opinion structure that is found in small communities is equal to the structure that is found in large systems of interconnected communities.

Self-similarity of the cluster size distribution would also imply that the size of the largest cluster C_{max} scales with the linear system size L according to:

$$C_{\max}(L) \sim L^{d_f} \tag{14}$$

where d_f is the (possibly) fractal dimension of the cluster [59,60]. In Fig. 8 (right) we observe the expected scaling for both lattices analysed and obtain the corresponding values for d_f. The quadratic lattice growth with a fractal dimension of 1.3 indicating a rough boundary, whereas the triangular lattice growth with a fractal dimension of 2, equal to the spatial dimension of the lattice, hence there is a high 'surface tension' and the clusters are more compact with a smooth boundary. One should note that this determination of d_f is not very accurate and one would need another approach to obtain a more exact version. This is, however, out of the scope of this manuscript and will be left for an upcoming publication.

With the results obtained, we are now in the position to explain mechanistically how clusters in our system are formed and how the observed self-similarity can be explained. Hereto, we draw on Fig. 9. If $q < q_c$ the system is in phase 1. There are too many fluctuations in the system such that no clusters of shared tables can emerge and possibly existing clusters are destroyed. For $q > q_m$ (phase 3) there is too little activity. Every single node behaves as a single nucleus of a new cluster and does not adapt to join other, possibly larger clusters. Upon a decrease of q below q_m neighbouring agents begin to form clusters. However, small nuclei of possibly incompatible clusters appear all over the system and grow (with dimension d_f) until they reach the boundary of other clusters. The result is a cluster-cluster competition between different incompatible clusters as it is also known for models like the naming game [54]. The smaller q the smaller is the probability that an initial nucleus appears. Close to q with $q > q_c$ there is only a very small probability for an initial nucleus, but once a first nucleus is stabilised it can grow over the full system without being limited by another growing cluster.
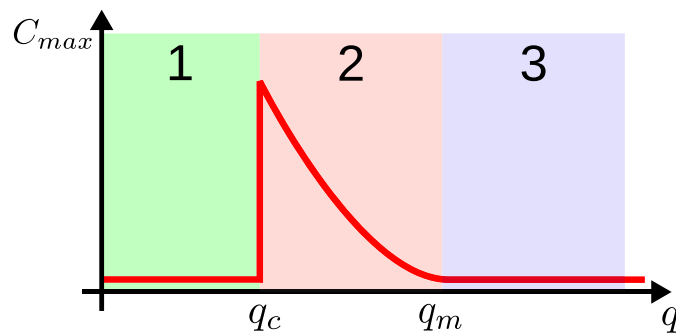
Figure 9. Phenomenological separation of three different phases. In phase 1 there is too much activity in the system. In phase 3 there is too little activity. Clusters can only grow in phase 2.

Discussion

We have shown that our simple model, the polycontextural network, has a phase where the world-views cluster globally, leading to a shared perception of signals, or global stable misunderstandings. In the following, we show how our model fits into the landscape of established models and discuss its implications.

Clustering of opinions. The clustering of opinions within human populations is an important and ongoing research topic [6]. Based on empirically validated mechanisms like 'homophily', different models tried to explain how opinion clustering might happen [61]. Most of these models are able to show a clustering of opinions. Depending on the detailed mechanisms of the model, adding noise can facilitate mono culture or maintains pluralism [6,62]. A prominent source of such noise is the misinterpretation of information. Starting from an initial configuration where no predefined definition of a right or wrong interpretation of signals is given, our polycontextural network shows the build-up of a shared perception of signals. Our model hence provides a framework to understand the basic mechanisms of how a first basic understanding of information might emerge. A prerequisite most current models of social sciences (implicitly) rely on.

