



OPEN

# When open mindedness hinders consensus

Hendrik Schawe &amp; Laura Hernández

We perform a detailed study of the Hegselmann-Krause bounded confidence opinion dynamics model with heterogeneous confidence  $\varepsilon_i$  drawn from uniform distributions in different intervals  $[\varepsilon_1, \varepsilon_0]$ . The phase diagram reveals a highly complex and non-monotonous behaviour, with a re-entrant consensus phase in the region where fragmentation into multiple distinct opinions is expected for the homogeneous case. A careful exploration of the phase diagram, along with an extensive finite-size analysis, allows us to identify the mechanism leading to this counter-intuitive behaviour. This systematic study over system sizes which goes well beyond those of previous works, is enabled by an efficient algorithm presented in this article.

Opinion formation and its dissemination within a society, a recurrent subject of study in social sciences, is often addressed through statistical analyses of data collected by the means of surveys or polls. By following the evolution of the statistical outcomes over time, it is possible to obtain some information about the dynamical processes underlying these phenomena. What we call *the opinion of a society* is a global property that characterises the society as a whole and has emerged from the repeated interactions among their agents. Defined in this way, opinion may be studied statistically. Although individual fluctuations are always present, it is very unlikely that they change the stabilized social opinion. This approach is the mathematical realization of the ideas introduced more than one century ago by E. Durkheim<sup>1</sup>, who coined the concept of *social facts*. A social fact is a property characterizing the whole society instead of the individuals, which emerges as an outcome of the dynamics governed by the interactions among them. Once the social fact is created, it is imposed on the members of the society who will find it very difficult to change it.

Early opinion dynamics studies mainly dealt with the case of a fully mixed population, which means that every agent may potentially interact with any other in the population, in other words, the interactions among agents were supposed to be long range. It is well known that this approach is equivalent to what is called in Physics, a mean-field approximation which neglects the structure of the interactions. However, if social opinion emerges from the interactions between the agents, the structure of these interactions may be relevant. In fact, the key role of interactions has been recognized long ago by works that collected data about the detail of social interactions in very small societies, using graph theory to represent them. J. Scott<sup>2</sup> gives a nice historical overview of network development in social sciences. Nowadays, the development of mathematical models along with the continuous growth of data collection on human activities that proliferated in particular due to the rise of online social networks, allow us to investigate different dynamical processes of opinion dynamics.

The rationale behind most opinion dynamics models is grounded in *Social Influence Theory*<sup>3,4</sup> which assumes that, as agents interact, they may influence each other making their opinions more alike. Accordingly, the dynamical rules that govern the evolution of the opinion of each agent, are often based on functions which aggregate the opinions of a given set of agents which are assumed to interact with it. Recently a complete mathematical classification of the possible outcomes of the dynamics based on aggregation functions in the case of discrete opinion variables has been obtained<sup>5,6</sup>.

The mathematical description of the structure of social interactions in terms of networks allows to identify the neighbourhood of a given agent (the set of other agents directly connected to it) and it is now clear that this structure is relevant. Different neighbourhood choices along with the type of influence they have on each agent have been considered, for example the adoption of the opinion of a randomly chosen neighbour<sup>7</sup>, the adoption of the neighbours' majoritarian opinion<sup>8</sup>, or of agents whose opinion on other topics are already near the one of the target agent<sup>9</sup>.

Among all the studied models<sup>10,11</sup>, those representing the agents' opinion by a continuous variable are well suited to describe situations where the opinion on a particular problem is gradually built through the exchanges among the agents. In particular, *bounded confidence* models consider that each agent will only interact with those

Laboratoire de Physique Théorique et Modélisation, UMR-8089 CNRS, CY Cergy Paris Université, Paris, France.  
e-mail: [hendrik.schawe@cyu.fr](mailto:hendrik.schawe@cyu.fr); [laura.hernandez@u-cergy.fr](mailto:laura.hernandez@u-cergy.fr)

agents whose opinions are already close to theirs and will not interact with the others. In the original models a given agent may interact with all the others *provided* that they fall within the confidence interval, meaning that the difference of opinions between potentially interacting agents is lower than a certain threshold that characterizes the society. This simple model is already different from a fully mixed population. Moreover as the opinions evolve in time, the individuals that may interact with a given one also change with time.

The best studied bounded confidence models are the *Deffuant-Weisbuch* (DW) model<sup>12</sup> and the *Hegselmann-Krause* (HK) model<sup>13</sup>. In both models the opinion of agent  $i$  is coded in a continuous variable and each agent may be influenced only by others whose opinion differs from his at most by a quantity called *confidence*  $\varepsilon_i$ . Unlike the DW model, which considers pairwise interactions, in the HK model agents are synchronously influenced by all others within their confidence range.

