Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org
pp. **X–XX**

OPINION DYNAMICS UNDER THE INFLUENCE OF RADICAL GROUPS, CHARISMATIC LEADERS, AND OTHER CONSTANT SIGNALS: A SIMPLE UNIFYING MODEL

Rainer Hegselmann

Bayreuth University
Universitaetsstrasse 30
95440 Bayreuth, Germany

Ulrich Krause

Bremen University
Bibliotheksstrasse 1
28359 Bremen, Germany

Abstract. By a simple extension of the bounded confidence model, it is possible to model the influence of a radical group, or a charismatic leader on the opinion dynamics of 'normal' agents that update their opinions under both, the influence of their normal peers, and the additional influence of the radical group or a charismatic leader. From a more abstract point of view, we model the influence of a signal, that is constant, may have different intensities, and is 'heard' only by agents with opinions, that are not too far away. For such a dynamic a Constant Signal Theorem is proven. In the model we get a lot of surprising effects. For instance, the more intensive signal may have less effect; more radicals may lead to less radicalization of normal agents. The model is an extremely simple conceptual model. Under some assumptions the whole parameter space can be analyzed. The model inspires new possible explanations, new perspectives for empirical studies, and new ideas for prevention or intervention policies.

1. Introduction.

In the following we define, analyze, and interpret a model, that tries to cover some important aspects of an opinion dynamics under the influence of a radical group or a charismatic leader.[1] The basic structure and central components are the following: There is an ongoing exchange of opinions among 'normal' agents. The normals are neither members of the radical group, nor are they charismatic. There is an ongoing opinion exchange among the normals. A dynamical system describes the normals' updating of their opinions. Radical groups or charismatic leaders influence that dynamical system with their 'very radical' and somehow 'very strong messages': Their messages are an *additional* input for the normals' updating of their opinions.

Our radicals and charismatic leaders are *highly stylized figures*. Each of them is characterized by just *one* assumption: A *radical group* has, compared to normal

2000 *Mathematics Subject Classification.* Primary: 90B10, 90B18, 91D30, 37N40; Secondary: 15B51, 39A20.

Key words and phrases. Opinion dynamics, bounded confidence, constant signals.

[1] A first version and short analysis of the model was given in [8].

agents, a comparatively stable in-group consensus on an extreme opinion. A *charismatic leader* counts for normal agents, that actually are under his/her influence, much more than other normal agents.

That are *informal* statements. To make them precise, requires, that we first explicitly formulate the dynamical systems that describes the normals' opinion dynamics.

1.1. **The BC model.** For the ongoing and underlying opinion dynamics of normals we will use the so called *bounded confidence model* (for short: *BC*-model). It is a high-dimensional dynamical system. The basic idea is: Agents take seriously those others, whose opinions are *not too far away* from their own opinion. Stated more precisely, the assumptions are:

1. There is a set of n agents; $i, j \in I$.
2. Time is discrete; $t = 0, 1, 2, \dots$.
3. Each individual starts with a certain opinion, given by a real number; $x_i(t_0) \in [0, 1]$.
4. The profile of opinions at time t is $X(t) = x_1(t), \dots, x_i(t), \dots, x_j(t), \dots, x_n(t)$.
5. Each agents i takes into account only 'reasonable' others. Reasonable are those individuals j whose opinions are not too far away, i.e. for which $|x_i(t) - x_j(t)| \leq \epsilon$, where ϵ is the *confidence level* that determines the size of the *confidence interval*.
6. The set of all others, that i takes into account at time t, is:
$$I\big(i, X(t)\big) = \{j \big| |x_i(t) - x_j(t)| \leq \epsilon\}. \tag{1}$$
7. The agents update their opinions. The next period's opinion of agent i is the average opinion of all those, which i takes seriously:
$$x_i(t+1) = \frac{1}{\#\big(I(i, Xt)\big)} \sum_{j \in I\big(i, X(t)\big)} x_j(t). \tag{2}$$

The *BC*-model was extensively analyzed in [10]. It received a lot of attention. For a survey see [15]. There are variants and alternatives (see [6], [12], [19], [3], [1]). The best understood and most used *linear* alternative is the so called *DeGroot*-model (see [7]). In that model the agents assign weights to others. The weights are *independent* of the opinion distance to others. Updating, then, is weighted averaging. – Throughout the following, we always assume the *BC*-dynamics. The effects of *other* underlying opinion dynamics should be studied. But we can't do that here.

1.2. **A group of radicals.** For the modified *BC*-model we now assume that there are *two* groups of agents: The first group, the *normals*, have opinions from the interval $[0, 1]$, they all have a strictly positive, constant, and symmetric $\epsilon > 0$. The second group, the *radicals*, have all the same opinion R, more or less close to the upper bound of the unit interval, e.g. $R = 0.9$, or even holding the most extreme position $R = 1.0$. (Alternatively, we might locate the radical position R close to the lower bound of our opinion space. Whatever we do, for the following it does not matter.) The radicals stick to their opinion: Their opinion is R, and that forever. The size of the radical group will matter. We refer to the number of radicals by $\#_R$.

Normals update according to equation (2). Now the modification comes: Whenever the radicals are in a normal agent i's confidence interval, i.e. whenever $|x_i(t) -$

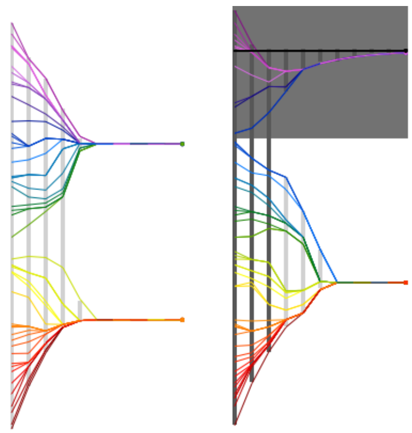

FIGURE 1. 50 normals, the same random start distribution of normals in both runs, $\epsilon = 0.2, R = 0.9$. *Left*: no radicals. *Right*: 5 radicals.

$R| \leq \epsilon$, then *the whole group* of radicals is in $I\big(i, X(t)\big)$. Since the radical group has $\#_R$ members, the radical position R is $\#_R$-times in $I\big(i, X(t)\big)$.

Figure 1 shows two single runs with the same uniform start distribution for 50 normal agents with $\epsilon = 0.2$. In the left figure there are no radicals. Light grey vertical lines between *neighboring* opinions indicate that their distance is not greater than ϵ. As a consequence they influence each other mutually. In the right figure a group of 5 radicals is added. Their opinion is $R = 0.9$. The black horizontal line is their trajectory. The dark grey area indicates that part of the opinion space, in which all normals, given the size of their confidence interval, are under the *direct* influence of the radicals. Dark grey vertical lines indicate the existence and length of a chain of direct *or indirect* influence of radicals on normals: Normals in the dark grey area are directly influenced by the radicals. But the radicals' influence does not end there. A normal j outside that area is indirectly influenced by a normal i inside the area of direct radical influence if $|x_i(t) - x_j(t)| \leq \epsilon$. Agent j, then, may influence other agents k outside the area of direct radical influence with opinions not further away than ϵ, and so forth. *Figure* 1 shows a far reaching indirect influence of the radical group for the first 3 periods: The chain of radical influence pervades the whole opinion profile, i.e. the radicals influence *all* normals. In period 4 that chain breaks. An upper part of the opinion profile converges towards the radical position. Below, the normals end up (obviously in finite time!) in a cluster. That cluster is far away from the radical position R. However, compared to the dynamics without the 5 radicals (see *Figure* 1 left) the lower cluster's final position is shifted in the direction of the radical position. Obviously indirect radical influence matters.

1.3. **Charismatic leaders and constant signals in general.** Charismatic leadership has many forms. It may use all sorts of communication channels. To be maintained, it may require continuous success (in terms of the given context). We abstract away all these facets, except for one: a charismatic leader counts for normal agents, that are under his/her influence, *much more* than other normal agents. Reduced to this effect, we do *not* need any further extension of the modified model; a *re-interpretation* is sufficient: We consider a radical group with $\#_R$ members as *one* person (a kind of 'super agent') that counts $\#_R$-times for normals, given the charismatic leader is within their confidence interval. Thereby, the radicals' group size $\#_R$ turns into a kind of *degree of charismaticity*. Since this ('reduced') type of charismaticity is not bound to extreme positions, R may now be any value in the interval $[0, 1]$.[2]

The ease of this re-interpretation is an eye-opener. There are much more possible applications of our conceptual model—once we look at it from a more abstract point of view. We can formulate that abstract point of view in a certain language: the *language of signals*. In that language the very essence of our framework is this:

1. Period by period there is a constant signal $R \in [0, 1]$.
2. In each period the signal can be sent $\#_R$ times; $\#_R = 1, 2, \dots$. We refer to $\#_R$ as the *strength* or *intensity* of the signal.
3. The signal is *received* by an agent i, if and only if the signal is within i's confidence interval, i.e. $|x_i(t) - R| \leq \epsilon$.
4. *If* the signal is received by an agent i, then, in addition to all opinions of other agents $j \in \{j \big| |x_i(t) - x_j(t)| \leq \epsilon\}$ the signal counts in i's updating procedure according to the strength of the signal, i.e. $\#_R$-times.
5. Updating is averaging over all what counts.

The dynamics of opinions under the influence of a radical group or a charismatic leader are *instances* of such a framework. But there are more. For instance, a *group of dogmatists*: They stick to their opinion forever, but their opinion R need not be an extreme opinion. Or a *campaign*: Using this or that channel of communication, a message R is sent with this or that campaign intensity. In abstract terms, the following is a study on the *effects of a distance depending constant signal of a certain intensity* on an underlying *BC*-dynamics.

Our study has both, a rigorous *analytical*, and an *experimental* part, that is based upon computer simulations. The analytical part consist of a theorem, the *Constant Signal Theorem*, and its proof. The theorem says, that for any confidence level ϵ and signal intensity $\#_R$, the system converges to a stable segregation of the normal agents into those, who approach the radical position R, and a clustering in finite time of all the other normal agents. The theorem is *fundamental* for what follows. Since the theorem holds, each single run, shown in a figure above or below, is an illustration.[3] However, the theorem doesn't say anything about the exact numbers of normals, that end up at R—everything is possible, from none to all. Nor does it say something about the cluster structure of all the normals, that 'escape the signal'. Though fundamental, we put the theorem and its proof in an appendix. The main focus of this paper is on *one decisive number*: the number of normals that end up at R. In terms of radicalization: Our focus is the number of normals that end up radicalized. Since no rigorous analytical results are available, simulation will be our method. To make the article a bit more vivid (and seductive), we describe and present *all* results in the language of *one* of the application instances: an opinion dynamics under the influence of a group of radicals. The transfer to other instances is left to the reader.

[2] If R actually is an extreme value, our conceptual model can also be interpreted as covering a situation of a radical group *with* a radical leader.

[3] For instance, figure 1: According to the Appendix, the chain of influence at period t is given by the set $J(t)$. Therefore, in figure 1, $J(t) = I$ for $t = 1, 2, 3$ and $J(4)$ is the upper part of the opinion profile. Figure figure 1 illustrates the segregation as stated in the *Constant Signal Theorem*, where $J = J(4)$ and $T = 4$.

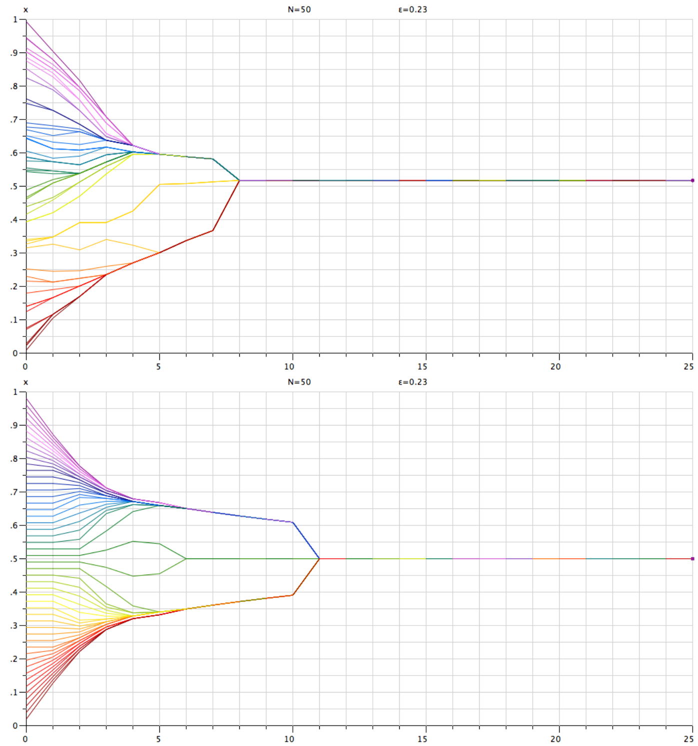

FIGURE 2. Dynamics for 50 agents and a confidence level $\epsilon = 0.23$. *Top*: Random uniform start distribution. *Bottom*: Expected value start distributions. For $t = 0$ the ith opinion is $i/51$.

There are other models, that study processes of radicalization (see [5] and the follow-up studies), or the incremental establishment of charismatic leaders (see [2]). Our approach differs: We take the group of radicals as *given*. Our model is extremely simple. Basically we will be able, to compute for the *whole* parameter space the number of finally radicalized normals. These numbers and the patterns therein are *macro* effects. Our approach will allow, to lay bare *in all detail* the cogs and wheels, working on an underlying *micro* layer, that – via surprising, but understandable mechanisms – bring about all the surprising macro-effects. To get an understanding, we go for simplicity—and that with the hope, to get 'a feeling' for what might go on in the real world.[4]

In the next section we will explain the details of our simulation strategy. Section 3 will give a complete overview for the case in which the radicals hold the most extreme position, i.e. $R = 1.0$. In section 4 we explore and explain the most surprising effects, that we found in section 3. Section 5 shows what happens, if R becomes less and less radical, and moves direction center of the opinion space. In section 6 we draw conclusions, discuss and interpret our results under different perspectives, and, finally, outline a research agenda.

