THESIS FOR THE DEGREE OF LICENCIATE OF PHILOSOPHY

Consensus formation in the Deffuant model

TIMO HIRSCHER

Division of Mathematics
Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
AND
UNIVERSITY OF GOTHENBURG
Göteborg, Sweden 2014

Consensus formation in the Deffuant model
Timo Hirscher

NO 2014:6
ISSN 1652-9715

Department of Mathematical Sciences
Chalmers University of Technology and
University of Gothenburg
SE-412 96 Göteborg
Sweden
Phone: +46 (0)31-772 10 00

Author e-mail: hirscher@chalmers.se

Printed in Göteborg, Sweden 2014

Consensus formation in the Deffuant model

Timo Hirscher

Department of Mathematical Sciences
Chalmers University of Technology
and
University of Gothenburg

Abstract

This thesis deals with a mathematical model used in the context of social interaction in large groups, introduced by Deffuant et al. in 2000. Each individual holds an opinion and shares it with others in random pairwise encounters. If the difference in opinions of two interacting agents is less than a given threshold, the discussion will lead to an update of their opinions towards a compromise. If the difference is too large, however, they will ignore each other and separate with their opinions staying unchanged.

Many results on long-time behavior of this opinion formation process – mainly dealing with whether a common consensus is reached or not – were established using computer simulations (for different underlying network topologies; interactions can only take place between neighboring individuals).

In the two papers this thesis is based on, we study the model on integer lattices analytically, using geometric arguments and probabilistic tools as well as concepts from statistical physics. While the first paper focusses on univariate opinions but considers also higher-dimensional lattices as well as infinite percolation clusters as underlying network graphs, the second one sticks to the infinite line graph as topology and deals with multivariate opinions instead.

Keywords: Deffuant model, consensus formation, opinion dynamics, sociophysics, vector-valued opinions, percolation

List of Appended Papers

A Olle Häggström and **Timo Hirscher**, Further results on consensus formation in the Deffuant model (published in *Electronic Journal of Probability*, Vol. 19).

B **Timo Hirscher**, The Deffuant model on $\mathbb{Z}$ with higher-dimensional opinion spaces (under review for *Latin American Journal of Probability and Mathematical Statistics*).

Acknowledgments

First of all, I would like to thank my advisor Olle Häggström for constantly providing me with a supportive and motivating blend of positive reinforcement and constructive criticism. He manages to give me enough guidance in order that I don't feel lost and enough freedom in order that I don't feel guided. Discussing mathematics (as well as other important aspects of life) with him is each time both pleasant and instructive.

Then I also want to express my gratitude to Jeff Steif for co-organizing the graduate course on 'Markov chains and mixing times' together with me, as this gave me the great opportunity to gain experience in and to get his advice on lecturing on graduate level.

I also would like to thank all colleagues at the mathematical department for creating this still professional but also homelike working atmosphere – day-to-day proving wrong the common preconception that mathematicians are socially incompetent, preoccupied human beings. Without degrading others, I want to explicitely name Matteo, who acted like a close friend already before we really got to know each other.

Finally, I want to dearly thank my family and Mareile for their unconditional and constant love and support.

Timo Hirscher
Göteborg, March 2014

To Mareile

Contents

Abstract **i**

List of Appended Papers **iii**

Acknowledgments **v**

1 Introduction **1**

2 Statistical physics and social dynamics **3**

2.1 A first connection: the Ising model 5
2.2 Opinion dynamics 8
2.3 Results for the Deffuant model 14
2.4 Cultural dynamics 18

3 Summary of appended papers **21**

References **25**

A Further results on consensus formation in the Deffuant model **29**

B The Deffuant model on $\mathbb{Z}$ with higher-dimensional opinion spaces **57**

1
Introduction

In the 19th century, when James Clerk Maxwell and Ludwig Boltzmann elaborated the ideas of Daniel Bernoulli to describe the kinetic dynamics in gases not by focussing on each single particle but by characterizing the whole system with a set of parameters and their distributions among the particles, the field of modern statistical physics was born.

There were two main reasons for other disciplines to get interested in this idea: The significant progress in the study of collective phenomena in physics and the fact that many complex systems are of statistical nature, i.e. a large number of similar microscopic elements form a macroscopic object, which has properties that are formed by the collective but the contribution of any individual particle is negligible.

In the second half of the 20th century social sciences became one of those disciplines. Even though the idea of statistics has its origin in the attempt to get a quantitative understanding of large groups of human beings (e.g. birth and death rates), it was not until the appearance of the article titled *'The value of statistical laws in physics and social sciences'* [19] in 1942, written by theoretical physicist Ettore Majorana, that the relevance of statistical laws for studies in social sciences was spelled out. However, the fact that the article was written in Italian and moreover found and published after Majorana's disappearance by his brother (as reported in [20]) resulted in very little attention by other scholars and so an additional delay of several decades for Majorna's ideas to be applied.

One of the first to do so was Wolfgang Weidlich [27] in 1971. He used statistical models that were originally designed for the description of the dynamic development of an ensemble of interacting particle spins to study the structure of opinion formation in a social group of individuals mutually influencing each other.

Throughout the last two decades more and more physicists and mathematicians started similar attempts to analyze the opinion dynamics in large groups of individuals by using a simplistic interaction model and applying qualitative and quantitative methods from statistical physics to it. The fact that new social phenomena which arose with the advancement of the internet – like e-mail correspondences for example – feature large groups of individuals, simple interactions and allow for a computational treatment of the corresponding large datasets contributed substantially to this evolution.

