

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/27336148>

Repeated Averaging and Bounded Confidence Modeling, Analysis and Simulation of Continuous Opinion Dynamics

Article · January 2007

Source: OAI

CITATIONS

37

READS

242

1 author:



Jan Lorenz

Jacobs University

77 PUBLICATIONS 3,162 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



BIGSSS Summer Schools in Computational Social Science [View project](#)



Continuous opinion dynamics and collective decisions [View project](#)

Repeated Averaging and Bounded Confidence

Modeling, Analysis and Simulation
of Continuous Opinion Dynamics

von Jan Lorenz

post@janlo.de

Dissertation

zur Erlangung des Grades eines Doktors der
Naturwissenschaften
– Dr. rer. nat. –

Vorgelegt im Fachbereich 3 (Mathematik & Informatik)
der Universität Bremen
im Februar 2007

Datum des Promotionskolloquiums: 5. März 2007

1. Gutachter: Prof. Dr. Dr. Ulrich Krause (Universität Bremen, Mathematik)
2. Gutachter: Prof Dr. Rainer Hegselmann (Universität Bayreuth, Philosophie)

Weitere Mitglieder des Prüfungsausschusses:

- Prof. Dr. Diederich Hinrichsen (Universität Bremen, Mathematik)
Prof. Dr. Peter Maaß (Universität Bremen, Mathematik)
Dr. Jan Nagler (Universität Bremen, Physik)
Dipl. Mathematikerin Sabine Pfarr (Studentin Universität Bremen, Psychologie)

Acknowledgments. I thank my advisor Ulrich Krause the chance of a very free dissertation project but also for giving me guidelines, inspiration and support whenever necessary. I thank the Friedrich-Ebert-Foundation for funding my doctoral studies and for providing the opportunity to take part in the German-Brazil dialogue project 2005. I thank Diemo Urbig and Dirk A. Lorenz for fruitful and funny work together. Some parts of this thesis and dissertation are only due to the joint work with each of them. I thank Frank Schweitzer, Dirk Helbing, Christophe Deissenberg and their helpers for organizing several events in the community of Complex Systems and Physics of Socio-Economic Systems where I could participate during the work on this thesis, especially for their organization of events and awards for young researchers. I thank Ulrich Hörning and Fabian Scholtes for organizing the FES German-Brazil dialogue project 2005 and all participants. It was a great opportunity to bring parts of this work for the first time to the political debate. In this context I especially thank Frederico Menino and Carsten Herzberg for good collaboration. I thank Annegret Piecha, Sally Welsch, Hanna Logemann, Kristin and Benjamin Haase for collaboration in an experiment of social psychology. For discussions, immediate help, and other scientific support I thank Rainer Hegselmann, the members of the ‘Seminar Positive Dynamische Systems’, Iain D. Couzin, Jan Nagler, Karsten Wenzlaff, Roger Waldeck, Tadeusz Płatkowski and Darald Hartfiel. I am also thankful for a good mathematical education at the University of Bremen and to all people who worked for the free supply with such extraordinary tool as L^AT_EX, the beamer.cls, arXiv.org or Wikipedia.

I thank my parents for their support in the last months and most of all Rebekka for enduring in the last months and keeping me and our baby alive.

Jan Lorenz
www.janlo.de

Abstract. This thesis is about dynamical systems of agents which perform repeated averaging under bounded confidence. It contributes to questions of their modeling, mathematical analysis and simulation.

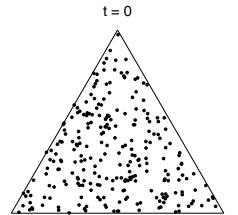
Modeling. The main modeling issue is continuous opinion dynamics. This includes dynamics of agents in a political opinion space as well as dynamics of collective motion in swarms of mobile autonomous robots. The existing bounded confidence models of Hegselmann-Krause and Deffuant-Weisbuch are presented under a common general framework. For both models density-based versions are introduced which give additional options for analysis. Surprising phenomena are presented, e.g. fostering consensus by lowering confidence or by introducing heterogeneity.

Mathematics. Conditions for convergence to consensus are derived for systems where dynamics is driven by very generally defined averaging maps. If averaging maps are linear then the fundamental part of the system is a product of row-stochastic matrices infinite two the left. Several conditions, examples and counter-examples for convergence of such products are given. Thereby, we extend a result about convergence to consensus: Intercommunication intervals need not be bounded but may grow very very slow. Finally, the sets of fixed points are characterized for the bounded confidence models.

Simulation. Interesting candidates for this class of systems are the models of opinion dynamics under bounded confidence, which are then analyzed via computer simulation. The density-based modeling approach is used to get a systematic overview for the case when agents' initial opinions are uniformly distributed in the opinion space. Dynamics always converges to situations with clustered opinions. Bifurcation diagrams for attractive clustered states are computed for homogeneous bounds of confidence as well as extended phase diagrams for the consensus transitions in populations with two different levels of confidence.

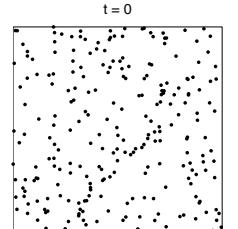
The thesis concludes with some advices on how to foster consensus in continuous opinion dynamics under bounded confidence and a list of open problems.

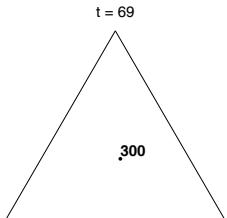
Keywords: social simulation, agent-based, general mean, density-based, interactive Markov chain, discrete master equation, convergence to consensus, infinite matrix products, nonnegative matrices, row-stochastic, scrambling, Gantmacher form, coefficient of ergodicity, averaging map, equiproper, sociophysics, bifurcation diagrams, extended phase diagrams, social psychology



Contents

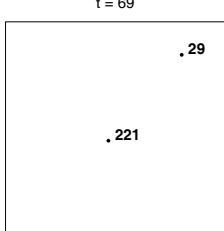
1	Introduction	7
	Mathematical Notations	11
2	Modeling	13
2.1	Opinion dynamics in the real world	14
2.2	An experiment of social psychology	16
2.3	Modeling vocabulary	19
2.3.1	Agents and densities of agents	19
2.3.2	The state spaces	19
2.3.3	Dynamics	21
2.3.4	Bounded confidence	22
2.3.5	Repeated averaging	23
2.3.6	Matrices and networks	26
2.4	The mathematical models	28
2.4.1	Repeated pooling of opinions	28
2.4.2	Agent-based bounded confidence models	29
2.4.3	Agent-based examples and phenomena	31
2.4.4	Density-based bounded confidence models	36
2.4.5	Convergence for large numbers of agents and classes	42
2.4.6	Density-based examples and phenomena	42
2.4.7	Continuous opinion dynamics and swarms	50
2.5	Some related models and some references	54
3	Mathematical Analysis	55
3.1	Averaging maps	55
3.1.1	The homogeneous case	55
3.1.2	The inhomogeneous case	56
3.1.3	Comparison with Moreau's theorem	62
3.2	Matrix-based analysis	64
3.2.1	The Gantmacher form for nonnegative matrices	65
3.2.2	Powers of a row-stochastic matrix	71
3.2.3	Homogeneous Markov processes and processes of opinion pooling	75
3.2.4	Infinite products of row-stochastic matrices	77
3.2.5	Convergence to consensus matrices	78
3.2.6	Convergence of the zero pattern	84
3.2.7	Convergence	89
3.2.8	Trying to apply the joint spectral radius	93

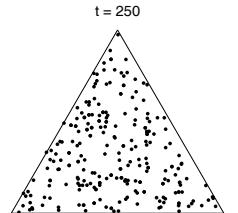




CONTENTS

3.2.9	Inhomogeneous processes of opinion pooling	95
3.3	Agent-based bounded confidence models	97
3.3.1	Convergence	98
3.3.2	Mean conservation in homogeneous gossip processes . . .	99
3.3.3	Scaling and translation invariance	99
3.3.4	The impact of the communication regime	99
3.3.5	The set of fixed points	100
3.4	Density-based bounded confidence models	102
3.4.1	The Deffuant-Weisbuch model	102
3.4.2	The Hegselmann-Krause model	104
3.5	References and relations	109
4	Simulation	111
4.1	Stabilization and clustering in time evolution	111
4.2	Bifurcation of clusters in the evolution of the bound of confidence	113
4.2.1	Aggregated data of multiple agent-based simulation runs .	114
4.2.2	The basis of bifurcation diagrams	114
4.2.3	Bifurcations in the DW model	116
4.2.4	Bifurcations in the HK model	119
4.2.5	Comparison between the DW and the HK model	122
4.2.6	Is there a universal scale for transition points?	123
4.3	Heterogeneous bounds of confidence	126
4.3.1	Extended phase diagram of the heterogeneous HK model	127
4.3.2	Extended phase diagram of the heterogeneous DW model	129
4.4	Results in accompanying papers	133
4.4.1	Multidimensional opinions	133
4.4.2	Communication strategies	134
4.4.3	Theoretic questions	134
5	Conclusion	135
5.1	How to foster or dilute consensus	135
5.2	Achievements	136
5.3	Open problems	137
Bibliography		139
Index		145





Chapter 1

Introduction

This thesis regards the modeling, the mathematical analysis, and the simulation of dynamical systems under the concepts of repeated averaging and bounded confidence. It takes *continuous opinion dynamics* as its main application.

In opinion dynamics we consider agents that may change their opinions if they hear the opinions of others. We consider opinion issues which can be expressed in terms of real numbers such that compromising in the middle is possible. Therefore, ‘continuous’ relates to the issue not to the time. Examples are

- experts discussing an issue for which they have imperfect information,
- voters discussing about their desired policy from left wing to right wing,
- customers that discuss prices,
- animals that try to move in a common direction in a swarm, flock, or herd,
- computers which communicate and process variables in a setting of decentralized computing.

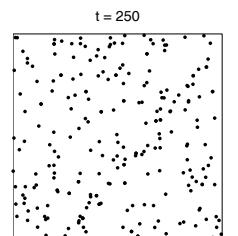
We study processes where agents adjust their own opinion by building an average of opinions of agents whom they trust. If this adjusting is repeated we call this *repeated averaging*.

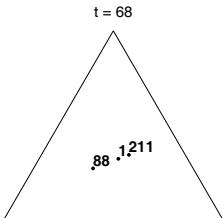
Agents may change their trust in the competence of other agents during the process. If they trust only agents which have opinions close to their own opinion we say that these agents have *bounded confidence*. Under this assumption the trust structure changes dynamically with the opinions.

In modeling repeated averaging and bounded confidence we draw on the model of Hegselmann and Krause [37, 46] and the model of Deffuant, Weisbuch and others [19, 88]. They serve as the main examples for agent-based models and deliver the heuristics for the definition of density-based models.

The thesis is divided into three main chapters about modeling, mathematical analysis and simulation with different focus.

Modeling means the transfer of a dynamical system observed in reality to a mathematical model which can be simulated and analyzed to deepen the understanding of the system. The kind of desired understanding can be diverse such as prediction, finding hidden driving forces or determining phases of attractive states in a qualitative manner. The same real system can be modeled in various ways depending on the desired kind of





1. Introduction

understanding. Especially we distinguish agent-based and density-based modeling.

Mathematical Analysis means deriving theorem which determine the behavior of the system – e.g. convergence, the set of fixed points and attractive states. The main problems here are, non-linearity, non-continuity and sensitivity to the initial conditions. An additional mathematical aim is to use models and simulations to get new ideas for deeper knowledge in pure mathematical fields. The fields regarded in this work are nonlinear difference equations, inhomogeneous infinite matrix products, nonnegative matrices and positive systems.

Simulation means the programming of the modeled system such that example runs and systematic simulations are possible. While the example runs can help to find interesting phenomena systematic simulations can give hints about the general behavior and attractive states of the system, which are not easy accessible with pure mathematics until now. The main problems here are a huge parameter space, numerical problems, and long computation times. Regarding the simulation part, this thesis is also a contribution to the interdisciplinary fields sociophysics and social simulation.

Our modeling framework is a discrete dynamical systems. Consider a *state space* M and a *self map* $f : M \times \mathbb{N} \rightarrow M$ such that the *process* of the discrete dynamical system is recursively defined for an initial state $x(0) \in M$ through

$$x(t+1) = f_t(x(t)).$$

The self-maps in this thesis are defined such that they model dynamics of repeated averaging and bounded confidence.

In an *agent-based* modeling approach we have agents $\{1, \dots, n\}$ which have opinions $x^1, \dots, x^n \in S$ with $S \subset \mathbb{R}^d$ being an appropriate continuous opinion space. The state space is thus $S^n \subset (\mathbb{R}^d)^n$.

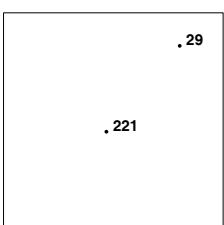
A core result of this thesis is the derivation of a theorem which ensures convergence to consensus for the case where all functions f_t are equicontinuous proper averaging maps but possibly nonlinear.

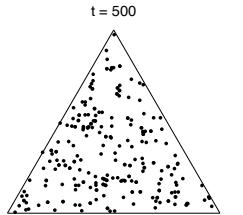
In a *density based model* the state space is the set of probability density functions on the opinion space S . So, the agents are only virtual and represented by their density in the opinion space. We will model density based systems only with a discretized opinion space. Then the probability density function reduces to a stochastic vector and the density-based model is an interactive Markov chain discrete in time and discrete and finite in state. It is called ‘interactive’ because the transition probabilities change in the process with respect to the actual state.

