A REJECTION MECHANISM IN 2D BOUNDED CONFIDENCE PROVIDES MORE CONFORMITY

Article *in* Advances in Complex Systems · November 2011

DOI: 10.1142/S0219525908001799 · Source: DBLP

3 authors:

Sylvie Huet
Inrae LISC

Guillaume Deffuant
French National Institute for Agriculture, Food, and Environment (INRAE)

Wander Jager
University of Groningen

Advances in Complex Systems

REJECTION MECHANISM IN 2D BOUNDED CONFIDENCE PROVIDES MORE CONFORMITY

HUET S.

Cemagref, Laboratoire d'Ingénierie des Systèmes Complexes
*63172 Aubière, France**
sylvie.huet@cemagref.fr

DEFFUANT G.

Cemagref, Laboratoire d'Ingénierie des Systèmes Complexes
63172 Aubière, France
guillaume.deffuant@cemagref.cf

JAGER W.

Faculty of Economics and Business, University
Groningen, The Netherlands
w.jager@rug.nl

This paper explores the dynamics of attitude change in 2 dimensions (2D) as a result of social interaction. We add a rejection mechanism into the 2D bounded confidence (BC) model proposed by Deffuant et al (2001). Individuals are characterised by two-dimensional continuous attitudes, each associated with an uncertainty u, supposed constant in this first study. Individuals interact by random pairs. If their attitudes are closer than u on both dimensions, or further than u on both dimensions, or closer than u on one dimension and not further than $u + \delta u$ on the other dimension, then the rules of the BC model apply. But if their attitudes are closer than u on one dimension and further than $u + \delta u$ on the other dimension, then the individuals are in a dissonant state. They tend to solve it by shifting away their close attitudes. The model shows metastable clusters, which maintain themselves through opposite influences of competitor clusters. Our analysis and first experiments support the hypothesis that, for a large range of uncertainty values, the number of clusters grows linearly with the inverse of the uncertainty, whereas this growth is quadratic in the BC model.

1. Introduction

Many behaviours, especially in conditions of a high involvement, can be understood as originating from underlying attitudes. For instance, one may vote for an extreme nationalist party because of a negative attitude towards immigrants, or consume organic food because of a positive attitude towards environmentally friendly agriculture. Hence attitudes motivate behaviour and exert selective effects at various stages

*Cemagref, Laboratoire d'Ingénierie des Systèmes Complexes, BP 50085, 63172 F-Aubière of information processing [1]. Consequently, to understand behavioural change it is essential to understand the underlying attitudinal dynamics that give rise to such a change. Attitude is here understood in its psychological meaning as a tendency to evaluate a particular entity with some degree of favour or disfavour. The dynamics of attitudes are closely related to social influence, which includes individual influence on feelings, beliefs and behaviours of others [2]. These dynamics are studied, by experiments in laboratory on individuals and small groups, and are the subject of a variety of theories and assumptions. The most common assumption is a tendency of attitudes to get closer to already similar ones (attraction). A less usual assumption is a tendency to reject the other's attitude if it is psychologically uncomfortable (rejection).

Whereas an abundance of studies have been published in social psychology on the processes leading towards attitude change, relative little attention has been devoted to the interactions between multiple attitudes in social interactions. Yet, the issue of interactions between attitudes in a social interaction context seems to be highly relevant. People often discuss different (unrelated) issues, and shifts on one attitude dimension may have an impact on other dimensions. For example, if a friend, who is having attitudes similar to yours on different issues, is speaking favourably about organic food, about which you have a negative attitude, the resulting dissonance may be resolved by either developing a more positive attitude on organic food as well, or by shifting away from the attitude of your friend on other issues. In contrast, if a person you disagree with on many issues also advocates for organic food, your attitudes are not likely to change as no dissonance is experienced.

This paper proposes a simple model, implementing individuals with both these opposite tendencies (attraction and rejection in some conditions), and studies through agent based computer simulations how a population initially uniformly distributed in the attitude space evolves towards different global patterns. Our main result is that we observe fewer clusters than in the case of dynamics only based on attraction for a large range of uncertainty values. Before going through this result in more detail, we briefly present related research in social simulation and social psychology.

To begin, we consider the assumption of homophily. It assumes that people, especially if they are uncertain about their capacity and knowledge to evaluate a particular object, are more likely to adopt opinions and attitudes of similar others. For example, [3] shows that people like to have opinions similar to the ones of people they interact with. Similarity between receiver and source has a strong impact on the influence level of Word of Mouth [4]. Additionally, [5] suggest that homophily facilitates the flow of information between people because of perceived ease of communication. Secondly, besides a perspective on what drives people's attitudes towards each other, some experiments and theories focus on the forces that may drive people's attitudes apart. At the individual level, the reactance theory [6], the balance theory [7], the motivation to protect oneself [8], and the social judgement theory [9] indicate that a persuasive effort can induce a rejection reaction: the behaviour, and/or the attitude changes in the direction opposite to the persuasion effort. In groups, the social identity theory [10], the self-categorization theory [11] and the optimal distinction theory [12] consider a capacity to differentiate from the individuals who are members of the same group by rejecting their opinions. This rejection is usually called the 'boomerang effect'. The conditions of its occurrence vary from one theory to another. Furthermore, some social psychologists admit that the boomerang effect remains poorly understood [13]. The social judgement theory states that uncertainty plays an important role in both attitude attraction and rejection. The social identity theory stresses that attitude rejection is linked to the salience, at a given time, of the individual social identity. At the individual level, the theories link attitude rejection to loss of control or freedom, or a negative relation with others. From these theories, we retain that attitude rejection occurs when several attitudes are implicitly or explicitly activated. Moreover, it is favoured by a 'dissonant' situation, such as agreement on some attitudes and disagreement on others. As an example, [14] reports about students who, informed that their attitudes regarding a particular issue are close to the one of the Ku Klux Klan, decide to reinterpret this issue and adopt an attitude further away from the one of the Ku Klux Klan.

Another group of interesting results for our purpose comes from the social influence paradigm which has exhibited two important group behaviours: the average consensus [15, 16] and the polarized consensus [17]. The average consensus occurs when the value of an object given by a group after discussion, is close to the average of the values given by individuals before discussion. The polarized consensus takes place when the value given by the group after discussion is significantly more extreme than the average of individual opinions before discussion. Following these studies, Nowak [18], in the social simulation domain, has recommended to investigate the tendency of individual attitudes to become more extreme (polarisation) as well as the tendency of individual to aggregate themselves in groups (clustering).