2

Statistical physics and social dynamics

Even though the idea became increasingly popular to model social interactions using simplified mathematical models derived from those used in statistical physics, there are glaring differences between the two fields of application. Possibly most important is the contrasting complexity of the elementary components: In physical applications the systems consist of relatively simple objects, usually atoms and molecules, the behavior of which is relatively well understood; hence the complex evolution of the collective arises from the interaction patterns. In social science, however, the collective consists of a large number of human beings and the behavior of each single individual is already the outcome of a complex interplay between physiology and psychology of which only very little is understood. Especially the fact that in all common models for social dynamics the individuals are presupposed to behave adaptively, that is reacting to influences, and not strategically, in other words following a certain plan they have in mind, seems to be an unrealistic assumption. So even though Majorana [19] tried to motivate these simplifications of human decision behavior by philosophical reflections about whether human consciousness and free will might be reducible to quantum effects in the brain, the reduction of humans to simplistic elements in a large system is still a controversial issue.

One might come to the conclusion that reducing the complexity on microscopic level to such an extent that the system makes a treatment using tools from statistical physics possible without changing the essential macroscopic phenomenology is a hopeless task.

Nevertheless, there is a striking structural similarity in the dynamics of opinion formation in a group and physical systems that suggests a meaningful relation between the two. Just like the spins in an ensemble of interacting particles, the individual opinions might be in a chaotic state at first – meaning that no large scale structure exists – but then gradually align through interaction and finally undergo a transition from disorder to order in the sense that the system exhibits large scale regularities in the long run (which in the physical context correspond to a state of low energy). When it comes to social interactions, the drive towards order is due to the tendency of interacting individuals to become more alike, an effect called *social influence*. If we stick to the metaphor, ordered low energy states in physics correspond to consensus or uniformity in the context of opinion dynamics and disordered states of higher energy in turn to fragmentation or disagreement. One of the main questions in social dynamics is – similarly to the situation in statistical physics – under which circumstances the microscopic interactions will lead to such a transition, since if there were no interactions, in both contexts heterogeneity would prevail.

There is another important argument that alleviates the problem of reducing humans to particles: In statistical physics most of the qualitative properties of a system on a larger scale do not depend on the microscopic details of the dynamics but instead on global properties like symmetries, dimensionality or conservation laws. In this respect it is at least justifiable that modelling a few of the most important properties of single individuals will capture the essential driving forces of the evolution and thereby give meaningful results when it comes to qualitative features of the model's large scale behavior.

2.1 A first connection: the Ising model

With these considerations in mind, it is no longer surprising that Weidlich suggested in a physics colloquium in 1969 to compare the interactions within a group of individuals holding opposing attitudes towards a certain topic with the magnetization process. In a ferromagnetic material the atoms have elementary magnetic dipoles, called *spins*, that are fluctuating at random at high temperature. If the temperature drops below a critical threshold, the cooperative interaction of the spins dominates and leads to first local then global alignment, finally magnetizing the material on a macroscopic scale. Two years later he published this idea in the article *'The statistical description of polarization phenomena in society'* [27] in which he elaborated how the mathematical model introduced by Ernst Ising [14] in 1925 to describe ferromagnetism in statistical mechanics can be interpreted in a sociological context.

In the *Ising model*, a collection of n atoms is considered, each holding a spin state which can be either $+1$ or -1. Each atom's spin is energetically pushed to align with the ones of its nearest neighbors in the following sense: The total energy of a spin configuration $\sigma \in \{-1,+1\}^n$ is defined by

$$H(\sigma) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j,$$

where the sum is taken over all pairs $\langle i,j \rangle$ of nearest neighbors in the atomic lattice and J models the interaction strength ($J > 0$ corresponding to a ferromagnetic interaction). One of the most common models for the dynamics is the so-called *Metropolis-Hastings algorithm*, a rejection sampling method, in which a single move is the flip of a randomly chosen spin that is accepted with probability $\min\{\exp(-\frac{\Delta H}{k_B T}), 1\}$, where ΔH is the invoked change of the total energy, k_B a (positive) physical constant and T the temperature (in degree Kelvin). In the ferromagnetic regime, flipping the spin at site v might be rejected if the majority of neighbors agrees with the current spin as this will lead to ΔH being positive. Hence a low temperature will drive the system towards one of the two homogeneous states – all spins being $+1$ or -1.

For this model, it was found that there exists a critical temperature T_c above which the spins only correlate on a small spatial scale. For a temperature lower than the critical value, however, the ordered domains grow in some kind of coarsening process until one of them is dominating and spreads over all sites. This type of phase transition is a characteristic feature of both many systems considered in statistical physics and collective dynamics of large social groups, in which a common language is formed or a consensus about a specific issue is reached.

Weidlich suggested to consider the spins as attitudes towards a given yes-no question, the parameter J to be reinterpreted as the willingness of an individual to adopt the attitude of the majority among its neighbors and the temperature as a model parameter for the social pressure exerted on each individual (low temperature corresponding to high social pressure). He already suggested natural extensions of this connection between social dynamics and statistical physics, for instance an external magnetic field (shaping some preference of one attitude over the other, shared by all individuals), the consideration of more than two possible attitudes and letting the transition probability to flip the spin at v depend not only on its current value, but also on its history – introducing a sense of tradition or stubbornness.

In 1982, Galam et al. [11] used the Ising model to shape the collective behavior in a plant where dissatisfied workers might start a strike. They rediscovered the phase transition described above and interpreted the regime of high temperature as an individual phase (mutual influences are very limited) and low temperature as a collective one (the group behaves coherently), separated by a critical phase in which small changes in the system can lead to drastic changes in the group. In contrast to the application of the Ising model to a collection of atoms forming a regular lattice, it is reasonable to consider the interaction pattern among workers in a plant to be all-to-all, meaning that every worker can actually influence all his fellow workers, which suggests the complete graph K_n as underlying interaction network.

Clearly the topology, i.e. the structure of the interaction network, is a very important aspect in social dynamics. Under the assumption that the interaction is all-to-all, often termed *homogeneous mixing*, it is possible to use another tool coming from statistical physics, the so-called *mean field approximation*. In most cases this makes an analytical treatment possible, in the sense that solving the corresponding differential equations will give insights about the long-term behavior. However, already in the globalized companies of today this assumption is hardly realistic – not to mention the extremely sparse networks of e-mail correspondences and the like.