---

[4]Our approach is very much in the spirit of what is called *analytical sociology*, or the study of *social mechanisms*.

2. **The simulation strategy.** It is very natural to think, that final numbers of radicalized normals crucially depend upon the number of radicals compared to the number of normals (hopefully the ratio), the confidence level ϵ, and the radicals' position R. Under this working hypothesis, our model has only few parameters and we should be able to answer our questions for major parts of the parameter space. One possible simulation strategy, then, is the following—and it is the one we will use: Let's assume we have 50 normal agents and the most extreme radical position that is possible, i.e. $R = 1.0$. Two parameters are left: The number of radicals, and the confidence level ϵ. Now we put a grid on the two dimensional parameter space: In 50 steps of size 0.01 the confidence level of normals increases from 0.01 to 0.5 on the x-axis. On the y-axis the number of radicals increases in 50 steps from 1 to 50 (then the group of radicals has as many members as the group of normals). We run simulations for each of the 50×50 parameter constellations $\langle \epsilon, \#_R \rangle$. Once a run has stabilized, we do the relevant statistics. A run is considered stabilized, iff the opinion profiles $X(t)$ and $X(t+1)$ are almost the same. More precisely, we stop a run if for *all* agents i it holds that $|x_i(t+1) - x_i(t)| \leq 10^{-5}$. As to statistics, we will focus on *one* number only: the number of normals that finally hold an almost radical position. And we consider a normal agent i's opinion as almost radical iff $|x_i(t) - R| \leq 10^{-3}$.

2.1. **Getting rid of randomness.** The opinion and influence dynamics given by the equation (2) is completely *deterministic*. But it seems to be clear, that, unavoidably, at a certain point randomness comes into our simulation design: To find for a parameter constellation $\langle \epsilon, \#_R \rangle$ the number of finally radicalized normals seems to require averaging over a sufficiently high number of repeated runs—and all the runs will have to start with a *random* start profile, e.g. based upon a uniform start distribution. – Natural as it is, we will *not* do that. For each of our 50×50 parameter constellations $\langle \epsilon, \#_R \rangle$ there will be just *one* run, and that run will *not* start with a random profile. All 50×50 parameter constellations will start with the *same* very special, but in a certain sense 'typical' start distribution of n opinions of the n normals (in our case always 50 normals): Let's call an opinion profile an *ordered* profile iff for all $i \leq (n-1)$ it holds that $x_i(t) \leq x_{i+1}(t)$. In *all* runs we use the specific ordered start profile $X(0)$ for which

$$x_i(0) = \frac{i}{n+1}, \forall i = 1, ..., n \tag{3}$$

holds. In such an ordered and equidistant start profile the ith opinion is exactly there, where it will be at the average over infinitely repeated uniform *random* distributions of n opinions. Or in other words: Equation 3 gives the *expected value* of the ith opinion of a uniform random distribution of n opinions. Consequently, it also realizes the expected distances between neighboring opinions of such a random distribution. We refer to that type of regular and equidistant start distribution as the *expected value distribution*.

In our case with always 50 normals, the expected value start distribution implies that for $t = 0$ the ith opinion is $i/51$. Figure 2, top, shows a dynamics starting with a random uniform distribution of 50 opinions. Figure 2, bottom, is the corresponding dynamics based upon the expected value distribution for 50 opinions. (The confidence level is in both cases $\epsilon = 0.23$.)

By using for all parameter constellations $\langle \epsilon, \#_R \rangle$ the same equidistant start profile $X(0)$, *we get completely rid of randomness*: Nowhere in our simulation design is any random element. Everything is *deterministic*.

On the one side, the expected value distribution is representative. On the other side, the equidistant structure of the profile may make us blind for important effects, that are caused by the variation of distances between neighboring opinions as they are typical for random distributions. However, there is a major advantage of using the expected value start distribution all the time: Imagine, we find in the grid of 50×50 parameter constellations $\langle \epsilon, \#_R \rangle$ some interesting effects, e.g. counter intuitive (non-)monotonicities. Then, the exclusive use of the *one* expected value distribution has a nice consequence: We can go straight into the *unique single runs* that generated the perplexing macro-effect. And the inspection of these runs may help to understand what is going on—at least that is the hope. Soon it will turn out, that this hope is not an illusion.

The next section presents for the grid of 50×50 parameter constellations $\langle \epsilon, \#_R \rangle$ the number of normals that ends up at the radicals' position R. It is assumed that R is the most extreme position, i.e. $R = 1$. But R may be less extreme. To get the complete overview, we will therefore – in another section – use the same 50×50 'grid approach' to analyze the effects if the radical position (or more general: the *signal*) R moves towards direction 0.5, i.e. direction center of the opinion space $[0, 1]$.

There are other models of radicalization. Normally they are more complicated than ours (cf. e.g. [5], [2]). That makes a rigorous and in depth analysis difficult. We go for extreme simplicity to get a complete understanding, including the micro level that brings about the figures on the macro level, like, e.g. the numbers of finally radicalized agents.

2.2. **An enemy within: The floating point arithmetic.** Before we can do, what we plan to do, we have to realize and to remove a major *technical obstacle*. Our computational tasks – time consuming, but in principle easy to solve for human beings – may well be *too difficult* for a computer: To compute the *BC*-dynamics requires – and that again and again – to decide questions of the type: Is $|x_i(t) - x_j(t)| \leq \epsilon$? The usual floating point arithmetic[5] approximates real numbers. Only a subset of real numbers can be exactly represented. The data format called "real" is *not* really a continuum—it has holes all over. As a consequence, if agent j's opinion is *exactly* the upper or lower bound of agent's i's confidence interval, that may cause a numerical error. Such an error is not only 'theory', it is reality: For $x_i(t) = 0.6$, $x_j(t) = 0.4$, and $\epsilon = 0.2$, the critical question is, whether or not $|0.6 - 0.4| \leq 0.2$. A question that a 10 years old child can answer—a computer program, that uses the data formats called *real*, *float*, or *double* and then makes use of the built-in and hard-wired floating point arithmetic *may get it wrong*! It may well be the case, that the answer to the question whether $|0.6 - 0.4| \leq 0.2$ *differs* from the answer to the question whether $|0.4 - 0.6| \leq 0.2$. The numerical results of algorithms that in the world of *real* real numbers are *logically equivalent*, may be *different*.

Figure 3, top, shows an obvious numerical disaster, produced by a *NetLogo* program of the *BC*-dynamics. The dynamics starts with the profile

$$X(0) = \langle 0, 0.2, 0.4, 0.6, 0.8, 1 \rangle.$$

[5]The details are defined by the IEEE 754 standard.

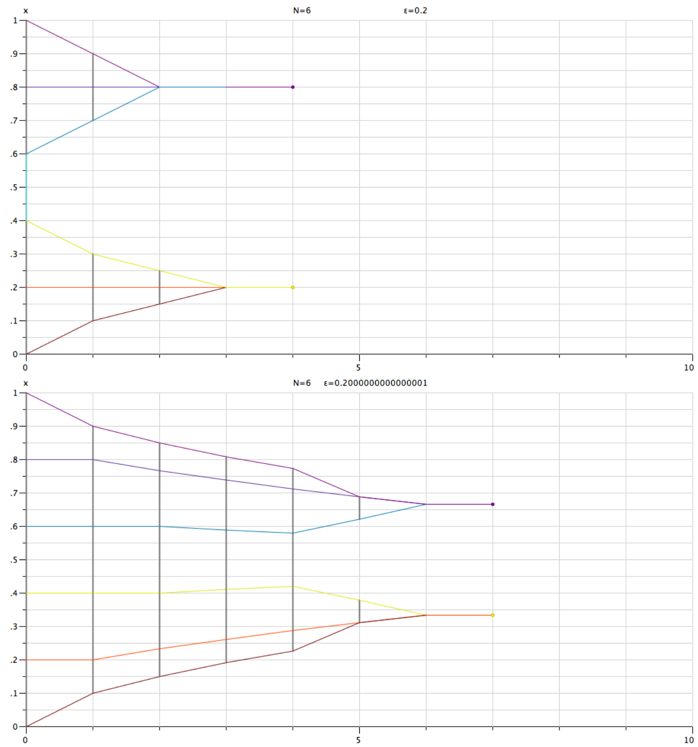

FIGURE 3. Dynamics for 6 agents and a confidence level $\epsilon = 0.2$. *Top*: Completely wrong trajectories. *Bottom*: Correct dynamics. For $t = 0$ the ith opinion is $i/(n-1)$.

We assume $\epsilon = 0.2$. Thus, the profile is an equidistant start profile in which ϵ actually *is* the equidistance. Given that start profile, in the first updating step, no opinion except for 0 and 1 should change. All *other* opinions $x_i(0)$ have exactly one opinion above in their confidence interval, namely $x_{i+1}(0) = x_i(0) + 0.2$. Additionally, below, the opinion $x_{i-1}(0) = x_i(0) - 0.2$ is in their confidence interval. Trivially, we should have $x_i(1) = x_i(0)$ in all these cases. But the program gets it wrong for $x_i(0) = 0.4$ and $x_i(0) = 0.6$. Due to obvious arithmetical errors the dynamics is corrupted by the very first updating step. The disaster is *not* a specific *NetLogo*-disaster. Any programing language, that uses floating point arithmetic, would get into trouble, though the errors may have another form.[6]

The disaster is caused by the fact that opinions are distributed in such a way that most of them are *exactly* at the lower or upper bound of the confidence interval of neighboring opinions above or below. In that situation a minor misrepresentation at an order of magnitude of $\leq 10^{-12}$ may have consequences at an order of magnitude of 10^{-1}. In figure 3, top, it has: For the opinions 0.4 and 0.6 the program gets it completely wrong and both change: 0.4 downwards, 0.6 upwards. Then in period 1 *another* error occurs in the lower segment of the profile where the lower cluster

[6]The *Programming Guide* of *NetLogo* is very explicit about these problems. It even shows the example that for $0.1 + 0.1 + 0.1$ one gets the result 0.30000000000000004. The problem is, that most readers and programmers do not understand the possibly dramatic consequences of such errors.

of finally stabilized opinions starts to evolve. *Given* that the *first* error has already happened, we now should have $|x_3(1) - x_1(1)| = |x_1(1) - x_3(1)| = 0.2 = \epsilon$. As a consequence, in the lower segment of the profile *all* opinions should be within the confidence interval of *all* others. Therefore, in the next period the segment should collapse into *one* opinion cluster, as it does in the corresponding upper segment of the profile. But that does not happen in the lower segment. Instead, there is a *second* numerical error of the same type as before: The two opinions $x_1(1)$ and $x_3(1)$ at the lower or upper bound of each other's confidence interval are not correctly classified as within their respective confidence intervals. As a consequence, the three opinions do *not* collapse into one single cluster in the next period.

These observations and considerations demonstrate an important point: Given, we have an expected value start distribution according to equation 3, then in a correctly computed dynamics there will always be a *mirror symmetry* with respect to the horizontal line $y = 0.5$. However, the second numerical error in our example above makes very clear, that mirror symmetry is only a *necessary*, but not a sufficient condition for the numerical correctness of the computation: Without error two, we would have mirror symmetry—but with a built-in error one.

Figure 3, bottom, shows the correct trajectories for our example. They are again generated by a *NetLogo* program. But this time we use a computational trick: All computations are based upon a confidence level ϵ that is infinitesimally bigger than 0.2. The increase is sufficiently big, that the floating point arithmetic detects correctly *all* elements that have to be elements of the sets $I\big(i, X(t)\big)$. At the same time it is small enough to guarantee, that *only* these opinions are in the sets. In short: By a slightly wrong ϵ, we get the right result.

Numerical errors caused by opinions right on the bounds of confidence, are frequent. To see that, one can use what we call *ϵ-diagrams*. Their very essence is to visualize for *one and the same* start distribution $X(0)$ the effects of an stepwise increasing ϵ on the *final, completely stabilized* cluster structure. Figure 4 is an ϵ-diagram: On the x-axis there are increasing values of ϵ. They increase in steps of 1/100 from 0 to 0.4. For each of the 41 ϵ-values the run starts with the same expected value start distribution. We assume 49 agents. Thus, the ith opinion in the start profile is always $i/50$. As an *intended* consequence we get, that, given the specific step size of the increasing ϵ values, again and again, and often in numerous cases, the start opinions of $X(0)$ are right on the bounds of confidence of other opinions. For each of the 41 ϵ-values we run the dynamics until it is stabilized in the sense $X(t+1) = X(t)$. The y-axes of an ϵ-diagram is used to indicate the *final stabilized* positions of the opinions in the run for the specific ϵ value. Since opinions cluster, there are always much less than 49 such positions. What looks like trajectories, are horizontal lines, that, step by step, connect the final positions of the ith opinion for the stepwise increasing confidence level ϵ. Colors indicate the ranks in the profile.