The most studied variant of the latter is the case of homogeneous confidence  $\varepsilon_i = \varepsilon$ , where large populations will always converge towards a uniform opinion, if the confidence is above a threshold  $\varepsilon_c \approx 0.2$ <sup>13,14</sup>. However, a society is not a homogeneous collection of individuals. Some of them are *open minded* and this property may be modelled by relatively large values of their confidence  $\varepsilon_i$ , accordingly the *closed minded* ones will have low values of their confidence  $\varepsilon_i$ . Therefore heterogeneity of the confidence of single agents is an obvious and well motivated addition to the model. Until now heterogeneity in the HK model has mainly been studied by introducing multiple sub-populations, each of them formed by a homogeneous set of agents characterized by a specific confidence value<sup>15–18</sup>. Some other works study systems with random confidence, with  $\varepsilon_i$  drawn from some distribution<sup>19–21</sup>, or from different distributions for multiple sub-populations<sup>22</sup>. Those studies are usually performed by the means of multi-agent simulations on small statistical samplings (50–100 samples), where each realization represents also a very small system (typically a few hundred agents). An alternative method, the multiple chain Markov model<sup>15</sup>, has been proposed in order to obtain the properties of the infinite system. However, in this model, opinions are discretized which is known to cause deviations from the HK model<sup>23</sup> and the probabilities of each opinion are derived from an initial set of agents. Nevertheless, these works seem to indicate that heterogeneity indeed leads to some surprising non-monotonous effects in the dynamical outcomes. In particular, the existence of consensus in regions where the confidence of the agents is below the critical threshold of the homogeneous case has been reported<sup>17</sup>.

In this article, we present a systematic study of the phase diagram of the HK heterogeneous model in the parameter space given by the possible lower and upper bounds of the confidence values of the agents,  $(\varepsilon_l, \varepsilon_u)$ . The system is neither a fully mixed population, nor a fixed network of contacts. Interestingly, it could be thought of as a non-symmetrical, dynamical interaction network, which may become particularly useful to define a method for building clusters of agents of similar opinions. We carefully sample the parameter space for large systems and we are able to obtain very good statistics for the studied quantities. Furthermore, we study the finite-size effects on the dynamics of the model, which reveal complex details of the consensus landscape that only become apparent at system sizes that are larger than those studied before. A careful study of the dynamical trajectories in the opinion space allows us to explain the re-entrant phase observed in the phase diagram. Finally, we describe the algorithm that allows us to study system sizes and sampling statistics that go well beyond the present state of the art.

## Models and Methods

We study the Hegselmann-Krause model (HK), which describes a compromise dynamic under bounded confidence. Each of  $n$  agents  $i$  is endowed of a dynamical continuous variable  $x_i(t)$  representing *opinion* and a fixed *confidence*  $\varepsilon_i$ , modeling the heterogeneous idiosyncrasies of the agents. The *neighbours* of agent  $i$  are all agents  $j$  whose opinions lie inside the interval  $[x_i - \varepsilon_i, x_i + \varepsilon_i]$ , i.e.

$$I(i, \vec{x}) = \{1 \leq j \leq n | |x_i - x_j| \leq \varepsilon_i\}. \quad (1)$$

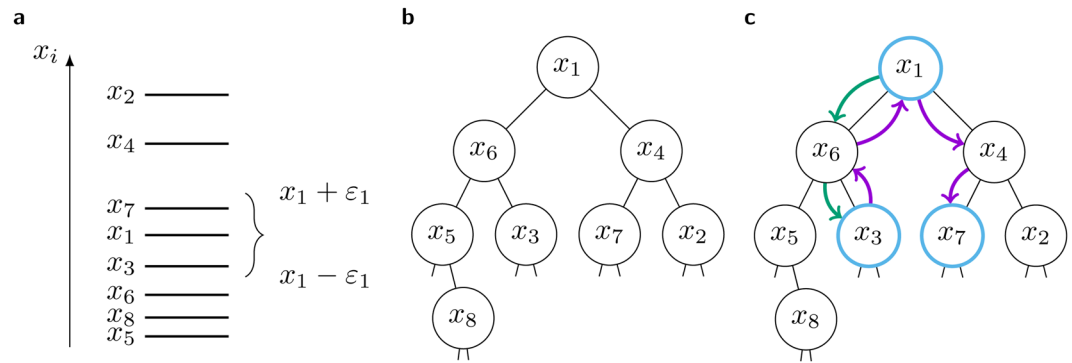
Note that every agent is a neighbour of itself. In each time step an agent  $i$  talks to all its neighbours  $j$  and adopts the average opinion of all neighbours, i.e.

$$x_i(t+1) = |I(i, \vec{x}(t))|^{-1} \sum_{j \in I(i, \vec{x}(t))} x_j(t). \quad (2)$$

This update is performed synchronously for all agents, although a sequential random update is possible. The latter typically leads to longer convergence times, as single agents are left behind and may, for small values of  $\varepsilon$ , persist as isolated clusters in the final state. However, besides these effects, the observations which we describe in the following sections are qualitatively robust against the update schedule.

The dynamics leads to a stable configuration, where the agents settle into one or several *clusters*, (a cluster being a group of agents with the same opinion). The situation in which only one giant cluster exists is called *consensus*, that with two clusters is called *polarization* and with more clusters it is called *fragmentation*. For the homogeneous case it is well known that above a critical  $\varepsilon_c \approx 0.2$  large systems will always converge to consensus and never reach consensus below this threshold<sup>10</sup>.

While the original HK model uses homogeneous confidence  $\varepsilon_i = \varepsilon$ , here we draw each  $\varepsilon_i$  from an i.i.d. uniform distribution bounded by two parameters  $\varepsilon_l, \varepsilon_u$  with  $\varepsilon_l \leq \varepsilon_u$ . The choice of the lower bound parameter,  $\varepsilon_l \geq 0$ , determines how closed minded the most closed minded agents are and the value of the upper bound parameter,  $\varepsilon_u \leq 1$ , determines how open minded the most open minded agents are.