For that reason, all of the models we are about to review in the next section were mainly considered on sparser networks, for example lattices or realizations of random graphs like the well-known *Erdős-Rényi model* or the *Barabási-Albert model*. The latter is based on preferential attachment – that means it is built incrementally from a core of m fully connected individuals by adding new nodes one by one, each choosing m older nodes to connect to with a probability proportional to their degree – and produces a scale-free network, which turns out to be a realistic model for e-mail networks or friendship graphs, both popular objects of study in the branch of social network analysis.

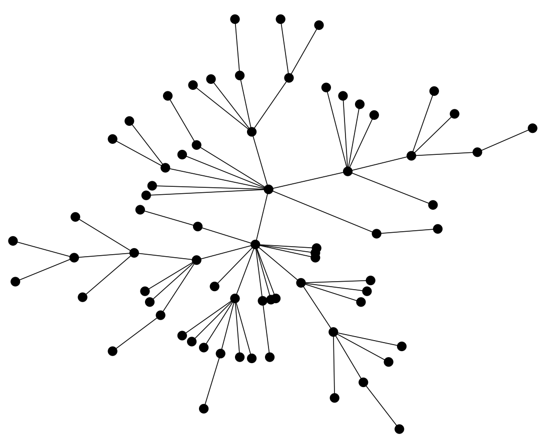

Figure 2.1: *A typical Barabási-Albert network for $m = 1$ of small size ($n = 70$).*

The lack of analytical means that could be applied to the common models for social dynamics as well as the increasing computational power resulted in numerous simulation-based analyses during the last two decades. On the one hand this surely complements the analytical study of such models based on tools from statistical physics, on the other hand this approach is limited to a rather small number of individuals and even if this seems to be sufficient for an examination of the opinion formation in social groups, the concept of order-disorder phase transitions is rigorously defined only in the limit of a system with infinitely many particles. A number of individuals that is not sufficiently large might therefore cause finite size effects that invalidate conclusions drawn from a comparison with analog systems in physics, in which the number of interacting particles is commonly by far larger than in a social group. In this respect it is of vital importance to be able to figure out which macroscopic features are robust with respect to changes in the number of interacting individuals by analyzing the used model for different orders of magnitude of the system's size.

2.2 Opinion dynamics

Since there are many situations in everyday life where it is necessary for a social group to reach a consensus in order to make a shared decision, it has always been a major focus of social science to understand the opinion formation process in larger groups. Inspired by statistical physics, and particularly Weidlich's sociological reinterpretation of the Ising model, various models for opinion dynamics arose in the sequel. All of them share similar ideas as well as the common aim to define opinion states of a population and to determine what kind of changes in the elementary interactions invoke a transition between such states. Depending on the nature of the random variables representing the opinion values, the interpretation of such variables varies from case to case: While a binary variable might stand for the attitude towards a given yes-no question or if the individual has been reached by a rumor or not, a continuously distributed opinion value on the positive real numbers might embody each individual's belief regarding the age of the universe, for instance. In all cases the dynamics tend to reduce the variability of the initial opinion values which may lead to a consensus state in the long run.

In this section we will shortly introduce a few of the most common models for opinion dynamics and refer the reader to the comprehensive survey article *'Statistical physics of social dynamics'* [3] by Castellano, Fortunato and Loreto for a more detailed discussion and further references.

(a) Voter model

This model was introduced by Clifford and Sudbury [4] in 1973 as a model for two competing species and later named for its natural interpretation in terms of opinion dynamics among voters. Its definition is very simple: Each individual holds an opinion given by a $\{-1,+1\}$-valued variable. At every time step, one individual is selected at random and will then adopt the opinion of another agent, picked uniformly among its neighbors. On regular lattices the evolution of this model is to some extent similar to the Ising model – in one dimension, that is on the line graph $\mathbb{Z}$, it actually corresponds exactly to the limiting case of the Ising model with zero temperature.

Variants include the *multitype voter model* (analyzed in [22] for example), in which more than two opinion values are possible, as well as the *constrained voter model* introduced by Vazquez et al. [26]: Each agent is in one of three states (left, right or center) and interactions as described above can only occur involving at least one centrist (as the extremists 'left' and 'right' are assumed not to interact with each other). This behavior is a discrete analog of the so-called *bounded confidence* principle (see below).

(b) Majority rule model

A finite collection of n individuals is considered, a fraction p_+ of which initially holds opinion $+1$, all others the opinion -1. At each iteration a random group of individuals is chosen, and all group members then adopt the majority opinion inside the group. In the simplest version, the size of the chosen groups is a fixed odd number. But there are various variants with random size and different ways to resolve a tie in a group consisting of an even number of individuals. The model was introduced in [10] and proposed to describe public debates.

Another model based on the majority rule is the so-called *majority-vote model* introduced by Liggett [17]. Just like in the Ising model, spins are updated one at a time. At each step, the spin to be updated takes on the value of the majority of its neighbors with probability $1-q$, the minority value with probability q and is chosen uniformly from $\{-1,+1\}$ if there is a tie. For $q=0$ this corresponds again to the Ising model with zero temperature. Several studies (focussing on different interaction networks) showed that the majority-vote model exhibits an order-disorder phase transition when q is increased.

(c) Hierarchical majority rule model

A structurally different model using the majority rule was proposed by Galam [9]: A group of $n=r^k$ individuals ($r,k\in\mathbb{N}$), which are equipped with identically distributed $\{-1,+1\}$-valued opinions, is considered. Let p_0 denote the common probability for the opinion to be $+1$.