Thinking it through, it is easy to see: *If* $X(0)$ is an expected value start distribution, then an ϵ-diagram for that start distribution has to be completely mirror symmetric with respect to $y = 0.5$. Additionally, for an *odd* number n of opinions, there is always an opinion $x_i(0) = 0.5$, the *center* opinion (with the rank $0.5 \cdot (n+1)$). If a run is computed correctly, then the center opinion never changes: It starts at 0.5, and stays there forever—whatever the value of ϵ may be. In the ϵ-diagram in figure 4 the center opinion is painted black. With the symmetry considerations in mind, figure 4 reveals a major numerical disaster: The computations *are* correct for

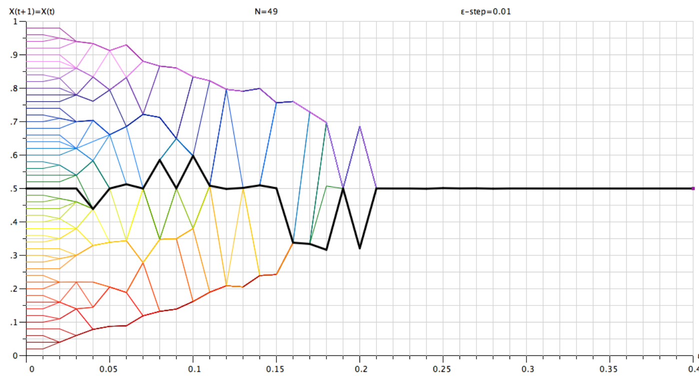

FIGURE 4. ϵ-*diagram* for an *expected value* start distribution with 49 opinions. *Black*: The final positions of the 25th opinion in the middle of the ordered profile. Any deviation of the black line from 0.5 and any deviation from a mirror symmetry around 0.5 of the whole diagram, is a *sufficient* condition, that numerically something somewhere went wrong in the computation of the underlying unique single runs based upon a certain ϵ.

very small ϵ values, when – given the distance 1/50 between neighboring opinions in $X(0)$ – ϵ is so small, that no one is in the confidence interval of anyone else, and, therefore, nothing happens. The computations *may* be correct for $\epsilon \geq 0.21$. But for the huge ϵ range in-between, the computer got it badly wrong—*and that for sure*. Bleak as it is, the floating point arithmetic is for the *BC*-model a kind of enemy within.[7]

How, then, can we dare to do, what we plan to do? We plan, first, to use one and the same equidistant expected value start distribution for all 50×50 $\langle \epsilon, \#_R \rangle$ parameter constellations. And, second, as to ϵ, we plan an increase by steps of 1/100. Isn't that a recipe for disaster like the one above? Not necessarily, but we have to be very, very cautious! Different from the disaster above, we will use an expected value start distribution based on $n = 50$ (and *not* $n = 49$ as before). For $n = 50$ we get an equidistance of $1/51 = 0.\overline{0196078431372549}$. Given that distance, none of the 50 opinions in $X(0)$ is for any $\epsilon = 0.01, 0.02, ..., 0.4$ at the bounds of confidence of any other opinion. That has consequences—and they are reassuring: Figure 5 shows the ϵ-diagram for an expected value start distribution with $n = 50$. As it seems, a *perfect symmetry*. Since n is an even number, there are now *two* center opinions: x_{25} and x_{26}, both painted black, and, again, both in perfect mirror symmetry with respect to the line $y = 0.5$. Of course, as shown above, symmetry is only a necessary condition. We can't be perfectly sure that nowhere – maliciously camouflaged by symmetry – the enemy within has done some numerical damage. But, additionally, it seems (by visual inspection) that there isn't *any* ϵ in figure 5 with a broken symmetry. We consider that as an indicator that the numerical risks

[7]Arguably there is *no* other model that is more vulnerable to that enemy: For n agents and under the assumption, that it is *not* critical to find out that each agent i's opinion is within i's confidence interval, the model still requires in *each* period $n(n-1)$ decisions of the type $|x_i(t) - x_j(t)| \overset{?}{\leq} \epsilon$. A major fraction of them has the potential for disaster. – As to other models cf. [17].

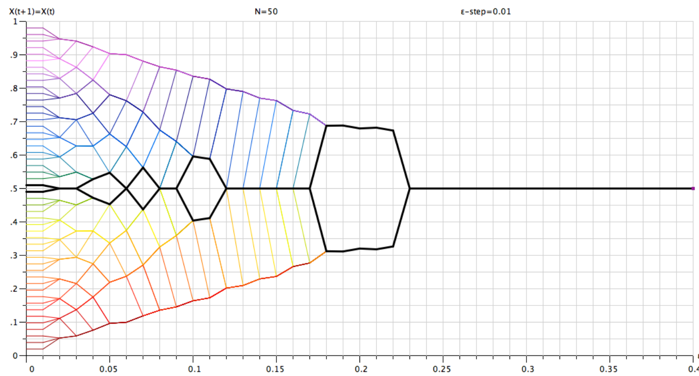

FIGURE 5. *ϵ-diagram* for an *expected value* start distribution with 50 opinions. Black: The final positions of the 25th and the 26th opinions. Mirror symmetry around 0.5 as a necessary, though not sufficient condition for numerical correctness, seems to be fullfilled.

of our plan are sufficiently small.[8] In the following we will take the remaining risk. That it is there, should be kept in mind all the time.

3. **Under the influence of most extreme radicals: An overview with some surprises.** Result of the computation of single runs for the 50×50 parameter constellations $\langle \epsilon, \#_R \rangle$ is an array of numbers. It is much easier to detect patterns and structures in *colored landscapes* rather than in an array of numbers (even if they are only integers—as in our case). Therefore (as a kind of phase diagram) *figure* 6 shows, indicated by color, the number of normals that finally end up at the radical position $R = 1.0$. For ease of reference, we refer by the capital letters A, B, C, D, E, and F to certain regions of the parameter space, as they are partitioned by the black lines (two vertical, one horizontal).

On the y-axes the number of radicals increases stepwise. Therefore, sudden dramatic color changes in vertical direction are dramatic changes in the number of radicalized normals that – *ceteris paribus* – are caused by just *one* more radical. Correspondingly, a dramatic color change in horizontal direction is – *ceteris paribus* – a dramatic change in the number of radicalized normals caused by a tiny increase of ϵ by 1/100.

In the following we inspect region-wise our parameter space (discretised by 50×50 grid of parameter constellations $\langle \epsilon, \#_R \rangle$). We start with the three vertical regions. There our question is, how the number of normals that end up at the radical position, depends upon the number of radicals. Then an inspection of the two horizontal region follows. Our question there, is, how the radicalization of normals depends upon confidence levels.

1. In the region $F \cup C$, i.e. a region with *higher* confidence levels ϵ, the number of radicalized normals *monotonically* increases as $\#_R$ increases. But for all $\epsilon <$

[8] There are other strategies to fight the numerical problems caused by floating point arithmetic. We could use the trick in figure 3. Random start distributions would minimize the problem as well. But random start distributions would, then, require *repeated* runs for each parameter constellations. For an understanding of all the strange effects, that we will find in the next chapter, we could *not* directly go into the unique single runs that produced them. As a consequence, understanding would become much more difficult.

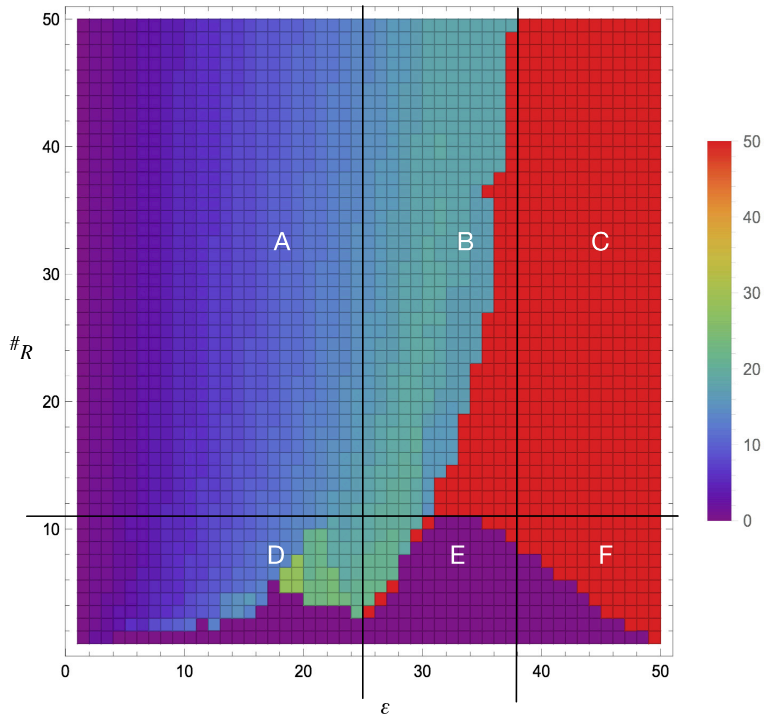

FIGURE 6. x-axis: the confidence level increases in 50 steps of size 0.01 from 0.01 to 0.5. y-axis: the number of radicals increases from 1 to 50. Colors indicate the number of normals that end up at the radical position which is here assumed to be $R = 1.0$. The total number of normals is always 50.

0.5, there is a sudden jump: One more radical, and the number of radicalized normals jumps from none to all. Obviously there is an ϵ depending threshold $\#_R^*$ of radicals, such that, first, for that threshold *no* normal ends up at the radical position, while, second, for $\#_R^* + 1$ *all* normals end up at the radical position.

2. In the region $E \cup B$, i.e. a region with *middle-sized* confidence levels we find jumps of all sorts and directions: In region E there are – again in vertical direction – jumps from none to all: One more radical, and, instead of none, all normals end up at the radical position R. But, additionally, in region E and B there are jumps in the opposite direction: One *more* radical, and instead of all, *significantly less* (about half of the normals, or even less) end up radical.

   By careful inspection of region $E \cup B$ in figure 6 one can verify: With the exception of one of the ϵ values (the exception will be discussed later), it holds for the *middle-sized* confidence levels in region $E \cup B$:

   (a) For all of them exists an threshold $\#_R^*$ for none-to-all jumps.
   (b) For all of them exists another ϵ depending threshold $\#_R^{**}$ of radicals, such that, first, for that threshold *all* normals end up at the radical position, while, second, for $\#_R^{**} + 1$ *significantly less* normals become radical.

Obviously, there is a second type jumps, now working into the opposite direction.

(c) In region E the second threshold $\#_R^{**}$ equals $\#_R^{*}+1$. As a consequence, two steps of adding just one more radical causes the dramatic change from none to all, and then back to about half of the normals being radicalized.

(d) For the thresholds $\#_R^{**}$ that are in region B, the jump from all to significantly less comes later—but it comes: There is always a number of radicals such that just one more reduces the number of radicalized normals from all to about less than the half.

There is one exception from observation (a) to (b): For step 28 on the y-axis ($\epsilon = 0.28$) there is a threshold $\#_R^{**}$, but no threshold $\#_R^{*}$. To be frank: We do not know the reason. May be it is simply as it is—in the non-linear BC-dynamics often minor differences matter. The missing threshold may be a hint, that the exact position pattern of the thresholds $\#_R^{*}$ is a more complicated issue than it looks under our 50×50 grid of $\langle \epsilon, \#_R \rangle$ parameter constellations. A grid that is finer with regard to ϵ could give an answer. And finally, bleak as it is: The missing threshold $\#_R^{*}$ may be the consequence of numerical problems that – despite of our precautionary measures – are still there.

Leaving the one exception aside and summing up: $E \cup B$, a region of *middle-sized* confidence levels, is a region with sudden ups (from none to all) and downs (from all to significantly less) radicalized normals. Along certain lines in the parameter space the sensitivity to tiny changes is extreme. The predominant phenomenon is, that the number of radicalized normals is *not* monotonically increasing with an increasing number of radicals. Just one more radical may lead to much less radicalization. In the more abstract language of signals: Intensifying a signal that convinces too little, may convince even less.

Even in the smooth *non*-red areas of region B the radicalization of normals is clearly *not* monotonically increasing with regard to $\#_R$. On the contrary: In the left part of that area the radicalization of normals is slightly *decreasing* as the number of radicals *increases*.

3. The region $D \cup A$ is the region of *smaller* confidence levels ϵ. Again, for an increasing $\#_R$, there are certain threshold values where jumps occur. But they are not jumps from none to all. Nevertheless, they are jumps from none to a significant proportion. In the right part of D the sudden increase is more drastic than in the left part. Again there is a striking effect: Above the jumps from none to a significant proportion, the number of radicalized normals clearly *decreases* as the number of radicals *increases.* In the language of another intended interpretation: In this part of the parameter space more charismaticity attracts less normals to the position of the leader. The more intensive campaign for R, convinces less and less of R. Or the other way round: Less would have more effect.
4. *Horizontally*, i.e. with regard to $\#_R$, we distinguish *two* regions. There is an upper region with a *major* number (or proportion) of radicals, the region $A \cup B \cup C$. It is a region with always more than 10 radicals, i.e. a radical group size of more than 1/5 of the number of normals, or, respectively, more than

1/6 of the whole population.[9] Given such a major $\#_R$, if ϵ increases, there always exists a threshold ϵ^* such that for $\epsilon^* + 0.01$ the number of radicalized normals jumps from about 1/3 to all. The upward jumps are compatible with monotonicity. However, careful color inspection of the area to the left of the thresholds ϵ^* clearly shows (especially clear for the middle sized confidence levels in region B) that, with increasing confidence levels, $\#_R$ – slightly and smoothly – first increases, but then decreases (again slightly and smoothly). In sum, for major numbers of radicals, as to the figures of radicalized normals, there is no general monotonicity with respect to the size of their confidence interval.

There are *two* regions in the whole parameter space, that behave very smooth: the regions A and C. Both belong to the upper horizontal region with a *major* number of radicals that we inspect right now. In region C, i.e. for *higher* confidence levels, for any $\#R$ all normals end up radicalized. In region A, i.e. for *smaller* confidence levels, it is the normals' confidence level that matters—not $\#_R$: The radicalization of normals increases as their ϵ increases. In the bottom right area of A the number of the radicals has a bit effect: The number of normals, that end up at R, slightly decreases as $\#_R$ increases.