**Figure 1.** (a) Example state of  $n = 8$  agents, with highlighted confidence interval of agent  $i = 1$ , which includes agents 3 and 7. (b) BST obeying the order property in respect to the opinions  $x_i$ . (c) Sketch of the in-order traversal of the range  $[x_1 - \varepsilon_1, x_1 + \varepsilon_1]$ . Green arrows indicate the steps necessary to find the lower bound. Violet arrows show the steps necessary for the traversal. Blue nodes are the ones collected by the algorithm and used to update the opinion of agent  $i = 1$ .

We are mainly interested in the effects of heterogeneity on consensus for the whole  $(\varepsilon_i, \varepsilon_u)$  space. To evaluate this, we look at the relative size of the largest cluster  $S$  averaged over all simulated realizations,  $\langle S \rangle$ . A value near 1 indicates consensus, smaller values indicate polarization or fragmentation.

**Tree-based updating scheme.** One difficulty in simulating the HK dynamics is that each time step, i.e. the update of all  $n$  agents, requires  $\mathcal{O}(n^2)$  operations such that the simulation of large system sizes quickly becomes infeasible. There are attempts to solve this problem with algorithms which are faster, but have some drawbacks, such as not simulating the actual dynamics, but generating an approximation for the  $n \rightarrow \infty$  case<sup>24</sup> or imposing a discretization of the opinion variables<sup>23</sup>.

Here, we apply an algorithm which is much faster than the naive approach for typical realizations, while preserving the continuous character down to the precision of the data types used. This enables us to simulate larger sizes with far better statistics than other contemporary studies of the HK model. Its fundamental idea is that to update agent  $i$  we do only have to look at the agents within its confidence interval  $[x_i - \varepsilon_i, x_i + \varepsilon_i]$ , which are typically far fewer than  $n$  for small values of  $\varepsilon_i$  – but in the order of  $\mathcal{O}(n)$  in the worst case.

To achieve this in an efficient way, we maintain a *binary search tree* (BST) [ref.<sup>25</sup>, p.~286] of all opinions  $x_i(t)$ . This data structure can be used to find all entries within a range of values efficiently [ref.<sup>26</sup>, p.~412]. Since, in order to update the opinion of agent  $i$ , we need to find all agents whose opinion is within the range  $[x_i - \varepsilon_i, x_i + \varepsilon_i]$ , this data structure is exactly what we need.

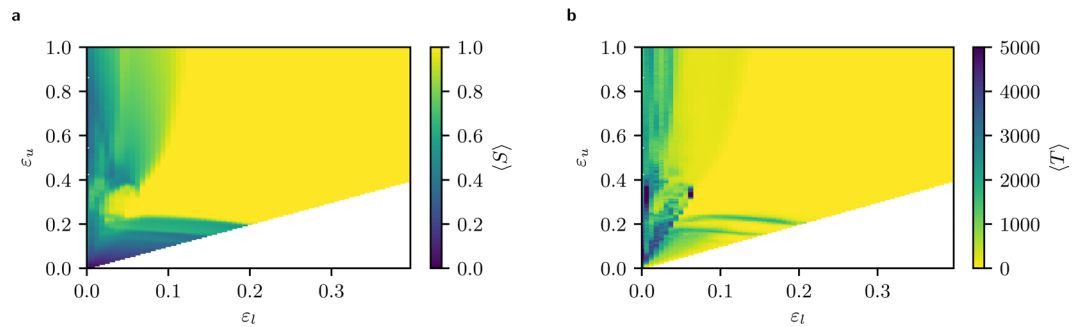
Let us recall the definition of a BST. A *tree* consists of a hierarchical arrangement of *nodes*. Each node is assigned a unique *key* (and possibly additional data). For simplicity, we will refer to a node as *greater* than another node if the key of the first node is greater than the key of the second node and analogously for *smaller*. All nodes, except the *root* ( $x_1$  in Fig. 1), have exactly one *parent* node one *level* above and at most two connected nodes called *children* one level below, as sketched in Fig. 1. The children are called *left* and *right* child. The defining property of a BST is the *order property* with respect to the nodes' keys: in a BST the left child is always smaller than its parent and the right child is always greater than its parent.

For our application, each node represents an agent  $i$  and its key is the opinion  $x_i$  of the corresponding agent. An example is sketched in Fig. 1, which should foster intuition about the algorithms explained in the following paragraphs. Note that removal and insertion of nodes into a BST can be done efficiently using standard algorithms<sup>25</sup>.

A BST with  $M$  levels can store at most  $2^M - 1$  nodes. A tree with  $n$  nodes is called *balanced*, if it only has  $m = \mathcal{O}(\ln n)$  levels. Since the order property allows us to find any node using only one comparison per level, this allows very efficient search, hence the name. Further, it can also be used to *traverse* all nodes whose keys are within a given range of values, sorted according to their keys. To traverse the range  $[l, u]$ , one starts at the root. We will call the node where we are at the *current* node  $c$ . The main idea of the range traversal is to cut whole subtrees from the tree, if all their nodes are smaller than the lower bound  $l$  or larger than the upper bound  $u$ . This can be expressed very compactly with a recursive function:

1. if  $c > u$ , it calls itself on its left subtree (otherwise all nodes of its left subtree would be smaller than  $l$ )
2. if its own key is inside the range, it stores its key to be handled later
3. if  $c < l$ , it calls itself on its right subtree (otherwise all nodes of its right subtree would be larger than  $u$ )

In other words, this algorithm will use the shortest (and only) path through the tree to find the smallest node larger or equal to the lower bound  $l$  and collect every node in ascending order until it encounters a node larger than the upper bound  $u$ . Consecutively, the  $M$  nodes which were collected during the traversal can be used to calculate the new opinion of the agent. During this procedure, we only had to inspect  $\mathcal{O}(\log n + M)$  nodes.