Polycontextural logic. Binary logic is a cornerstone of western thought and technology and hence an important component in our decision strategies and opinion formation processes. As a consequence, social dynamics around opinion formation are challenged by situations where facts can not easily be mapped to global and objective 'true' and 'false', as is e.g. the case in modern social phenomena, like the emergence and prevalence of fake news or the counter-phenomenon of fact-checking entities: *What is obvious to us may be a lie from hell to our neighbour* [52]. When facts are not perceived equally by two distinct observers, each of whom adheres to binary logic, we are in a situation, which the philosopher Gotthard Günther attempted to capture with his theory of polycontexturality [48]. He argues that each of the observers is applying the framework of Aristotelian logic consistently within this observer's own realm of observation, called *contexture*. Each contexture has its own set of factual embedding of true and false. This alignment of their own subjective understanding with the community requires that each individual is able to self-reflect their own understanding. Based on Hegel's dialectics, Günther formally analysed how living beings with only subjective perception can interact and how they can become aware of their own subjectivity [48,63]. He assumed that every ordered combination of an observer (a subject) and the observed object form a contexture that each has its own classical two-valued logic. He found the mutual interaction between *three* contextures to be a crucial requirement for successful communication. The three contextures are arranged in a structure as shown in Fig. 10, which Günther termed *proemial relation*. In the first contexture (C1) an object (O) is observed by the subject S1. This contexture can become the observed object of a second subject (S2). Thereby, S2 observes the object of the first contexture *as observed* by the observer S1 from the first contexture [64]. Subsequently, within a third contexture S2 can compare the original object with its subjectified version. The proemial relation hence allows the single observers to reflect their own understandings of the world.

The dynamics of our polycontextural network can be interpreted in this manner: Let us assume two nodes A and B. The colour (the fact) of node A corresponds to the object O in the first contexture. This colour can subjectively be observed by node B (S1), which colours itself according to the result of this observation. Now, in the second contexture, node A (S2) can observe the colour of B (S1). Last (the third contexture), node A compares its own colour (O) with node B's colour (as observed by A itself) which is (based on A's standpoint) the own colour through the eyes of another [65]. Following this dynamics, node A is able to notice a possible misalignment between its own and node B's world-view (translation table). Our results hence indicate how the subjective observation of observers enables a self-reflection that can lead to the emergence of shared signals and provides a numerical example of Günther's and Hegel's philosophy.

Within social systems, communication and influence often lead to each of the social entities gathering 'followers', supporters of their particular interpretation of a given set of facts and, hence, of their respective contexture. Contextures in this way become entrenched in society. Similarly, in our model, we observe the growth of clusters of similar world-views.
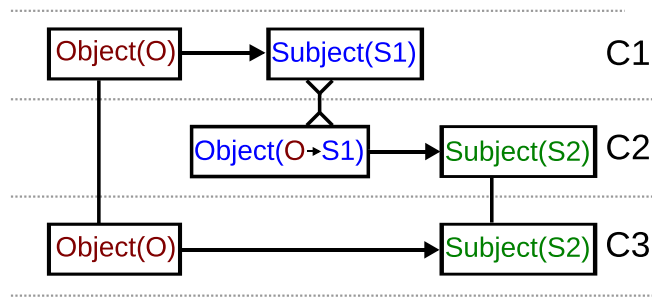
Figure 10. Sketch of the proemial relation. In each contexture (C1, C2, C3) a subject observes an object as indicated by the arrows (The direction of the arrows indicates the flow of information; Object → Subject reads: Subject observes Object). Figure and caption taken from [63].

Stable misunderstandings. Misunderstandings are a commonplace in communications [66]. Often, these misunderstandings arise from different interpretations of a message [67,68], whereby the interpretations are influenced by factors like personality and values [69]. Within our model, we observe the situation of *stable misunderstandings*, where two agents with different world-views interpret messages differently but in a mutually compatible form. One might argue that such a permanent misunderstanding is an artificial situation and just an artefact of our model's design. However, misunderstandings are often subtle, persistent and difficult to uncover [70]. In this regard, a nice example are students' misunderstandings of quantum mechanical wave functions [71]. Our model can be understood as a minimal model of such misunderstandings.

Conclusion

In this manuscript, we presented a new model to explain if and how consensus can appear between agents that can only judge based on a subjective understanding of the world. Focusing on two regular lattices, the triangular and the quadratic lattice, as well as random regular graphs we observed the emergence of a system-wide (and then by definition objective) understanding of signs within the network. This emergence depends on a single parameter that controls the volatility: If the agents are too volatile they change their convictions too often to form clusters of shared understandings. If the agents are not volatile enough they do not adapt to majority opinions and locally separated clusters of different convictions appear. Both phases are connected by a phase transition and only at the transition point the growth of a spanning cluster is possible.