However, a warning side remark: Figure 6 shows the number of normals that become radical. We consider a normal agent i as "radical","radicalized", "ending up at the radical position" etc. iff $|x_i(t) - R| \leq 10^{-3}$. Therefore, even if $\#_R$ has (almost) no effect on the number of – in this sense – radicalized normals, it may nevertheless have (and often has) a major effect on the mean or median opinion of the normals' opinions, the cluster structure etc., which we do not analyze here.

In region A and C the number of radicals has very little or no effect on the radicalization of normals. In region B that is different, and $\#_R$ seriously matters: The exact location (though not the existence) of the threshold ϵ^* depends upon the number of radicals: As they become more, the jumps occur more to the right, i.e. they require higher confidence levels.

5. The lower horizontal region, i.e. $D \cup E \cup F$, is a region with *minor* numbers of radicals (not more than 1/6 of the whole population of normals *plus* radicals). In terms of jumps it is the wildest region: In E and F we find (as in B) thresholds ϵ^* such that for $\epsilon^* + 0.01$ the number of radicalized normals jumps from less than a half or even none to all. But, additionally, there are values ϵ^{**}, such that for $\epsilon^{**} + 0.01$ the number of radicalized normals jumps from all to zero. The most striking point is, that in E both threshold values are horizontally next to each other, i.e. $\epsilon^{**} = \epsilon^* + 0.01$. Obviously, the radicalization of normals reacts in E, i.e. an area with both, a minor number of radicals *and* a middle-sized confidence level, extremely sensitivity with regard to both initial conditions, the confidence level *and* the number of radicals.

   However, there is again a conspicuity in area E, now in vertical direction: For all $\#_R > 2$ *except* for $\#_R = 6, 7, 8$ there exists a threshold ϵ^* (as defined above). Again we do not know the reason. The same considerations, as mentioned above in the corresponding case for $\epsilon = 0.28$, apply (see the second observation).

[9] If one counts the cells up to the horizontal line, the result is 10. Note: The y-axis' origin is 1 (and the x-axis origin is 0.01).

In area D we find for an increasing ϵ a lot of jumps in both directions: from some to none and from none to some. The jumps are less dramatic than in region E, but they are there. For a description one might introduce thresholds that correspond ϵ^* and ϵ^* but have reduced requirements.

To sum up, with respect to the confidence level, the whole region $D \cup E \cup F$ is a region of non-monotonicity, jumps up and jumps down.

What are the main results of our inspection? There are two 'smooth' areas in the parameter space, the areas A and C. But in *all* other areas we find the thresholds ϵ^*, ϵ^{**}, $\#_R^*$, and $\#_R^{**}$ (in D we observe corresponding thresholds with reduced requirements). We can draw a line through the parameter constellations that are ϵ- or $\#_R$- thresholds of the defined sort. To account for details that we may have missed by the discretization of ϵ, we draw the line with a pen, that is *not very much sharpened*. As to the two cases of conspicuity, we do some interpolation. And we draw the line 'by hand', meaning using a graphics program, but without any numerical calculation of the line. The result is given in figure 7. The thick white lines indicate, where in the parameter space – and that in a completely *deterministic* dynamical system – the number of radicalized normals is *highly sensitive to the initial conditions*, either ϵ, or $\#_R$, or both. (In other words: The line displays, where in the parameter space the so called *butterfly effects* occur.) If extremely high sensitivity to initial conditions is the hallmark of *chaos*, then it is clear: Along the thick white lines we face chaos. We will call these lines *sensitivity-lines*.

Let's call a region of our partitioned $\langle \epsilon, \#_R \rangle$-parameter space *wild* iff, first, we have in that region a non-monotonicity (both, decreasing *and* increasing) with respect to one or both parameters, and, second, the region is pervaded by a sensitivity-line. Under that definition we can distinguish *two* wild regions: In *vertical* direction the region $E \cup B$, a region of middle-sized confidence levels; in *horizontal* direction the region $D \cup E \cup F$, a region of comparatively small numbers of radicals.

For all regions, wild or not, immediately "Why is it, that ...?"- questions arise. Why is it, that in region C neither the number of radicals, nor the confidence level has any effect on the number of radicalized normals? Whatever the specific parameter constellation in that region, *all* normals end up radical—but why? Why is it, that in region A, a region where the number of radicals is above $1/6$ of the whole population, radicalization of normals is *not very much* influenced by the number of radicals. Obviously, it is the confidence level of normals that matters—but why?

In the following we will *not* answer question about the smooth regions. We will focus on the wild ones. We will make expeditions into the vertical and the horizontal wild region and try to understand, how 'in the deep' some completely deterministic mechanisms create the wild '*radicalization landscape'*, that figure 6 displays.

4. **Expeditions into the wild parameter regions.** It now comes to the point, where our approach starts to pay off: First, we computed *one and only one* run for each of our 50×50 grid of parameter constellations $\langle \epsilon, \#_R \rangle$. Second, each of the 50×50 runs started with the *same* expected value start distribution of 50 normals, distributed in the opinion space according to equation (3). That approach (strange as it may look at first glance) has *two* consequences: First, whenever we want to understand what causes certain effects in our radicalization landscape, we can go into *unique single runs.* No statistical analysis of 100 or so randomly started runs is necessary. Second, since, additionally, all the single runs start with the same start distribution of normals, we can, by comparison of single runs, *on the level of single*

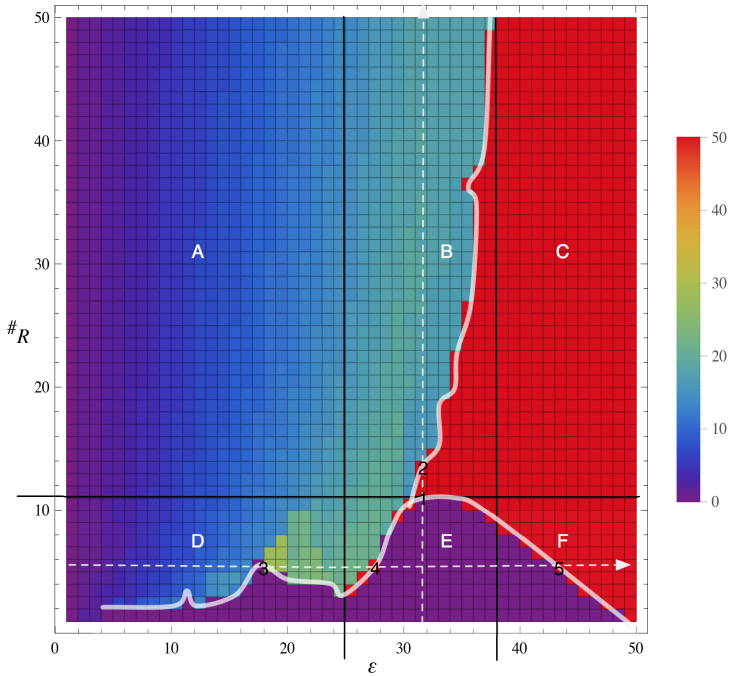

FIGURE 7. *Thick and curvy white line:* The sensitivity-line. *Dashed white lines:* Route of the northern and eastern expedition into the wild parameter space regions. *Black numbers in the landscape:* Positions of the *explananda*, that we discuss in chapter 4.

agents directly observe the effects of changes of ϵ or $\#_R$. Especially, if we inspect single-run-sequences of *small stepwise changes*, we directly observe the working of the 'forces in the deep' that generate the surface of our radicalization landscape—and that should be a good starting point for an identification and understanding of the mechanisms that bring about the puzzling landscape.

In the following we will go on *two expeditions into the wild regions* of our parameter space: One direction north in $E \cup B$, one direction east in $D \cup E \cup F$. The dashed white lines in figure 7 are our routes. In both directions we will try to understand, at least, to get a feeling for the mechanics underneath.

4.1. **Expedition 1: Going north in $E \cup B$.** Our first expedition starts at $\epsilon = 0.31$. It goes direction north. The vertical white dashed line is the route, on which we will cross two sensitivity lines. For each of the points that we pass on the 50×50 grid of $\langle \epsilon, \#_R \rangle$ constellations, we have unique single runs with the same start distribution. 50 periods turn out to be sufficient to know the final stabilized pattern and the number radicalized normals therein. Therefore, we can generate a sequence of 50 pictures, one for each $\langle \epsilon, \#_R \rangle$ constellation that we pass. Each of the pictures displays the trajectories of all 50 agents. The time scale on the x-axis is always the same: 50 periods. Going strictly north implies, that ϵ is kept constant. Therefore, whatever is changing in the sequence of pictures, it is the consequence of *one* factor only: the number of radicals.

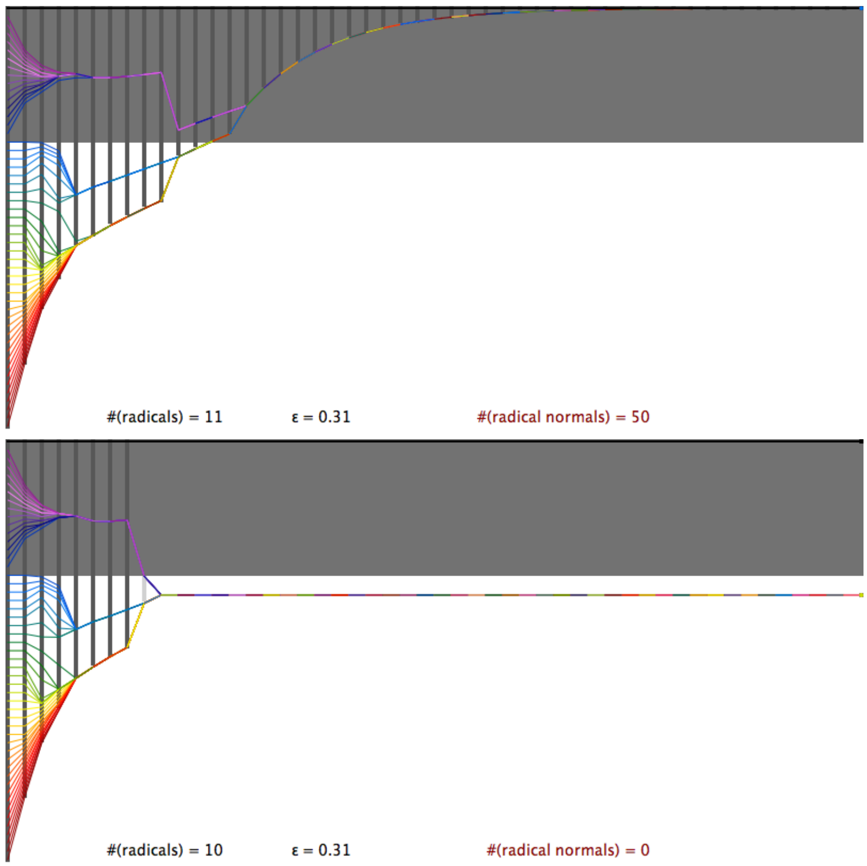

FIGURE 8. Explanandum 1. *Bottom* $\#_R = 10$. *Top*: $\#_R = 11$. $\epsilon = 0.31$. One more radical causes a jump from none to all of the normals ending up at the radical position.

It is easy to visualize the effects of an increasing $\#_R$ by an animation. In this article we do not have the space to show the pictures for all $\#_R$ on the route north. But we can show, what is going on underneath, when we pass the sensitivity-lines—and especially these passages ask for explanation.

Figure 8 displays the *first* explanandum: The jump from none to all normals being radicalized, if the number of radicals increases from 10 to 11. For an explanation we start in figure 8 bottom, i.e. the dynamics under the influence of 10 radicals. As the dark grey vertical lines indicate, there is up to period 7 (we start with period 0) a chain of direct or indirect influence of the 10 radicals even on the most distant normals. Soon 3 opinion clusters emerge among the normals. The cluster in the middle functions as a bridge between the upper and the lower cluster. The upper cluster is in the dark grey area of the opinion space, and that is the area of direct influence of the radicals. Thus the upper cluster of normals is a bridge between the radical group and other two clusters of normals, which, period by period, move direction R. But that works only for a while: As the lower cluster of normals moves upward, at a certain point the upper normal cluster has both, the middle and the lower cluster within its confidence interval. Their combined influence on the upper cluster is strong enough to pull the upper cluster completely out of the area of direct influence of the radicals. The consequence is, that no bridge between radicals and normals exists any longer. However, the radicals had an effect on the normals: Without radicals the normals would end up at a 0.5-consensus. With 10 radicals (1/5 of the number of normals, 1/6 of the whole population) it is about 0.64.

Now we add just one radical, analyze the trajectories in figure 8 top, and compare it with what we see in figure 8 bottom: With the one more radical, again, the three clusters of normals evolve. However, it takes a few periods more until the upper cluster gets under the direct influence of both clusters below, which, therefore, both are moving direction R a bit longer, before, then, that period comes. When it comes, the upper cluster makes – as in the case with one radical less – a steep move away from the radical's position. But different from the case with $\#_R = 10$, the upper cluster does *not* get out of the area of direct influence of radicals. From that moment onwards everything is lost: Though now further away from the radical position, the upper cluster *continues* to function as a bridge between all other normals and the radicals. After two more periods the bridge isn't necessary any longer: Now *all* normals are under the *direct* influence of the radicals—and that is a point of no return: From now on (and as one cluster) all normals irreversibly move and converge – though in infinite time – to the radical position.