Instead of forming a consensus by interacting, they elect a representative for the whole group in the following way: In the first round, all individuals are randomly divided into groups of size r. In every group a representative is chosen among the members sharing the majority opinion of the group – uniformly among all members if r is even and there is a tie in the group. This procedure is then iterated among the representatives until a single leader is elected in the kth round. If p_i denotes the probability that a representative on hierarchical level i holds opinion $+1$, the recursion is given by

$$p_{i+1}=\sum_{l=\frac{r+1}{2}}^{r}\binom{r}{l}p_i^l\,(1-p_i)^{r-l}\quad\text{if } r \text{ is odd and}$$

$$p_{i+1}=\frac{1}{2}\binom{r}{\frac{r}{2}}p_i^{\frac{r}{2}}\,(1-p_i)^{\frac{r}{2}}+\sum_{l=\frac{r}{2}+1}^{r}\binom{r}{l}p_i^l\,(1-p_i)^{r-l}\quad\text{if } r \text{ is even.}$$

(d) Sznajd model

There are different versions of this model sharing the same basic interaction principle. The following is not the one originally introduced by Sznajd-Weron and Sznajd [25] although the most popular variant. The individuals are considered to occupy the sites of a graph (shaping the interaction network) and to hold again $\{-1,+1\}$-valued opinions. A pair of neighboring agents is picked and if they agree, all their neighbors adopt this opinion as well. If they disagree, however, nothing happens. The Sznajd model is designed to incorporate the typical human behavior to be influenced more easily by a group of people that agree on a certain topic compared to the influence of single individuals. Variants of the Sznajd model have in fact been applied in order to model and understand voting behavior in elections.

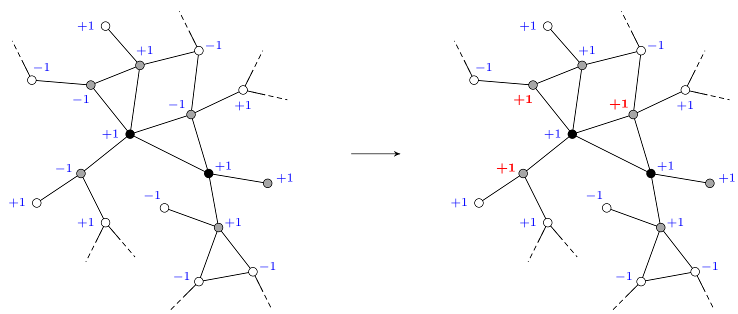

Figure 2.2: *Update rule in the Sznajd model: If the two neighbors picked (black) agree, they impose their opinion on all other individuals linked to them (gray).*

The three models considered so far feature binary variables, partly having generalizations to a finite number of possible opinion values. As the dynamics have to be defined in such a way that this property is maintained, all three do not preserve local averages when elementary particles interact – a principle that is found in many physical systems. In that sense considering continuous opinions can be quite different, also because concepts like 'equality of opinions' or 'majority opinion' do not have equivalents in the continuous setting.

Unlike the models that are closely related to the Ising model, the so-called *bounded confidence models* involve a rational aspect in the interaction behavior that can not be found in the interaction of physical particles: When two individuals meet, they will only influence one another if their present opinion values are not too far apart from each other. In other words, there exists a parameter $\theta \geq 0$ shaping the *tolerance* of the individuals: If the current opinion value of an agent is $\eta \in \mathbb{R}$, other agents holding opinions outside of the interval $[\eta - \theta, \eta + \theta]$ will just be ignored. The bounded confidence models below have also been reviewed in [18].

(d) Deffuant model

Besides the tolerance θ, this model features another parameter $\mu \in (0, \frac{1}{2}]$ that embodies the willingness of an individual to approach the opinion of the other in a compromise. Encounters always happen in pairs, so if agents u and v meet at time t, holding opinions $a, b \in \mathbb{R}$ respectively, the update rule reads as follows:

$$(\eta_t(u), \eta_t(v)) = \begin{cases} (a + \mu(b-a), b + \mu(a-b)) & \text{if } |a-b| \leq \theta, \\ (a, b) & \text{otherwise,} \end{cases}$$

where $\eta_t(u)$ denotes the opinion of agent u at time t. The idea behind this is simple: When two individuals interact and discuss the topic in question, they will only rate the opinion encountered as worth considering if it is close enough to their own personal belief. If this is the case, however, they will have a constructive debate and their opinions will symmetrically get closer to each other – in the special case $\mu = \frac{1}{2}$ they will separate having come to a complete agreement at the average of the opinions they hold before the interaction.

In this manner, groups of compatible agents concentrate more and more around certain opinion values (their initial average) and once each such cluster of individuals is sufficiently far from the others, the final opinions are formed and all groups will from then on only become more homogeneous by internal interactions.

When Deffuant et al. introduced this model in [6], it was considered on a finite number of agents having i.i.d. initial opinions uniformly distributed on $[0, 1]$. The encounters occurred in discrete time by picking at each time step a pair of agents uniformly at random from the edge set of the underlying interaction network graph. They ran computer simulations in order to figure out for which values of the parameters θ and μ the group will end up in one opinion cluster (corresponding to a consensus) or split into several clusters (fragmentation).

Stauffer et al. [24] introduced a discretized version of the model, in which the opinions can take on values from the set $\{1, 2, \dots, q\},\ q \in \mathbb{N}$ and are rounded to the nearest integer after an update of the form described above. There have also been attempts to analyze the model with the tolerance parameter varying from individual to individual, revealing that in such a generalization it is the individuals with largest tolerance that ultimately determine the system's behavior.

(e) Hegselmann-Krause model

The model introduced in [13] is quite similar to the Deffuant model, just the rule for encounters (again they happen in discrete time) is different: Given a graph (modelling the interaction network) at every time step each individual interacts with all its compatible neighbors at once and takes the average as its new opinion. If we let $\sim$ denote the reflexive adjacency relation, i.e. $u \sim v$ if $u = v$ or u and v are neighbors in the graph, and $\eta_t(u)$ once more the opinion of agent u at time t, we can write the update rule as follows:

$$\eta_{t+1}(v) = \frac{1}{N_t(v)} \sum_{\substack{u \sim v \\ |\eta_t(u) - \eta_t(v)| \leq \theta}} \eta_t(u) \quad \text{for all } v,$$

where the sum runs over the set of agents that consists of v plus its compatible neighbors and $N_t(v) = \big|\{u : \ u \sim v, \ |\eta_t(u) - \eta_t(v)| \leq \theta\}\big|$ is the size of this set at time t. Note that in contrast to the Deffuant model, the mean opinion is not conserved over time.