The findings of our model add to several ongoing discussions in social science as well as philosophy and computer science. Obviously, the study of our model on different types of networks is not exhausted. In this manuscript, we have restricted our analysis to the regular triangular and square grid lattices which already showed—especially in terms of the cluster composition—two quite different behaviours. The next natural step could hence be to observe how fast a consensus can be reached on random networks like ER or BA graphs. Here, it might also be interesting to introduce a degree-dependent threshold parameter such that nodes with a high degree are more convinced of their position and change their tables rarer. Additionally, it could be worthwhile to transform the dynamics of the model to a mechanism comparable to the q-voter model where each node observes q neighbours and only changes its table if a given fraction of these neighbours have a wrong colour. To model people's behaviour a little more realistically an interesting modification of the model would also be to reduce the memory of the agents [72]. Instead of the possibly infinite memory, one could restrict the agents to only remember the last X observations. This would create more fluctuations and could avoid the creation of small but stable minority clusters. Additionally, it is interesting to increase the number of possible colours C and thereby the number $C!$ of different world-views. This leads to a larger variety of possible misunderstandings and the possibility of stable and partially compatible world-views.

References

1. Kearns, M., Suri, S. & Montfort, N. An experimental study of the coloring problem on human subject networks. *Science* **313**, 824–827. https://doi.org/10.1126/science.1127207 (2006).
2. Kearns, M., Judd, S., Tan, J. & Wortman, J. Behavioral experiments on biased voting in networks. *Proc. Natl. Acad. Sci.* **106**, 1347–1352. https://doi.org/10.1073/pnas.0808147106 (2009).
3. Helbing, D., Yu, W. & Rauhut, H. Self-organization and emergence in social systems: modeling the coevolution of social environments and cooperative behavior. *J. Math. Sociol.* **35**, 177–208. https://doi.org/10.1080/0022250X.2010.532258 (2011).
4. Gavrilets, S., Auerbach, J. & van Vugt, M. Convergence to consensus in heterogeneous groups and the emergence of informal leadership. *Sci. Rep.* **6**, 29704. https://doi.org/10.1038/srep29704 (2016).
5. Flaxman, S., Goel, S. & Rao, J. Filter bubbles, echo chambers, and online news consumption. *Public Opin. Q.* **80**, 298–320. https://doi.org/10.1093/poq/nfw006 (2016).
6. Mäs, M., Flache, A. & Helbing, D. Individualization as driving force of clustering phenomena in humans. *PLoS Comput. Biol.* https://doi.org/10.1371/journal.pcbi.1000959 *(2010).*
7. Ehrlich, P. R. & Levin, S. A. The evolution of norms. *PLoS Biol.* **3**, e194. https://doi.org/10.1371/journal.pbio.0030194 (2005).
8. Hawkins, R. X., Goodman, N. D. & Goldstone, R. L. The emergence of social norms and conventions. *Trends Cogn. Sci.* **23**, 158–169. https://doi.org/10.1016/j.tics.2018.11.003 (2019).