**Figure 2.** Left: Average relative size of the largest cluster  $\langle S \rangle$ . Right: Average convergence time  $\langle T \rangle$ . These data are collected for 8224 pairs of  $(\varepsilon_l, \varepsilon_u)$  for a system of  $n = 16384$  agents and averaged over 1000 realizations. Note that three parameter pairs at  $\varepsilon_l = 0$  did not converge in reasonable computing time for all realizations. To avoid selection bias, they are therefore omitted and marked in white.

Since BSTs can only store elements with unique keys, we have to handle the case of multiple agents having the same opinion within the precision of the used datatype. Therefore, we save at each node of the tree, additionally to the opinion, a *counter*, which keeps track of the number of agents which have this opinion and are therefore represented by this node. Note that this additional information has no influence on the order of the nodes. Correspondingly, this counter has to be handled as a weight when calculating the new opinion of the active agent.

When the opinion of an agent is updated, the tree must be updated by removing its former opinion (or decreasing its counter by one) and inserting its current opinion (or increasing its counter by one). Both operations can be performed in time  $\mathcal{O}(\log(n))$ , and have to be performed once for each agent in every time step (except  $x_i(t) = x_i(t + 1)$ ). This algorithm therefore has a typical time complexity for one step between  $\Omega(n \log(n))$  in the best case, if each agent has at most  $\mathcal{O}(\log(n))$  neighbours with distinct opinions (i.e. when almost all agents condensed into clusters), and  $\mathcal{O}(n^2)$  in the worst case, i.e. in the beginning when each agent has  $\mathcal{O}(n)$  neighbours.

Note that the performance of this algorithm benefits from two independent effects. First, for agents with small confidence  $\varepsilon_i$  we profit from the reduced number of opinions we need to look at to determine the average opinion of its neighbours. Second, the HK model tends to form clusters (within the chosen numerical precision) quickly, especially for large  $\varepsilon_i$ . Since clusters are represented in the BST as a single node with a high counter, the computational cost to calculate the average opinion is greatly reduced.

A commented minimal implementation and pseudocode for the tree range traversal is provided in supplementary information SI1<sup>1</sup>. There, to further speed up this algorithm, we use a B-tree [ref.<sup>25</sup>, p.~484] instead of a BST.

## Results

We simulated the system for a broad range of system sizes  $64 \leq n \leq 131072$  and we carefully explored the parameter space of confidence intervals. For each run, we have drawn the confidence parameter of each agent  $\varepsilon_i$  uniformly from an interval bounded by  $(\varepsilon_l, \varepsilon_u) \in [0, 0.35] \times [0, 1]$ . The simulations are run until the opinions, represented with 32 bit IEEE 754 float data-types, converge. The convergence criterion we use here requires that the sum of the changes over all the agents is below a threshold, i.e.  $\sum_{i=1}^n |x_i(t) - x_i(t-1)| < 10^{-4}$ .

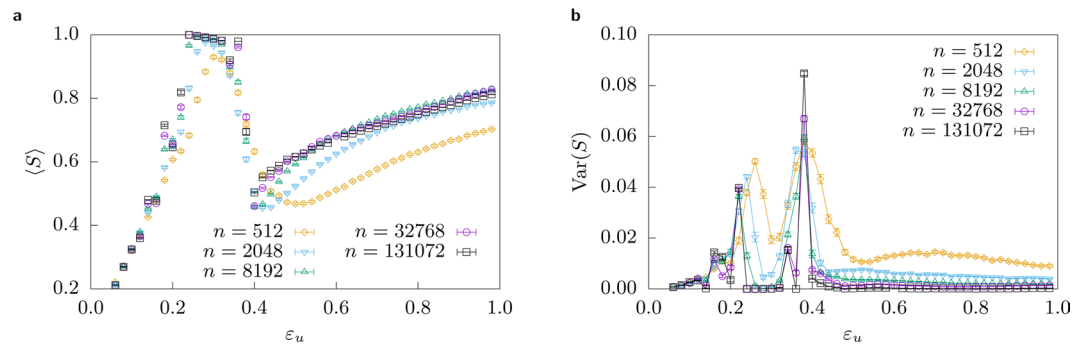
Clusters are composed of agents holding the same opinion within a tolerance range of  $10^{-4}$ . We have checked that the results are robust when using different clustering criteria, e.g. binning the opinion space, which is briefly shown in supplementary information SI2.

Figure 2 shows a heat map of the average size of the largest cluster  $\langle S \rangle$  for each of the 8224 points that we have simulated in the parameter space. The results shown by the heat map at each point  $(\varepsilon_l, \varepsilon_u)$  correspond to the average over 1000 simulated samples, which differ in their initial conditions although all represent the same society, where the heterogeneous confidence parameters of the agents  $\varepsilon_i$  have been uniformly drawn from the same interval. Note that the aspect ratio is not unity. While we concentrate on uniformly distributed confidences, supplementary information SI2 shows that qualitatively the same effects occur for other distributions, including heavy tail ones.