If the self-map f_t is determined by the matrix $A(x(t), t)$, then a process of repeated pooling of opinions is determined for an initial state $x(0) \in S^n$ by

$$x(t+1) = A(x(t), t)x(t).$$

The matrix must be row-stochastic to produce averaging dynamics. Now conditions on the matrices $A(x(t), t)$ are of interest which drive the opinion dynamics process to stabilization or even to consensus. We will collect known results and





derive some new ones. The bounded confidence models of Hegselmann-Krause and Deffuant-Weisbuch fit into this framework.

An interactive Markov chain is determined for an initial state $p(0)$ (as a row vector) by

$$p(t+1) = p(t)B(p(t))$$

with $B(p(t))$ being the transition matrix. We define interactive Markov chains that follow the communication heuristics of the Hegselmann-Krause model and the Deffuant-Weisbuch model. We are then able to derive the sets of fixed points for these density-based systems. Thereby evolves a pretty good analogy to master equation approaches on agent-systems inspired by systems of particles in statistical physics.

While mathematical analysis aims at a full understanding of models it is sometimes slow in delivering results. So in this thesis the simulation is on a par with mathematical analysis.

We take the interactive Markov chain as the dominating equation for agent-based systems with enough agents and regard the initial opinions to be uniformly distributed in the opinion space. Under this conditions we derive bifurcation diagrams of the final clustering with respect to the bound of confidence by simulation.

Further simulations are made for heterogeneous bounds of confidence which lead to extended phase diagrams which give an overview about the possibilities of reaching consensus.

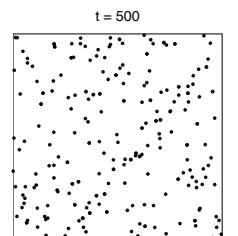
Some results – especially simulation results in agent-based models – are not part of this thesis but part of the dissertation. The following four publications accompany this dissertation thesis:

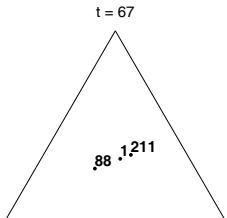
1. Opinion Dynamics of Multidimensional Allocation Problems under Bounded Confidence: More Dimensions Lead to Better Chances for Consensus. To appear in *European Journal of Economic and Social Systems*, 2007. [57]
2. Fostering Consensus in Multidimensional Continuous Opinion Dynamics under bounded Confidence. To appear as chapter in *Managing Complexity*, Editor Dirk Helbing, Springer, 2007. [58]
3. Jan Lorenz and Diemo Urbig. About the Power to Enforce and Prevent Consensus. To appear in *Advances in Complex Systems*, 2007. [61]
4. Consensus Strikes Back in the Hegselmann-Krause Model of Continuous Opinion Dynamics under Bounded Confidence. *Journal of Artificial Societies and Social Simulation*, 9(1), 2006. [55]

They are all accepted and have already undergone their final revision if not published already. In the last publication – which is an online publication – are movies which show some dynamics.

Finally, we derive some advices how an opinion dynamics process guided by repeated averaging and bounded confidence may be designed to foster the reaching of consensus.

We finish the introduction with some guidelines for the reader: In most cases numbered definitions are omitted to enable a quick reading. In exchange the text is quite strict with the convention that terms written in *italics* are defined and indexed. So, the index at the end can serve as a glossary for definitions.

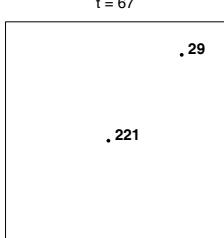


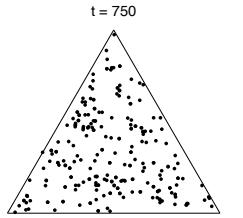


1. Introduction

Basic definitions (but perhaps sometimes nonstandard) are summarized in a preliminary section on mathematical notations.

There are plenty of figures in this thesis. All numbered figures are also referenced in the text. So, if one gets attracted by a figure it is not worthless to search for the part in the main text where additional information and the broader context is given. But there are also sometimes marginal figures besides the main text. These figures are never referenced or mentioned in the text but they may give hints for understanding.





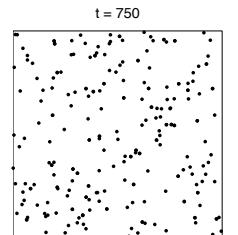
Mathematical Notations

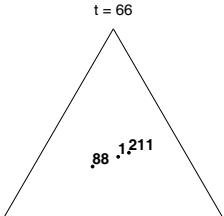
General. \mathbb{N} are the natural numbers including zero. $\mathbb{R}_{\geq 0}, \mathbb{R}_{>0}$ are the non-negative respectively positive reals. For $n \in \mathbb{N}$ it is $\underline{n} := \{1, \dots, n\}$. For a set I there is $\#I$ the number of elements of I . The brackets $\lfloor \cdot \rfloor$ ($\lceil \cdot \rceil$) mean rounding to the lower (upper) integer. For $x, y \in \mathbb{R}$ with $x \leq y$ we define intervals $[x, y],]x, y[, [x, y[,$ and $]x, y]$ with the outward bracket defining an open bound and the inward bracket a closed bound. A *discrete interval* is a subsequent sequence of integers. For bounding integers $i, j \in \mathbb{Z}$ with $i \leq j$ we denote $\{i, \dots, j\}$. A set is *relatively compact* if its closure is compact. If the set is a subset of \mathbb{R}^d then relatively compact is equivalent to bounded. The p -norm is defined as usual. For a set $M \subset \mathbb{R}^d$ the diameter of M is defined $\text{diam } M := \sup\{\|x - y\|_2 \mid x, y \in S\}$. For two sets $S, P \subset \mathbb{R}^d$ the addition $S + P$ is defined as $\{x + y \mid x \in S, y \in P\}$.

Entries, indices, subvectors and submatrices. We denote the components of vectors of vectors $x \in (\mathbb{R}^d)^n$ as $x^i \in \mathbb{R}^d$ and $x_j^i \in \mathbb{R}$. Let $A \in \mathbb{R}^{m \times n}$. If the context is simple we denote entry ij by a_{ij} . For more sophisticated terms we may switch to $A_{ij}()$. The notation $(\cdot)_{[i,j]}$ is used to denote an entry in a longer matrix term. $A_{[i,:]}$ denotes the i th row and $A_{[:,j]}$ the j th column of A . For distinct indices $i_1, \dots, i_k \in \underline{m}$ and $j_1, \dots, j_l \in \underline{n}$ we denote by $A_{[i_1, \dots, i_k | j_1, \dots, j_l]} = B$ a $k \times l$ matrix with $B_{[a,b]} = A_{[i_a, j_b]}$. If $\mathcal{I} = (i_1, \dots, i_k) \in \underline{n}^k$ is a tuple of distinct indices then $A_{[\mathcal{I}, :]}$ is a $n \times k$ matrix with the columns of A through indicated by \mathcal{I} . $A_{[:, \mathcal{J}]}$ and $A_{[\mathcal{I}, \mathcal{J}]}$ analog. If $\mathcal{I} = \{i_1, \dots, i_k\} \in \underline{n}^k$ is a set of distinct indices then the definition should also hold with the implicit assumption that $i_1 < \dots < i_k$. The same holds for vectors.

Functions and operations on matrices and vectors. Generally, functions and operations should be applied componentwise unless otherwise state. Explicit examples follow. For $c \in \mathbb{R}$ and $A \in \mathbb{R}^{m \times n}$ we define $A < c \Leftrightarrow a_{ij} < c$ for all i, j . For $B \in \mathbb{R}^{m \times n}$ we define $A < B \Leftrightarrow a_{ij} < b_{ij}$ for all i, j . For $>, \leq, \geq, =$ analog. For $x, y \in \mathbb{R}^d$ with $x \leq y$ we define a *d-dimensional interval* as $[x, y] := \{z \in \mathbb{R}^d \mid x \leq z \leq y\}$. Let $S \subset \mathbb{R}^d$ be a set of vectors. Then we define the infimum of S componentwise $\inf S := (\inf\{x_1 \mid x \in S\}, \dots, \inf\{x_d \mid x \in S\})$. For supremum $\sup S$, minimum $\min S$, and maximum $\max S$ analog. Notice, that in this case the minimum and the maximum need not be elements of the set S . Convergence of matrices and vectors is meant to be componentwise.

Stochasticity and convexity. A vector $p \in \mathbb{R}^n$ is called *stochastic* if it is nonnegative and all components sum up to 1. The set of all n -dimensional stochastic vectors is called the $(n - 1)$ -dimensional *unit simplex* $\Delta^{n-1} := \{p \in \mathbb{R}^n \mid p \geq 0 \text{ and } \sum_{i=1}^n p_i = 1\}$, sometimes just called *simplex*. Let $x_1, \dots, x_k \in \mathbb{R}^d$ and $\alpha^1, \dots, \alpha^k \in \mathbb{R}_{\geq 0}$ with $\sum_{i=1}^k \alpha^i = 1$ be *convex coefficients* then $\sum_{i=1}^k \alpha^i x_i$ is called a *convex combination* (of length k) of x_1, \dots, x_k . A stochastic vector $p \in \mathbb{R}^n$ defines the coefficients of a convex combination of length n . A set $C \subset \mathbb{R}^d$ is called *convex* if all convex combinations of elements of C are again in C . For a set $M \subset \mathbb{R}^d$ the *convex hull* $\text{conv } M$ is defined as the set of all convex combinations of elements in M . An affine combination is analog to a convex combination but the affine coefficients need not be positive. The *affine hull* $\text{aff } M$





1. Introduction

is the set of all affine combinations. It is the intersection of all affine spaces that cover M . The *relative interior* $\text{ri}M$ of M is the interior of M restricted to its affine hull. So, $x \in M$ is in $\text{ri}M$ if there is an ε -ball $B_\varepsilon(x)$ around x such that $B_\varepsilon(x) \cap \text{aff } M \subset M$.

Matrices. E is the quadratic unit matrix. $\mathbf{1}$ is the vector with all entries equal to one of the appropriate dimension. A^T is the transpose of matrix A . (Attention. T is also rarely used as variable name for an integer.) A *permutation matrix* is a square matrix with all entries zero except for exactly one entry in each row and each column which is one.

Let $p \in \mathbb{R}^n$ be a vector, then Ap and pA is the usual matrix times vector respectively vector times matrix. For pA the vector has to be regarded as a row vector. We will not point this out explicitly if it is clear from the context.

A nonnegative matrix is *row allowable* if each row contains at least one positive entry. A row allowable matrix is *row-stochastic* if each row is a stochastic vector. For a row allowable matrix A we define its *row-stochastified* matrix $\text{sto}(A)$ as $\text{sto}(A)_{[i,:]} := \frac{A_{[i,:]}}{\|A_{[i,:]}\|_1}$ for each row $i \in \underline{n}$.

For a directed graph $G = (V, L)$ with the vertex set $V := \underline{n}$ and a certain link set $L \subset V \times V$ we define its *adjacency matrix* N which is an $n \times n$ $\{0, 1\}$ -matrix which entries are defined $n_{ij} = 1$ if $(i, j) \in L$ and zero otherwise. For an agent $i \in \underline{n}$ and an adjacency matrix $N \in \{0, 1\}^{n \times n}$ we define the *neighbors* of i in N as $\text{nb}(i, N) := \{j \in \underline{n} \mid n_{ij} = 1\}$.

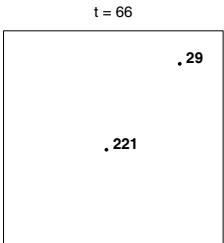
For a sequence of square matrices $(A(t))_{t \in \mathbb{N}}$ and natural numbers $t_0 < t_1$ we denote $A(t_0, t_1) := A(t_0)A(t_0 + 1) \cdots A(t_1 - 2)A(t_1 - 1)$ as the *forward accumulation* from t_0 to t_1 and $A(t_1, t_0) := A(t_1 - 1)A(t_1 - 2) \cdots A(t_0 + 1)A(t_0)$ as the *backward accumulation* from t_0 to t_1 . Consistent with this it holds for $t \in \mathbb{N}$ that $A(t, t + 1) = A(t + 1, t) = A(t)$ and $A(t, t) = E$.

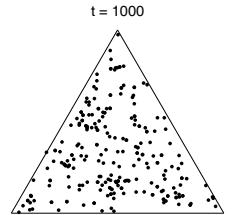
For $A, B \in \mathbb{R}^{m \times n}$ the *Hadamard product* $A \bullet B$ is defined as componentwise multiplication $(A \bullet B)_{ij} = a_{ij}b_{ij}$.

For $x \in (\mathbb{R}^d)^n$ and a $n \times n$ matrix A the product Ax is computed regarding the d -dimensional vectors x^1, \dots, x^n as components of x . A fruitful interpretation is to see x as a matrix $X \in \mathbb{R}^{n \times d}$ with the x^1, \dots, x^n as rows. The computation transforms to the matrix product AX .

A *matrix norm* should be a norm for which submultiplicativity $\|AB\| \leq \|A\|\|B\|$ holds for all A, B . The *spectral radius* of a square matrix A is defined as $\rho(A) := \max\{|\lambda| \mid \lambda \text{ is an eigenvalue of } A\}$. If λ is an eigenvalue of A then $\text{eig}(A, \lambda)$ denotes its eigenspace.

Attention, there are non consistent functional definitions! Once $A(t, x)$ means a matrix function dependent on time t and state x , in another context $A(s, t)$ means the accumulation of matrices form time steps s to t or $A(x, \varepsilon)$ means the confidence matrix derived from opinion profile x for bound of confidence ε . The type is always clear through the context. A help can be that s, t are nearly always variables for the time. And i, j are nearly always variables for agents or indices.