A large number of computer models are based on homophily. They postulate the existence of an attractive force between agents having close attitudes, which can be formulated using thresholds that determine when agents move towards each other's position [19-23] (see [46] for an interesting review on opinion dynamics). This attraction threshold, also called uncertainty, can be fixed or dynamic [24, 25].

Other models, less numerous and more recent, also include a rejection mechanism in addition to assimilation. In formalising the Social Judgement Theory [26, 27, 35], an individual has two thresholds on an attitude dimension: a first for assimilation and a second one for rejection (the second is assumed higher than the first). In [28], based on the theory of self-categorisation and the meta-contrast principle, an individual tends to minimise the distance to a prototypical opinion which defines his own group and, at the same time, he maximises the distance to an external group. Moreover, a rejection effect appears in [29, 30] as an emerging effect of homophilic individual interactions. This effect is due to the fact that getting closer in the 2-dimensional attitude space may in some cases result in a shift away on the global attitude (which is a weighted sum of the attitudes).

Another form of rejection mechanism can be found in the 'contrarians' of Galam [38, 39] who tend to adopt an attitude which is opposite to the one of the majority (attitudes are supposed binary). The stochastic Sznajd model [40, 41] also includes individuals who oppose to the majority. In this model, a social temperature implies with a probability p the application of the appropriate Sznajd rule for the opinion choice of an agent, and the application of the opposite rule with the probability $1 - p$. Both of these models consider one-dimensional binary attitudes and tend to a particular final state, due to the "contrarians" effect, for which 50% of the population adopts one opinion, and the other 50%, the other opinion.

The attitude dynamic model we propose postulates multidimensional attitudes, like in [27, 29-34, 36, 43-44]. Considering two dimensional attitudes with an equal importance, our main assumption is that, if you strongly disagree with someone on attitude x_1, and are close on attitude x_2, you tend to solve the dissonance by shifting away on attitude x_2. More precisely, when attitudes are both far or both close from each other, we follow the hypotheses of bounded confidence (BC) models [19, 21, 23-25] ([45] for a review): when both are close, the attitudes tend to get closer, when both are far apart, there is no influence. Two models are usually identified as bounded confidence models: the Deffuant-Weisbuch model [19] and the Hegselmann-Krause model [21]. These two models differ regarding their communication regime. Agents of the Hegselmann-Krause model adopt the average opinion of all agents which lie in her area of confidence. Agents of Deffuant-Weisbuch model meet in random pairwise encounters after which they comprise or not. The model presented in this paper follows the Deffuant-Weisbuch communication regime. Therefore, our model is similar to a multi-dimensional bounded confidence model, except that we added the rejection mechanism when people are close on one attitude and far apart on the other.

The next part of this paper describes the model in a simplified version of the ODD framework [37] which is a protocol for describing individual and agent based models in three blocks (Overview, Design concepts, and Details). Following that, we present examples of simulation runs for different parameters, which lead to the hypothesis that the number of clusters grows linearly with the inverse of the uncertainty for a large range of uncertainty values. Other example show that higher uncertainty values tend to less consensus than the classical two dimensional bounded confidence model. Then, we show results of a systematic exploration of the parameter space which support these hypothesis. Finally we will discuss the results and conclude.

2. Overview of the model

2.1. *Purpose of the model*

The purpose of the model is to test the collective effects of a particular rejection mechanism in 2-dimensional bounded confidence models which are based on individual attraction mechanisms. The rejection takes place when individuals are close on one attitude and far on the other.

2.2. *State variables and scales*

We consider a population of N individuals, each having a 2-dimensional attitude or two different attitudes x_1 and x_2, represented by real numbers between -1 and $+1$, and the related uncertainties u_1 and u_2. Uncertainty is a term used for convenience, because this variable may represent confidence in one's own attitude position as well as the motivation to comply with other's attitude positions (social susceptibility). It corresponds also to the latitude of acceptance of the Social Judgement Theory and represents the level of ego-involvement in the value of the attitude. In the following experiments, all individuals have the same uncertainties U on both attitudes $u_1 = u_2 = U$).

2.3. *Process overview and scheduling*

At each time step, we choose a pair of individuals A and B at random, and they may influence each other. More precisely, at each time step, the algorithm is as follows:

```
N times repeat:
- choose couple of individuals (A,B) at random;
- B may influence A.
```

The influence depends on the conditions describing the values of attitudes and uncertainties. Suppose A has attitudes a_1 and a_2 with uncertainties u_1 and u_2, and B has attitudes b_1 and b_2 with uncertainties v_1 and v_2. We only studied the case where all individuals have the same uncertainty for all their attitudes. Thus, for sake of simplicity, we used U instead of u_1 and u_2 in the following since $u_1 = u_2 = U$. Then, A compares its attitudes with the ones of B. Three cases arise.

2.3.1. *Case 1: B is close to A on both attitudes:*

$$\left|a_1^t - b_1^t\right| \leq U \text{ and } \left|a_2^t - b_2^t\right| \leq U \tag{1}$$

Then both attitudes of A get closer to the ones of B:

$$a_1^{t+1} = a_1^t + \mu\left(b_1^t - a_1^t\right) \text{ and } a_2^{t+1} = a_2^t + \mu\left(b_2^t - a_2^t\right) \tag{2}$$

Here μ is a kinetic parameter of the model, representing the velocity of the attraction or the rejection. In our following study, μ has the same value for all individuals.

2.3.2. *Case 2: B is far from A on both attitudes:*

$$\left|a_1^t - b_1^t\right| > U \text{ and } \left|a_2^t - b_2^t\right| > U \tag{3}$$

Then, there is no influence of B on A.