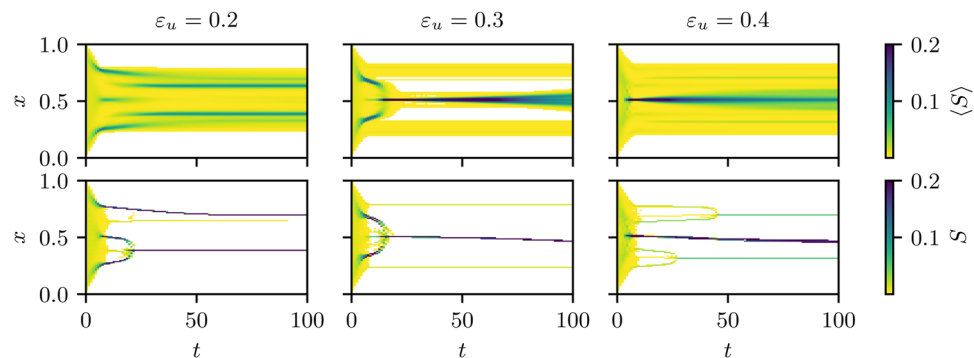
The region not shown here  $([0.35, 1] \times [0, 1])$  corresponds to values of confidence intervals that always lead to consensus. This is expected, as both bounds are far above the critical value of the homogeneous case. The white triangle in the lower right corner consists of impossible intervals, where the lower bound would be higher than the upper one. The diagonal elements correspond to the homogeneous case, i.e.  $\varepsilon_l = \varepsilon_u$ . One can see that on this diagonal,  $\langle S \rangle$  changes from 1 to approximately 1/2 around  $\varepsilon_l = \varepsilon_u = 0.2$ , showing the transition from consensus to polarization that has been found in the homogeneous HK model<sup>10</sup>.

The interesting results are situated on the left of the map, where a *re-entrant consensus region* occurs around the point  $(0.05, 0.3)$ . This reveals a non-monotonous behaviour of the system as the fraction of confident agents increases in the population. Let us examine the map by considering a fixed value of  $\varepsilon_l = 0.05$  while varying the upper value  $\varepsilon_u$ . We start at the homogeneous case  $\varepsilon_u = 0.05$ , where we find fragmentation (dark). As  $\varepsilon_u$  increases,

<sup>1</sup>see also [https://github.com/surt91/hk\\_tree](https://github.com/surt91/hk_tree)



**Figure 3.** Left: Average size of the largest cluster  $\langle S \rangle$ . Fixed  $\varepsilon_l = 0.05$  for varying values of  $\varepsilon_u \in [\varepsilon_l, 1]$ ; this corresponds to a straight vertical line in Fig. 2 through the re-entrant consensus phase. The consensus region is robust with increasing system sizes. Also note that the behaviour is highly complex as the behaviour with increasing system size is not always monotonous, e.g. at  $\varepsilon_u = 0.2$  or  $\varepsilon_u = 0.42$ . Right: Variance of the size of the largest cluster  $\text{Var}(S)$ . The transition points into consensus and out of consensus can be located at the peaks of the variance. Using the finite-size behaviour, we estimate them for  $\varepsilon_l = 0.05$  at  $\varepsilon_u = 0.22(1)$  and  $\varepsilon_u = 0.38(1)$ . Lines are just guides to the eye.



**Figure 4.** We look for fixed  $\varepsilon_l = 0.05$  and  $n = 16384$  at three values of  $\varepsilon_u \in \{0.2, 0.3, 0.4\}$  (from left to right) corresponding to one value left of the peak, one at the peak and one right of the peak of Fig. 3. The images show which fraction of agents have a given opinion  $x$  for each time step. Note that the colour scale is truncated at 0.2 to generate images with a good contrast in the interesting region at small times, such that the darkest colour can also represent values larger than 0.2. The time axis is also truncated to focus on the most interesting region of small times, i.e. not the whole range until reaching a stationary state is visualized. Top: aggregated statistics over 10000 realizations of the initial conditions. Bottom: single trajectories.

a light region is encountered showing strong consensus. This is expected as more agents with an increasing confidence (having a larger number of neighbours) enter the system, contributing to integrate agents with low confidence into the consensus group. However, as  $\varepsilon_u$  increases further, consensus surprisingly disappears (colour gets darker) although the fraction of confident agents and the magnitude of their confidence is even larger.

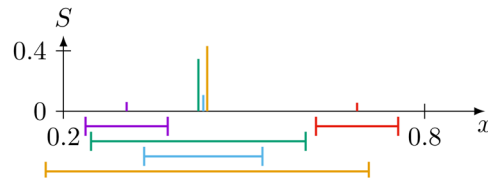
The left panel of Fig. 3 shows the evolution of the order parameter  $\langle S \rangle$  with  $\varepsilon_u$  for a vertical cut in the region of the re-entrant phase of the heat-map shown in Fig. 2 for different system sizes. For a system that contains very closed minded agents, when the fraction of open minded agents increases, the size of the largest cluster also increases leading to consensus, however as the fraction of open minded agents along with their corresponding confidence values increases further, consensus is lost.

An additional small peak, which grows with the system size, is visible on the left of the consensus peak, this reflects the rich structure observed in the phase diagram between  $\varepsilon_l \approx 0.13$  and  $\varepsilon_l \approx 0.18$ . Although these fine structures call for more inspection, here we focus on the dominant structure, the re-entrant consensus phase.