Chapter 2

Modeling

In general a *model* is a simplified image of a system which only contains the elements of the original system which the modeling person regards as relevant. A model is a *mathematical model* when all elements, relations and functions of the model are clearly defined in mathematical terms and are thus accessible for mathematical analysis.

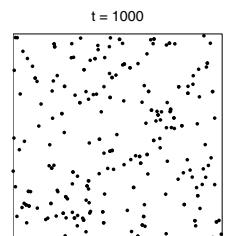
In this chapter three issues are of interest: First, how do we make an image of reality – so, what do we regard as relevant? Second, how this is transformed into mathematical terms and formalisms. And third, does this produce interesting phenomena in simulation? The art of modeling is to hold the balance between displaying the complex nature, making useful simplifications, and keeping mathematical strength with the aim of gaining interesting scientific results.

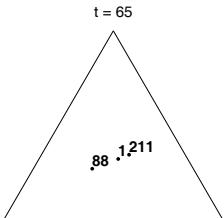
The definition of ‘interesting results’ can be diverse. One can distinguish:

1. Reality and prediction. One tries to model reality as exactly such that prediction is possible when measuring of the state of nature is possible at an acceptable level. The existence of deterministic chaos or self organized complexity makes this often a very hard task with the weather forecast problem as a famous example.
2. Abstraction and qualitative analysis. One tries to simplify a real world system to a simple model which nevertheless shows real world phenomena to find hidden driving forces which are not visible directly by intuition but may play an important role. Further on, robustness, degrees of complexity and universality classes can be of interest.
3. Efficiency. One tries to find rules such that a system selforganizes to a desired state (e.g. consensus) as quick as possible. This is interesting for designing systems.

This thesis follows mainly the second aim in the list. But we also touch the third aim when we are looking for good conditions for reaching consensus.

We use two modeling approaches in this thesis. In the agent-based approach each agent is modeled explicitly while in the density-based approach we model a probability distribution of the agents’ positions in the opinion space. Models can be formulated with both approaches and the same heuristics. The definition of density-based dynamics is in analogy to the formulation of a dominating





2. Modeling

equation in particle physics (also called master equation). The discrete master equations of our systems are derived in Section 3.4 from the interactive Markov chains which we define in Section 2.4.4 in this chapter. From the perspective of the master equation agent-based simulations are a *Monte-Carlo analysis* of the system.

We outline the rest of the chapter. We will give some examples of continuous opinion dynamics and bounded confidence in the real world and some links to social psychology in Sections 2.1 and 2.2. Then, we provide the modeling vocabulary in Section 2.3 – descriptive and mathematical – to ensure that the interdisciplinary reader gets the chance to cope with the results of the mathematical parts and on the other hand to help the mathematical reader to find mathematical strength when we use common words. Further on, we define and discuss in Section 2.4 all mathematical models that will be covered in the math and the simulation chapter. We will also give examples which give a first impression about interesting phenomena.

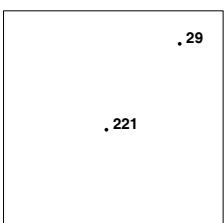
2.1 Opinion dynamics in the real world

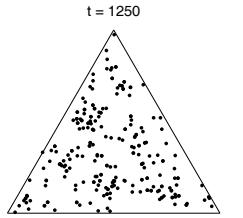
First, we want to distinguish three different problems where continuous opinion dynamics between agents may happen in the real world.

One possible problem is **fact-finding**. A famous example is the weight judging competition on the fat stock and poultry exhibition which Francis Galton attended and analyzed 1907 in [28]. People were invited to estimate the weight of a fat ox. To Galton's surprise it turned out that the average over all estimates was one of the best estimates. Similar competitions have been played over a thousand times on television in the popular game show 'The Price is Right!'. Another example for a fact-finding problem is a commission of experts that should determine macroeconomic indicators. The problem in a group of such experts is that often there is a lack of information. So experts have to estimate. But then they may change their estimates if they hear about the estimates of others because they might suspect that other experts have better information or better intuition.

One level deeper is a **prediction** problem, where it is clear that the value which is to estimate is not measurable at the moment but only in the future. So the estimate is not only difficult because of imperfect information but also because it may be influenced until it is measurable. This can produce feedbacks from the discussion process to the evolution of the fact itself. The information that the economy will be booming might significantly change the behavior of the acting economic agents which may again influence the fact that economy is booming or not. If a widely published prediction comes true it is suspect of having been a self-fulfilling prophecy and if not of having been a self-destroying prophecy. It is obvious, that an estimate which is consensual in a huge group of experts has in many cases a bigger chance to get a self-fulfilling prophecy. So, there is another reason why changing opinions towards a consensus might be attractive for experts in a commission which has to deliver forecasting results. An example for a group of agents with forecasting problems is the commissions for tax revenue estimation. But there are numerous others. A recent study of decision making in committees of experts is Visser and Swank [85].

The first observation of Galton that the average opinion of a huge crowd can





2.1. Opinion dynamics in the real world

be better than most experts has later been called the *Delphi effect* and inspired the development of the Delphi method (initially in the RAND Corporation). The Delphi method is the organization of a process of repeated judgments of experts which communicate their results in rounds – often anonymously. The idea is that such groups may come to better estimates if one controls the process e.g. to avoid groupthink. More details and many empirical work is provided in the book [51]. Nowadays, the idea that experts are dispensable if one has a crowd and the right process or aggregation rule is known under the popular catch word of the ‘wisdom of crowds’ [77] due to ‘swarm intelligence’ [10]. New popular tools to extract the wisdom of crowds are imperfect information markets (see e.g. [12]). Where contracts like ‘pay \$1 if Hillary Clinton will be the next president of the United States, get \$1 if not’ can be traded. Several such markets are now set up e.g. under the names Policy Analysis Market, Iowa Electronic Markets or Hollywood Stock Exchange (to determine the chance of a movie to become a blockbuster). They often perform better than classical prediction with opinion polls.

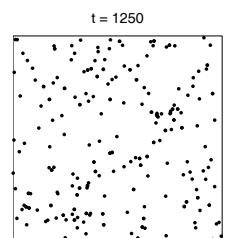
Obviously, the border between fact-finding and prediction is not sharp. Actually, the weighing of the ox is an event in the future which might be manipulated and is thus unsure. On the other hand the weather forecast regards prediction but it is at least questionable if a crowd of people running around with umbrellas does significantly change the probability for rain. But on the other hand again we know that a flap of a butterfly may do.

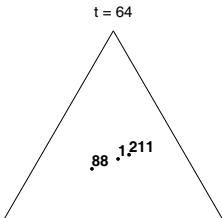
One level deeper again are problems of **negotiation**. The issue in a negotiation problem is of a nature such that there are structurally few arguments for a correct answer. Classical negotiation problems are the bargaining about prices on a market or the negotiations of wages between trade unions and employer associations. In negotiation problems agents are often strongly interested in reaching agreement, because it is a prerequisite for coming into contract. So, consensus can be a good in itself – regardless of the existence of true or optimal solutions. Problems to reach agreement can be different interests of agents or problems in communicating opinions especially if there are many agents. Naturally, also strategic play may evolve due to agents which try to manipulate the outcome, e.g. by starting with more radical opinions as they really have.

Besides monetary examples the whole political debate can also be seen as a huge opinion dynamics process with often continuous opinions – e.g. the determination of tax rates as a practical example or the perceived continuum from left wing to right wing as a more metaphoric example. Opinion dynamics under bounded confidence as it will be presented in this thesis can be seen as an evolution of political parties out of a society of agents with uniformly distributed opinions.

Political opinion dynamics with many agents may get more important in the future because of the spread of participation processes [2, 17, 59] in democratic societies (often supported by new communication technologies). This broadens the political debate and perhaps dilutes a little bit the influence of mass media. Then the question arises if the design of the discussion, participation and decision process has an impact on the chances for consensus.

A quite different example of a negotiation process is the finding of a common direction in a flock of birds, which has also similarities to the problem of keeping track on macro-parameters in environments of distributed computing. Here appropriate consensus-protocols have become increasingly of interest





2. Modeling

[43, 62, 64, 79].

The border between prediction and negotiation may be not sharp, too. An example are agents in a commission of experts mixed with politicians which have the mission to reach a joint estimate. So, their discussions will be a mixture of discussing facts and negotiating what might be an appropriate joint result to sell to the public.

Probably most real world opinion dynamics process are mixtures of fact-finding, prediction and negotiation processes. In each isolated process there are reasonable arguments why agents may revise their opinions towards the opinions of others. So, the formal analysis of continuous opinion dynamics guided by repeated averaging could lead to numerous applications.

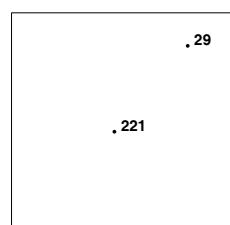
Social psychology distinguishes two main driving forces for change in opinion: informational reasons and normative reasons. Informational reasons rely on the fact that agents regard other agents as competent and therefore adjust their own opinion. Normative reasons arise due to a feeling of group pressure. In many groups it is necessary to have more or less the opinion of the others. Some basic empirical work here goes back to Sherif and Hovland [71] and is known as Social Judgment Theory. But following this theory agents will only adjust towards opinion which lie within their latitude of acceptance. This concept is resembled here under the name ‘bounded confidence’. Additionally there are repulsive forces assumed if agents differ even more in opinion which is neglected in this thesis. Another link from the bounded confidence assumption to social psychology is the Theory of Self-Categorization introduced by Turner and others [80] which states that individuals in a large group tend naturally to perceive them self as members of a subgroup which changes dynamically – e.g. a person from Bremen and a Bavarian may feel to be in different groups while standing in a pub watching football but may perceive themselves as part of the same group surrounded by Japanese. The same may happen to a mathematician and a physicist on a sociology conference. So, the bounded confidence assumption may also have some real-world background.

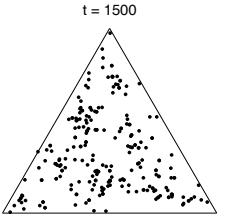
2.2 An experiment of social psychology

Here is a brief summary of experimental work done together with undergraduate students of psychology [60]. The question was: How do humans behave in an environment of continuous opinion dynamics?



Description of the experiment. A small group of test persons (2-6) sits in front of a glass which is filled up with 470 noodles. Each test person has a stack of papers in front and a pen. They cannot see each other. The instructor gives the order to estimate the number of noodles in the glass and to write it down. They are told that everyone who is not more than 5% away from the real value will get a prize. This was round one. Afterwards, the instructor collects all answers and writes them on a blackboard. Now, the persons are ordered to estimate again. So, people can freely decide to adjust their estimates by using the information of the other estimates. They write a new (or the old) estimates down. The instructor publishes the new estimates. This is repeated until four rounds are performed. Then all persons which do not differ more than 5% of the 470 get a prize. One should emphasize that this is not a game where one plays





2.2. An experiment of social psychology

against each other. Everyone fights alone ‘against’ the truth and everyone has his individual chance to win.

Results. The experiment has been done nine times with different groups of test persons. So, the dataset contains nine groups. Group k has $m(k) \in \{2, 3, 4, 5, 6\}$ test persons and for each test person her published opinions for time steps $t = 1, 2, 3, 4$. We denote the opinion of test person i of group k at time t with $x_i^k(t)$ and $x^k(t)$ denotes the vector of opinions in time step t . The arithmetic mean of the opinions of all test persons in group j at time t is defined $M(x^k(t))$.

The vast majority of the test persons underestimated the number of noodles. This was not intended when the experiment was constructed.

Figure 2.1 shows for group $k = 1, \dots, 9$ the evolution of the opinions $x^k(t)$ together with the evolution of $M(x^k(t))$ through the four rounds.

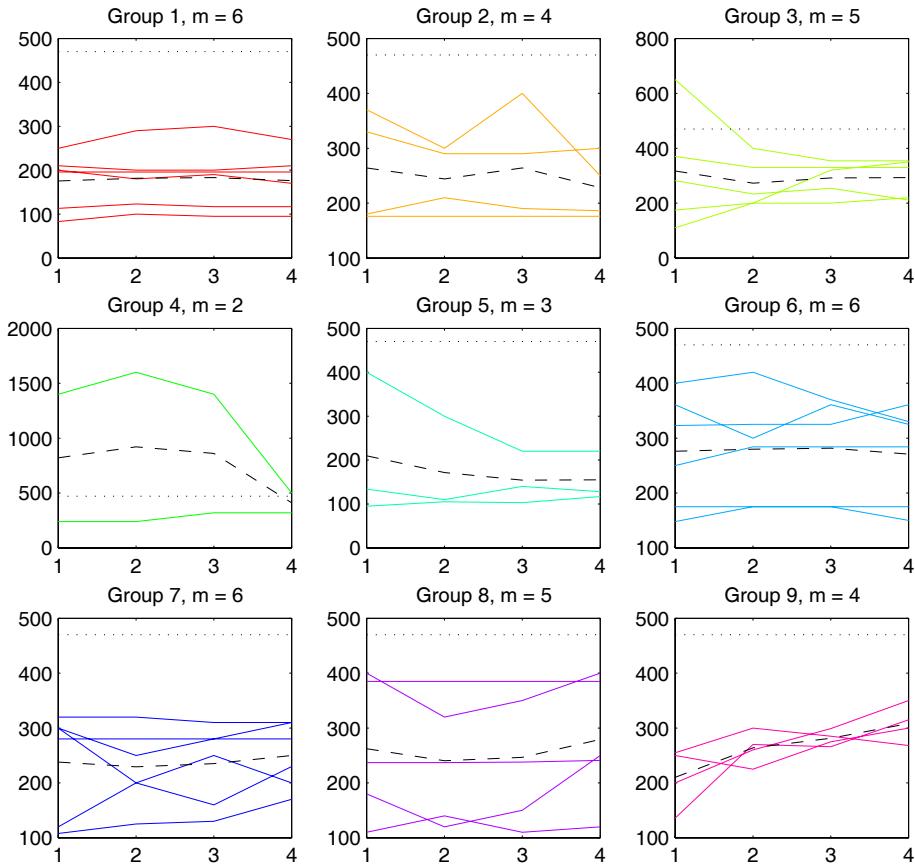
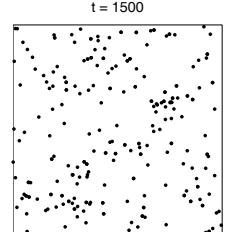
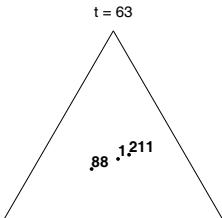


Figure 2.1: Evolution of opinion dynamics $x^k(t)$. The black dashed line is the arithmetic mean in the group $M(x^k(t))$. The black dotted line is the correct value 470.