The inspection of the underlying dynamics resolves the puzzling effect in the radicalization landscape: The sudden jump from none to all at the sensitivity-line is due to what we might call a *positive bridging effect*: For $\#_R = 10$ a cluster evolves that for a while functions as a bridge between the radicals and all other normals. But then the bridge breaks down and, additionally, the upper cluster gets out of the area of direct radical influence. For $\#_R = 11$ the bridge to the radicals *continues* to function until it becomes superfluous. The functioning of the bridge is critical and accounts for the difference between none or all of the normals ending up radical.

Figure 9 displays the *second* explanandum: The sudden *drop down* from all normals being radicalized to only 1/3, and that by *increasing* the number of radicals from 13 to 14. We start our analysis in figure 9, bottom. What we see there, is very similar to figure 8, top. Based on what we saw there, we understand the positive bridging effects that are (still) at work in the dynamics in figure 9, bottom. But the one more radical in figure 9, top, causes dramatically different trajectories: the middle cluster of normals (blue in color) is somehow 'blown up': One of the 'former members' joins the upper cluster, all others join the lower cluster. For some periods the radicals still influence even the most distant normals. But there is no evolution of a bridging cluster *in-between* the two clusters of normals. The distance between the two clusters enlarges. As a consequence, the radicals' chain of influence breaks and becomes very short afterwards. More than 2/3 of the normals form a cluster at about 0.45, i.e. even slightly *lower* than the center of the opinion space.

Decisive for the sudden jump downwards is, that the one more radical causes a rupture in a segment of the normals' opinion profile that, without the additional radical, would have become a bridging cluster. – Obviously, there are not only positive bridging effects. In our explanation of the second explanandum a *negative* bridging effect is at work: Under one more radical a former bridge to the radicals ceases to exist. And that causes a dramatic reduction in terms of radicalized normals.

We can't analyze here the details of the mechanics, that works underneath the radicalization landscape on our route further north, i.e. once we have passed a

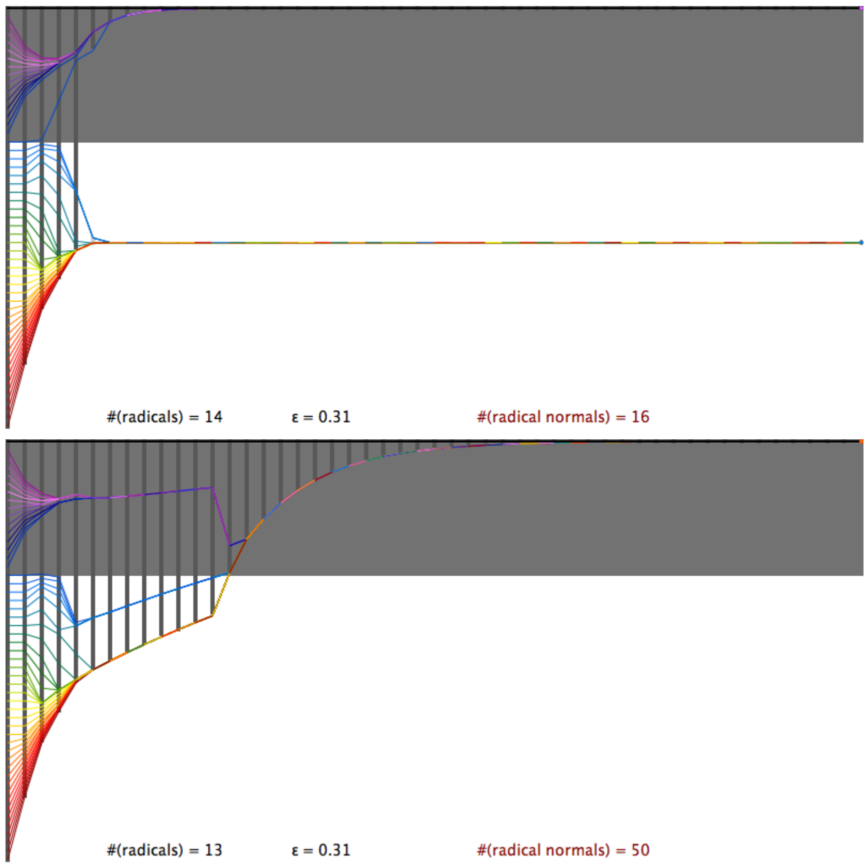

FIGURE 9. Explanandum 2. *Bottom*: $\#_R = 13$. *Top*: $\#_R = 14$. $\epsilon = 0.31$. One more radical causes a jump down from all to less than one third of the normals ending up at the radical position.

second time a sensitivity-line. Only so much: The number of radicalized normals goes up to a maximum of 19, fluctuates for a while between 18 and 19, and ends for $\#_R = 50$ with 18 radicalized normals. *All* explanations of these figures, their small range, and their fluctuations, they are all about the details of how the blue cluster, that for $\#_R = 13$ functions as a bridge between the upper and the lower cluster of normals (see figure 9, bottom), is 'blown up', disassembled, and ruptured into pieces, if we add one more radical, and another one, and so forth.

4.2. **Expedition 2: Going east in** $D \cup E \cup F$**.** Our eastward expedition into the wild region $D \cup E \cup F$ starts at $\#_R = 5$. The horizontal dashed line in figure 7 is the route. Going strictly east, implies, that the number of radicals is kept constant. Whatever is changing now, it is again the consequence of *one* factor only: the normals' confidence level ϵ. As a consequence, further east in the region, the *direct* influence of radicals reaches *further down* in the opinion space. The region is pervaded by sensitivity-lines. Three times we pass them. Again we will focus on the mechanics, that is working underneath there, where we pass the lines. In the subsection above we had two explananda; now we get three *three* more.

Explanandum 3 are the sudden jumps for $\epsilon = 0.16, 0.17, 0.18$: As ϵ *increases* by a tiny 0.01, the number of normals that end up radical, *drops down* from 13 to zero. If ϵ *increases* by another 0.01, then that number *jumps up* to 28, i.e. more than 1/2 of all normals. Figure 10 shows the unique single runs that cause the butterfly effects.

For an explanation we start with the trajectories in figure 10, bottom, i.e. the case $\epsilon = 0.16$. Three clusters of normals evolve from the start distribution. Only the upper cluster converges to the radical position. The chain of radical influence breaks twice. First the 5 radicals loose their influence on the lower cluster. Then, one period later, they loose their influence on the cluster in the middle. With an $\epsilon = 0.17$ the radicals, again, soon loose their influence on the lower cluster, but they continue to influence that segment of the opinion profile, from which before the cluster in the middle evolved. In-between the upper and the former middle cluster two more cluster evolve. The upper cluster functions as the decisive bridge between the radicals above and three clusters of normals below. The 5 radicals are about as many as the members of the cluster of normals, that evolves right below the upper cluster. For a while that is – given the confidence level – the only cluster of normals with an influence on the upper cluster. For a while the upward and downward forces, that pull on the upper cluster and the cluster right below, are equally strong. As a consequence, both clusters move *almost horizontally*. But nevertheless, both are decisive bridges in the chain of radicals' influence on normals. At the same the two lower clusters, that still are under the influence of radicals, move upwards, join, and become a cluster of a major size. One period later the upper cluster, that still is a decisive bridge in the chain of radical influence, gets into the confidence interval of the joint lower cluster, which, therefore, makes a steep upward move. But confidence is a symmetric matter. As a consequence, the upper cluster makes a steep downward move. By that move the upper cluster leaves the area of direct influence of the radicals. The decisive bridge for the radicals' influence ceases to exist. Instead of 13 as before, none of the normals ends up radicalized. Without any radicals, the normals would have ended up in three clusters. With a group of 5 radicals, they end up polarized.

If ϵ increases by another 0.01 the cluster, that is second from above, gets early under an ongoing direct radical influence, moves upward, and pulls the two clusters below upward as well. Again, all clusters in the upper segment join. But under the increased ϵ that happens now in the region of direct radical influence. As a consequence all 28 members of the upper segment end up radicalized.

Explanandum 4 are the sudden jumps for $\epsilon = 0.26, 0.27, 0.28$: For an increasing ϵ the number of finally radicalized normals jumps from 22 to all and than to zero. Figure 11 shows the underlying sequence of single runs. For $\epsilon = 0.26$ the chain of radical influence breaks in period 8, when the one agent, that before was a bridge to the evolving upper cluster, joins the lower cluster. For $\epsilon = 0.27$ exactly that agent forms together with his next neighbor above a cluster, that is a bridge to the upper cluster of normals, which itself is a bridge to the radicals. All normals finally join into one cluster. They do that in the area of direct radical influence. Figure 11 shows the dynamics up to period 50. It will take some more time until all normals have become radicalized. But the point of no return was passed long before and they all converge to R. In all three runs in figure 11 the upper cluster in the first periods predominantly moves downward, *away* from R. For the ϵ values 0.26 and 0.27 that downward movement is stopped and then reversed. For $\epsilon = 0.28$

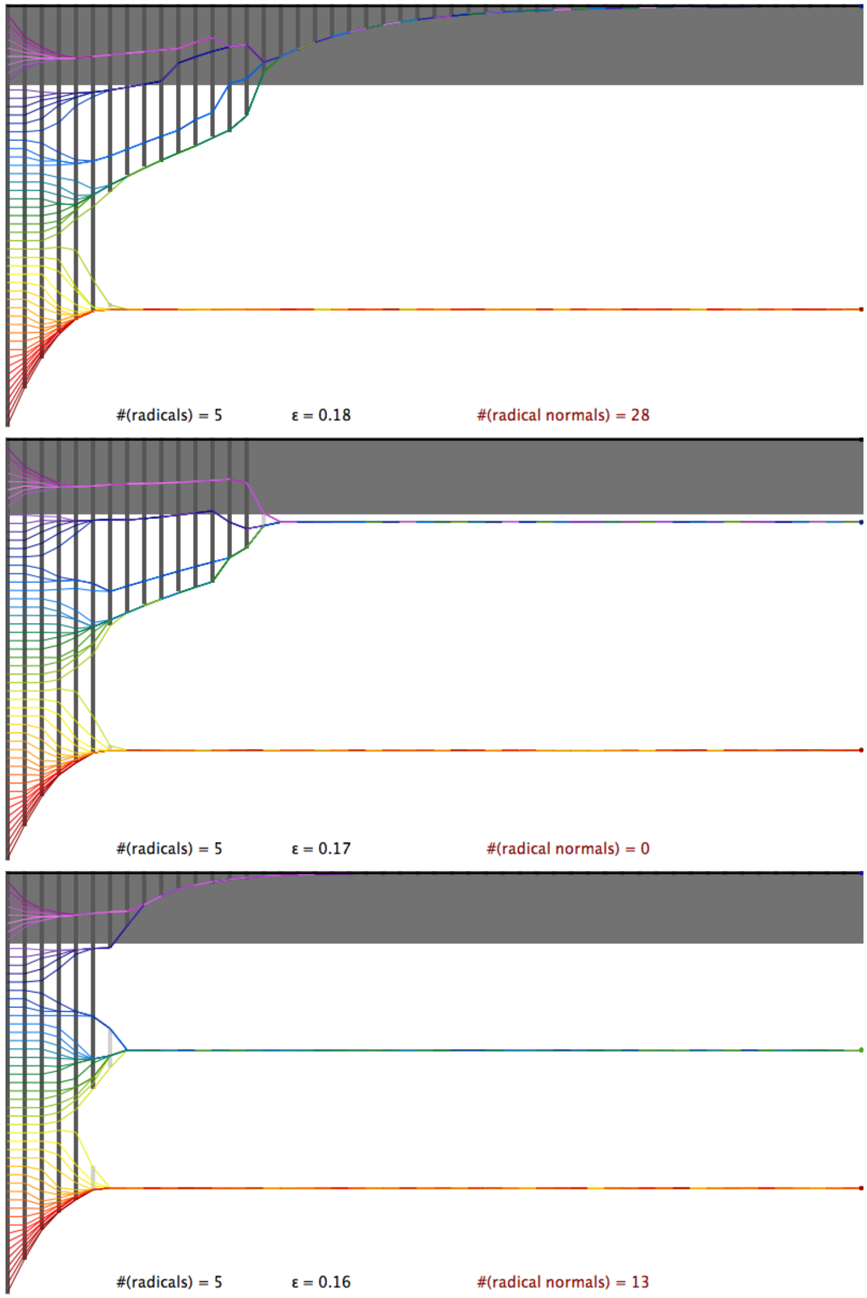

FIGURE 10. Explanandum 3. *Bottom* $\epsilon = 0.16$. *Middle*: $\epsilon = 0.17$. *Top*: $\epsilon = 0.18$. In all runs $\#_R = 5$. A tiny increase of ϵ has major effects.

that does not work any longer: The tiny increase of ϵ makes the pull downward the stronger force. As a result the upper cluster leaves the area of direct radical influence. Thereby the decisive bridge to the radicals ceases to exist and none of the