The right of Fig. 3 shows the fluctuation of the order parameter  $\langle S \rangle$ . There are two visible peaks, which indicate the location of the two transitions, in and out of the re-entrant phase. Interestingly, the finite-size behaviour signals a real double transition as the peaks do not decrease and separate with increasing size, suggesting that this is not just a finite-size effect, but that it remains in the thermodynamic limit, i.e. the  $n \rightarrow \infty$  limit. Using these data, we estimate the re-entrant phase to span the interval  $\varepsilon_u \in [0.22(1), 0.32(1)]$  for fixed  $\varepsilon_l = 0.05$ .

In order to explain this paradoxical result, a careful examination of the dynamic behaviour of the agents' opinions is necessary. In Fig. 4 we show the early evolution of opinions for systems corresponding to three different values of  $\varepsilon_u$  in the re-entrant region. The top row corresponds to averages over 10000 realizations of the initial conditions, while the bottom row illustrates the evolution of opinions with single example realizations.





**Figure 5.** Final state of the single trajectory evolution at  $\varepsilon_u = 0.4$  shown in Fig. 4. Each cluster is shown in a different color, the height corresponds to its size  $S$  and a typical confidence  $\varepsilon_i$  of a member of that cluster is shown below the plot. The split of the central strand into three clusters is clearly visible (see text for an explanation taking the confidences into account).

The mechanism leading to the observed behaviour is rooted in the different characteristic times that open and closed minded agents need to join a majoritarian opinion strand. The latter take more time to reach the consensus opinion as they need to meet agents within their narrow confidence interval. This is at the origin of the bell shaped structures in the bottom left panel of Fig. 4

These structures enable closed minded agents from the whole region over which the bell spans, to join a strand and evolve. If the strand counts enough open minded agents, they can pull it towards a consensus opinion, bringing with them the closed minded agents that have an opinion close to theirs. This is what happens in the bottom middle panel of Fig. 4 at  $\varepsilon_u = 0.3$ . Also the signature of this structure is clearly visible in the average over many realization in the top middle panel. The two smaller strands visible in this panel contain only 3% of the agents. As it is very frequent that the central cluster converges to one of them, the averaged picture shows an apparent diffusion of the central opinion for long times. Moreover we expect the small clusters to vanish in the thermodynamic limit, since the peak of Fig. 3a shows an upward trend for increasing system sizes.

Note that agent  $i$  can in each time step move at most  $\varepsilon_i$ , such that closed minded agents need to see other agents for at least a few iterations to be able to change their opinion from one extreme to the consensus opinion at 0.5. More open minded agents (like those that appear when increasing  $\varepsilon_u$ ) will evolve very quickly to a central opinion, because they can interact with a large fraction of the other agents and therefore are able to jump directly into the centre. As a consequence closed minded agents are left behind and are not able to join the consensus opinion. This is the situation depicted in the right panels of Fig. 4.

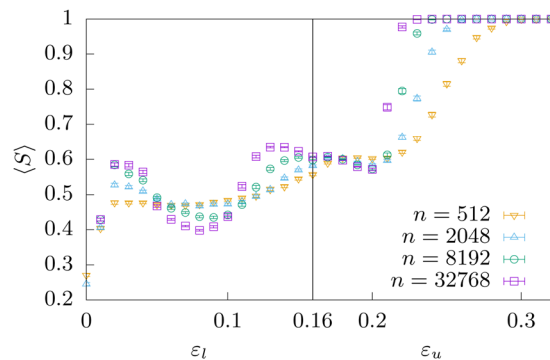
The case of  $\varepsilon_u = 0.4$ , shown on the right of Fig. 4, seems to show that the central strand contains about 90% (the two secondary clusters have a size of  $\approx 5\%$  each), which is an apparent contradiction with the mean cluster size  $\langle S \rangle \approx 0.5$  shown in Fig. 2. The solution to this discrepancy is that the central strand located at an opinion  $x \approx 0.44$  splits in the stationary state to three very close clusters, therefore there are 5 clusters in total, which are composed of different groups of agents. This final state is visualized in Fig. 5.

The largest cluster of the central strand contains about 50% of the agents (yellow in Fig. 5), mainly having large confidence intervals, able to interact with agents of the small strands located at  $x \approx 0.32$  (violet) and  $x \approx 0.68$  (red). Another cluster contains about 30% of the agents (green), which can interact only with the bottom strand at  $x \approx 0.32$  (and all agents of the central strand), but not with the upper one at  $x \approx 0.68$ , such that it will converge to a slightly lower opinion. The last cluster contains about 10% of the agents (blue) which can only interact with those in the central strand (their opinion being in the middle of the other two clusters forming the central strand). Therefore, the size of the largest cluster of this realization is  $S \approx 0.5$ , the value observed in the valley of Fig. 3. Note that this effect can easily be missed when using a more discrete clustering criterion, like binning with too few bins. Also note that in the  $\varepsilon_u = 0.3$  case the secondary clusters, which are already significantly smaller in this example, become smaller in the thermodynamic limit. In contrast, at  $\varepsilon_u = 0.4$  the size of secondary clusters stays roughly constant with increasing system size, such that this effect will not vanish in the thermodynamic limit. That means even if we count the central strand as a single cluster,  $\langle S \rangle$  would not grow beyond  $\langle S \rangle \approx 0.9$ , in contrast to  $\varepsilon_u = 0.3$  where it convincingly converges to  $\langle S \rangle \rightarrow 1$ . A demonstration of this by using a different method to determine clusters, which will always treat the central strand as a single cluster, is shown in the supplementary information SI2 and exhibits the same telltale phase transition signature at the same threshold between  $\varepsilon_u = 0.3$  and  $\varepsilon_u = 0.4$ .