Figure 2.2 shows a scattered plot for each test person. Plotted is for person i in group k the opinion distance to the mean after the first round $M(x^k(1)) - x_i^k(1)$ against the change of opinion to the second round $x_i^k(2) - x_i^k(1)$. The colors of the dots coincide with the colors in Figure 2.1.





2. Modeling

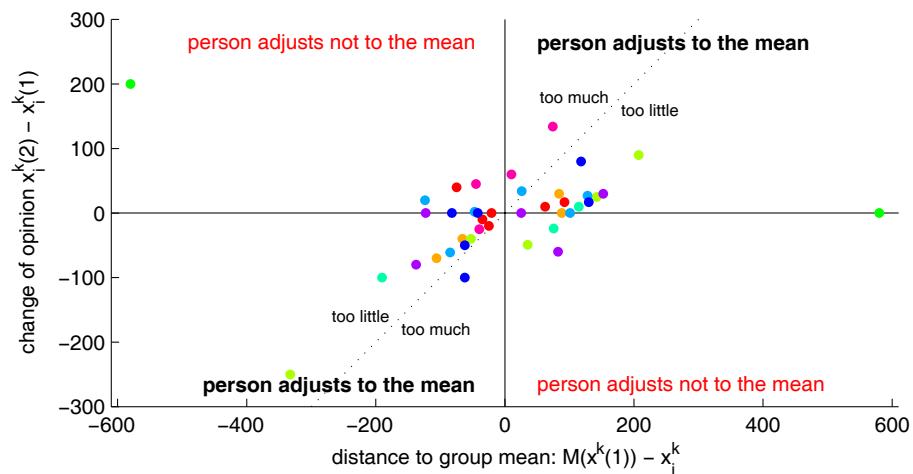


Figure 2.2: The behavior of all test persons after the first round is displayed here with respect to the mean opinion in their group. There is a tendency that test persons adjust towards the mean but ‘careful’.

The plot shows the opinion change of the test persons with respect to the mean of all opinions. The plot divides into the regions where persons adjust to the mean (first and third quadrant) and where they did not (second and forth quadrant). The diagonal dotted line divides the adjusting people into the ones who do it too little or too much. Persons on the abscissa did not change.

From the 41 test persons 8 ($\approx 20\%$) did not change, 8 ($\approx 20\%$) adjusted away from the mean and 25 ($\approx 61\%$) adjusted towards the mean. Form this 25 person who adjust towards the mean 21 ($\approx 84\%$) adjusted too little and 4 ($\approx 16\%$) adjusted too much.

The opinion dynamic processes in figure 2.1 show that a variety of phenomena may appear, e.g. opinion dynamics that bring the group (represented by their mean) closer to the truth (groups 4,8,9), opinion dynamics that brings the ones with good estimates on the wrong track (groups 2,5), opinion dynamics that gets consensual (groups 3,5), stays diverse (groups 7,8), or a common drift towards the truth (group 9). (On the topic of truth their is in continuous opinion dynamics under bounded confidence there is a work of Hegselmann and Krause [39] regarding simulation and mathematical analysis.)

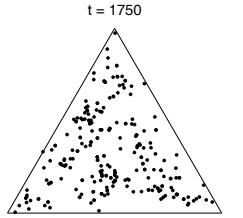
Interpretation. The idea of the experiment was to construct an opinion dynamics environment, where people can use the estimates of others to refine their own without effects of group pressure or opinion leaders.

The results in figure 2.2 show that there is a strong tendency that people adjust their opinions towards the mean. But most times they change less as it would be necessary to reach the mean. We call this ‘slow drift to the center’. If we regard that the test persons internal processing is similar to building a weighted arithmetic mean, then the slow drift to the center can be caused e.g. by a higher weight on the own opinion. But it could also be possible that the averaging person neglects the opinions which are far away from their own, which will also slow down the drift towards the center.

t = 63

.29

.221



2.3. Modeling vocabulary

The results of this small experiment give strong evidence that repeated averaging plays a role in real-world opinion dynamics processes. Further on, the bounded confidence assumption may also play a role, but this is not as strongly validated, because the reasons for change were not the scope of the experiment.

Further questions. It would be quite interesting to repeat this experiment with a larger sample of people and larger group sizes to gain hints on questions like:

- Under which condition will the group reach a consensus or split into more groups with internal consensus?
- Under what conditions does opinion dynamics help to approach the truth?
- Is there an optimal level how long opinion dynamics should last when the aim is that the group should get close to the truth?

2.3 Modeling vocabulary

This section defines our modeling vocabulary and serves as a glossary in further reading. It helps to describe the models defined in section 2.4 mathematically as well as regarding real world. It should help to hold the link between mathematics in its pure generality and the real world systems that, on the one hand, inspired the models, and on the other hand, could be analyzed by mathematical and simulation results derived with the models.

2.3.1 Agents and densities of agents

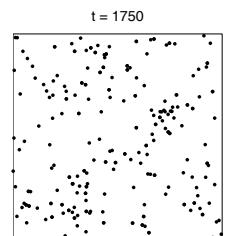
The basic entity of our models is a set of *agents*. The agent may stand for an expert, citizen, customer, politician, bird, fish, insect, robot, particle, computer program and so on. Thus, the whole set of agents may describe a commission, society, crowd, parliament, flock, school, swarm, herd and so on. In model descriptions and theory we will speak of agents, but while discussing examples or real world implications we may use the other words. Mathematically, they are synonyms.

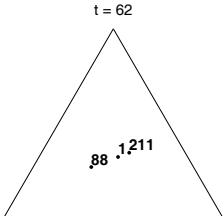
We will use two modeling concepts. The elementary one is the *agent-based* concept. In an agent-based model we have a finite number of $n \in \mathbb{N}$ agents indexed $\{1, \dots, n\}$ each holding an opinion. States and dynamics are defined for each agent with respect to the other agents.

A more elaborated concept is *density based* modeling. In a density based model we deal with population distributions on the opinion space. This concept is more elaborated because one uses the heuristics for agents behavior to describe dynamics of the population distribution.

2.3.2 The state spaces

The *state* of a dynamical model captures all relevant information to describe the system at a step in time. The *state space* contains all possible states of the model.





2. Modeling

State spaces in agent-based models. In agent-based models we call the number of agents n the *agent dimension* of the model. For one agent the relevant information is his *opinion*. We treat sets $S \subset \mathbb{R}^d$ as *opinion spaces*. One can think of compact sets, often convex. In the following we will say *appropriate opinion space* which means an opinion space such that dynamics in this space is welldefined.

The dimension $d \in \mathbb{N}$ is called the *opinion dimension*. The set $\{1, \dots, d\}$ is called the set of *opinion issues*. For one specific time step $t \in \mathbb{N}$ we call the collection of the opinions of all agents $x(t) \in S^n \subset (\mathbb{R}^d)^n$ the *opinion profile* at time t . Each state of an agent-based model of continuous opinion dynamics is an opinion profile. Figure 2.3 shows an example how to visualize an opinion profile. The cube $S = [0, 1]^n$ and the simplex $S = \Delta^n$ stand for different types of multidimensional opinions. While the issues in the cube are independent of each other the opinions in a simplex are under *budget constraints*. One can only raise one issue if one reduces another. Figure 2.5 shows these opinion spaces.

For agent $i \in \underline{n}$ we denote its opinion vector as $x^i \in S$ and for $j \in \underline{d}$ we denote by $x_j^i \in \mathbb{R}$ the opinion of agent i about opinion issue j .

An opinion profile $x \in S^n$ is called a *consensus* if $x^1 = x^2 = \dots = x^n$.

State spaces in density-based models. Given an opinion space $S \subset \mathbb{R}^d$ for an agent-based model. The state space of a density based model is the set of probability density function on S

$$\{p : S \rightarrow \mathbb{R}_{\geq 0} \mid \int_S p = 1\}.$$

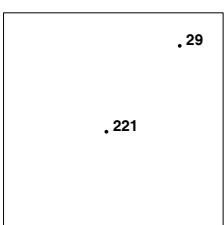
To simulate the evolution of a probability density function on S in a computer one has to discretize S . We discretize by partitioning the opinion space into a finite number of equally sized parts, which we call *opinion classes*. We limit this work to $d = 1$ for density-based models. Thus, S is an interval which has to be discretized. So, we discretize it with equidistant subintervals. We map the set of subintervals to the set $\{1, \dots, n\}$ which we call the set of opinion classes. In a density-based setting n determines the number of opinion classes, not the number of agents.

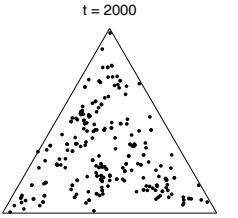
A vector $p \in \mathbb{R}^n$ is an *opinion distribution* if it is stochastic. So the state space is the unit simplex Δ^{n-1} . The entry p_i stands for the proportion of the agent population holding opinion $i \in \underline{n}$. For a specific time step $t \in \mathbb{N}$ the vector $p(t)$ determines the distribution of the (imaginary infinite) agent population to the opinion classes $\{1, \dots, n\}$ at time t . An opinion distribution p is a ‘probability vector’ because if we pick an agent, then p_i determines the probability that he holds opinion i . For convenience we define $p_i = 0$ for all $i \notin \underline{n}$.

In density-based models we may consider n as a parameter how accurate a continuous opinion can be communicated, e.g. how many steps do we allow on the continuous scale from minimum to maximum opinion.

The accuracy topic has been simulated as agent-based model by Urbig [81] with links to communication and marketing theory about selective attention. He distinguishes attitudes (which are continuous) and opinions as verbalized attitudes which can only be discrete. In our words he claims that via verbalization one transforms a continuous attitude into a discrete opinion class.

Figure 2.4 shows an example how to visualize an opinion distribution.





2.3. Modeling vocabulary

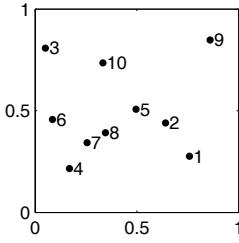


Figure 2.3: Example opinion profile for $n = 10$ agents in the opinion space $S = [0, 1]^2$.

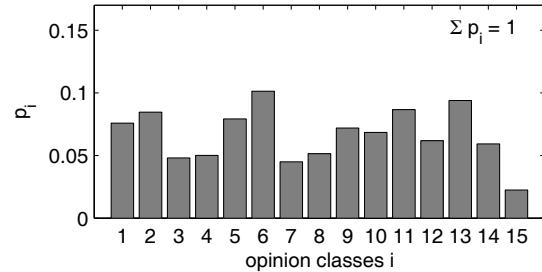


Figure 2.4: Example opinion distribution for $n = 15$ opinion classes.

2.3.3 Dynamics

A discrete dynamical system is determined by the state space and a sequence of self-maps in the state space, which are applied iteratively. It is called *dynamic* because it contains a time evolution. While considering only discrete time we regard change in opinion as a discrete event and avoid problems of discretization of time, which we would have when simulating continuous time models.

Agent-based dynamics We need for each time-step $t \in \mathbb{N}$ a self-map in the state space

$$\begin{aligned} f_t : & S^n \rightarrow S^n \\ & x \mapsto f_t(x) \end{aligned}$$

to define for an initial profile $x(0) \in S^n$ the *agent-based process* (or *trajectory*, or *solution*) as the sequence $(x(t))_{t \in \mathbb{N}}$ recursively through

$$x(t+1) = f_t(x(t)). \quad (2.1)$$

The pair $(S^n, (f_t)_{t \in \mathbb{N}})$ describes the non-autonomous *discrete dynamical system*.

The component-function $(f_t)_i : S^n \rightarrow S$ computes a new state of agent i out of the states of all agents.

Notice that the term ‘solution’ comes from the theory of differential equations and one may think about the problem of existence and uniqueness of a solution. But discrete dynamical systems clearly always have a unique solution for each initial profile $x(0) \in S^n$.

If the self-map $f_t = f$ for all time-steps then the system is called *independent of time*, otherwise it is called *time-dependent*.

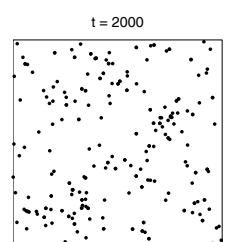
If there is a matrix valued function

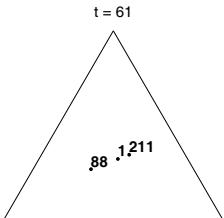
$$\begin{aligned} A : & \mathbb{N} \times S^n \rightarrow \mathbb{R}^{n \times n} \\ & (t, x) \mapsto A(t, x). \end{aligned}$$

for which it holds

$$f_t(x(t)) = A(t, x(t))x(t)$$

then the system is called *linear*. Respectively, it is called *state-dependent* and *time-dependent*. If for all $x \in (\mathbb{R}^d)^n$ it holds $A(t, x) = A(t)$ then it is linear and





2. Modeling

time-dependent. If $A(t, x) = A$ it is linear and independent of time and state. If for all $t \in \mathbb{N}$ it is $A(t, x) = A(x)$ then it is called linear state-dependent and independent of time.