2.3.3. *Case 3: B is far from A on one attitude and close to A on the other.*

$$\left|a_1^t - b_1^t\right| \leq U \text{ and } \left|a_2^t - b_2^t\right| > U \tag{4}$$

We only describe the case where people are close to each other on the attitude 1 and far from each other on the attitude 2 because the case where people are close on the attitude 2 and far on the attitude 1 is analog with a_1 and b_1 interchanged with a_2 and b_2.
Then two cases arise, depending on whether A and B differ strongly on attitude 2. We introduce the positive parameter δ, ruling the intolerance threshold which globally depends on the uncertainty:

Case 3.1: A and B do not differ strongly on attitude 2

$$\left|a_2^t - b_2^t\right| \leq (1+\delta)U \tag{5}$$

Then, the disagreement is not strong enough to trigger the rejection. A approaches B on attitude 1 and ignores B on attitude 2:

$$a_1^{t+1} = a_1^t + \mu\left(b_1^t - a_1^t\right) \tag{6}$$

Case 3.2: A and B differ strongly on attitude 2

$$\left|a_2^t - b_2^t\right| > (1+\delta)U \tag{7}$$

Then, A shifts away from B on attitude 1. The movement is proportional to the distance needed to get b_1 out of A's range of uncertainty around a_1. The specific form of the equation express that people move their own average attitude in order to put the average attitude of the unacceptable other out their own attitude segment. This means people try to adopt a new attitude in such a way that they don't judge themselves similar to the unacceptable other.

$$a_1^{t+1} = a_1^t - \mu \text{ psign } \left(b_1^t - a_1^t\right)\left(U - \left|b_1^t - a_1^t\right|\right) \tag{8}$$

Where psign() is a particular sign function, which returns -1 if its argument is strictly negative, or $+1$ otherwise. The particularity compared with the standard sign function is that when the argument is 0 psign returns $+1$. Moreover, we confine the attitude within the bounds (-1, $+1$) of the attitude space:

$$\text{If } \left|a_1^{t+1}\right| > 1 \text{ then } a_1^{t+1} := \text{sign}(a_1^{t+1}) \tag{9}$$

The following figures illustrate the different types of interactions (attraction, rejection or indifference).

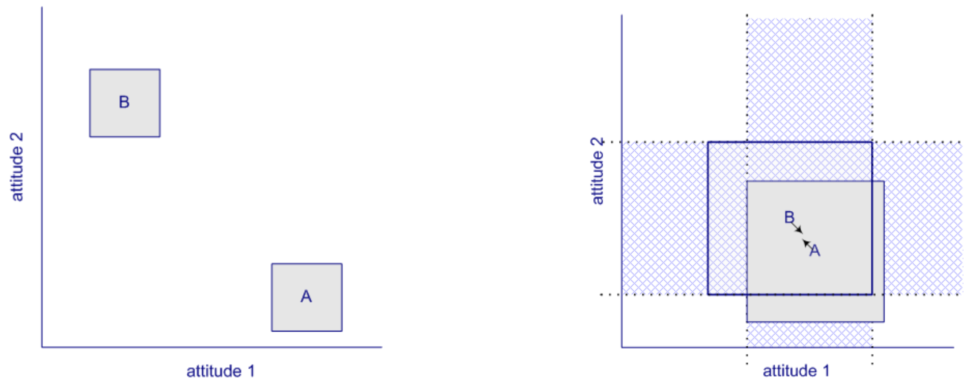

Fig. 1. A and B in a situation of no influence on both dimensions (left) and in situation of attraction (right)

Figure 1 shows on the left the case where A is not influenced by B: they are far from each other on both dimensions. On the right, Figure 1 shows the case where A is attracted by B and vice-versa because they are close to each other. This means each one has his attitude in the other's acceptance zone.

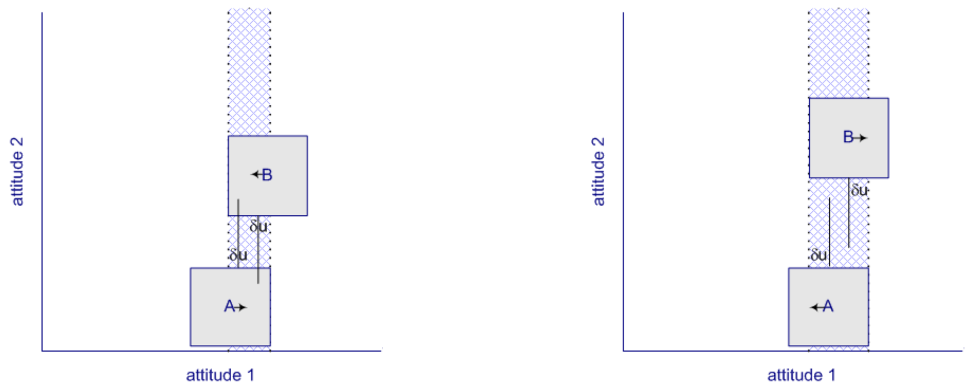

Fig. 2. Left: A and B in a situation of attraction on one dimension (on attitude 1 dimension here) and indifference on the other dimension. Right: A and B in a situation of rejection on one dimension (on attitude 1 dimension here) and indifference on the other dimension

Figure 2 left shows another case where people are close to each other on only one dimension. People are far from each other on one dimension but not far enough to consider the proximity on the other dimension as unacceptable. Thus, they attract each other on the dimension where they are close. On the contrary, Figure 2 right shows the cases where people are far enough from each other on one dimension. The proximity on the other dimension is perceived as unacceptable. Thus, they move away from each other on this dimension.

2.4. *Initialisation*

When we do not vary the population size, we consider a population of 1000 individuals with two attitudes. On each dimension, the attitude is randomly initialised following a uniform distribution comprised between -1 and +1. Uncertainty U is constant and identical on each dimension.

3. Analysis of several examples

In this section, we observe several simulation examples. Their analysis leads in particular to formulate the hypothesis that the number of clusters is a linear function of $1/U$ for weak to average values of uncertainty. Higher uncertainties globally exhibit close final states from those of the 2-dimensional bounded confidence model.