Figure 3b confirms these results: average convergence time in the region of re-entrant consensus is much larger than on neighbouring regions, illustrating the extra time needed by closed minded agents to join the consensus strand *pulled* by open minded agents.

Interestingly, the bell shapes observed in the evolution of opinions, which help to integrate isolated agents into a single strand, are observed for all the parameters shown here. However, when  $\varepsilon_u$  is either too low or too high, these structures are mainly formed by agents with very low confidence, in the first case just because they are the majority of the population and in the second because agents with large confidence have already joined the main strand. For intermediate values of  $\varepsilon_u$ , these bell structures contain both open and closed minded agents, and the former may bring the latter into the final consensus strand.

Now, that we understood the main structure of the phase diagram, we can take a more detailed look. There are apparently large horizontal structures visible in Fig. 2 closely below  $\varepsilon_u \approx 0.2$ . In vertical slices, these correspond to small local minima and maxima, which can be recognized in Fig. 3 in the left increasing flank. Note that the maxima increase with increasing system size, while the rest of the flank does not, leading to local minima. This is even better visible in the corresponding variance, where small maxima below  $\varepsilon_u = 0.2$  arise and persist for larger



**Figure 6.** Average size of the largest cluster  $\langle S \rangle$ . The left half shows  $\varepsilon_u = 0.16$  fixed and  $\varepsilon_l$  varying, the right half shows  $\varepsilon_l = 0.16$  fixed and  $\varepsilon_u$  varying. The position marked by the vertical line is the homogeneous case  $\varepsilon_l = \varepsilon_u = 0.16$ . This corresponds to a straight line in Fig. 2 reflected at the diagonal.

system sizes. This hints that these structures do not vanish in the thermodynamic limit and are properties inherent to this model. This suggests a very rich behavior that would be worthwhile to study in the future.

In a previous work, where the HK model was studied using an “interacting Markov chain” instead of agent-based simulations<sup>15,27</sup>, an effect called there “consensus strikes back” was reported. This effect consisted in a consensus phase observed in the homogeneous case for a very narrow range of the confidence parameter  $0.152 \leq \varepsilon \leq 0.174$ .

Since one of the aforementioned local maxima appearing in Fig. 3 along with its corresponding fluctuation is located at  $\varepsilon_u \approx 0.16$  for all sizes, we have investigated our model in order to see whether we observe such effect. Figure 6 shows finite size behaviour of  $\langle S \rangle$  for different systems where one bound (either  $\varepsilon_l$  or  $\varepsilon_u$ ) is fixed to 0.16 while the other bound is indicated in the horizontal axis. In spite of the fact that, for large systems, there is a slight increase in  $\langle S \rangle$  as one approaches  $\varepsilon \approx 0.16$  with the lower or the upper bound of the interval, the value of the largest cluster is far from consensus so as to conclude about the presence of the aforementioned effect.

## Conclusions

We have performed a very detailed characterization of the phase diagram of the heterogeneous Hegselmann-Krause model by means of an efficient algorithm that allows the simulation of large samples. In this way we were able to obtain very good statistics and to investigate finite-size effects overcoming the size limitations of previous works. Our results reveal a non-monotonous behaviour with a re-entrant consensus phase in the region of the parameter space where fragmentation is expected.

Previously, the HK model or closely related models with two different values of the bounded confidence coexisting in the population have been studied, also revealing a non-monotonous behaviour depending on the amount of open and closed minded agents<sup>15,17</sup>. Here, by the means of the simulation of large systems, we characterize the rich behaviour of the fully heterogeneous HK model, and we identify the region  $[\varepsilon_l, \varepsilon_u]$ , where the consensus phase clearly enters in the region, in which one would not expect consensus based on the behaviour of the homogeneous case. In particular, we find that increasing the proportion of open minded agents may lead to a loss of consensus, provided closed minded agents are still in the system. We were able to explain this counter-intuitive observation with a careful study of the opinion evolution. Its origin is the slow movement of closed minded agents in the opinion space. When the system contains a large fraction of very open minded agents, who converge very quickly to a majoritarian opinion, the closed minded agents are left behind and full consensus is precluded. On the contrary it is easier to reach consensus when the closed minded agents coexist with a fraction of moderately open minded ones, who converge to the majoritarian opinion much slower. This relative slow convergence allows for the interaction with the closed minded agents who are slowly dragged to the majoritarian opinion.

We have performed several robustness tests that show that our results qualitatively hold for different probability distributions of confidences, including heavy tail ones, and also for different ways of computing the size of the largest opinion cluster. In particular, we observe that although the main mechanism leading to the dramatic loss of consensus we observe is a split-up of a central strand through interaction with small secondary clusters far away, a sharp transition, of a smaller amplitude is still observed when ignoring this split (details can be found in SI2).