Density-based dynamics For dynamics in a density-based state space we define a self-map in the unit-simplex Δ^{n-1} . Let $p \in \Delta^{n-1}$ be an opinion distribution. We define the self-map on Δ^{n-1} with a state-dependent row-stochastic *transition matrix* $B(p) \in \mathbb{R}^{n \times n}$. For the initial opinion distribution $p(0) \in \Delta^{n-1}$ the density-based process is then defined as the sequence $(p(t))_{t \in \mathbb{N}}$ recursively through

$$p(t+1) = p(t)B(p(t)). \quad (2.2)$$

So, the process (2.2) is an *interactive Markov chain* with discrete time and a finite number of discrete states. It is called interactive because the transition matrix depends on the actual state. We do not consider time-dependent transition matrices. The entry $B_{ij}(p(t))$ determines the fraction of agents moving from opinion class i to opinion class j . The explicit definition of the interactive transition matrix is the core of a density-based model. This is done in Section 2.4.

2.3.4 Bounded confidence

Bounded confidence in agent-based state spaces. One assumption on opinion dynamics in this thesis is that the agents have *bounded confidence*. We call $\varepsilon_i \in \mathbb{R}$ a *bound of confidence* of agent i if it is greater than zero. If $\varepsilon_1 = \dots = \varepsilon_n = \varepsilon$, then we call ε a *homogeneous bound of confidence*. Further on, we need a norm $\|\cdot\|$ in the opinion space. The idea is that agent i 's *area of confidence* is the unit ball of the norm scaled by ε_i and translated to the agent's actual opinion. Usually, we will use the p -norms with $p = 1, 2, \infty$.

Thus, for $d = 1$ we have intervals as areas of confidence. Figure 2.5 gives impressions for two-dimensional areas of confidence.

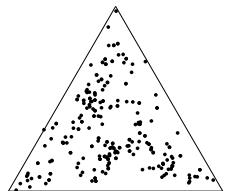
The norm displays the way how agents define and judge their distance in opinion to other agents.

It is important for understanding of model assumptions to have an idea what the different norms mean for the confidence judgment of the agents. So here is a brief interpretation for $p = 1, 2, \infty$.

An agent who judges distances in the 1-norm is looking at the sum of the absolute values in all dimensions. So, he is willing to accept a higher absolute value in one dimension if this is compensated by a lower absolute value of the same magnitude in another dimension. We describe this judgment as *compensating*.

An agent who judges distances in the 2-norm is looking at the Euclidean distance. Thus, this agent regards the distances like opinions sitting in a space with Euclidean geometry. This distance is especially useful when modeling bounded perception in 2- or 3-dimensional spaces, because it regards distances in the way we measure them with a ruler. It is of interest in modeling collective motion in swarms.

An agent who judges distances in the ∞ -norm is looking at the maximum distance over all dimensions. So, he judges the magnitudes of all dimensions independently but he demands for all dimensions the same maximal deviation and is not willing to accept any exceptions. We call this judgment *non-compensating*.



2.3. Modeling vocabulary

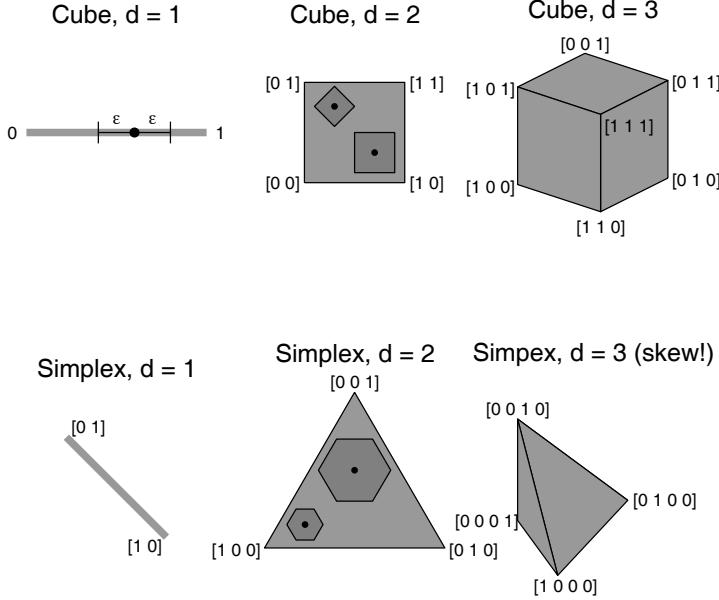


Figure 2.5: The opinion spaces cube $[0, 1]^n$ and simplex Δ^n for $n = 1, 2, 3$. For $d = 2$ we display example areas of confidence for 1- and ∞ -norm.

An agent with bounded confidence is regarded as an agent who is only capable or willing to make interactions with agents within his area of confidence respectively perception. We define the *confidence set* of agent $i \in \underline{n}$ in dependence of the opinion profile $x \in S^n$, the norm parameter p and the bound of confidence $\varepsilon \in \mathbb{R}_{>0}$ as $I_\varepsilon(i, x) := \{j \in \underline{n} \mid \|x_j - x_i\|_p \leq \varepsilon\}$.

Bounded confidence in density-based state spaces. Regarding the opinion classes $\underline{n} = \{1, \dots, n\}$ bounded confidence for a *discrete bound of confidence* $\varepsilon \in \mathbb{N}$ means the following. Regarding two opinion classes $i, j \in \underline{n}$ and an opinion distribution $p \in \Delta^{n-1}$, it must hold for the transition probability that

$$B_{ij}(p) = B_{ij}(p_{[i-\varepsilon, i+\varepsilon]}).$$

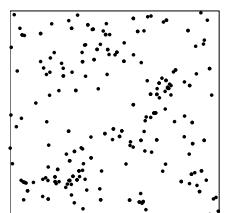
There are only transitions between classes if their distance is below ε .

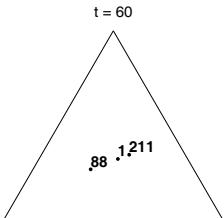
2.3.5 Repeated averaging

Averaging one-dimensional opinions. First we consider $d = 1$. So an appropriate opinion space S is an interval in \mathbb{R} . A *general mean* is a function $g : S^n \rightarrow S$ such that the *sandwich inequality*

$$\min_{i \in \underline{n}} x^i \leq g(x) \leq \max_{i \in \underline{n}} x^i \quad (2.3)$$

holds. A central assumption in continuous opinion dynamics is that each agent computes his new opinion as a general mean of the opinions of all agents. An





2. Modeling

example is the *power mean* for a real parameter $p \neq 0$

$$\begin{aligned} P_p : (\mathbb{R}_{>0})^n &\rightarrow \mathbb{R}_{>0} \\ x &\mapsto \left(\frac{1}{n} ((x^1)^p + \cdots + (x^n)^p) \right)^{\frac{1}{p}} \end{aligned}$$

The power mean includes the popular *arithmetic* ($p = 1$) and *harmonic* ($p = -1$) means. The *geometric mean* $\sqrt[n]{x^1 \dots x^n}$ is approached for $p \rightarrow 0$, the maximum $\max\{x^1, \dots, x^n\}$ for $p \rightarrow \infty$ and the minimum $\min\{x^1, \dots, x^n\}$ for $p \rightarrow -\infty$. For power means and their 'limit means' for $p = -\infty, 0, \infty$ it holds that $p < q$ implies that for a given appropriate opinion profile x which is not a consensus it holds $P_p(x) < P_q(x)$.

Further on, there are weighted means. For nonnegative numbers $\alpha_1, \dots, \alpha_n$ which sum up to one there is the *weighted arithmetic mean* $\sum_{i=1}^n \alpha_i x^i$ and the *weighted geometric mean* $\prod_{i=1}^n x_i^{\alpha_i}$.

Another generalization is the *f-mean*. For a continuous and injective function $f : S \rightarrow \mathbb{R}$ (which is thus invertible) the *f-mean* of x^1, \dots, x^n is

$$f^{(-1)}\left(\frac{1}{n} \sum_{i=1}^n f(x^i)\right).$$

The power mean is represented here as $f(x) = x^p$, the geometric mean as $f(x) = \log(x)$.

The *a-mean* is defined for a given vector $a \in \mathbb{R}^n$ by $\frac{1}{n!} \sum_{\sigma} (x^{\sigma_1})^{a_1} \dots (x^{\sigma_n})^{a_n}$, where the summation is over all permutations of n . The arithmetic mean appears for $a = (1, 0, \dots, 0)$, the geometric mean for $(\frac{1}{n}, \dots, \frac{1}{n})$.

If we regard the opinions x^1, \dots, x^n as outcomes in samples of a random variable, we can use notions of descriptive statistics to define further general means, like the expected value, the median or other quantiles. Descriptive statistics serves as an interpretation of opinion dynamics when we imagine a commission of n experts which has to estimate a certain quantity. As data they have only the estimates of all their colleagues and their own from a former round. So each expert estimates again with a certain statistical technique where he may decide on using miscellaneous estimators for the real value, miscellaneous assumptions on the underlying distribution leading to different general means. He may regard estimates of others as outliers and neglect them. One possible method to select the outliers is to take all agents which are not in his area of confidence.

In the case of outliers (e.g. by restriction to a set of confidence) the general mean is performed only on a subset of the agents. Thus, such means have been called *partial abstract means* in [38] which also contains some more discussion about general means.

Other means are the Lehmer means or the Heronian mean [1]. Further on, means can be defined by means of means.

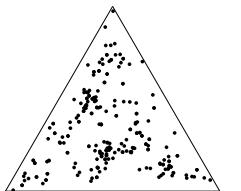
A main topic in the mathematical analysis in Chapter 3 is convergence to consensus when agents repeatedly apply various partial abstract means. But beforehand we extend to means of vectors.

Averaging multidimensional opinions. Here we want to extend the definition of a *general mean* to d -dimensional opinion vectors for $d \geq 2$. So let $S \subset \mathbb{R}^d$ be an appropriate opinion space.

t = 60

.29

.221



2.3. Modeling vocabulary

All concrete means mentioned for the case $d = 1$ can be generalized to more dimensions by taking them componentwise. Further on, one may define different one-dimensional means in each component. We define two different multidimensional equivalents for the sandwich inequality (2.3).

First, a function $g : S^n \rightarrow S$ is called a *general convex hull mean* if it holds for all $x \in S^n$ that $g(x) \in \text{conv}\{x^1, \dots, x^n\}$. Obviously, the componentwise weighted arithmetic mean fulfills this property. But e.g. the componentwise geometric mean does not (see Figure 2.6).

So, for $x \in S^n$ we define $\text{cube}_{i \in n} x^i := [\min_{i \in n} x^i, \max_{i \in n} x^i]$. Notice that max and min are componentwise and that the interval is multidimensional. So cube represents the smallest closed hypercube in \mathbb{R}^d which covers all vectors x^1, \dots, x^n . Figure 2.6 shows an example of the convex hull and the cube of a set of points in \mathbb{R}^2 .

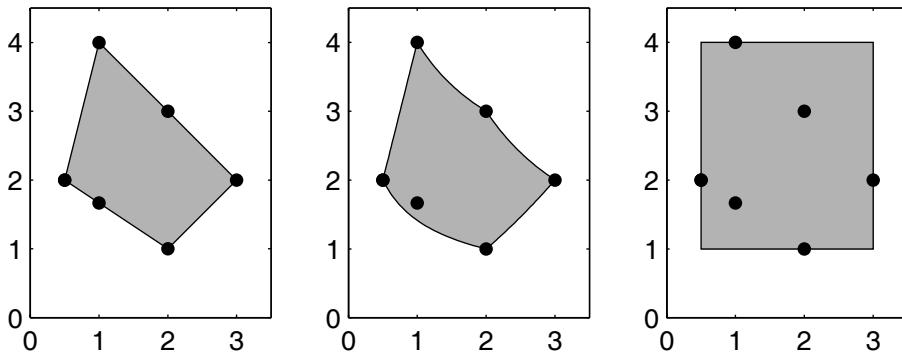


Figure 2.6: A set of six points in $\mathbb{R}^2_{\geq 0}$. The gray area is their convex hull (left hand), all possible weighted geometric means (center) and their cube (right hand).

Another description of cube is

$$\text{cube}_{i \in n} x^i = \left\{ \sum_{i=1}^n a^i \bullet x^i \mid (a^1, \dots, a^n) \in (\mathbb{R}_{\geq 0}^d)^n \text{ with } \sum_{i=1}^n a^i = \mathbf{1} \right\}$$

We call the vectors a^1, \dots, a^n *componentwise convex coefficient vectors* and a sum $\sum_{i=1}^n a^i \bullet x^i$ a *componentwise convex combination*. The $d \times n$ matrix $A := [a^1 \dots a^n]$ is row-stochastic. An ordinary convex combination is then the special case with each a^i having equal entries (thus with A being a consensus matrix as defined in Subsection 2.3.6).

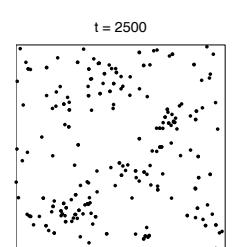
A function $g : S^n \rightarrow S$ is called a *general cube mean* if it holds for all $x \in S^n$ that $g(x) \in \text{cube}\{x^1, \dots, x^n\}$.