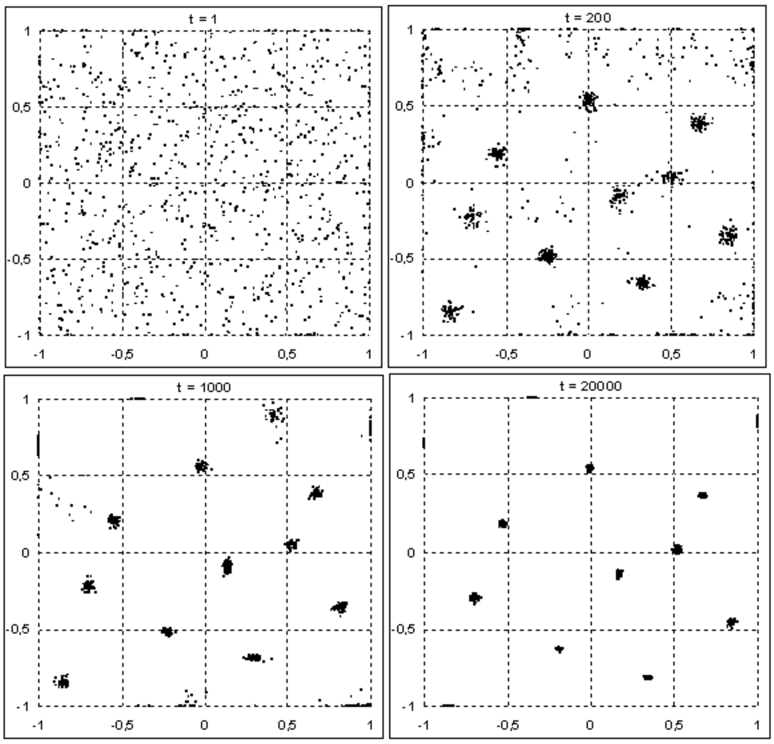

Fig. 3. Initial population uniformly distributed in 2D attitude space. U=0.2, μ=0.3, δ=0. We observe the emergence of metastable clusters, with remaining fluctuations of individuals within the clusters. Moreover, some flat clusters are located on the borders of the attitude domain, containing radicalised individuals

On Figure 3, both attitude axes are represented; black spots indicate the attitude position of individual agents. On this figure, we observe that the population is progressively organized into several clusters. The clusters are not regularly organised on horizontal and vertical lines, as observed with the classical bounded confidence model. They rather tend to be located on oblique lines, which are not strictly regular. Moreover, the individuals fluctuate in the clusters, with constant amplitude of fluctuation, leading to a permanent diversity within the cluster. The reason is that individuals are pushed away from the cluster by other clusters, located close on one dimension and far enough on the other. These movements compensate each other because generally there are several neighbouring clusters that reject the cluster in opposite directions. Moreover, individuals are attracted by the cluster itself, especially if the cluster includes many individuals. Therefore one can say that the clusters are metastable, because if there is a strong perturbation (deletion of a neighbouring cluster) this may dramatically modify the equilibrium. This is a big difference with the classical bounded confidence model, in which after a while, clusters keep concentrating with time, each independently from the others. A second important difference with the classical BC model is that we get clusters on the border of the attitude domain. With the BC model, the first clusters are always inside the attitude domain, on a distance which is about the uncertainty U. On the contrary, with this model, it appears that there are always some clusters which have an extremer attitude position than any individual at the initialisation. These border clusters are flat, because their neighbouring clusters tend to push them away outside the attitude domain. This is a polarisation phenomenon in the sense of Nowak: a part of the individuals gets more extreme. If we removed the constraint to remain within the bounds of the attitudes, the global range of attitude would grow, and we would finally end up with stable clusters, not disturbing each other, in a significantly larger attitude domain.

3.1. *Evolution with uncertainty U=0.2 and intolerance threshold with δ=0*

Figure 3 shows an example of evolution for uncertainties U=0.2, and intolerance parameter δ=0, and the kinetic parameter μ=0.3. The number of time steps t appears on the top of each picture.

3.2. *Spatial organisation of the clusters and hypothesis of linearity of their number with 1/U*

The spatial organisation of the clusters can be further analysed. In this particular case where $\delta = 0$, we note that there is only one cluster on a horizontal or vertical line. Indeed, two clusters on the same horizontal or vertical line is an unstable situation. If the clusters are far, they tend to push each other from the line. If the clusters are close, they tend to merge. This can be checked by considering the histogram of presence of the individuals on each axis on Figure 4. We note that 13 clusters appear on the projection of both axes. Moreover, the distance between the clusters is small enough to trigger rejection (11 clusters is the maximum, to provide a distance of at least U between two consecutive clusters), which explains why the individuals fluctuate in the clusters.

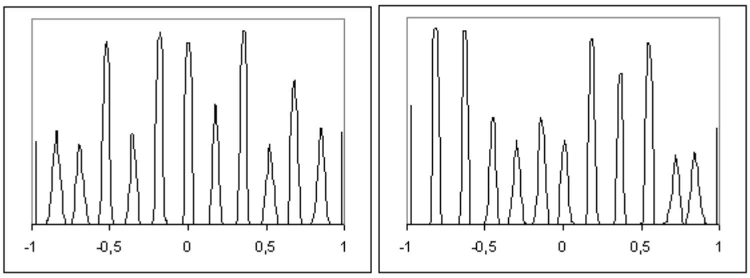

Fig. 4. Kernel density estimator on horizontal axis (left) and vertical axis (right), for the final situation of Figure 3 (U=0.2, t=20000). One notes that the 13 final clusters are regularly distributed on each axis.

In this case, the number of clusters can be analysed on a single axis: there should be a minimum interval between the clusters on each axis which is about the value of U. As we have seen, because of the metastability, it is possible to get slightly smaller intervals. Nevertheless, one can expect a number of clusters varying linearly with $1/U$.

3.3. *Influence of intolerance threshold $\delta > 0$*

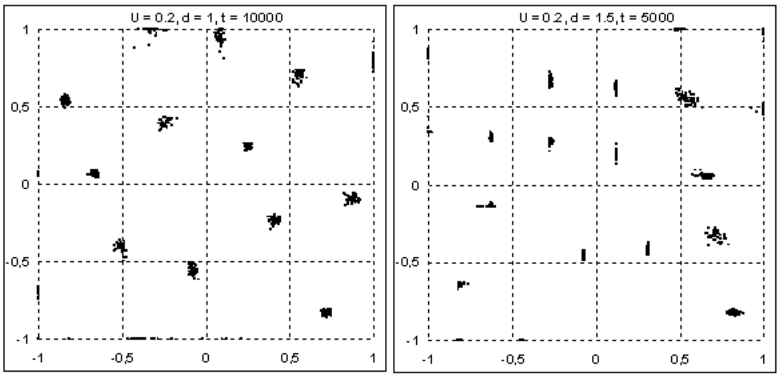

Fig. 5. Example of final configuration for $U = 0.2$, μ=0.3, $N = 1000$, δ=1 (left), δ=1.5 (right). It is possible to get 2 clusters on the same horizontal and vertical line, which is unstable when δ=0. Moreover, for δ=1.5, some clusters are flat inside the attitude domain.

When the intolerance threshold gets higher, the conditions for rejection are more restricted: the disagreement on one attitude must be higher. Figure 5 shows two examples of final attractors, for U=0.2, δ=1 (left) and δ=1.5 (right). The number of clusters appears to increase with δ .