When it comes to simulations of the model, the major disadvantage of the Hegselmann-Krause model compared to the one introduced by Deffuant et al. is that for a dense network graph averages of large groups of agents have to be calculated. This makes the running time until a meaningful structure – in order to decide whether the system approaches consensus or fragmentation – emerges rather long. On the other hand, for a finite number of individuals the system converges to a stable state in finite time: Once the opinion clusters are formed and all agents in one fixed cluster are compatible with one another, they will completely agree after one more time step making further changes of their opinions impossible.

2.3 Results for the Deffuant model

Having put the model proposed by Deffuant et al. into the broader context of other common models for opinion formation processes, we now want to give a short overview of the results that have been achieved in the analysis of the Deffuant model.

The findings in the original paper [6] were threefold. The authors tried to shed light on the influence of the model parameters θ and μ as well as the underlying network topology: In their computer simulations of the homogeneous mixing case (that means the interaction network is the complete graph) with $n = 1000$ individuals – initially holding i.i.d. unif$([0, 1])$ opinions – Deffuant et al. noted that for a confidence parameter $\theta = \frac{1}{2}$ the system most likely converges to a consensus at the mean opinion $\frac{1}{2}$, whereas $\theta = \frac{1}{5}$ leads to a fragmentation into two finally homogeneous groups, whose opinion values lie roughly at 0.25 and 0.75. Besides this dichotomy of regimes, by varying the model parameters they found that the convergence parameter μ and the number of individuals n influence the convergence time only, not the qualitative dynamics which primarily depend on θ. The persistent opinions were arranged equidistantly and their number scaled roughly like $\lfloor \frac{1}{2\theta} \rfloor$.

When they tried to track the opinion evolution of single agents from their initial opinions to one of the several persistent ones in the fragmentation case, they observed that the overlap of ranges (in terms of initial opinions) that finally led to one of the persistent opinions strongly depends on μ. For $\mu = \frac{1}{2}$ agents holding initial opinions in regions between two persistent ones could end up in either one of the groups, while for smaller values (e.g. $\mu = \frac{1}{20}$) almost every individual joined the group, whose final opinion was closest to its initial opinion value. So in a sense, the parameter μ determines how conservative the individuals are – both in microscopic interactions and overall.

In addition, they simulated the model also for agents occupying the sites of a square lattice (of size 29×29). Here, essentially the same qualitative behavior was found: for $\theta > 0.3$ a large group consensus around $\frac{1}{2}$ with few extrem-valued outliers and no consensus for smaller values of θ. In the fragmentation case, however, the variety of scattered opinions was way bigger than in the setting of homogeneous mixing as groups of individuals holding opinions that are close together could be separated spatially and in that way be prevented from interacting.

In another article published by essentially the same group of authors [28] they added an investigation concerning heterogeneous confidence bounds. Simulating the homogeneous mixing case with 200 individuals, 8 of which have a confidence bound of $\theta = \frac{2}{5}$, all the rest $\theta = \frac{1}{5}$ instead, revealed an interesting combination of the fragmentation and consensus case over the course of time: On the short run clustering depends on the lower confidence bound, on the long run it depends on the higher bound. First two opinion clusters at distance larger than $\frac{1}{5}$ were formed, then the few 'open-minded' agents acted as mediators between the groups which finally led to a global consensus – not at $\frac{1}{2}$ though as distinct confidence bounds cause pairwise interactions in which the mean of opinions is not preserved leading to such a drift. The transition time from one regime to the other depended largely on the proportion of individuals with larger confidence bound.

In addition to that, Deffuant et al. also simulated the model with confidence bounds decreasing in time (which can be seen to describe the realistic process of positions hardening with time). In the simplest fragmentation case this led to major opinion clusters at values of about 0.60 and 0.42 – closer to each other than in the case of constant confidence bounds. Clearly, this arises from the fact that the opinions gather in a convergence movement first before the confidence bounds become to small and they split.

A different approach to the original model with fully mixed population, i.e. everybody interacts with everybody else, was pursued by Ben-Naim et al. [2]. Using the *mean field approximation* – which is called *thermodynamical limit* in statistical physics – they considered not the agent based model, but (assuming that the number of individuals is large) a density based model, in which $P(x,t)$ describes the density of agents at opinion x at time t. If μ is fixed to be $\frac{1}{2}$, the following rate equation arises:

$$\frac{\partial}{\partial t} P(x,t) = \iint\limits_{|x_1 - x_2| \leq \theta} P(x_1,t)\, P(x_2,t) \left[\delta(x - \tfrac{x_1+x_2}{2}) - \delta(x - x_1)\right] \mathrm{d}x_1\, \mathrm{d}x_2,$$

where $\delta(.)$ denotes the Dirac delta function. For i.i.d. unif$([0,1])$ initial opinions and $\theta > \frac{1}{2}$ they showed that the density converges to a delta function at the initial mean $\frac{1}{2}$. For $\theta < \frac{1}{2}$, however, the rate equation is no longer analytically solvable. By numerically solving it, Ben-Naim and his co-workers discovered interesting facts about the persistent opinion clusters in the case of fragmentation. In this regime the density converges to a finite weighted sum of delta functions, i.e.

$$P(\infty, x) = \sum_{i=1}^{r} m_i \, \delta(x - x_i),$$

where $x_1, \ldots, x_r$ are the persistent opinions (at pairwise distance larger than θ) and $m_i,\ 1 \le i \le r$, the masses of (that is the fraction of agents ending up in) the corresponding clusters. The conservation laws (for mass and mean) obviously force

$$\sum_{i=1}^{r} m_i = 1 \quad \text{and} \quad \sum_{i=1}^{r} m_i \, x_i = \tfrac{1}{2}.$$

Furthermore, they found that there occur three types of persistent opinion clusters: major (mass $> \theta$), minor (mass $< \frac{\theta}{100}$) and a central cluster located at opinion value $\frac{1}{2}$. All of them are generated (and the central cluster annihilated) by a periodic sequence of bifurcations as θ is decreased. Actually they considered $\theta = 1$ to be fixed, the initial opinions instead to be i.i.d. unif$([-\Delta, \Delta])$ with variable Δ, but a simple rescaling translates their results to the original model.