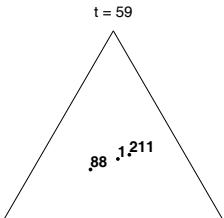
A function is a *general mean* if it is a general cube or convex hull mean.

A function $f : S^n \rightarrow S^n$ is called *averaging map* if all the component functions f_i are general means. An averaging map is called *convex hull averaging map* if all component functions are general convex hull means. The *cube averaging map* is defined analog.

It holds for a cube averaging map f and all $x \in S^n$ that

$$\text{cube}\{f_1(x), \dots, f_n(x)\} \subseteq \text{cube}\{x^1, \dots, x^n\}. \quad (2.4)$$





2. Modeling

For a convex hull averaging map f and all $x \in S^n$ it holds respectively that

$$\text{conv}\{f_1(x), \dots, f_n(x)\} \subseteq \text{conv}\{x^1, \dots, x^n\}. \quad (2.5)$$

An averaging map f is called *proper* if it holds

$$x \text{ is not a consensus} \Rightarrow \text{conv}\{f_1(x), \dots, f_n(x)\} \neq \text{conv}\{x^1, \dots, x^n\}.$$

Respectively, the same with conv replaced by cube. So the cube or the convex hull of all opinions has to shrink after application of an averaging map. The contraposition is also of interest

$$\text{conv}\{f_1(x), \dots, f_n(x)\} = \text{conv}\{x^1, \dots, x^n\} \Rightarrow x \text{ is a consensus.}$$

Respectively, the same with conv replaced by cube.

It is easy to see that every convex hull averaging map is also a cube averaging map. But a proper convex hull averaging map is not necessary a proper cube averaging map.

In this thesis we study the discrete dynamical system defined in (2.1) where the self-maps are averaging maps which might be proper.

In the linear framework $f(x) := Ax$ it is easy to see, that f is a convex hull averaging map if and only if A is row-stochastic.

Averaging in density-based dynamics Let us regard the opinion classes $\underline{n} = \{1, \dots, n\}$. The *support* of $p \in \Delta^{n-1}$ is the discrete interval from the minimal class with positive mass to the maximal class with positive mass. Taking the heuristics of averaging maps from the agent based settings it is natural to define that

$$B_{ij}(p) = 0 \quad \text{if } p_{[0, \pm 1, j]} = 0 \text{ or } p_{[j, \pm 1, n]} = 0$$

So, there are no transitions to classes which are outside of the support of p .

Given a *discrete bound of confidence* $\epsilon \in \mathbb{N}$. Then a property of averaging is that there can be no transitions to classes which are not within the support in the interval $\{i - \epsilon, \pm 1, i + \epsilon\}$.

These concepts are stated as ideas for further work, we do not deepen them in the following.

2.3.6 Matrices and networks

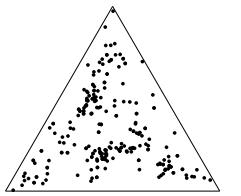
As already mentioned, a row-stochastic matrix describes a convex hull averaging map. We call a matrix $A \in \mathbb{R}^{n \times n}$ a *confidence matrix* if it is row-stochastic. For a confidence matrix the index set \underline{n} is the set of agents. The entry a_{ij} stands for the *confidence weight* that agent i gives to agent j . Row i in a confidence matrix represents a weighted arithmetic mean what agent i uses to compute his new opinion out of all other opinions. Agent i computes a convex combination of x^1, \dots, x^n with coefficients a_{i1}, \dots, a_{in} . The explicit definition of a confidence matrix with respect to the actual profile or time step lies at the heart of each agent-based model in this thesis.

A communication *network* is a graph with the set of agents as the vertex set. In this thesis we will not stress graph theory and thus we will think of a network as its adjacency matrix $N \in \{0, 1\}^{n \times n}$. So, the restriction of a confidence matrix

t = 59

.29

.221



2.3. Modeling vocabulary

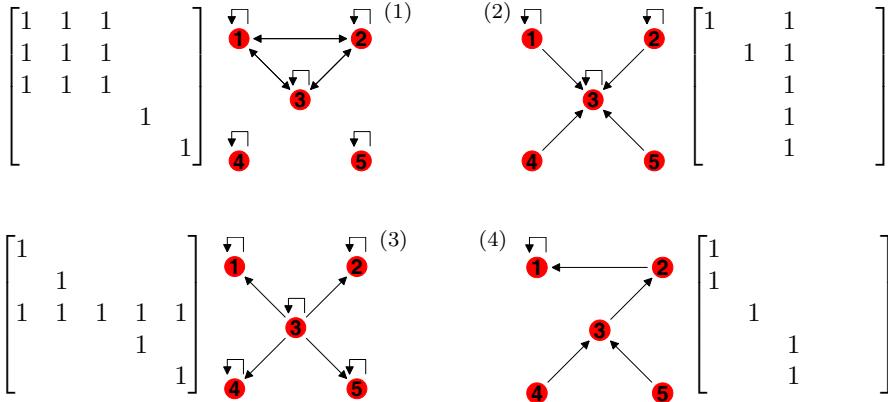


Figure 2.7: Examples for communication networks. Here, an arrow means ‘listens to’ and $i \rightarrow j$ coincides with a positive entry in ij . Attention, other authors interpret arrows the other way round as ‘influences’, which is also plausible and leads to transposition of the adjacency matrix. The examples may be interpreted as (1) a group meeting of 1,2 and 3; (2) 1, 2, 4 and 5 watching television 3, with 1 and 2 having an own opinion and 4 and 5 just adopting television’s opinion; (3) politician 3 gathers information from the independent experts 1, 2, 4 and 5; (4) a hierarchy where 2 obeys 1 and rules 3 and while 3 rules 4 and 5.

A by a network can be expressed by $A \leq N$. A graph with self-links (i, i) corresponds to a network with a positive diagonal.

A communication network which changes in time is called a *communication regime*. Formally, it is a sequence of adjacency matrices $(N(t))_{t \in \mathbb{N}}$. It restricts communication in each time step by the current network.

A network $N(t)$ can display a meeting of certain participants or influence by mass media or an order to adopt. Examples are in figure 2.7. The communication regime can display a time schedule of such diverse events.

We say i is *connected* to j if $n_{ij} = 1$. If we display N as a directed graph this is in analogy to an arrow from i to j . It means that i takes j ’s opinion into account (gives a positive weight to it) and not vice versa. Some other authors (e.g. Moreau [62]) define ‘connected’ the other way round. This is also plausible because one can argue that j influences i . In that respect one could say j is connected to i . This would lead to matrix transposition and is not meant by ‘connected’ in this thesis.

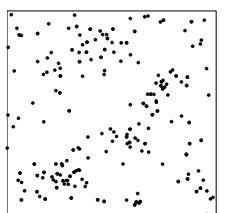
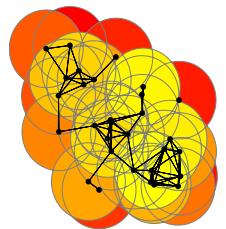
Let $x \in S^n$ be an opinion profile, $\varepsilon_1, \dots, \varepsilon_n \in \mathbb{R}_{>0}$ bounds of confidence and N a restricting network coming form a communication regime. Then, the *bounded confidence network* $N^{BC}(x, \varepsilon_1, \dots, \varepsilon_n) \in \{0, 1\}^{n \times n}$ is defined as $n_{ij}^{BC} = 1$ if $\|x_j - x_i\| \leq \varepsilon_i$ and zero otherwise. The bounded confidence network restricted by N is $N^{BC}(x, \varepsilon_1, \dots, \varepsilon_n) \bullet N$.

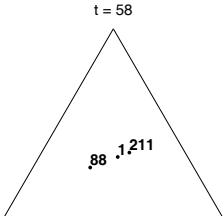
We define the *bounded confidence matrix*

$$A(x, \varepsilon_1, \dots, \varepsilon_n, N) = \text{sto}(N^{BC}(x, \varepsilon_1, \dots, \varepsilon_n) \bullet N).$$

If there is no network restriction we write $A(x, \varepsilon_1, \dots, \varepsilon_n)$. If we got a homogeneous bound of confidence ε for all agents we abbreviate $A(x, \varepsilon, N)$. Further on, N is omitted if there are no restriction by a communication regime.

A *consensus matrix* is a row-stochastic matrix with rank one. Thus, all its rows are equal due to row-stochasticity. If K is a consensus matrix and x an opinion profile, it is easy to see, that Kx is a consensus.





2. Modeling

2.4 The mathematical models

In this section we define all models that will be covered in this thesis. For each model we present a mathematical definition, and examples that present the basic dynamical behavior as well as some interesting phenomena which serve as starting points for more detailed mathematical and simulation analysis in Chapters 3 and 4.

2.4.1 Repeated pooling of opinions

We begin with the definition of a very general process of continuous opinion dynamics, which includes all the following processes as special cases.

Model Definition 2.4.1 (Process of opinion pooling). Let $S \subset \mathbb{R}^d$ be an appropriate opinion space and $(N(t))_{t \in \mathbb{N}}$ a communication regime. Consider $x(0) \in S^n$ to be an initial opinion profile with n agents holding d -dimensional continuous opinions and let A be a matrix valued function

$$\begin{aligned} A : \quad \mathbb{N} \times S^n \times \{0, 1\}^{n \times n} &\rightarrow \mathbb{R}^{n \times n} \\ (t, x, N) &\mapsto A(t, x, N) \end{aligned}$$

such that for all (t, x, N) it holds $A(t, x, N)$ is row-stochastic and $A(t, x, N) \leq N$. Then we call the sequence $(x(t))_{t \in \mathbb{N}}$ recursively defined through

$$x(t+1) = A(t, x(t), N(t))x(t)$$

a *process of opinion pooling*.

The term *opinion pool* for building weighted arithmetic means goes back at least to Stone [76].

The simplest case is, when there is for all x, t and N one fixed confidence matrix $A = A(t, x, N)$. The matrix A represents a fixed reputation structure between the agents regardless of time and of the opinions of the agents. The idea of modeling opinion dynamics as repeated pooling of opinions with a fixed confidence matrix goes back to DeGroot [21] who analyzed the model about conditions for reaching consensus. In Berger [7] necessary and sufficient conditions for consensus have been presented. We give these conditions in subsection 3.2.3 in the context of the powers of an arbitrary row-stochastic matrix.

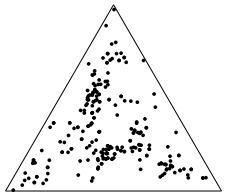
The same model with a fixed confidence matrix appears also in Lehrer and Wagner [50] as the elementary method for reaching rational consensus in science and society. They argue for repeated arithmetic averaging to determine consensus in such broad fields as politics, social choice, justice, epistemology, science, intuition and common sense and language. Their fundamental claim is, that there is no such thing as the truth. So, aggregation of information by weighted arithmetic means is the rational way to achieve consensus. That is presented as a normative theory for consensual rationality. There is also an axiomatic analysis in the flavor of the social choice theory which characterizes all aggregation methods which aggregate opinions which lie in the unit simplex of dimension larger or equal to two as weighted arithmetic means.

Collignon and Al-Saddon [13] presented recently basically the same model with a fixed confidence matrix to model the evolution of the so-called ‘stochastic consensus’ in a huge institution such as the European Union.

t = 58

.29

.221



2.4. The mathematical models

The model above is much more general because it allows ever changing confidence matrices. But nevertheless several results about convergence and convergence to consensus are possible with some assumptions on the matrices. This is subject to Section 3.2.

2.4.2 Agent-based bounded confidence models

We begin with a general agent-based bounded confidence model, which includes the basic versions of the following models as special cases.

Model Definition 2.4.2 (General bounded confidence process). Let $S \subset \mathbb{R}^d$ be an appropriate opinion space and $(N(t))_{t \in \mathbb{N}}$ a communication regime. Let $\varepsilon_1, \dots, \varepsilon_n > 0$ be bounds of confidence and $\|\cdot\|_p$ be the p -norm. For a certain opinion profile $x \in S^n$ and a certain time step $t \in \mathbb{N}$ let $A(t, x) := A(x, \varepsilon_1, \dots, \varepsilon_n, N(t))$ be the *bounded confidence matrix*¹. Explicitly

$$A_{ij}(t, x) := \begin{cases} \frac{1}{\#I(i, t, x)} & \text{if } j \in I(i, t, x) \\ 0 & \text{otherwise,} \end{cases}$$

with $I(i, t, x) := I_{\varepsilon_i}(i, x) \cap \text{nb}(i, N(t)) = \{j \in \underline{n} \mid \|x^j - x^i\|_p \leq \varepsilon_i \text{ and } N_{ij}(t) = 1\}$. For each initial profile $x(0) \in S^n$ we define the *bounded confidence process* $(x(t))_{t \in \mathbb{N}}$ recursively by

$$x(t+1) = A(t, x(t))x(t).$$

Let us regard the agents as mass points of unit mass. Then, in each time step each agent moves to the center of mass of all the agents in his confidence set (restricted by the communication regime).

The network determined by the areas of confidence is changing dynamically with the state of the system, while the communication regime is an external restriction for communication.

We continue with the definition of the two most cited bounded confidence models of Hegselmann-Krause (*HK*) [37, 46] and Deffuant-Weisbuch (*DW*) [19, 88]. Both can be unified under the framework of Model 2.4.2 and differ mainly in their communication regime. We define them separately to deliver independent definitions.

Model Definition 2.4.3 (Hegselmann-Krause model, communication in repeated meetings). Let there be $n \in \mathbb{N}$ agents and an appropriate opinion space $S \subset \mathbb{R}^d$.