We observe that, for these values of δ, it becomes possible to get two clusters on the same horizontal or vertical line, when they are not too far apart (they remain in the tolerance zone). This explains why there are more clusters. Nevertheless, we can hypothesise that this number should still vary linearly with $1/U$, but with a higher coefficient.

Moreover, for $\delta = 1.5$, we observe flat clusters inside the attitude domain, whereas this did not take place for $\delta = 1$. Such a flat cluster appears when all the neighbour clusters are on the same line in the tolerance zone, or far on both attitudes. The rejection interactions are therefore only in one direction.

3.4. *Different values of uncertainty U with intolerance threshold δ=0*

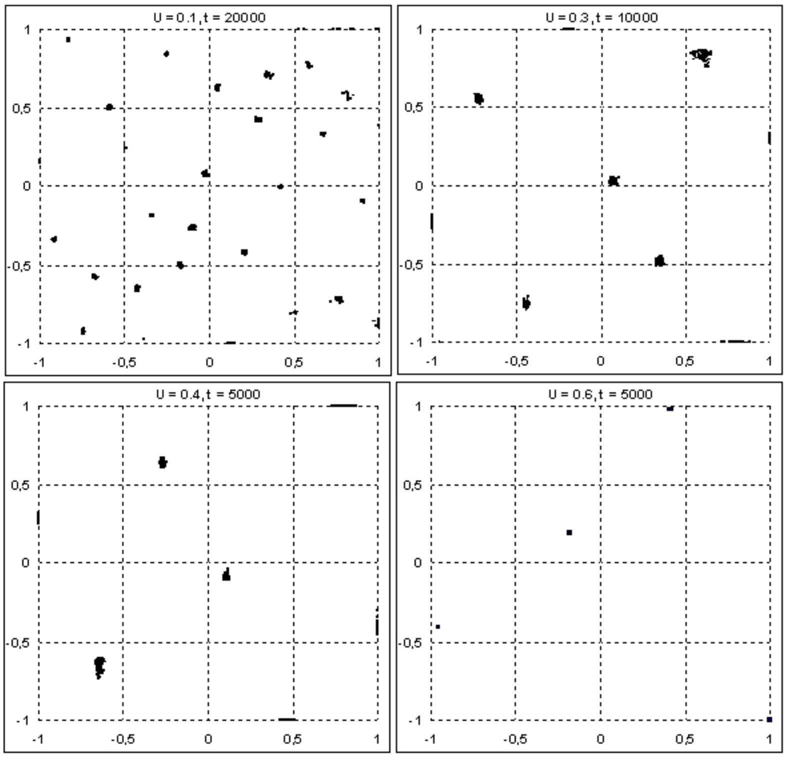

Fig. 6. Examples of attractor configurations for different values of uncertainty U and intolerance parameter $\delta = 0$, $\mu = 0.3$. Population size $N = 1000$.

Figure 6 shows several attractor configurations for different values of uncertainty

U. This first exploration suggests that the number of clusters decreases with U, like with the BC model. The observations made on our first simulation extend to these cases: Oscillations of individuals remain, with higher oscillations when U increases, and spatially organised to avoid two clusters on the same horizontal or vertical line. In each case, we get flat clusters with the maximum value for one attitude (polarisation).

For $U = 0.6$, we observe that the clusters become very concentrated, like in the simple BC model even if, for the same uncertainty value, the simple BC model has only one cluster for a population of 1000 individuals. The reason is that with 4 clusters, the intervals between the clusters on a same horizontal or vertical line can easily be higher than U, and therefore avoid to generate a competition between the attraction in the cluster and the rejection from the neighbouring clusters.

4. Systematic analysis of the number of clusters

We are interested in comparing the final number of attitude clusters with the one generated by the standard BC model proposed by [19]. First we describe how we compute the final number of clusters. Then, we analyse this final number of cluster regarding two different behaviour zones of the BC model (see [19, 47] for more details). The first zone is a zone for which the population is organized in several clusters; it is the object of our second point. The second zone is a zone for which the wide majority of people go in one cluster; it is the object of our third point.

4.1. *Computing the number of clusters*

From the individual-based simulations, we collect the average, minimum and maximum final number of clusters. To compute the number of clusters, we define a minimum distance ϵ between attitudes, below which we consider that they belong to the same cluster. We compute the clusters as groups of agents such that between any couple of agents of opinions x and x' in the group, there is a list of agents in the group of opinions $(x_1, x_2, \ldots, x_k)$ making a chain of couples distanced from each other of at most an Euclidian distance lower than ϵ. The following pseudo-code can be used to compute the clusters; `necessaryToLookAt` is a table containing the identification number of each individual for all the population:

```
for all i of the population
 if necessaryToLookAt[i] > 0
   currentCluster.add(i)
   compt++;
   necessaryToLookAt[i] = 0
   while currentCluster.isNotEmpty()
     for all j of the population
       if necessaryToLookAt[j] > 0
         if distance(pop[currentCluster.get(0)],pop[j]<epsilon)
```

```
            necessaryToLookAt[j] = 0
            currentCluster.add(j)
            compt++
      currentCluster.remove(0)
    nbClusters++
  if compt = populationSize then i = populationSize
```

In practice, we chose ϵ=0.2U and we neglected the clusters of size lower or equal to 3 individuals. The simulations are stopped after 1,000,000 iterations. They can be stopped before if the number of clusters has not changed after 100,000 iterations. Even though [47] have demonstrated the interest of minor clusters in wide populations and for high value of μ [48], we focus only on major clusters in this first study.

4.2. *Final number of clusters on the 'multiple clusters zone' of the BC model*

The BC model, in one dimension, yields a final number of clusters n_c in a population initialised with a uniform law on an attitude space of width $2M$, with all the same uncertainty U, which can be approximated by:

$$n_c \approx \frac{M}{U} \tag{10}$$

In the 2-dimensional case, when both attitude axes are adjusted independently and all have the same uncertainty U on both attitude dimensions, this rule is repeated on all lines of the space, therefore we get:

$$n_c \approx \left(\frac{M}{U}\right)^2 \tag{11}$$

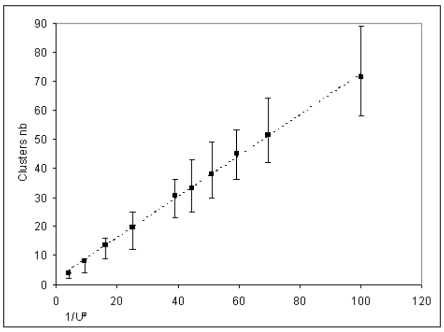

Fig. 7. Average final number of clusters of the 2D bounded confidence model as a function of $1/U^2$. Error bars indicates minimum and maximum obtained on 30 replicas.