9. Lipowska, D. & Lipowski, A. Emergence of linguistic conventions in multi-agent reinforcement learning. *PLoS ONE* **13**, e0208095. https://doi.org/10.1371/journal.pone.0208095 (2018).
10. Barr, D. J. Establishing conventional communication systems: is common knowledge necessary?. *Cogn. Sci.* **28**, 937–962. https://doi.org/10.1207/s15516709cog2806_3 (2004).
11. Spike, M., Stadler, K., Kirby, S. & Smith, K. Minimal requirements for the emergence of learned signaling. *Cogn. Sci.* **41**, 623–658. https://doi.org/10.1111/cogs.12351 (2017).
12. Oliphant, M. & Batali, J. Learning and the emergence of coordinated communication. *Center Res. Lang. Newslett.* **11**, 1–46 (1997).
13. Erez, A., Byrd, T. A., Vennettilli, M. & Mugler, A. Cell-to-cell information at a feedback-induced bifurcation point. *Phys. Rev. Lett.* **125**, 048103. https://doi.org/10.1103/PhysRevLett.125.048103 (2020).
14. Grolik, S. *Information logistics. Decentralized approaches of information allocation in information exchange networks* (ibidem Press, Hannover, 2012).
15. Anthony, R. J. Emergence: a paradigm for robust and scalable distributed applications. In *International Conference on Autonomic Computing, 2004. Proceedings.*, 132–139, https://doi.org/10.1109/ICAC.2004.1301356 (2004).
16. Kirby, S. Natural language from artificial life. *Artif. Life* **8**, 185–215. https://doi.org/10.1162/106454602320184248 (2002).
17. Staab, S., Santini, S., Nack, F., Steels, L. & Maedche, A. Emergent semantics. *IEEE Intell. Syst.* **17**, 78–86. https://doi.org/10.1109/5254.988491 (2002).
18. Steels, L. Evolving grounded communication for robots. *Trends Cogn. Sci.* **7**, 308–312. https://doi.org/10.1016/S1364-6613(03)00129-3 (2003).
19. Clifford, P. & Sudbury, A. A model for spatial conflict. *Biometrika* **60**, 581–588. https://doi.org/10.1093/biomet/60.3.581 (1973).
20. Holley, R. A. & Liggett, T. M. Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* **3**, 643–663 (1975).
21. Redner, S. Reality-inspired voter models: a mini-review. *C. R. Phys.* **20**, 275–292. https://doi.org/10.1016/j.crhy.2019.05.004 (2019).
22. Axelrod, R. The dissemination of culture: a model with local convergence and global polarization. *J. Conflict Resolut.* **41**, 203–226. https://doi.org/10.1177/0022002797041002001 (1997).
23. Shao, J., Havlin, S. & Stanley, H. E. Dynamic opinion model and invasion percolation. *Phys. Rev. Lett.* **103**, 018701. https://doi.org/10.1103/PhysRevLett.103.018701 (2009).
24. Gambaro, J. P. & Crokidakis, N. The influence of contrarians in the dynamics of opinion formation. *Phys. A* **486**, 465–472. https://doi.org/10.1016/j.physa.2017.05.040 (2017).
25. Klemm, K., Eguíluz, V. M., Toral, R. & San Miguel, M. Nonequilibrium transitions in complex networks: a model of social interaction. *Phys. Rev. E* **67**, 026120. https://doi.org/10.1103/PhysRevE.67.026120 (2003).
26. Nyczka, P., Sznajd-Weron, K. & Cisło, J. Phase transitions in the \$q\$-voter model with two types of stochastic driving. *Phys. Rev. E* **86**, 011105. https://doi.org/10.1103/PhysRevE.86.011105 (2012).
27. Lewis, D. *Convention: a philosophical study* (Wiley, Oxford, 2002).
28. Steels, L. A self-organizing spatial vocabulary. *Artif. Life* **2**, 319–332. https://doi.org/10.1162/artl.1995.2.3.319 (1995).
29. Baronchelli, A. A gentle introduction to the minimal naming game. *Belgian J. Ling.* **30**, 171–192. https://doi.org/10.1075/bjl.30.08bar (2016).
30. Golub, B. & Sadler, E. Learning in social networks https://doi.org/10.1093/oxfordhb/9780199948277.013.12 (2016).
31. Degroot, M. H. Reaching a consensus. *J. Am. Stat. Assoc.* **69**, 118–121. https://doi.org/10.1080/01621459.1974.10480137 (1974).
32. Dankulov, M. M., Melnik, R. & Tadić, B. The dynamics of meaningful social interactions and the emergence of collective knowledge. *Sci. Rep.* **5**, 12197. https://doi.org/10.1038/srep12197 (2015).