Given an initial profile $x(0) \in S^n$, bounds of confidence $\varepsilon_1, \dots, \varepsilon_n > 0$ and a norm $\|\cdot\|$ we define the *repeated meeting process* $(x(t))_{t \in \mathbb{N}}$ recursively through

$$x(t+1) = A(x(t), \varepsilon_1, \dots, \varepsilon_n)x(t), \quad (2.6)$$

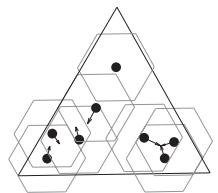
with $A(x, \varepsilon_1, \dots, \varepsilon_n)$ being the confidence matrix defined

$$A_{ij}(x, \varepsilon_1, \dots, \varepsilon_n) := \begin{cases} \frac{1}{\#I_{\varepsilon_i}(i, x)} & \text{if } j \in I_{\varepsilon_i}(i, x) \\ 0 & \text{otherwise,} \end{cases}$$

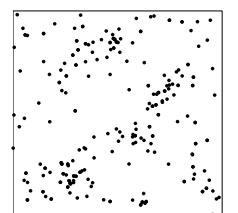
with $I_{\varepsilon_i}(i, x) := \{j \in \underline{n} \mid \|x^i - x^j\| \leq \varepsilon_i\}$.

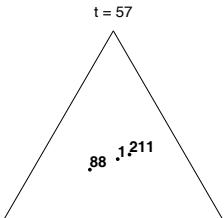
If $\varepsilon_1 = \dots = \varepsilon_n$ we call the model *homogeneous*, otherwise *heterogeneous*.

meeting step



¹One should not get confused by the different variables in $A(\cdot)$. For full clarity $A(x, \varepsilon_1, \dots, \varepsilon_n, N(t))$ is correct, but $A(t, x)$ is a common abbreviation.

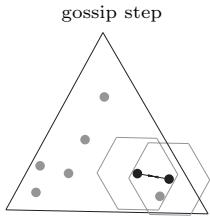




2. Modeling

In the framework of the general model 2.4.2 the Hegselmann-Krause model is determined by $N(t)$ being the ones matrix in every time step.

The term 'meeting' should display that every agent needs to know all other opinions to decide which is his confidence set. Of course, this need not in every case happen via a meeting. A repeated publication of all opinions will work either. Further on, not every real-world meeting may lead to the situation where everyone knows the opinions of every other. But in this thesis 'meeting' stands for a situation where an agent decides for a new opinion with the full knowledge of all opinions. To model a real-world meeting it would be probably more realistic to incorporate a speakers list and thus have a growing set of agents with published opinions. Dynamics could then be defined only on this growing set. This idea has been suggested by Hegselmann [36]. In the interpretation of Stauffer [74] the agents in the Hegselmann-Krause model are opportunists, because they want to be in the barycenter of all agents which are close to them. Thus they minimize their distance to all agents they trust.



Model Definition 2.4.4 (Deffuant-Weisbuch model, gossip communication). Let there be $n \in \mathbb{N}$ agents and an opinion space $S \subset \mathbb{R}^d$ convex. Given an initial profile $x(0) \in S^n$, bounds of confidence $\varepsilon_1, \dots, \varepsilon_n > 0$, a *cautiousness parameter* $0 < \mu \leq 0.5$ and a norm $\|\cdot\|$ we define the *gossip process* as the random process $(x(t))_{t \in \mathbb{N}}$ that chooses in each time step $t \in \mathbb{N}$ two random agents i, j which perform the action

$$x^i(t+1) = \begin{cases} x^i(t) + \mu(x^j(t) - x^i(t)) & \text{if } \|x^i(t) - x^j(t)\| \leq \varepsilon_i \\ x^i(t) & \text{otherwise.} \end{cases}$$

The same for $x^j(t+1)$ with i and j interchanged.

If $\varepsilon_1 = \dots = \varepsilon_n$ we call the model *homogeneous*, otherwise *heterogeneous*. If no cautiousness parameter is given then $\mu = 0.5$ should be default.

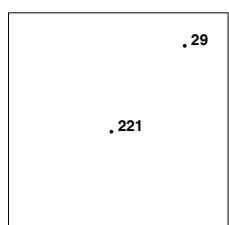
The communication regime of the Deffuant-Weisbuch model is a sequence $(N(t))_{t \in \mathbb{N}}$ with each network being a unit matrix with two additional off-diagonal symmetric entries equal to one. The position of the off-diagonal entries ij and ji is chosen randomly for every t . We call it 'gossip' because in one time step only two agents exchange information.

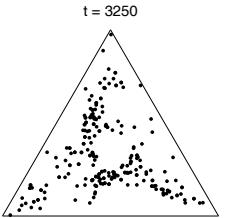
The cautiousness parameter μ gives the proportion an agent moves to his communication partner (if they are close enough to each other). For $\mu = 0.5$ an agent moves to the center of mass of the two agents. If $\mu < 0.5$ he is more cautious.

For $\mu = 0.5$ the Deffuant-Weisbuch model fits in the framework of the general model with the mentioned communication regime. For $\mu < 0.5$ the confidence matrix for a Deffuant-Weisbuch process is

$$(2\mu)A(x, \varepsilon_1, \dots, \varepsilon_n, N(t)) + (1 - 2\mu)E$$

Nearly always in the following we study only the standard value $\mu = 0.5$. The parameter has shown to have a low effect on dynamics on a first view (just on convergence times, see [19, 88]). But when extensions come in the parameter my 'wake up' and shows interesting interrelations with other parameters (see [61] which is also part of the dissertation or [18]).





2.4. The mathematical models

2.4.3 Agent-based examples and phenomena

Some examples should give a first insight how dynamics evolve.

In a first step we see in Figure 2.8 processes which start with a random initial opinion profile of fifty agents with opinions random and equally distributed chosen from the opinion space $[0, 1]$. On the left hand side repeated meeting processes of the HK model, on the right hand side gossip processes of the DW model. Each process for different homogeneous bounds of confidence ε . The first observation is that processes converge to a certain formation of opinion clusters. Low ε leads to a plurality of opinion clusters, while high ε leads to consensus. In between is a phase where opinions polarize into two more or less equally sized clusters. But notice that in the gossip process sometimes single agents survive as ‘outliers’.

We continue with a brief explanation how dynamics under a uniform initial distribution on a compact and convex opinion space evolve generally in time: Dynamics start at the border of the opinion space. Agents at the extremes who do a meeting step will move on average a step of $\varepsilon/2$ towards the center because the barycenter of their area of confidence lies there if we assume a uniform initial distribution. Agents at the extremes who do a gossip step will move towards the center because the probability that their communication partner has a more central opinion is positive while their probability to meet an agent with a more extreme opinion is zero. Due to this, agents move successively closer to the center of the opinion space in both processes. In the center where agents have an area of confidence uniformly filled with agents we observe only random fluctuations. This leads to a higher concentration of agents in the regions which lie at a distance of $\varepsilon/2$ to ε from the border of the opinion space towards the center. These high density regions attract agents also from the center in successive time steps. This leads to the formation of a cluster which gets disconnected from the center. This leads to a similar situation as in beginning but closer to the center. Where a new cluster evolves. This evolution of clusters continues until we reach the center where some interference of evolving clusters coming from more sides may occur.

This description is also valid for more dimensional opinion spaces with the additional feature that contraction occurs on all face levels at the same time. But close to lower dimensional faces of the opinion space (e.g. at the corner of a two-dimensional square) we reach higher concentrations of agents after the first steps because agents coming from different higher dimensional faces (e.g. the sides of the square) meet.

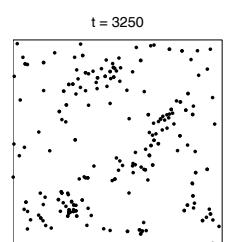
In the corners of the pages of this thesis there are flip books for the time evolution of two-dimensional bounded confidence processes, which give some insight to this phenomenon.

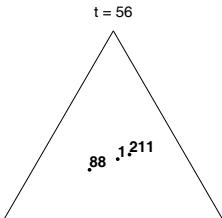
Each flip book process has 250 agents and initial opinions random and equally distributed in the respective opinion space. The bound of confidence is always $\varepsilon = 0.2$ and the norm is $\|\cdot\|_\infty$.

There is in the

upper corner on even pages Repeated meeting process with the two-dimensional simplex Δ^2 as opinion space. Time goes forward when page numbers go downward.

upper corner on odd pages Gossip process with the two-dimensional simplex Δ^2 as opinion space. Time goes forward when page numbers go for-





2. Modeling

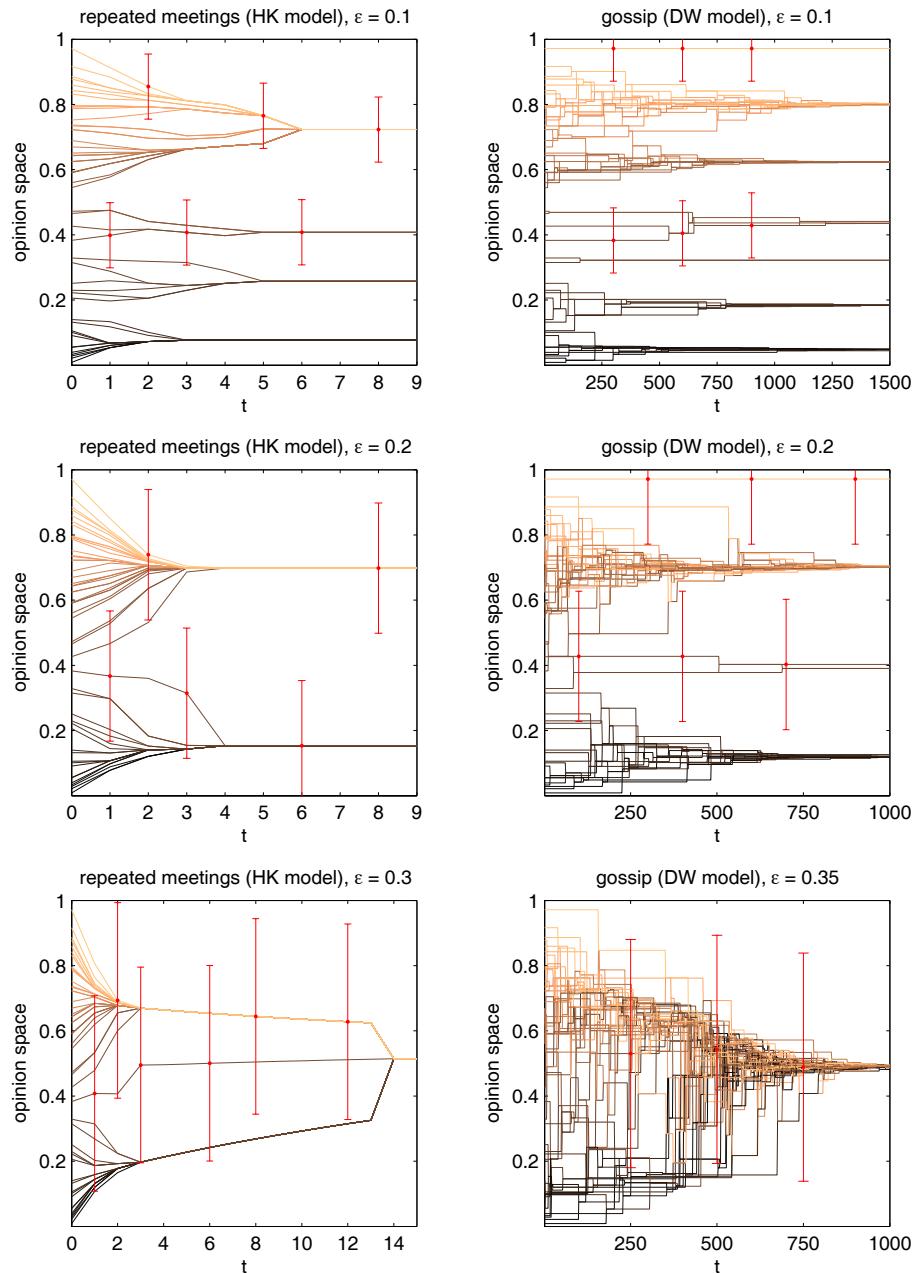
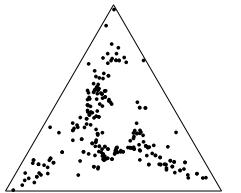


Figure 2.8: From plurality via polarization to consensus with rising bound of confidence ε . All subplots start with the same initial opinion profile $x(0) \in S^{50}$ opinion space $S = [0, 1]$. The homogeneous areas of confidence were drawn around selected agents at selected time steps. Notice some interesting similarities and differences in the clustering under repeated meetings and gossip communication: (a) the broad picture of number and location of evolving opinion clusters are quite similar, (b) in the example for consensus under repeated meetings one agent is crucial to bring them all together, (c) ε in the consensus example under gossip is higher as under repeated meetings, (d) some agents remain 'outliers' under gossip communication

$t = 56$

.29

.221



2.4. The mathematical models

ward.

lower corner on even pages Repeated meeting process with $[0, 1]^2$ as opinion space. Time goes forward when page numbers go downward.

lower corner on odd pages Gossip process with $[0, 1]^2$ as opinion space as opinion space. Time goes forward when page numbers go forward.

While the general tendency in the clustering after stabilization goes with rising ε from plurality via polarization to consensus, exceptions are possible. Figure 2.9 shows repeated meeting processes with the same initial opinion profile for successively higher ε . We notice especially that consensus occurs for $\varepsilon = 0.185$ while polarization occurs for the higher $\varepsilon = 0.2$ in contrast to the general trend. Structural reasons for this phenomenon are subject to Chapter 4.