This result is confirmed by Figure 7 which presents on abscissa $1/U^2$ and on y-axis, the average number of clusters obtained on 30 replicas.

We focus on the zone where BC model exhibits a final state including several clusters. For our attitudinal domain (attitudes beteween -1 and +1), it goes from $U = 0$ to $U = 0.54$. Figure 8 shows the number of clusters obtained with rejection dynamics, for different values of U and δ. These results confirm the hypothesis of linearity of the number of clusters with $1/U$ for $\delta = 0$ and $\delta = 0.5$ (left) on this zone.

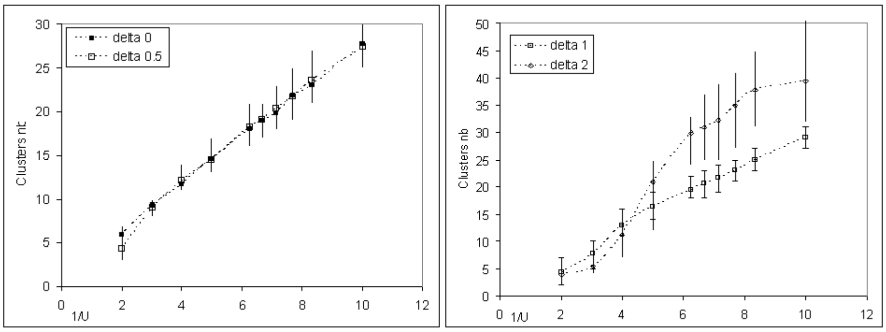

Fig. 8. Mean final number of clusters for the model with rejection as a function of $1/U$, for various values of δ. N=1000 and δ=0.3. The error bars are the minimum and maximum numbers met in 30 replicas. On the left, for δ=0 and δ=0.5, the number of clusters seems linear with $1/U$. On the right, the behaviour is not linear for large U.

For $\delta = 1$, 1.5, 2 and 3, there is a non-linearity for U larger than 1 (only 1 and 2 are presented on the figure). When U is larger than 0.3, and δ is large, the conditions for rejection are much constrained by the size of the domain: two individuals must be at both sides of the domain. Most of the interactions correspond therefore to the standard BC, and the curve is therefore quadratic. When U decreases ($1/U$ grows), the rejection becomes more common and the curve becomes linear.

Let's now verify if these results are robust when the population size varies (100, 1000 and 5000 individuals). Figure 9 shows the number of clusters obtained with rejection dynamics for $\delta = 0$ and different values of U and population size N. To be able to compare the different population sizes, we count all the clusters (no threshold). From Figure 9, we note that the population size does not change our conclusion: the final number of clusters tends to be linear with $1/U$.

4.3. *Final number of cluster on the "one major cluster" zone of the BC model*

We now focus on the zone where the BC model exhibits one single major cluster. For our attitudinal domain, it begins for $U > 0.54$. Since for $U > 1$, the rejection mechanism cannot work (all attitudes are at a distance which is within the attraction

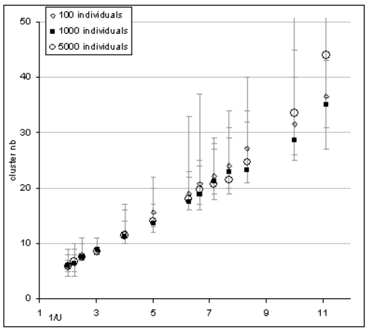

Fig. 9. Mean final number of clusters for the model with rejection as a function of $1/U$, for $\delta = 0$ and various value of population size N (100, 1000, 5000) and U (from U=0.09 to U=0.5). The error bars are the minimum and maximum numbers met in 30 replicas. The number of clusters seems also linear with $1/U$.

range), we only study the U value range from 0.54 to 1 (indeed, for $U > 1$, all people go in one unique central cluster, exactly as in the BC model). Figure 10 shows the results in this zone for different population sizes. We immediately see that the final number of clusters is not linear with $1/U$.

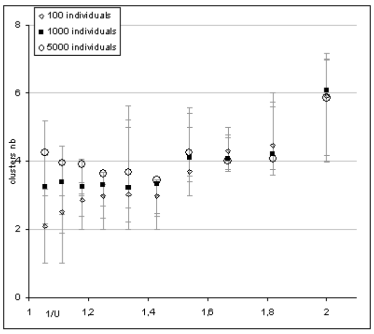

Fig. 10. Mean final number of clusters for the model with rejection as a function of $1/U$, for $\delta = 0$ and various value of N (100, 1000, 5000) and U (from U=0.55 to U=0.95). The error bars are the minimum and maximum numbers met in 30 replicas. The number of clusters is not linear with $1/U$.

In the standard BC model proposed by [19], the final state for this zone is one major cluster containing a large majority of the population with, in some cases, several very minor clusters when the population is very large [46]. Figure 11 shows that our model has also, on the zone of U values, one major cluster containing a majority of the population (from a part of half to the whole population depending on the parameter value). In our model with a rejection mechanism, we finally obtain between two and six final clusters as shown on Figure 10. Are the non-major clusters

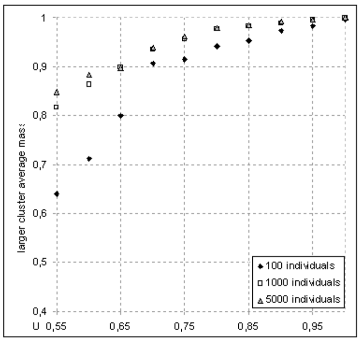

Fig. 11. Average mass of the final larger cluster for $\delta = 0$ and various value of N (100, 1000, 5000) and U (from U=0.55 to U=0.95).

the same as those of the BC model ? From [48], we know the very small clusters of the standard BC model are very numerous and do not exist for low values of μ. In our attraction-rejection model, minor clusters are not numerous, from one to five on average, and larger than those of the classical BC model. Moreover, they remain when we run simulation with a value of μ equal to 0.01. Finally, our population size is not large enough to really observe 'minor clusters' in the sense used by [47].