We believe that the introduction of quenched disorder by increasing the complexity of the HK model offers a lot of potential for further studies. The high-resolution phase diagram we calculated shows fine structures that we did not explain here, such that open questions remain. An example is the evidence of a tendency towards consensus leading to horizontal structures in the phase diagram at  $\varepsilon_u \approx 0.18$  and  $\varepsilon_l \approx 0.13$ . Another one is the “peninsula-like” structure showing that fragmentation appears into the re-entrant consensus region visible in Fig. 2a around (0.18, 0.39), and much more evident in the maximum of convergence time (Fig. 2b). Moreover, variations of the model that introduce a network of social ties, which limits the possible neighbourhoods of any agent, or the integration of cost that the agents must bear for changing their opinion, are work in progress.

Received: 26 February 2020; Accepted: 14 April 2020;

Published online: 19 May 2020

## References

1. Durkheim, E. *Les Règles de la Méthode Sociologique* [http://classiques.uqac.ca/classiques/Durkheim\\_emile/regles\\_methode/regles\\_methode.html](http://classiques.uqac.ca/classiques/Durkheim_emile/regles_methode/regles_methode.html) (Les Presses universitaires Paris, France, 1967).
2. Scott, J. Social network analysis: developments, advances, and prospects. *SOCNET* **1**, 21–26 (2011).
3. Scott, J. Compliance, identification, and internalization three processes of attitude change. *Journal of Conflict Resolution* **2**, 51–60 (1958).
4. study of normative, A. & informational social influences upon individual judgment. Deutsch, m. and Gerard, h. b. *Journal of Abnormal and Social Psychology* **51**, 629–636 (1955).
5. Grabisch, M. & Rusinowska, A. A model of influence based on aggregation functions. *Mathematical Social Sciences* **66**, 216–330 (2013).
6. Poindron, A. Influence négative dans les réseaux sociaux. Ph.D. thesis, UNIVERSITE PARIS 1 PANTHEON-SORBONNE (2019).
7. Holley, R. A. & Liggett, T. M. Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* **3**, 643–663 (1975).
8. Galam, S. Minority opinion spreading in random geometry. *The European Physical Journal B - Condensed Matter and Complex Systems* **25**, 403–406 (2002).
9. Axelrod, R. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution* **41**, 203–226 (1997).
10. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**, 591–646 (2009).
11. Sirbu, A., Loreto, V., Servedio, V. D. P. & Tria, F. *Opinion Dynamics: Models, Extensions and External Effects*, 363–401 (Springer International Publishing, Cham, 2017).
12. Deffuant, G., Neau, D., Amblard, F. & Weisbuch, G. Mixing beliefs among interacting agents. *Advances in Complex Systems* **03**, 87–98 (2000).
13. Hegselmann, R. & Krause, U. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation* **5** (2002).
14. Hegselmann, R. & Krause, U. Opinion dynamics driven by various ways of averaging. *Computational Economics* **25**, 381–405 (2005).
15. Lorenz, J. Heterogeneous bounds of confidence: Meet, discuss and find consensus! *Complexity* **15**, 43–52 (2010).
16. Kou, G., Zhao, Y., Peng, Y. & Shi, Y. Multi-level opinion dynamics under bounded confidence. *Plos one* **7**, 1–10 (2012).
17. Pineda, M. & Buendía, G. Mass media and heterogeneous bounds of confidence in continuous opinion dynamics. *Physica A: Statistical Mechanics and its Applications* **420**, 73–84 (2015).
18. Han, W., Huang, C. & Yang, J. Opinion clusters in a modified Hegselmann-Krause model with heterogeneous bounded confidences and stubbornness. *Physica A: Statistical Mechanics and its Applications* **531**, 121791 (2019).
19. Lorenz, J. Opinion dynamics under heterogeneous bounds of confidence for the agents. In *Proceedings of the Workshop on Economics with Heterogeneous Interacting Agents, Kiel, Germany* (2003).
20. Liang, H., Yang, Y. & Wang, X. Opinion dynamics in networks with heterogeneous confidence and influence. *Physica A: Statistical Mechanics and its Applications* **392**, 2248–2256 (2013).
21. Shang, Y. An agent based model for opinion dynamics with random confidence threshold. *Communications in Nonlinear Science and Numerical Simulation* **19**, 3766–3777 (2014).
22. Fu, G., Zhang, W. & Li, Z. Opinion dynamics of modified Hegselmann-Krause model in a group-based population with heterogeneous bounded confidence. *Physica A: Statistical Mechanics and its Applications* **419**, 558–565 (2015).
23. Fortunato, S. The Krause-Hegselmann consensus model with discrete opinions. *International Journal of Modern Physics C* **15**, 1021–1029 (2004).
24. Lorenz, J. Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C* **18**, 1819–1838 (2007).
25. Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C. *Introduction to algorithms* (MIT press Cambridge, 2009).
26. Sedgewick, R. & Wayne, K. *Algorithms* (Addison-Wesley, 2011).
27. Lorenz, J. Consensus strikes back in the Hegselmann-Krause model of continuous opinion dynamics under bounded confidence. *Journal of Artificial Societies and Social Simulation* **9**, 8 (2006).

## Acknowledgements

This work was supported by the OpLaDyn grant obtained in the 4th round of the TransAtaltic program Digging into Data Challenge (2016–147 ANR OPLADYN TAP-DD2016) and Labex MME-DII (Grant No. ANR reference 11-LABEX-0023).

## Author contributions

L.H. planned research, H.S. developed code, ran simulations and prepared the code for open access sharing, both authors analysed results and wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-64691-0>.

**Correspondence** and requests for materials should be addressed to H.S. or L.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

## Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

[onlineservice@springernature.com](mailto:onlineservice@springernature.com)