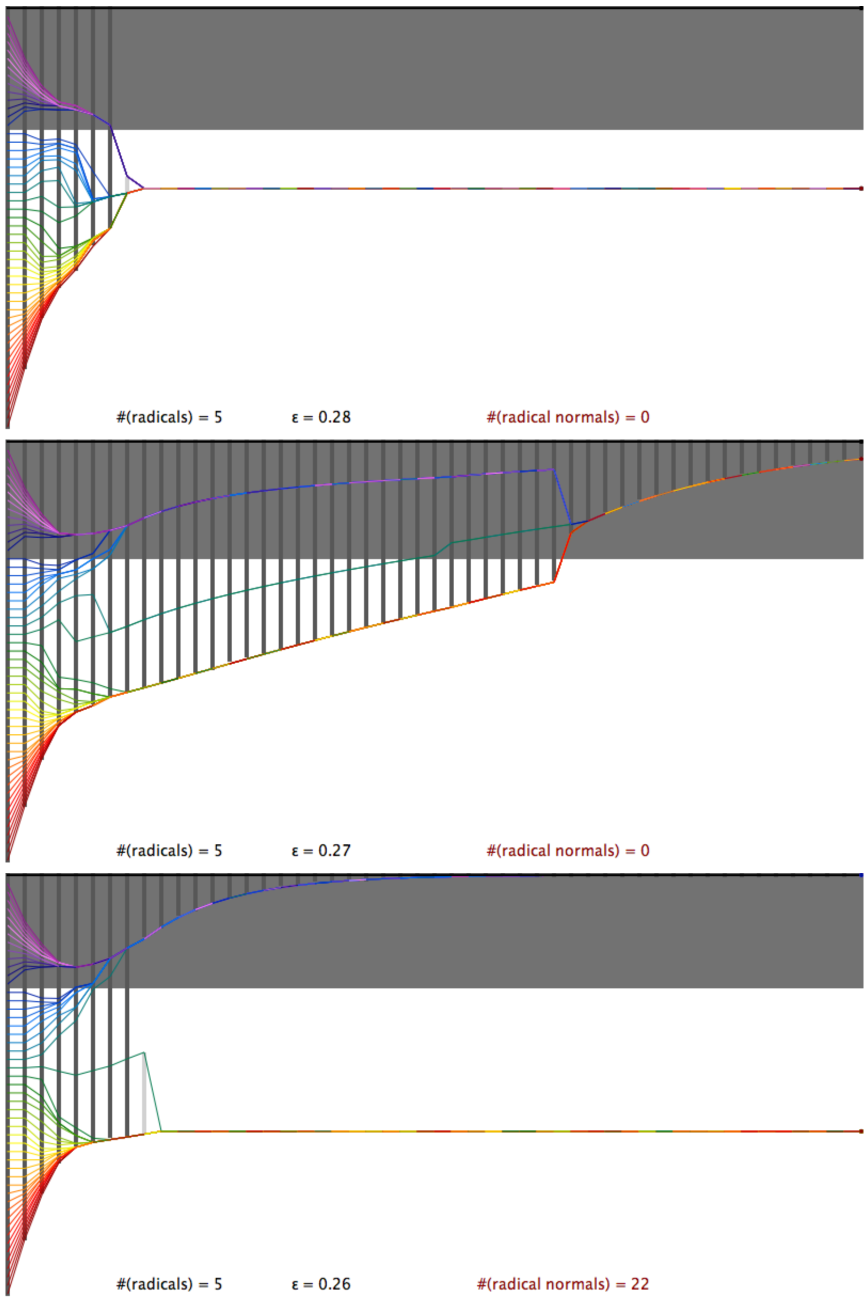

FIGURE 11. Explanandum 4. *Bottom* $\epsilon = 0.26$. *Middle*: $\epsilon = 0.27$. *Top*: $\epsilon = 0.28$. In all runs $\#_R = 5$. A tiny increase of ϵ has major effects.

normals ends up radicalized. It is remarkable how little impact the 5 radicals have on the final consensus of the normals: Without any radical it would be a consensus at 0.5; with 5 radicals it is about 0.58.

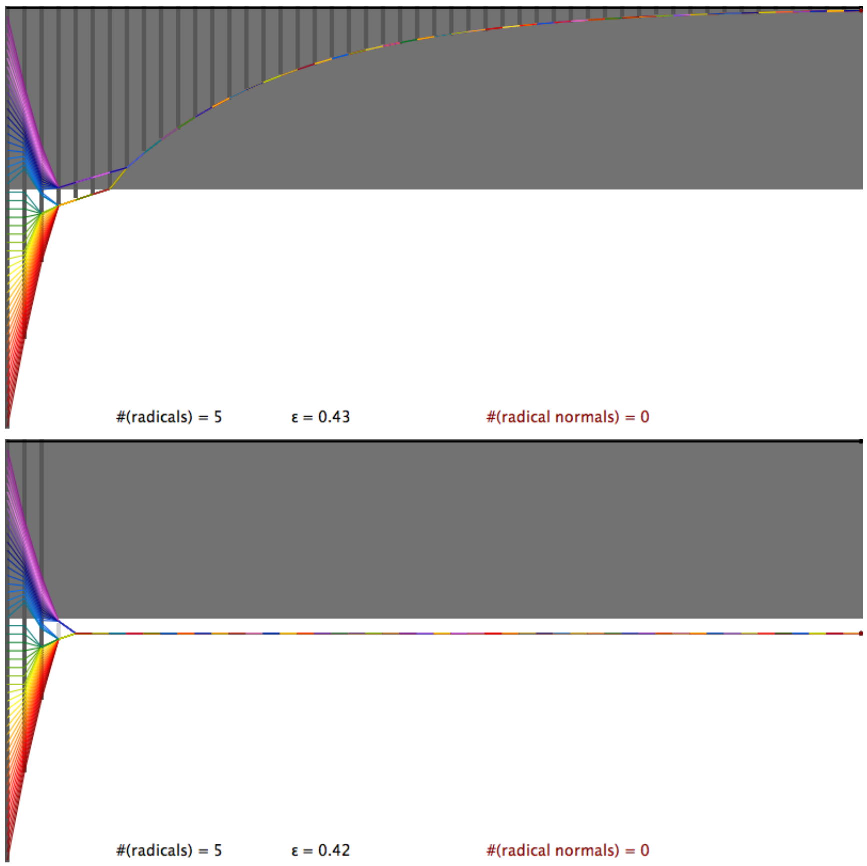

FIGURE 12. Explanandum 5. *Bottom* $\epsilon = 0.42$. *Top*: $\epsilon = 0.43$. In both runs $\#_R = 5$. A tiny increase of ϵ has major effects.

Explanandum 5 is the sudden jump from none to all for the ϵ-values 0.42 and 0.43. Figure 12 shows the single runs. In both runs, due to their now very high ϵ-values, all 50 opinions are very soon very close to each other. Though a bit compressed, we can again distinguish and observe several evolving clusters. The upper one makes a steep move downwards in both runs. For $\epsilon = 0.42$ the upper cluster *leaves* the area of direct radical influence in period 3 by a tiny margin. But the effect is major: The bridge to the radicals is broken. For $\epsilon = 0.43$ the upper cluster is kept, again by a tiny margin, *within* the area of radical influence, that now reaches a bit further down. The sudden and sharp change to an upward movement of the upper cluster in period 3 is an interplay of two factors: First, due to the very high ϵ-value, all 50 opinions are already very close to each other. Therefore, second, the relative impact of the 5 radicals, which directly influence the upper, but not the lower cluster, is high. Kept in the area of the radicals' direct influence, the upper cluster continues to function as a bridge to the radicals. Soon after that bridge becomes superfluous: all normals are in the area of the radicals' direct influence. As in figure 11, middle, in period 50 of figure 12, top, the normals do not yet count as radicalized in the technical sense that $|x_i(t) - R| \leq 10^{-3}$. But soon they will.

Further east, for even higher ϵ-values, it is basically the same as for $\epsilon = 0.43$: For a while the upper cluster is a bridge between all other normals and the radicals. Once all normals have passed the point of no return, and entered the area of direct radical influence, the bridge becomes superfluous.

What carry we home from our expeditions? At five locations in the radicalization landscape we dissected the cogs and wheels, that – working in the depth – bring about the puzzling surface structure. We laid open and made visible the micro layer. At each location we gave an explanatory sketch for the macro effects right there. What are the lessons? The most important are these:

1. A prominent role in all explanations of sudden jumps of the number of finally radicalized normals have *bridges from normals to radicals.* They require, as a kind of pier, normals (cluster or single) that, given a confidence level ϵ,
   (a) are themselves *inside* the area of direct radical influence,
   (b) are *within* the confidence interval of *other* normals, that are *outside* the area of direct radical influence.

   Let's call such bridges "type-R bridges". They are decisive for any influence of radicals outside their limited area of direct influence (which is determined by the normals' confidence level). Type-R bridges allow for *indirect* influence of radicals on normals.

   There is a second type of bridges, *bridges from normals to normals.* They require (again as a kind of pier) normals (cluster or single), that, given a confidence level ϵ,
   (a) are themselves *inside* the confidence interval of at least *two other* normals (cluster or single),
   (b) that themselves are *outside* each other's confidence interval.

   Let's call this type of bridge "type-N bridge".

   What we have seen, then, is, that via an uninterrupted chain of bridges, starting with a type-R bridge and then prolongated by a number of type-N bridges, the radicals may have an influence even on normals that are far away from the radical position. But, except for the group of radicals, the piers of our bridges *can move over time*—and that may *destroy* a bridge, whether of type-R or type-N. At the same time *new* piers for *new* bridges may evolve. *Lesson:* Understanding the radicalization landscape is an understanding of *types, evolution, and breakdown of bridges in a dynamical network.*
2. The probably most striking puzzles are sudden jumps *down* from all to none, or significantly less radicalized normals—and that caused by an *increasing* $\#_R$ or ϵ. Such jumps occur in the *explananda* 2, 3, and 4. In *explanandum* 2 the effect is due to one more radical; in *explanandum* 3 and 4 it is due to a tiny increase of ϵ.

   In explanandum 2 the one more radical causes a pull upwards, that, *via* a type-R bridge, disrupts a former and essential type-N bridge. The movable pier of the type-N bridge moves steeply directions R. Thereby the bridging capacity, given by ϵ, is over-streched, and the type-R-bridge breaks down (see figure 9). In explanandum 3 and 4 – cases with increased confidence levels – it is different: The increased confidence level is sufficient to pull the movable pier of the type-R bridge out of the area of direct radical influence. And that means: breakdown as a type-R bridge (see figure 10, bottom and middle; figure 11, middle and top). Obviously, the additional radical and the increasing ϵ damage an otherwise functioning chain of bridges in *different* ways. There is a reason behind: An increasing number of radicals causes a stronger pull upwards. What an increasing confidence level does, is directly visible in the ϵ-diagram [10] shown in figure 5—exactly our case of 50 normals, but without radicals, i.e. without an upward pull: Obviously an increasing ϵ causes a stronger *contraction* direction center of the opinion space. The upper half of the opinion profile is pulled downwards, the lower half is pulled upwards. But note (and that makes understanding often difficult): The two forces are *interlinked*: ϵ controls the strength of the contracting force, but it also controls what is in the *direct* reach of the force pulling upwards, whatever the strength of that force may be.

An increasing number of radicals may have the effect, that the upward pull disrupts piers of type-N bridges and/or attracts too fast the pier of a type-R bridge. An increase of ϵ causes a stronger contraction. That may sweep along a former pier of a type-R bridge, and the former pier gets outside the area of direct radical influence. In both cases the breakdowns of bridges depends upon thresholds. Therefore they are sudden events. In both cases the breakdown of bridges may stop the radicals' influence on major fractions of normals. As to the numbers of radicalized normals, even jumps from all to none are possible. *Lesson:* We can explain the sudden jumps downward in terms of effects on type-R and type-N bridges in an opinion profile, that is exposed *to two interlinked forces*, that get stronger: the first pulls upwards, the second contracts the range of the profile. If the forces get *stronger*, they may *destroy* decisive bridges of influence.

3. Sudden jumps upwards is another puzzling effect. Such jumps occur in the *explananda* 1, 3, 4, and 5. In *explanandum* 1 the effect is due to one more radical; in *explananda* 3 to 5 it is due to a tiny increase of ϵ. In *explanandum* 1 the pier of a type-R bridge is no longer pulled downwards outside the area of direct radical influence (see figure 8). In *explanandum* 3 it is the same (see figure 10, middle and top). In *explanandum* 4 the tiny increase of ϵ lets a type-N-bridge evolve (see figure 11, bottom and middle). In *explanandum* 5 the pier of a type-N bridge no longer leaves the area of direct radical influence. *Lesson:* Obviously we can explain the sudden jumps downwards in terms of effects of the interlinked forces for type-R or type-N bridges. If the forces get stronger, bridges that before broke down, may keep functioning, or piers for new bridges may evolve.

That are some lessons. Many questions are left open, for instance with regard to the special shape of sensitivity-lines. To answer them will require further future expeditions into the wild region of the parameter space.

5. **If R moves direction center.** So far we assumed that $R = 1.0$. What, if the radical position is less radical, or, as a kind of limiting case, even in the center,

[10] ϵ-diagrams are a powerful instrument to better understand the effects of an increasing ϵ for a given start distribution. Using that instrument, one can easily *proof by example* the following proposition: Given arbitrary start profiles of a *BC*-dynamics, it is *not* generally true, that the final range of the profile or the final number of clusters monotonically decreases as ϵ increases. – Often it is true, but sometimes not. It is possible, that we get consensus for a certain ϵ, and then polarization again for a slightly bigger ϵ—and that pattern can go on several times, until, above a certain ϵ-threshold, the final result is always consensus. It is easy to find such examples for expected-value start distributions. But they exist as well for random distributions. The proposition *refutes* a belief and intuition that is very common.

i.e. $R = 0.5$? Probably such a position would no longer count as "radical". But recall, that the formal description of our model allows for *different* interpretations—radicals and radicalization processes is *only one*. Other interpretations regard a charismatic leader, a groups of dogmatists, campaigning and advertising of all sorts, in short: anything, where a constant signal, with this or that intensity, influences an ongoing opinion dynamics driven by the BC-mechanism.

A very natural way to get an overview, is, to do for *all* values $R = 1.0, 0.99, \ldots, 0.5$, what we have done above for $R = 1.0$. As a result we get 51 phase diagrams of a type as we know it from *figure* 6. The sequence of phase diagrams can be visualized by an animation, that – it is a pity – we can't show here. However, figure 13 shows the phase diagrams for R =1.0, 0.95, 0.9, 0.8, 0.65, 0.5.