5. Discussion and conclusion

In the model of 2-dimensional attitude dynamics we propose, an agent shifts away from a close attitude on one axis when the interlocutor is far on the other axis. We assume that this is a way to solve the dissonance between the attitude axes. The distance threshold to trigger rejection depends on the intolerance parameter δ and on the uncertainty U, which may define a non-commitment zone, in which the dissonance is tolerated. When the conditions of rejection are not met, that is when we exclude the case where two individuals differ strongly on one attitude and are similar on the other attitude, the model behaves exactly like the 2D bounded confidence (BC) model.

The first explorations of this model, in the simple case where all uncertainties are the same, show several striking results, in comparison with the 2D BC model:

- When the uncertainty is lower than 1, allowing the rejection to occur, the dynamics leads to several metastable clusters, which are generally in competition and tend to reject each other. The stability is due to contradictory rejections from neighbouring clusters, which compensate each other. If one of its neighbouring clusters is removed, the position of a cluster changes significantly, and it may even disappear. Moreover, individuals belonging to a cluster are in constant fluctuation around the cluster center, with amplitudes depending on the cluster size and on the proximity of competing clusters. In this respect, the configuration is very different from the one obtained with simple BC model where, after a while, clusters keep concentrating with time, each independently from the other.

- Several clusters are moving towards the limits of the attitude domain. This may be interpreted as a radicalisation of a part of the population, which reaches the maximum absolute value of one of the attitudes. This never happens with the 2D BC model.
- In the case where the intolerance threshold $\delta = 0$, two clusters cannot be maintained on the same horizontal or vertical line. Therefore, the clusters tend to occupy points of the space where they are as far as possible from other clusters on each axis. This analysis suggests a number of clusters growing linearly with $1/U$ for values of U for which the 2D BC model exhibits several clusters called 'major' and 'central' clusters by [47]. However, for the 2D BC model, for this same range of U values, the cluster number grows quadratically with $1/U$ in the 2D BC model. When δ grows, configurations with more than one cluster on a line may be stabilised, but this number is limited by the size of the tolerance zone. Therefore, the growth of the cluster number should still be linear, but with a factor growing with δ. First systematic experiments support this statement, for different population sizes.
- For values of U for which the 2D BC model exhibits only one cluster called 'central' in [47], our model does not follow the same law and tends to have less consensus than the 2D BC model. Indeed, depending on the parameter value, it exhibits from two to six clusters on average, with one major cluster containing a majority of people. The other clusters are generally on the limits of the attitude domain. [47] and [48] show that the 2D BC model has, for a subpart of this zone of U values, numerous very minor clusters when μ is high and when the population size is wide. However these very minor clusters, even if they are located close to the bound of the attitudinal space, are different from the minor clusters of extremists of our model.

These results suggest several points to discuss.

- The metastability of the clusters is due to the bounds we impose on the attitude values. Indeed, without these bounds, the attitudes grow until the distance between the clusters is higher than the uncertainty in all directions. Then, the clusters do not influence each other, and they keep concentrating as in the BC model. First simulations performed on the same model with an unbounded attitude domain indicate that the final number of clusters is close to the one obtained with the bounded domain. However, the unbounded case should be the object of a particular study. In any case, the metastability of the clusters is an particular feature of this model, which better fits real group dynamics than the perfect similarity obtained without a bound (or by a standard BC model).
- Even without bounds, we obtain a global result which shows strong simi-

larities with social identity and self-categorization theories. Our individuals tend to minimise their in-group distance and maximise their out-group distance (to competing groups). We also get some polarized groups (which have more extreme opinions than all the individuals initially). This reminds of the results from Moscovici and Zavalloni [17]. Therefore, with a model considering only paired interactions, we get group dynamics which seem to make sense in a social psychology perspective.

- However, the model remains highly simplified, and a challenge that remains is checking if these interesting properties last when adding more sophisticated hypotheses. In particular, in our model, all attitudes are considered to have the same weight on the behaviour, whereas one expects that only disagreements on attitudes deeply related to social identity can lead to rejection. To take this aspect into account, we should consider attitudes of different types.
- We have chosen the particular communication regime of the Deffuant-Weisbuch model. Considering the communication regime as a parameter of the bounded confidence model as in the formulation proposed by [42], would be worth investigating. Moreover, it would be a logical extension to relate the chance of interacting to the attitude similarity between the agents, thus reflecting principles of preferential attachment.

In future research, we plan to continue to explore the properties of this model. In particular, introducing extremists like in [24] could produce unexpected effects.

6. Acknowledgement

We thanks the anonymous reviewers for their suggestions of improvements and their relevant remarks.

References

[1] Eagly, A.H. and S. Chaiken, *The Psychology of Attitudes* (Fort Worth, TX: Harcourt Brace Jovanovich, Thomson/Wadsworth ed. 1993).
[2] Mason, W.A., F.R. Conrey, and E.R. Smith, Situating Social Influence Processes: Dynamic, Multidimensional Flows of Influence Within Social Networks. *Personality and Social Psychology Review*, 2007(11): pp. 279-300.
[3] Festinger, L., *A Theory of Cognitive Dissonance* (Stanford, CA: Stanford University Press ed., 1957).
[4] Moschis, G.P., Social Comparison and Informal Group Influence. *Journal of Marketing Research*, 1976. 13(3): pp. 237-244.
[5] Price, L.L. and L.F. Feick, The role of interpersonal sources in external search: an informational perspective. *Advances in Consumer Research*, 1984(11): pp. 250-253.
[6] Brehm, J.W., *A theory of psychological reactance* (New York Academic Press, 1966).
[7] Heider, F., Attitudes and Cognitive Organization. *Journal of Psychology*, 1946(21): pp. 107-112.
[8] Rogers, R.W. and S. Prentice-Dunn, *Protection motivation theory*, in Handbook of health behavior research, D. Gochman, Editor (New York:Plenum, 1997) pp. 113-132.