33. Moreira, A. A., Mathur, A., Diermeier, D. & Amaral, L. A. N. Efficient system-wide coordination in noisy environments. *Proc. Natl. Acad. Sci.* **101**, 12085–12090. https://doi.org/10.1073/pnas.0400672101 (2004).
34. Granovsky, B. L. & Madras, N. The noisy voter model. *Stoch. Process. Appl.* **55**, 23–43. https://doi.org/10.1016/0304-4149(94)00035-R (1995).
35. Lim, C. C. & Zhang, W. Noisy naming games, partial synchronization and coarse-graining in social networks. In *Proceedings of the 2011 IEEE network science workshop*, 25–29, https://doi.org/10.1109/NSW.2011.6004654 (2011).
36. Baronchelli, A., Dall'Asta, L., Barrat, A. & Loreto, V. Nonequilibrium phase transition in negotiation dynamics. *Phys. Rev. E* **76**, 051102. https://doi.org/10.1103/PhysRevE.76.051102 (2007).
37. Skyrms, B. Evolution of signalling systems with multiple senders and receivers. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 771–779. https://doi.org/10.1098/rstb.2008.0258 (2009).
38. Skyrms, B. & Pemantle, R. A dynamic model of social network formation. *Proc. Natl. Acad. Sci.* **97**, 9340–9346. https://doi.org/10.1073/pnas.97.16.9340 (2000).
39. Huttegger, S., Skyrms, B., Tarrès, P. & Wagner, E. Some dynamics of signaling games. *Proc. Natl. Acad. Sci.* **111**, 10873–10880. https://doi.org/10.1073/pnas.1400838111 (2014).
40. Zollman, K. J. S. Talking to neighbors: the evolution of regional meaning. *Philos. Sci.* **72**, 69–85. https://doi.org/10.1086/428390 (2005).
41. Baronchelli, A., Dall'Asta, L., Barrat, A. & Loreto, V. Topology-induced coarsening in language games. *Phys. Rev. E* **73**, 015102. https://doi.org/10.1103/PhysRevE.73.015102 (2006).
42. Lou, Y., Chen, G., Fan, Z. & Xiang, L. Local communities obstruct global consensus: naming game on multi-local-world networks. *Phys. A* **492**, 1741–1752. https://doi.org/10.1016/j.physa.2017.11.094 (2018).
43. Cave, J. A. K. Learning to agree. *Econ. Lett.* **12**, 147–152. https://doi.org/10.1016/0165-1765(83)90126-X (1983).
44. Parikh, R. & Krasucki, P. Communication, consensus, and knowledge. *J. Econ. Theory* **52**, 178–189. https://doi.org/10.1016/0022-0531(90)90073-S (1990).
45. Heifetz, A. Comment on consensus without common knowledge. *J. Econ. Theory* **70**, 273–277. https://doi.org/10.1006/jeth.1996.0087 (1996).
46. Groeber, P., Lorenz, J. & Schweitzer, F. Dissonance minimization as a microfoundation of social influence in models of opinion formation. *J. Math. Sociol.* **38**, 147–174. https://doi.org/10.1080/0022250X.2012.724486 (2014).
47. Cialdini, R. B. & Goldstein, N. J. Social influence: compliance and conformity. *Annu. Rev. Psychol.* **55**, 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015 (2004).
48. Günther, G. *Idee und Grundriss einer nicht-Aristotelischen Logik* (Meiner, Hamburg, 1991), 3. edn.
49. Günther, G. *Beiträge zur Grundlegung einer operationsfähigen Dialektik* (Felix Meiner Verlag, 1976).
50. Di Biagio, A. & Rovelli, C. Stable facts, relative facts. *Found. Phys.* **51**, 30. https://doi.org/10.1007/s10701-021-00429-w (2021).
51. Thayer-Bacon, B. J. The nurturing of a relational epistemology. *Educ. Theory* **47**, 239–260. https://doi.org/10.1111/j.1741-5446.1997.00239.x (1997).
52. Sire, J. W. *Naming the elephant: worldview as a concept* (IVP ACADEMIC, Downers Grove, Illinois, 2015), revised edition edn.
53. Kluback, W. & Weinbaum, M. *Dilthey's philosophy of existence: introduction to weltanschauungslehre* (Vision, 1957).
54. Baronchelli, A. The emergence of consensus: a primer. *R. Soc. Open Sci.* **5**, 172189. https://doi.org/10.1098/rsos.172189 (2018).
55. Baronchelli, A., Felici, M., Loreto, V., Caglioti, E. & Steels, L. Sharp transition towards shared vocabularies in multi-agent systems. *J. Stat. Mech. Theory Exp.* **2006**, P06014–P06014. https://doi.org/10.1088/1742-5468/2006/06/P06014 (2006).