By careful visual inspection of the sequence it becomes clear:

1. As the radical position (or: the signal) R moves direction center of the opinion space, both, the horizontal and the vertical wild region, *shrink*. The horizontal wild region finally (i.e. for $R \leq 0.65$) *disappears completely*. From the vertical wild region $E \cup B$, a very narrow vertical stripe of middle sized ϵ-values is left (the remains of region B, nothing is left from E). There one *more* radical (or: an increase in the intensity of the signal by one unit) causes a sudden jump *down* in the numbers of radicalized normals (or: normals convinced by the signal).
2. As R moves direction center, the landscape is more and more *vertically structured*, i.e. more and more it is the size of ϵ, that really matters—whatever the number of radicals (or: the intensity of the signal).
3. The landscape becomes more and more *monotonic*. The number of radicalized normals (or: normals convinced by the signal) is more and more monotonically *increasing* with respect to the confidence level ϵ. With respect to $\#_R$ (the number of radicals or the intensity of the signal) the number of radicalized (or: convinced) normals is finally *decreasing* in a narrow vertical stripe of middle-sized *epsilon*-values and constantly low to left, or constantly 50, i.e. *all* normals, to the right of that stripe.
4. As R moves direction center, the landscape is finally shaped by a *big jump* that happens along a more and more vertical line for ever smaller confidence levels ϵ.

All these observations ask for explanations. There is reason to believe, that *all* explanations can be found by the type of exploration and analysis, that we did for $R = 1$. But we can't do that here.

6. **Conclusions and perspectives.** This paper investigates the impact of a constant signal R on a group of agents ('*normals*') who interact by bounded confidence. None of the agents does influence the signal, that constantly has the value R. But the signal has a *direct* influence on agent i at time t, iff $|R - x_i(t)| \leq \epsilon$. Thus, the signal is only 'heard' directly, if it is not too far away; the signal has a *distance depending* direct effect. The signal comes with a certain intensity: if it is sent simultaneously $\#_R$ times in t, then it is $\#_R$ times in the set $I(i, X(t))$ of an agent i who is under its direct influence, i.e. not too far away. – This is an abstract description in terms of signals. More concrete interpretations can be given in terms of radical or dogmatic groups, charismatic leadership, campaigning, or advertising. That are different interpretations—the formal model is always the same.

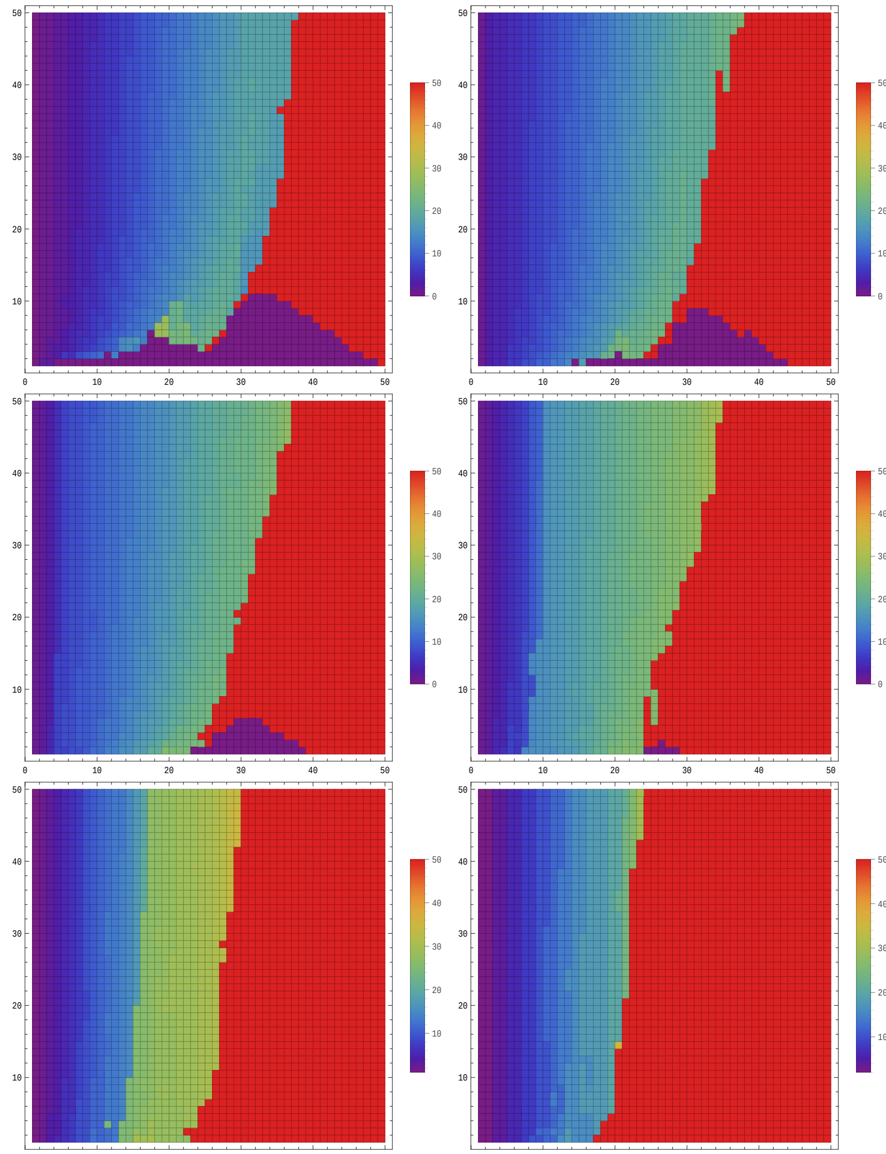

FIGURE 13. R moves direction center of the opinion space. *Top* row: $R = 1.0$ (left), $R = 0.95$ (right). *Middle* row: $R = 0.9$ (left), $R = 0.8$ (right). *Bottom* row: $R = 0.65$ (left), $R = 0.5$ (right). x-axis: the confidence level ϵ increases in 50 steps of size 0.01 from 0.01 to 0.5. y-axis: $\#_R$, i.e. the number of radicals, increases from 1 to 50.

We analyzed in detail the combined effects of the *intensity* of the signal, the *size* of the confidence interval, and the *position* of the signal. The main result for a signal R at the upper bound of the opinion space ('a radical position') is: The parameter space has wild regions. In the region of *low signal intensity* (small values of $\#_R$) and in the region of *middle sized confidence levels*, the number of agents, that end up at the signal's position, depends highly sensitive on the specific confidence/intensity combination $\langle \epsilon, \#_R \rangle$. Tiny increases/decreases in one or both dimensions, and the numbers jump up or down dramatically—even from all to none or none to all. The *intersection* of low signal intensity and middle sized confidence intervals, is especially wild.

We laid open and made visible the *micro* layer that generates deterministically the puzzling macro effects. They can be understood. It is an understanding of a dynamical network in terms of certain types and chains of bridges, their evolution, and their sudden breakdown under different forces that pull on their piers.

Under the perspective of intervention and control (cf.s [9], [18]), the results may illustrate, why both, policies to convince people, and policies to prevent people from getting convinced, *often fail badly.* Suppose a radicalization dynamics as above is going to start. There is a social planner, who, within certain and possibly narrow limits, can increase or decrease both, the normals' confidence interval, and the number of radicals. The radical position can't be changed and is $R = 1.0$. The social planner has perfect knowledge about the landscape of figure 6 and wants to minimize the number of finally radicalized normals. What should the planner do? *If* the planner knows the position of his *actual* society in the landscape, then the task is easy. Obviously, there are constellations $\langle \epsilon, \#_R \rangle$ where increasing the number of radicals would be the best policy to fight radicalization of normals. At some locations in the parameter space an increase of ϵ, at others a decrease would be best. And in some parts of the parameter space the social planner would give up—nothing would help. But that is not the typical epistemic situation of a social planner. The normal epistemic situation is, that the planner has *only a rough knowledge* where the actual society is located. If it is in the wild regions, then the prospects are bleak: Tiny differences may matter and decide about extreme success or extreme failure. Without knowledge of the *exact* position of the actual society, it is easy to fail, while, at the same time, success is not due the planner's expertise—it is just luck in lottery. – The transfer of these types of considerations to other interpretations of the model is easy. A campaigner, for instance, who wants to maximize the number of convinced agents, and now has to decide about the *intensity* with which he sends a certain signal (his 'messages'), may face a very difficult task.

In our analysis the signal has a *distance depending* direct effect: It is 'heard' only by agents with opinions not too far away. But one could think of a receptivity for, and, therefore, an influence of the signal *independent* of the distance to the signal. Some or all of the agents have that receptivity. Again the signal is a constant. The effects of such a mechanism was studied in our paper [11]. The decisive equation is (4): An updated opinion is a convex combination of two components. The first component is the distance independent attraction of the signal. It is determined by a weight α_i with $0 \leq \alpha_i \leq 1$. The second component with the weight $(1 - \alpha_i)$ is the usual BC-opinion dynamics. Agents i with an $\alpha_i > 0$ are attracted by the signal. In agents with $\alpha_i = 0$, only the usual BC-driven dynamics is at work.

$$x_i(t+1) = \alpha_i \cdot R + (1 - \alpha_i) \left(\frac{1}{\#\big(I\big(i, Xt)\big)} \sum_{j \in I\big(i, X(t)\big)} x_j(t) \right). \tag{4}$$

In [11] we interpreted R as the true value in the opinion space. α_i represented in a summary way agent i's capabilities as a truth seeker. Under that interpretation we 'told a story' about the chances of a society to get to the truth, if the truth is $R = 0.1$ (radical), 0.3, or 0.5 (center), the frequency of truth seekers is 10, 50 or 90 percent, the quality of truth seekers is $\alpha_i = 0.01, 0.02, ..., 1.0$, and the confidence levels are $\epsilon = 0.01, 0.02, ..., 0.4$. Different from above, the start distributions were always random. The truth seeker status was randomly assigned. The main lesson was: Wherever the truth is located, whether there is a tiny minority or even an overwhelming majority of truth seekers, under suitable values for the confidence level and the strength of the attraction of truth, the whole society will end up with a consensus on the truth. However, if the truth seekers are 'too good' and converge too fast in the direction of the truth, they may leave behind them – and often far distant from the truth – major fractions of their not truth seeking fellow citizens.

That is a *nice* result under a *nice* interpretation of equation 4. But there are others—some of them *not nice at all*: Equation 4 could be interpreted as modeling a campaigning or advertising context, in which a *campaigner* can manipulate the frequency of receptive agents and/or the strength with which receptive agents get the signal. Another interpretation could be a context of *conspiracy*: There is a group of agents which all have the opinion R. They are not very much. But they want to become more: all should belief that R. They agree, *not* to reveal their true opinion. Instead they assign to each group member a randomly chosen start opinion, that the member then will pretend to have. Finally the accomplices agree upon a joint or individual α_i, that they will apply period by period in the ongoing infiltration process. – Under that interpretation the *translated* main lesson says: Under a lot of circumstances the conspiracy would work—at least in the BC-model-world. (Probably some *real-world* agents, or "*merchants of doubt*" know that since long (see [16]). Real-world citizens should know it as well.)

These re-interpretations of our *old* results, the other way round, suggest a further interpretation of our *new* results about distance *dependent* signals: it is an interpretation in terms of *truth approximation*. Under that interpretation, the model might be about a scientific community under the direct influence of the truth, but embedded in a society of others, more distant to the truth, while an ongoing exchange allows for an indirect influence of the truth. *If* such an interpretation (or an interpretation close by) makes sense, then the translation of our 'radicalization story' into the truth-approximation context, is a contribution to *social epistemology*.

If one *compares* our results about the distance *depending* influence with our earlier results about a distance *independent* influence of a constant signal, then there is one main impression: In the distance depending case the radicalization landscapes are *wild* for extreme values of R. In our earlier analysis of a distance independent influence of a truth-signal, there are *corresponding* landscapes that show the final root mean square deviations of all opinions with respect to the truth-signal. But *whatever* the specific value of the truth-signal, the landscapes are always very *smooth* (though not plane). Of course, a landscape, that shows the final numbers of agents that end up with opinions equal to the signal, is *not* the same as a landscape showing the root mean square deviations of opinions with respect to the signal. But if the former is wild, and the latter smooth, then we have a *strong indicator*, that distance depending and distance independent influence of a constant signal *work very different*. – With that indicator a question comes up: If so, how and why do they differ so much? We do not know the answer—not yet.

Finally, as to the perspectives for rigorous proofs: There are many other relevant ways a constant signal can influence agents. For example, instead by positive integers, one may express the intensity of the signal by *any* positive number (and correcting the averaging correspondingly). In the distance independent case, the intensity is nonnegative constant depending on i only. Using different methods, this case has been delt with already in our earlier paper [11]. Generalizations were proven in [14] and [4]. Or, the intensity is neither constant nor dependent on distance but dependent on *time*. The interesting question then is, whether for the above cases also a segregation as in the Constant Signal Theorem does hold. This, actually, can be proven under quite natural assumptions in a way similar to the one in the Appendix. This, however, is the task for a subsequent paper.