[9] Sherif, M. and C.I. Hovland,*Social Judgement* (New Haven, CT: Yale University Press, 1961).
[10] Tajfel, H., *Differentiation between social groups: Studies in the social psychology of intergroup relations* (London: Academic Press, 1978).
[11] Turner, J.C., et al., *Rediscovering the social groups: A self-categorization theory* (Cambridge, MA: Basil Blackwell, 1987).
[12] Brewer, M.B., The Social Self: On Being the Same and Different at the Same Time. *Pers. Soc. Psychol. Bulletin*, 1991. 17(5): pp. 475-482.
[13] Eagly, A.H. and S. Chaiken, *Process Theories of Attitude Formation and Change: Attribution Approaches and Social Judgment Theory*, in The Psychology of Attitudes, A.H.C. Eagly, S., Editor (Fort Worth, TX: Harcourt Brace Jovanovich, Thomson/Wadsworth, 1993) pp. 351-388.
[14] Wood, W., et al., Self-definition, defensive processing, and influence: the normative impact of majority and minority groups. *Journal of Personality and Social Psychology*, 1996(71): pp. 1181-1193.
[15] Allport, G.W., *Social Psychology* (Boston: Houghton Mifflin, 1924).
[16] Sherif, M., A study of some social factors in perception. *Archives of Psychology*, 1935. 27(187).
[17] Moscovici, S. and M. Zavalloni, The Group as a Polarizer Of Attitudes. *Journal of Personality and Social Psychology*, 1969. 12(2): pp. 125-135.
[18] Nowak, A. and R.R. Vallacher, *Dynamical Social Psychology* (The Guilford Press, 72 Spring Street, New York, NY 10012. 318, 1998).
[19] Deffuant, G., et al., Mixing beliefs among interacting agents. *Advances in Complex Systems*, 2001(3): pp. 87-98.
[20] Galam, S. and S. Moscovici, Towards a theory of collective phenomena: consensus and attitude changes in groups. *J. Soc. Psychology*, 1991(21): pp. 49-74.
[21] Hegselmann, R. and U. Krause, Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation. *Journal of Artificial Societies and Social Simulation*, 2002. 5(3).
[22] Latané, B., Dynamic Social Impact: The Creation of Culture by Communication. *Journal of Communication*, 1996. 46(4): pp. 13-25.
[23] Weisbuch, G., et al., Meet, discuss and segregate ! *Complexity*, 2002. 7(3): pp. 55-63.
[24] Deffuant, G., et al., How can extremism prevail ? A study based on the relative agreement interaction model. *Journal of Artificial Societies and Social Simulation*, 2002. 5(4).
[25] Weisbuch, G., G. Deffuant, and F. Amblard, Persuasion dynamics. *Physica A*, 2005(353): pp. 555-575.
[26] Jager, W. and F. Amblard, A dynamical perspective on attitude change. *NAACSOS Conference 2004*. June 27-29, 2004, Pittsburgh PA http://www.casos.cs.cmu.edu/naacsos/.
[27] Jager, W. and F. Amblard. Multiple Attitude Dynamics in Large Populations. *Agent 2005 Conference on Generative Social Processes, Models, and Mechanisms*. 2005: cosponsored by Argonne National Laboratory and The University of Chicago.
[28] Salzarulo, L., A continuous Opinion Dynamics Model Based on the Principle of Meta-Contrast. *Journal of Artificial Societies and Social Simulation*, 2006. 9(1).
[29] Urbig, D. and R. Malitz. Dynamics of structured attitudes and opinions. *Third Conference of the European Social Simulation Association*. 2005. September 5-8, Koblenz, Germany.
[30] Urbig, D. and R. Malitz. Drifting to more extreme but balanced atttiudes: Multidimensionnal attitudes andselective exposure. in *4th Conference of the European Social*

Simulation Association. 2007. Toulouse, 10-14 September.
[31] Axelrod, R., The Dissemination of Culture - A model with local convergence and global polarization. *Journal of Conflict Resolution*, 1997. 41(2): pp. 203-226.
[32] Deffuant, G. and S. Huet, Propagation effects of filtering incongruent information. *Journal of Business Research*, 2007. 60(8): pp. 816-825.
[33] Fortunato, S., et al., Vector Opinion Dynamics in a Bounded Confidence Consensus Model. *International Journal of Modern Physics*, 2005. 16(10): pp. 1535-1551
[34] Huet, S. and G. Deffuant. When do Interactions Increase or Decrease Primacy Effect? in *Model-2-Model*. 2007. Marseille.
[35] Jager, W. and F. Amblard, Uniformity, Bipolarization and Pluriformity Captured as Generic Stylized Behavior with an Agent-Based Simlation Model of Attitude Change. *Computational and Mathematical Organization Theory*, 2004(10): pp. 295-303.
[36] T. Carletti, D. Fanelli, A. Guarino, F. Bagnoli and A. Guazzini, Opinion Dynamics in an Open Community, ArXiv.org, Vol. 0712.1024, 2008.
[37] Grimm, V., et al., A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 2006(198): pp. 115-126.
[38] S. Galam, Contrarian deterministic effects on opinion dynamics: 'the hung elections scenario'. *Physica A*, 2004, 333: pp. 453-460.
[39] C. Borghesi and S. Galam, Chaotic, staggered and polarized dynamics in opinion forming: the contrarian effect. *Physical Review E* , 2006, 73: 9 pages.
[40] M. S. de La Lama, J. M. Lopez and H.S. Wio, Spontaneous emergence of contrarian-like behaviour in an opinion spreading model. *Europhys. Lett.*, 2005, 72: pp. 851-7.
[41] H. S. Wio, M. S. de la Lama and J. M. Lopez, Contrarian-like behavior and system size stochastic resonance in an opinion spreading model. *Physica A*, 2006, 371: pp. 108-11.
[42] D. Urbig and J. Lorenz, Communication regimes in opinion dynamics: Changing the number of communicating agents, Second Conference of the European Social Simulation Association (ESSA), (Valladolid, Spain (September 16-19): 2004).
[43] J. Lorenz, Continuous opinion dynamics of multidimensional allocation problems under bounded confidence: More dimensions lead to better chances for consensus, *European Journal of Economic and Social Systems*, EJESS, 2006(19): pp. 213-27.
[44] J. Lorenz, Fostering Consensus in Multidimensional Continuous Opinion Dynamics under bounded Confidence, in Helbing Dirk ed., *Managing Complexity: Insights, Concepts, Applications* , (Budapest Springer Berlin / Heidelberg, Understanding Complex Systems), 2008: 321-334.
[45] J. Lorenz, Continuous Opinion Dynamics Under Bounded Confidence: A Survey, *International Journal of Modern Physics C*, 2007(18), 12: pp. 1819-1838.
[46] C. Castellano, S. Fortunato and V. Loreto, Statistical physics of social dynamics, (2007), http://arxiv.org/abs/0710.3256
[47] E. Ben-Naim, P. L. Krapivsky and S. Redner, Bifurcations and Patterns in Compromise Processes, *Physica D: Non linear phenomena*, 2003(184): pp. 190-204.
[48] M.F. Laguna, G. Abramson and D. H. Zanette, Minorities in a Model for Opinion Formation, *Complexity* , 2004(9): pp. 31-6.