Laguna et al. [15] discovered another feature of the long-term behavior in the Deffuant model with homogeneous mixing which is governed by the convergence parameter μ: The fraction of agents that end up in the two most extreme opinion clusters (which Ben-Naim et al. already showed to be minor but of larger order compared to the other minor clusters) is scaling with μ. For $\theta < \frac{1}{2}$ and larger values of μ, the formation of central opinions is fast enough to seclude many agents holding extreme initial opinions from the unification process. If μ is comparably small, however, those extremists have enough time to become more moderate in order to be included in one of the major opinion clusters later on. In this sense, even if it may seem counterintuitive, for $\theta < \frac{1}{2}$ consensus formation in the population actually benefits from a slower pace in the dynamics.

Stauffer and Meyer-Ortmanns [23] were among the first ones to follow up on the idea by Deffuant et al. to consider the model with an interaction topology other than the homogeneous mixing. They used random graphs generated by the Barabási-Albert model as underlying network – the usual undirected version (described above in Section 2.1) as well as a directed one. Their computer simulations suggest that the transition from fragmentation to consensus happens for the value of θ being about 0.4 (on both the directed and undirected network). Unlike the case of a fully mixed population, the number of persistent opinions in the non-consensus case not only depends on θ but also on n, the number of individuals (for the same reason as in the case of a square lattice). They estimated the dependence of the number of clusters on n (with θ fixed) to be linear.

In 2004, Fortunato [7] finally investigated the threshold for *complete consensus* among the agents – as opposed to previous notions of consensus describing the formation of a widely adapted main stream opinion neglecting single outliers. He simulated the Deffuant model on a complete graph, a square lattice with perodic boundary conditions as well as two random graphs – those originating from the Barabási-Albert and the Erdős-Rényi model. In the latter, a graph on n vertices is built by including each of the $\binom{n}{2}$ possible edges independently with probability p. As this leads to an average degree of $(n-1)p$, Fortunato chose to adapt the parameter p in order to keep np constant for different values of n.

He made two central observations: Firstly, the critical value for θ above which a complete consensus is formed is in all four social topologies $\frac{1}{2}$, irrespectively of μ. Secondly, on each of the four networks the probability of complete consensus converges to a step function at $\theta = \frac{1}{2}$ if the number of individuals is increased. However, it has to be mentioned that he performed update steps as ordered sweeps over the population: In each round every individual gets – one after the other – the opportunity to interact with a randomly selected neighbor. This is different from the original update rule where the edge along which the next potential interaction takes place is picked uniformly at random.

The first result for the Deffuant model on an infinite graph was published by Lanchier [16] in 2011. Using quite intricate geometric arguments he proved the following result for the standard Deffuant model on the infinite line graph:

Theorem 2.1. *Consider the Deffuant model on the graph $G = (V, E)$, where $V = \mathbb{Z}$ and $E = \{\langle v, v+1\rangle,\ v \in \mathbb{Z}\}$. If $\mu \in (0, \frac{1}{2}]$ is arbitrary but fixed, the initial opinions are i.i.d.* unif$([0,1])$ *and $\{\eta_t(v)\}_{v\in\mathbb{Z}}$ denotes the opinion profile at time t, then the following holds:*

(i) For $\theta > \frac{1}{2}$, all neighbors are eventually compatible in the sense that for all $\epsilon > 0,\ v \in \mathbb{Z}$:

$$\lim_{t\to\infty} \mathbb{P}(|\eta_t(v) - \eta_t(v+1)| \le \epsilon) = 1.$$

(ii) For $\theta < \frac{1}{2}$, with probability 1 there will be infinitely many $v \in \mathbb{Z}$ with

$$\lim_{t\to\infty} |\eta_t(v) - \eta_t(v+1)| > \theta.$$

Häggström [12] used different techniques to reprove and slightly sharpen this result – showing that in the consensus regime (i), all opinions actually converge to the mean $\frac{1}{2}$ of the initial distribution. The crucial idea in his proof resides in the connection of the Deffuant dynamics to a non-random interaction process on $\mathbb{Z}$ which he dubbed *Sharing a drink* procedure.

This idea could in fact be employed to tackle initial opinion distributions other than unif$([0,1])$ as was done in Paper A (see below) and by Shang [21] simultaneously and independently.

2.4 Cultural dynamics

In parallel to the advances in the field of opinion dynamics, a growing interest in the natural extension to vector-valued opinions arose. Axelrod [1] was one of the first who published an article focussed on higher-dimensional opinions as opposed to earlier publications considering opinions to be scalar variables. He coined the notion of *cultural dynamics* interpreting the opinion vector as 'culture' of an individual, comprising "the set of individual attributes that are subject to social influence". The border between opinion and cultural dynamics is not sharp and many similarities exist. However, there are models featuring multidimensional opinions that do not have counterparts in opinion dynamics and are therefore qualitatively different from the ones we have presented so far.

(a) Axelrod model

The model proposed in [1] was actually the first one introducing the concept of bounded confidence. However here, the probability of interaction decays gradually with respect to the distance of the two opinions involved:

Think of the individuals again as nodes of a network, each endowed with an opinion vector in $\{1, 2, \ldots, q\}^d$, where each coordinate is understood to represent one of d cultural features and q is the number of possible traits per feature. In that sense, the opinion vector $\boldsymbol{\eta}(i) = (\eta_1(i), \ldots, \eta_d(i))$ is modelling the current beliefs and attitudes of agent i with respect to d interrelated topics.