Appendix. As before let $I = \{1, 2, \ldots, n\}$ be the set of normal agents. For simplicity let the group of radicals consist of just one radical r (see the remarks below) and let $I^* = I \cup \{r\}$ the set of all agents. For $i \in I$ let $I(i, X(t)) = \{j \in I^* \mid |x_i(t) - x_j(t)| \leq \epsilon\}$ and for $j \in I^*$

$$a_{ij}(t) = \frac{1}{\#\big(I(i, X(t))\big)} \text{ if } j \in I(i, X(t)) \text{ and } a_{ij}(t) = 0 \text{ otherwise.}$$

Then the interaction among agents is for $t = 0, 1, 2, \ldots$ given

$$\begin{array}{llll} \text{for} & i \in I & \text{by} & x_i(t+1) = \sum_{j \in I^*} a_{ij}(t) x_j(t) \\ \text{and for} & i = r & \text{by} & x_r(t) = R, R \in [0, 1]. \end{array} \tag{5}$$

Concerning the behavior of agents as observed in the simulations, the following analytical result holds:

Main Theorem (Constant Signal Theorem)**.** *The following holds:*

(i) *There exists a (possibly empty) set J of normal agents and a point in time T such that on each agent J there exists a chain of influence from the radical r at each $t \geq T$.*
(ii) *The opinions of all agents in J converge to the opinion of the radical.*
(iii) *The opinions of all normal agents not in J converge in finite time to opinions different from that of the radical.*

Proof. Defining $a_{rj}(t) = 1$ for $j = r$ and $a_{rj}(t) = 0$ otherwise, the interaction of all agents can be described by

$$x_i(t+1) = \sum_{j \in I^*} a_{ij}(t) x_j(t) \text{ for all } i \in I^*$$

or, for the profile $X^*(t) = (x_r(t), X(t))$, compactly by

$$X^*(t+1) = A(t) X^*(t) \text{ for } t = 0, 1, 2, \ldots \tag{6}$$

where $A(t)$ is the $(n+1) \times (n+1)$–stochastic matrix with entries $a_{ij}(t)$ for $i, j \in I^*$, the latter ordered as $(r, 1, 2, \ldots, n)$. The matrix $A(t)$ can be specified furthermore as follows. Let $J(t)$ be the set of normal agents on which the radical has (direct or indirect) influence on at time t. That is, $i \in J(t)$ if and only if there exists a chain $(i_1, i_2, \ldots, i_k)$ of normal agents (also called an ϵ–chain) such that

$$a_{ii_1}(t) > 0, a_{i_1 i_2}(t) > 0, \ldots, a_{i_k, r}(t) > 0.$$

Obviously, for a normal agent $j \notin J(t)$ one has that $a_{jr}(t) = 0$ and, if $i \in J(t)$ in addition, $a_{ij}(t) = 0$. Thus, the matrix $A(t)$ is of the form

$$A(t) = \begin{bmatrix} B(t) & 0 \\ 0 & C(t) \end{bmatrix}$$

where $B(t)$ consists of the entries $a_{ij}(t)$ for $i, j \in \{r\} \cup J(t)$ and $C(t)$ consists of the entries $a_{ij}(t)$ for $i, j \in I \setminus J(t)$. Thus, it may happen that the above *block* structure changes with time.

The proof now proceeds in 4 steps as follows.

1. In a first step we show there exists T such that $J(t) = J$ for all $t \geq T$. For this let's first see that $J(t+1) \subseteq J(t)$ for all t. Without loss we may number the normal agents for fixed t as $x_1(t) \leq x_2(t) \leq \ldots \leq x_n(t)$. Suppose $i \notin J(t)$ and $x_i(t) \leq x_r(t) = R$. Then there exists some $j \in I$ such that

$$x_i(t) \leq x_j(t) < x_{j+1}(t) \leq x_r(t) \text{ with } x_{j+1}(t) - x_j(t) > \epsilon.$$

By the ordering of the $x_{\cdot}(t)$ we must have that

$$a_{jh}(t) = 0 \text{ for } j+1 \leq h, h = r \text{ and } a_{j+1,h}(t) = 0 \text{ for } h \leq j.$$

Therefore

$$x_j(t+1) = \sum_{h \in I^*} a_{jh}(t) x_h(t) \leq x_j(t)$$

and

$$x_{j+1}(t+1) = \sum_{h \in I^*} a_{j+1,h}(t) \geq x_{j+1}(t).$$

This implies $x_{j+1}(t+1) - x_j(t+1) \geq x_{j+1}(t) - x_j(t) > \epsilon$ and taking into account that $x_1(t+1) \leq x_2(t+1) \leq \ldots \leq x_n(t+1)$ we conclude that $i \notin J(t+1)$. An argument as above applies also to the case of $i \notin J(t)$ and $x_r(t) \leq x_i(t)$. This shows $J(t+1) \subseteq J(t)$ for all t. Now, being a descending sequence of finite sets the sequence of the $J(t)$ must become constant, that is there exists T such that $J(t) = J$ for $t \geq T$, where J may be empty. This proves, in particular, part (i) of the theorem.

2. By step 1 we obtain from the decomposition of $A(t)$ in $B(t)$ and $C(t)$ that for $t \geq T$ the interaction of the agents in $J^* = \{r\} \cup J$ and those in $I \setminus J$ is uncoupled, that is

$$x_i(t+1) = \sum_{j \in J^*} a_{ij}(t) x_j(t) \text{ for } i \in J^* \tag{7}$$

and

$$x_i(t+1) = \sum_{j \in I \setminus J} a_{ij}(t) x_j(t) \text{ for } i \in I \setminus J. \tag{8}$$

To these different subsystems we apply two different methods.

3. Considering system (7) we apply a result from [13, Corollary 8.5.10] which yields convergence to consensus provided the matrix $A(t)$ (restricted on J^*) has a positive diagonal, the minimal positive entry is bounded from below by a positive constant and $A(t)$ is coherent. From this result it follows that all agents in J converge to the opinion of the radical because of $r \in J^*$.

   Considering the assumptions on $A(t)$, the diagonal is positive because $a_{ii}(t) \geq \frac{1}{n+1}$ for $i \in I$ and $a_{rr}(t) = 1$. Also $a_{ij}(t) \geq \frac{1}{n+1}$ provided $a_{ij}(t) > 0$. It remains to show that $A(t)$ (restricted to J^*) is coherent. The latter means that any two nonempty saturated sets for $A(t)$ have a nonempty intersection.

Thereby, $\emptyset \neq M \subseteq J^*$ is saturated for $A(t)$ if $a_{ij}(t) > 0$ and $i \in M$ implies $j \in M$. Consider nonempty subsets M, M' of J^* saturated for $A(t)$ and let $i \in M, j \in M'$. If $i = r = j$, then $r \in M \cap M'$. If $i = r, j \neq r$, then $j \in J = J(t)$ for $t \geq T$ and, hence, there exists a chain $(i_1, \ldots, i_k)$ in I such that $a_{ji_1}(t) > 0, a_{i_1 i_2}(t) > 0, \ldots, a_{i_k,r}(t) > 0$.

Since M' is saturated it follows consecutively $i_1 \in M', i_2 \in M', \ldots, r \in M'$. Thus $r \in M \cap M'$. Similarly for $i \neq r, j = r$. Finally, let $i \neq r, j \neq r$ that is $i, j \in J = J(t)$ for $t \geq T$. Then there exists a chain from i to r as well as from j to r and, both M and M' being saturated, we get $r \in M \cap M'$. This proves $A(t)$ is coherent for all $t \geq T$. Therefore, by the result mentioned, part (ii) of the theorem holds true.

4. Since system (8) is the common unmodified BC–model from the fundamental result on fragmentation part (iii) of the theorem does follow. (For fragmentation see [10, Appendix D] and, for a generalization, [13, Theorem 8.5.20]). This proves the theorem.

This completes the proof. □

Remark 1. Whereas in the above theorem convergence among normal agents not in J holds in finite time, this need not be the case for the agents in J. This can be seen already from the case of just one normal agent influenced by one radical. If there is no influence then J is empty.

Remark 2. The case of several radicals sharing the *same* opinion R is covered by the theorem, too. If H is the set off radicals then the interaction of a normal agent i can be written as

$$x_i(t+1) = \sum_{j \in I} a_{ij}(t) x_j(t) + \left(\sum_{h \in H} a_{ih}(t) \right) R.$$

Selecting any $r \in H$ and defining $\tilde{a}_{ir}(t) = \sum_{h \in H} a_{ih}(t)$ one has $\tilde{a}_{ir}(t) = 0$ iff $|x_i(t) - x_r(t)| > \epsilon$. By this H is represented by the single radical r with a weight of $|H|$.

Remark 3. The theorem above bears some relationship to [11, Theorem 2 and Appendix] where the "truth", like a radical, is not influenced by normal agents. The two cases are, however, rather different in that the influcence of the "truth" is constant, whereas the influence of a radical depends on the distance to normal agents.

Acknowledgments. The model was defined in 2013 by Rainer Hegselmann during his year as a senior fellow at the Alfried Krupp Wissenschaftskolleg in Greifswald. The model is the direct result of intensive discussions with Gregor Betz, Michael Baurmann and Rainer Cramm. He profited a lot from the discussions and the stimulating intellectual atmosphere in the Alfried Krupp Wissenschaftskolleg. In 2014 Rainer Hegselmann was a visiting international fellow of the department of sociology at Surrey University. Corinna Elsenbroich, Jen Badham, and Peter Johnson (all of them members of CRESS, the Centre for Research in Social Simulation, based in the sociology department) helped a lot to analyze the model, to fix bugs, and to develop ideas for future modifications. A discussion with the statistician Walter Olbricht (Bayreuth University) about expected value start distributions was extremely helpful. Finally, with all their questions and constructive proposals, the participants of the seminar "Information and Control", regularly hold at Bayreuth University by Torsten Eymann, Sascha Kurz, Joerg Rambau, and Rainer Hegselmann, were a major help and a major source of inspiration. Many thanks to all.

REFERENCES

1. Daron Acemoglu and Asuman Ozdaglar, *Opinion dynamics and learning in social networks*, Dynamic Games and Applications **1** (2011), no. 1, 3–49.
2. Michael Baurmann, Gregor Betz, and Rainer Cramm, *Meinungsdynamiken in fundamentalistischen Gruppen - Erklärungshypothesen auf der Basis von Simulationsmodellen*, Analyse and Kritik **36** (2014), no. 1, 61–102.
3. Vincent D. Blondel, Julien M. Hendrickx, and John N. Tsitsiklis, *Continuous-time average-preserving opinion dynamics with opinion-dependent communications*, SIAM Journal on Control and Optimization **48** (2010), no. 8, 5214–5240.
4. Bernard Chazelle, *The total s-energy of a multiagent system*, SIAM Journal of Control and Optimization **49** (2011), no. 4, 1680–1706.
5. Guillaume Deffuant, Fréderik Amblard, Gerald Weisbuch, and Thierry Faure, *How can extremism prevail? A study based on the relative agreement interaction model*, Journal of Artificial Societies and Social Simulation **5** (2002), no. 4.
6. Guillaume Deffuant, David Neau, Fréderik Amblard, and Gerald Weisbuch, *Mixing beliefs among interacting agents*, Advances in Complex Systems **3** (2000), 87–98.
7. Morris H DeGroot, *Reaching a consensus*, Journal of the American Statistical Association **69** (1974), no. 345, 118–121.
8. Rainer Hegselmann, *Bounded confidence, radical groups, and charismatic leaders*, Advances in Computational Social Science and Social Simulation. Proceedings of the Social Simulation Conference 2014 Barcelona, Catalunya (Spain), September 15 (Barcelona) (Francisco J. Miguel, Frédéric Amblard, Joan A. Barceló, and Marco Madella, eds.), Autònoma University of Barcelona, (DDD repository http://ddd.uab.cat/record/125597), 2014, pp. 217–219.
9. Rainer Hegselmann, Stefan König, Sascha Kurz, Christoph Niemann, and Jörg Rambau, *Optimal opinion control: The campaign problem*, Journal of Artificial Societies and Social Simulation (JASSS) (resubmitted) (2015).
10. Rainer Hegselmann and Ulrich Krause, *Opinion dynamics and bounded confidence: Models, analysis and simulation*, Journal of Artificial Societies and Social Simulation **5** (2002), no. 3.
11. ———, *Truth and cognitive division of labour: First steps towards a computer aided social epistemology*, Journal of Artificial Societies and Social Simulation **9** (2006), no. 3.
12. Sylvie Huet, Guillaume Deffuant, and Wander Jager, *Rejection mechanism in 2d bounded confidence provides more conformity*, Advances in Complex Systems **11** (2008), no. 4, 1–21.
13. Ulrich Krause, *Positive dynamical systems in discrete time. Theory, models, and applications*, De Gruyter, Berlin, 2015.
14. Sascha Kurz and Jörg Rambau, *On the Hegselmann-Krause conjecture in opinion dynamics*, Journal of Difference Equations and Applications **17** (2011), no. 6, 859–876.
15. Jan Lorenz, *Continuous opinion dynamics under bounded confidence: a survey*, International Journal of Modern Physics C **18** (2007), no. 12, 1819–1838.
16. Naomi Oreskes and Erik M. Conway, *Merchants of doubt - How a handful of scientists obscured the truth on issues from tobacco smoke to global warming*, Bloomsbury Press, 2010.
17. G. Gary Polhill, Luis R. Izquierdo, and Nicholas M. Gotts, *The ghost in the model (and other effects of floating point arithmetic)*, Journal of Artificial Societies and Social Simulation **8** (2005), no. 1, http://jasss.soc.surrey.ac.uk/8/1/5.html.
18. Suttida Wongkaew, Marco Caponigro, and Alfio Borzí, *On the control through leadership of the Hegselmann–Krause opinion formation model*, Mathematical Models and Methods in Applied Sciences **25** (2015), no. 03, 565–585.
19. Haoxiang Xia, Huili Wang, and Zhaoguo Xuan, *Opinion dynamics: A multidisciplinary review and perspective on future research*, International Journal of Knowledge and Systems Science **2** (2011), no. 4, 72–91.

Received December 1, 2014; revised March 6, 2015.

E-mail address: rainer.hegselmann@uni-bayreuth.de; krause@math.uni-bremen.de