56. Zöllner, D. & Streitenberger, P. Normal grain growth: monte carlo potts model simulation and mean-field theory. In Bertram, A. & Tomas, J. (eds.) *Micro-macro-interaction: in structured media and particle systems*, pp. 3–18, https://doi.org/10.1007/978-3-540-85715-0_1 (Springer, Berlin, Heidelberg, 2008).

57. Mullins, W. W. The statistical self-similarity hypothesis in grain growth and particle coarsening. *J. Appl. Phys.* **59**, 1341–1349. https://doi.org/10.1063/1.336528 (1986).

58. Abbott, A. *Part 1. Self-similarity in social science* (University of Chicago Press, 2010).

59. Lesne, A. & Lagües, M. *Scale invariance* (Springer, Berlin heidelberg, 2012).

60. Tsakiris, N., Maragakis, M., Kosmidis, K. & Argyrakis, P. Percolation of randomly distributed growing clusters: finite-size scaling and critical exponents for the square lattice. *Phys. Rev. E* https://doi.org/10.1103/PhysRevE.82.041108 *(2010)*.

61. Dandekar, P., Goel, A. & Lee, D. T. Biased assimilation, homophily, and the dynamics of polarization. *Proc. Natl. Acad. Sci.* **110**, 5791–5796. https://doi.org/10.1073/pnas.1217220110 (2013).

62. Sîrbu, A., Pedreschi, D., Giannotti, F. & Kertész, J. Algorithmic bias amplifies opinion fragmentation and polarization: a bounded confidence model. *PLoS ONE* **14**, e0213246. https://doi.org/10.1371/journal.pone.0213246 (2019).

63. Falk, J., Eichler, E., Windt, K. & Hütt, M.-T. Physics is organized around transformations connecting contextures in a polycontextural world. *Found. Sci.* https://doi.org/10.1007/s10699-021-09814-0 *(2021)*.

64. Klagenfurt, K. *Technologische Zivilisation und transklassische Logik: Eine Einführung in die Technikphilosophie Gotthard Günthers* (Metropolis, 2016), first edn.

65. Günther, G. *Das Bewußtsein der Maschinen* (Agis-Verlag, 1964).

66. Edwards, R., Bybee, B. T., Frost, J. K., Harvey, A. J. & Navarro, M. Thats not what i meant: how misunderstanding is related to channel and perspective-taking. *J. Lang. Soc. Psychol.* **36**, 188–210. https://doi.org/10.1177/0261927X16662968 (2017).

67. Berger, C. R. Miscommunication and communication failure. In *The new handbook of language and social psychology* (eds Robinson, W. P. & Giles, H.) 177–192 (Wiley, Hoboken, 2001).

68. Miller, H. *et al.* "blissfully happy" or "ready to fight": Varying interpretations of emoji. In *Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016*, pp. 259–268 (AAAI press, 2016).

69. Edwards, R. Listening and message interpretation. *Int. J. List.* **25**, 47–65. https://doi.org/10.1080/10904018.2011.536471 (2011).

70. Stanton, A. H. Problems in analysis of therapeutic implications of the institutional milieu. In *Symposium on Preventive and Social Psychiatry* (US Government Printing Office, 1958).

71. Porter, C. D. & Heckler, A. F. Graduate student misunderstandings of wave functions in an asymmetric well. *Phys. Rev. Phys. Edu. Res.* **15**, 010139. https://doi.org/10.1103/PhysRevPhysEducRes.15.010139 (2019).

72. Shoham, Y. & Tennenholtz, M. On the emergence of social conventions: modeling, analysis, and simulations. *Artif. Intell.* **94**, 139–166. https://doi.org/10.1016/S0004-3702(97)00028-3 (1997).

Author contributions

E.E. initiated the project. J.F., M-T.H. developed the model, J.F. ran the simulations and analysed the data. J.F. and M-T.H. wrote the manuscript. E.E., K.W. and M-T.H. supervised the project. All authors discussed the results and implications and commented on the manuscript at all stages.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.