In an elementary step of the dynamics an individual i and a neighboring one, say j, are randomly selected and interact with probability

$$p_{i,j} = \frac{1}{d} \sum_{k=1}^{d} \mathbb{1}_{\{\eta_k(i)=\eta_k(j)\}},$$

which is scaling with the number of shared attitudes. If they interact, one of the features in which they disagree (i.e. k such that $\eta_k(i) \neq \eta_k(j)$) is chosen uniformly at random and the neighbor j assumed to be convinced by the arguments of i, in other words $\eta_k(j)$ is set to be equal to $\eta_k(i)$.

This model became quite popular among social scientists for the fact that it includes two principles that were found to be typical in cultural assimilation: *social influence*, i.e. interacting makes people more alike, and *homophily* – humans tend to interact more frequently with others that share essential beliefs, attitudes and behaviors. Obviously, there are two kinds of absorbing states: Either all opinions are the same (consensus) or any two neighboring opinions do not share one single trait (fragmentation).

Following the initial paper of Axelrod [1] – who focussed on i.i.d. initial opinion vectors being uniform on $\{1, 2, \ldots, q\}^d$ and finite square lattices as network – several analyses based on numerical simulations have been performed and show that the value of q determines whether the final state reached will be consensus or fragmentation, for different networks and initial distributions.

In the original Axelrod model, the actual values of the coordinates are mere labels: It does not make a difference if two neighbors have traits that differ by 1 or $q-1$. In [5] a more metric variant of the model has been considered in the sense that the interaction probability is changed to

$$p_{i,j} = \frac{1}{d}\sum_{k=1}^{d}\Big(1 - \frac{|\eta_k(i) - \eta_k(j)|}{q-1}\Big).$$

Yet another variant of the Axelrod model was suggested in the paper by Deffuant et al. [6] as a multidimensional counterpart of the Deffuant model: They considered the traits to be binary variables (corresponding to $q = 2$ above) and neighbors interact only if the number of features they disagree on does not exceed a given threshold. So the interaction probability becomes a step function at some given confidence bound. Also the interaction itself was defined slightly different, since once the random feature i and j disagree on is selected, j is not convinced of $\eta_k(i)$ by default but adapts with probability $\mu \in (0, \frac{1}{2}]$.

(b) Multivariate Deffuant and Hegselmann-Krause model

The models introduced by Deffuant et al. as well as Hegselmann and Krause for opinion dynamics, as described in Section 2.2, can be transferred to vector-valued opinions without any further changes – only the notion of distance has to be specified in order to determine the confidence ranges around a given opinion vector. The vectorial version of both models was studied for instance in [8] – on the complete graph with opinion vectors from the unit square $[0, 1]^2$, using both square and circular confidence ranges. The generalization of the Deffuant model on the line graph $\mathbb{Z}$ to higher-dimensional opinion spaces using the Euclidean as well as other metrics as notions of distance is the object of investigation in Paper B (see below).

3

Summary of appended papers

Paper A: Further results on consensus formation in the Deffuant model

(co-authored with Olle Häggström)

The contribution of this paper to the analysis of long-time behavior in the Deffuant model on infinite graphs can be broken down into three parts. The first one – as alluded to in Section 2.3 – is the extension of the statement from Theorem 2.1 to more general initial distributions. It turns out that for i.i.d. initial opinions in the model on $\mathbb{Z}$ being distributed according to the law $\mathcal{L}(\eta_0)$, there exists a critical value θ_c (which can be infinite) for the parameter θ that marks a sharp transition from fragmentation (different persistent opinions) to complete consensus (all opinions converge almost surely to $\mathbb{E}\,\eta_0$). This holds under the weak assumption that not both $\mathbb{E}\,\eta_0^+$ and $\mathbb{E}\,\eta_0^-$ are infinite. The value of θ_c depends on the distribution $\mathcal{L}(\eta_0)$ only, more precisely on the radius of and gaps in its support. We point out, how these results still hold for special ergodic sequences of initial opinions that are not necessarily i.i.d.

In the second part, the model is considered on higher-dimensional integer lattices $\mathbb{Z}^d,\ d \geq 2$. Although the idea of proof from dimension one does not transfer to higher dimensions, elaborating some of its arguments allows to prove at least a partial result, namely that if the marginal distribution of the i.i.d. initial opinions is bounded and θ large enough (strictly larger than $\frac{3}{4}$ in the case of initial opinions that are uniform on $[0,1]$), the opinion of any agent will still a.s. converge to the mean of the initial distribution. In addition to this, on the one hand we show that the opinions converge in distribution for any θ and on the other hand discuss a generalization to transitive, amenable graphs.

In the last part, we consider the Deffuant model on the infinite cluster of supercritical i.i.d. bond percolation on $\mathbb{Z}^d,\ d \geq 2$. In this setting one can retrieve the results derived for the full grid and on top of that we were able to show that for small values of θ the opinions of the agents belonging to the infinite cluster cannot converge to one fixed value. Neighboring individuals could, however, still become finally compatible without their opinions converging to a deterministic limit.

Paper B: The Deffuant model on $\mathbb{Z}$ with higher-dimensional opinion spaces

This paper deals with the generalization of the Deffuant model on $\mathbb{Z}$ to vector-valued opinions – as mentioned in Section 2.4. First we generalize the findings for univariate opinions from Paper A to multivariate opinions – which however requires a more involved reasoning based on geometric arguments – taking the Euclidean norm as natural replacement for the absolut value (which was taken to measure the distance between two opinions in the case of real-valued opinions). In the course of this we gather information about the support of the opinion distribution $\mathcal{L}(\eta_t)$ for times $t > 0$, based on the properties of the initial distribution. Especially the notion of a gap in the support of $\mathcal{L}(\eta_0)$ has to be properly defined and analyzed in higher dimensions in order to play the same role as for univariate distributions.

In the second part, we allow for more general metrics ρ to be employed as measures of distance – determining if the opinions of two agents are close enough for them to interact. We are able to transfer the results from the Euclidean setting if ρ satisfies appropriate extra conditions. By considering several examples we show the necessity of those additional assumptions.

References

[1] AXELROD, R., *The dissemination of culture: A model with local convergence and global polarization,* The Journal of Conflict Resolution, Vol. 41 (2), pp. 203-226, 1997.

[2] BEN-NAIM, E., KRAPIVSKY, P.L. and REDNER, S., *Bifurcations and patterns in compromise processes,* Physica D: nonlinear phenomena, Vol. 183, pp. 190-204, 2003.

[3] CASTELLANO, C., FORTUNATO, S. and LORETO, V., *Statistical physics of social dynamics,* Reviews of Modern Physics, Vol. 81, pp. 591-646, 2009.

[4] CLIFFORD, P. and SUDBURY, A., *A model for spatial conflict,* Biometrika, Vol. 60 (3), pp. 581-588, 1973.

[5] DE SANCTIS, L. and GALLA, T., *Effects of noise and confidence thresholds in nominal and metric Axelrod dynamics of social influence,* Physical Review E: statistical, nonlinear and soft matter physics, Vol. 79 (4), 2009.

[6] DEFFUANT, G., NEAU, D., AMBLARD, F. and WEISBUCH, G., *Mixing beliefs among interacting agents,* Advances in Complex Systems, Vol. 3, pp. 87-98, 2000.

[7] FORTUNATO, S., *Universality of the threshold for complete consensus for the opinion dynamics of Deffuant et al.,* International Journal of Modern Physics C – Computational Physics and Physical Computation, Vol. 15 (9), pp. 1301-1307, 2004.

[8] FORTUNATO, S., LATORA, V., PLUCHINO, A. and RAPISARDA, A., *Vector opinion dynamics in a bounded confidence consensus model,* International Journal of Modern Physics C – Computational Physics and Physical Computation, Vol. 16 (10), pp. 1535-1551, 2005.

[9] GALAM, S., *Social paradoxes of majority rule voting and renormalization group,* Journal of Statistical Physics, Vol. 61 (3-4), pp. 943-951, 1990.

[10] GALAM, S., *Minority opinion spreading in random geometry,* The European Physical Journal B - Condensed Matter and Complex Systems, Vol. 25 (4), pp. 403-406, 2002.

[11] GALAM, S., GEFEN, Y. and SHAPIR, Y., *Sociophysics: A new approach of sociological collective behaviour. I. Mean-behaviour description of a strike,* Journal of Mathematical Sociology, Vol. 9, pp. 1-13, 1982.

[12] HÄGGSTRÖM, O., *A pairwise averaging procedure with application to consensus formation in the Deffuant model,* Acta Applicandae Mathematicae, Vol. 119 (1), pp. 185-201, 2012.

[13] HEGSELMANN, R. and KRAUSE, U., *Opinion dynamics and bounded confidence: models, analysis and simulation,* Journal of Artificial Societies and Social Simulation, Vol. 5 (3), 2002.

[14] ISING, E., *Beitrag zur Theorie des Ferromagnetismus,* Zeitschrift für Physik, Vol. 31 (1), pp. 253-258, 1925.

[15] LAGUNA, M.F., ABRAMSON, G. and ZANETTE, D.H., *Minorities in a model for opinion formation,* Complexity, Vol. 9 (4), pp. 31-36, 2004.

[16] LANCHIER, N., *The critical value of the Deffuant model equals one half,* Latin American Journal of Probability and Mathematical Statistics, Vol. 9 (2), pp. 383-402, 2012.

[17] LIGGETT, T.M., "Interacting Particle Systems", Springer, 1985.

[18] LORENZ, J., *Continuous opinion dynamics under bounded confidence: A survey,* International Journal of Modern Physics C – Computational Physics and Physical Computation, Vol. 18 (12), pp. 1819-1838, 2007.

[19] MAJORANA, E., *Il valore delle leggi statistiche nella fisica e nelle scienze sociali,* Scientia, Ser. 4, pp. 58-66, 1942.

[20] MANTEGNA, R.N., *Majorana's article on "The value of statistical laws in physics and social sciences",* Proceedings of Science (EMC2006)011, 2006.

[21] SHANG, Y., *Deffuant model with general opinion distributions: First impression and critical confidence bound,* Complexity, Vol. 19 (2), pp. 38-49, 2013.

[22] SUZUKI, Y., *Invariant measures for the multitype voter model,* Tokyo Journal of Mathematics, Vol. 14 (1), pp. 61-72, 1991.

[23] STAUFFER D. and MEYER-ORTMANNS, H., *Simulation of consensus model of Deffuant et al. on a Barabási-Albert network,* International Journal of Modern

Physics C – Computational Physics and Physical Computation, Vol. 15 (2), pp. 241-246, 2004.

[24] STAUFFER D., SOUSA, A. and SCHULZE, C., *Discretized opinion dynamics of the Deffuant model on scale-free networks,* Journal of Artificial Societies and Social Simulation, Vol. 7 (3), 2004.

[25] SZNAJD-WERON, K. and SZNAJD, J., *Opinion evolution in closed community,* International Journal of Modern Physics C – Computational Physics and Physical Computation, Vol. 11 (6), pp. 1157-1165, 2000.

[26] VAZQUEZ, F., KRAPIVSKY, P.L. and REDNER, S., *Constrained opinion dynamics: freezing and slow evolution,* Journal of Physics A: Mathematical and General, Vol. 36 (3), pp. L61-L68, 2003.

[27] WEIDLICH, W., *The statistical description of polarization phenomena in society,* British Journal of Mathematical and Statistical Psychology, Vol. 24 (2), pp. 251-266, 1971.

[28] WEISBUCH, G., DEFFUANT, G., AMBLARD, F. and NADAL, J.-P., *Meet, discuss, and segregate!,* Complexity, Vol. 7 (3), pp. 55-63, 2002.