The gossip process is sensitive to the realization of the random pairwise collection of communication partners. Figure 2.10 shows processes with the same initial opinion profile and the same bound of confidence but different random pairwise realizations of the communication regime. We observe polarization, polarization with a missed last chance for consensus, a vast majority consensus with a minor cluster and total consensus.

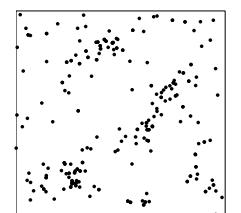
Figure 2.11 shows a phenomenon which happens in repeated meeting processes. A process with a regular equidistant initial opinion profile in $[0, 1]$ leads to three clusters while the same setting with a higher number of agents leads to the evolution of two very small mediator groups which bring the group to consensus but very slowly. An opinion profile which converges very slowly is called a *metastable* state. The evolution of a metastable state will turn out to be a generic feature of the repeated meeting process. This phenomena might have inspired Hegselmann to his conjecture in a footnote in [35] that one might find for an arbitrary low ε a certain very high number of agents such that the process must converge to consensus. The analysis and some evidence that the Hegselmann conjecture is wrong are subject to [55] which is also part of the dissertation.

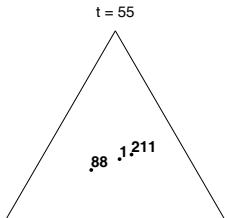
Phenomena get richer when heterogeneous bounds of confidence come in. Figure 2.12 shows two typical phenomena in repeated meeting processes. There we have a large group of closed-minded agents with a low ε and a small group of open-minded agents. This may lead to a situation where the open-minded agents are able to pull all the closed minded groups together to find a consensus but it may also happen that the open-minded get between two clusters of closed-minded agents where they get stuck ‘between the chairs’.

Interesting in the example where consensus is reached is that initially the open-minded agents had the opposite opinion compared to the final consensual opinion. (This is plausible even in a two agent example with different bounds of confidence.) There happened a severe drift to one side. This drifting is a generic possibility when heterogeneous bounds of confidence are involved.

Figure 2.13 shows the paradox fact that when drifting occurs in repeated meeting communication and leads to consensus, then the initially smaller group of closed minded agents has a bigger impact on the location of the final consensual value.

Finally, we show in Figure 2.14 two examples about what may happen in gossip dynamics under heterogeneous bounds of confidence. We take two equally sized groups of agents one with $\varepsilon = 0.22$ and the other with $\varepsilon = 0.11$. The figure shows two possible outcomes. One is that a vast majority reaches a consensus in





2. Modeling

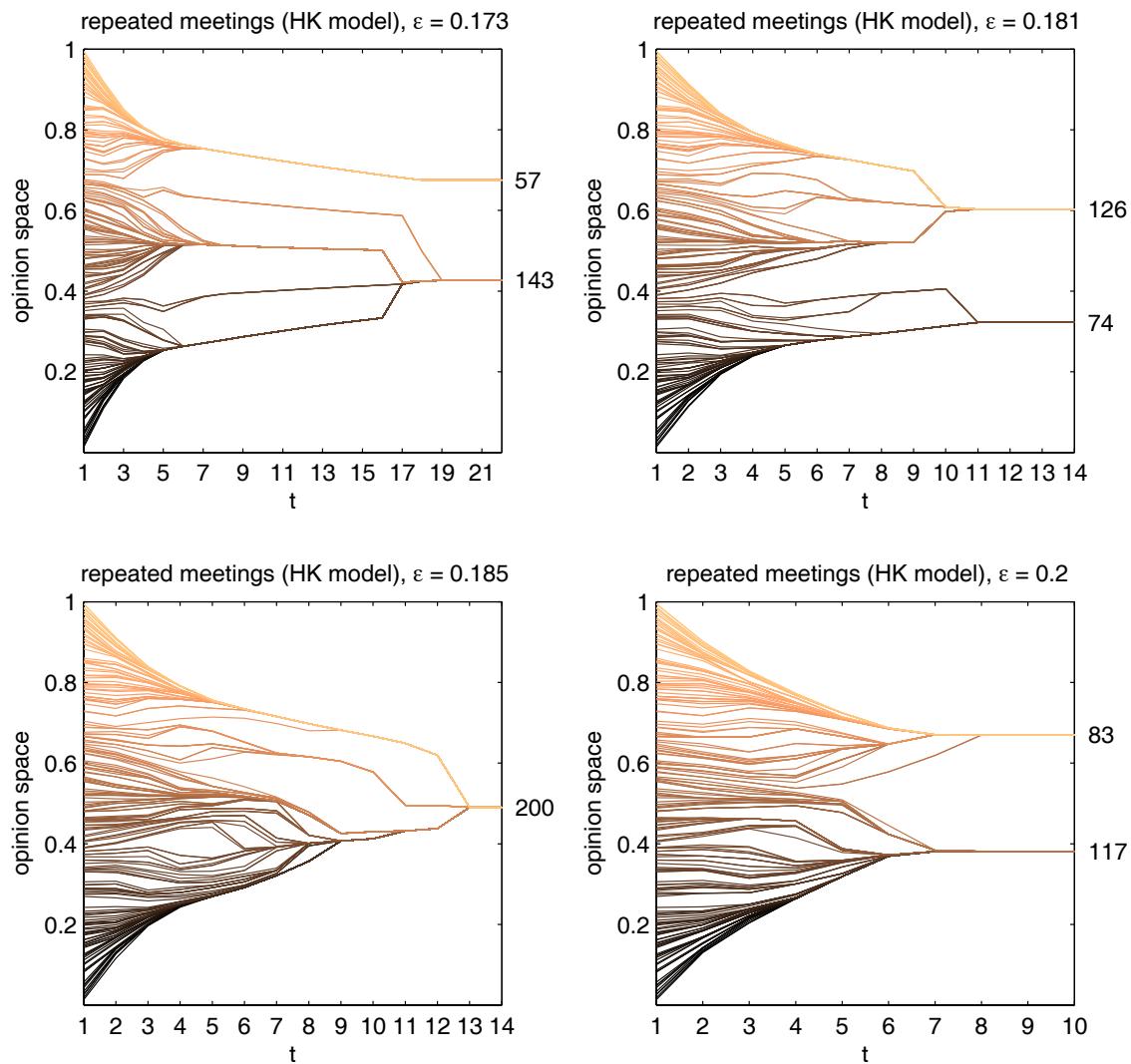
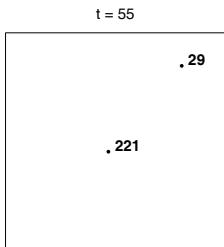
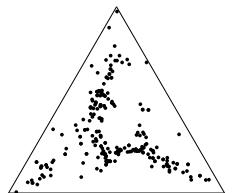


Figure 2.9: The HK model is very sensitive to changes in ε . These four runs do all start with the same initial profile $x(0) \in [0, 1]^{200}$. Notice that the number of evolving clusters and the size of the biggest cluster does not change monotonously with ε .





 2.4. The mathematical models

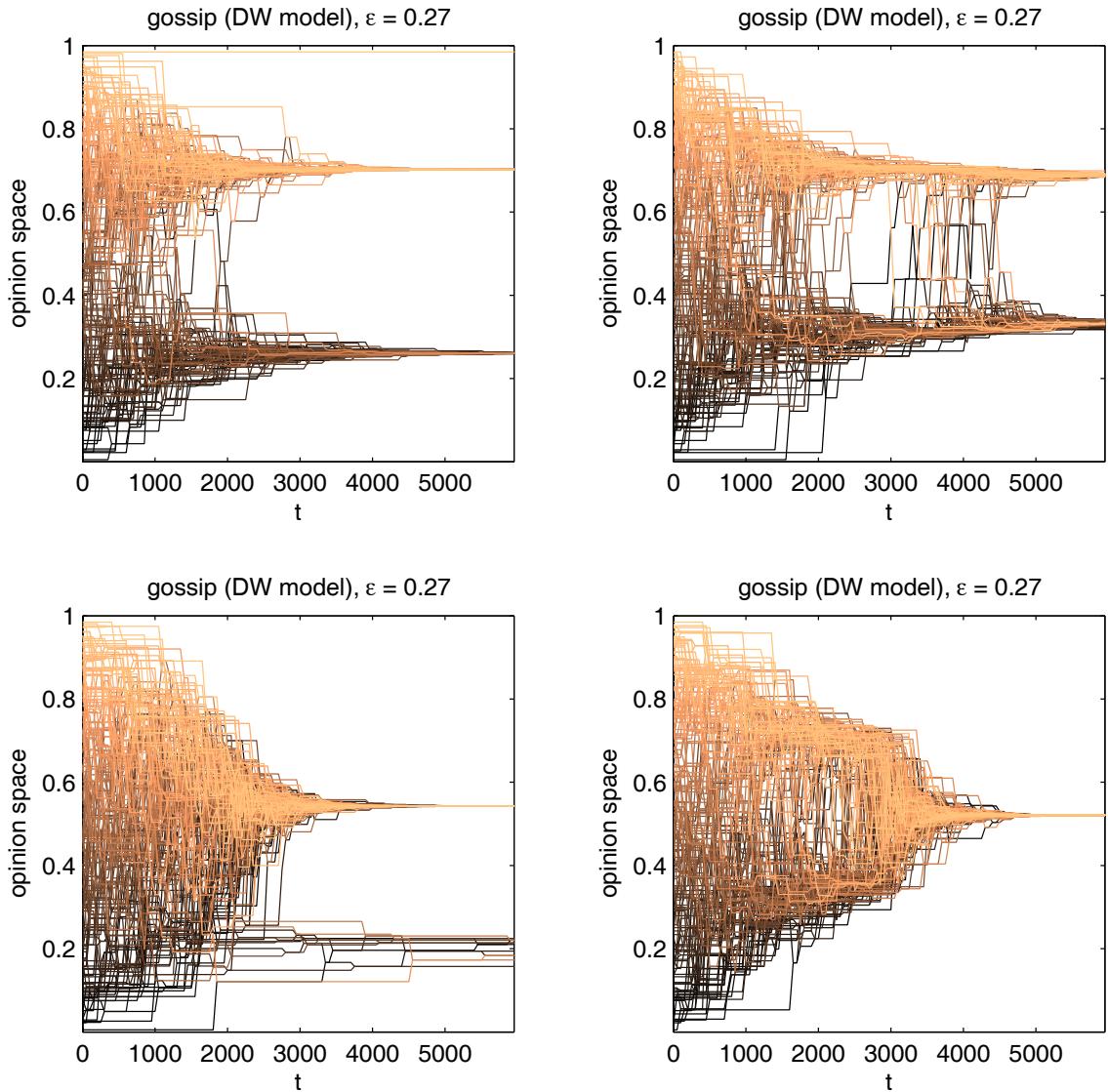
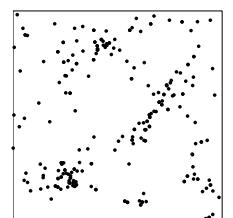
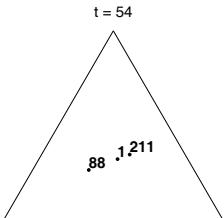


Figure 2.10: The DW model is very sensitive for the realization of the random pairwise communication regime. These runs have all been computed with the same ε and the same initial opinion profile $x(0) \in [0, 1]^{200}$





2. Modeling

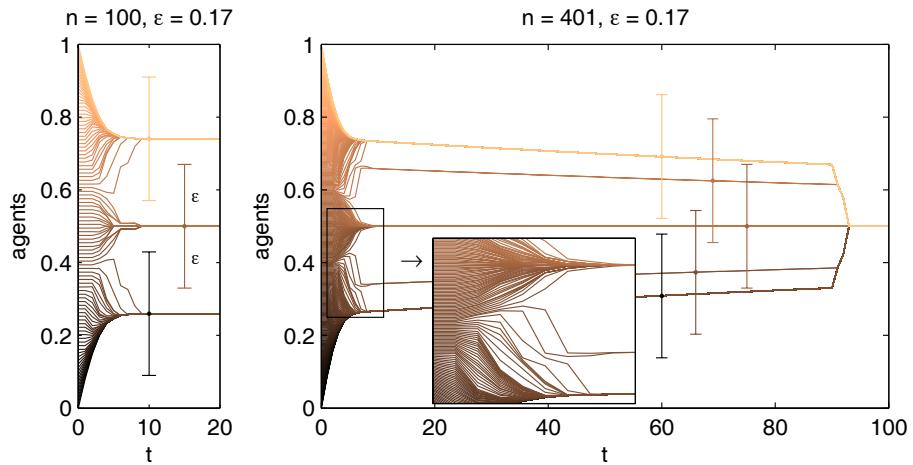


Figure 2.11: More agents can lead to situation where some mediators balance the group to consensus. This phenomenon occurs under repeated meetings. In both subplots the initial opinion profile is the equidistant profile in the opinion space $S = [0, 1]$ ($x_i := \frac{i-1}{n-1}$). For $n = 100$ the group splits into three clusters. For $n = 401$ there evolve small mediator group at each side of the central cluster which attracts the outer big groups in a very slow approximation process.

the middle and the other is that two big cluster evolve which both drift towards the same side. The surprise here is that if we study processes with homogeneous bound of confidence $\varepsilon = 0.22$ consensus never occurs under random pairwise communication. So, here the closed minded agents helped the society to build a vast consensual majority.

On the other hand we see that drifting also may happens in gossip processes with heterogeneous bounds of confidence.

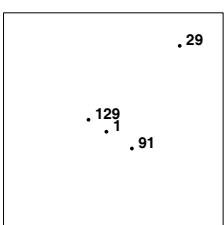
Another example of drifting under gossip is the bunch of work about propagation of extremism [3, 18, 20, 87]. In these models one starts with a stylized initial opinion profile with open-minded agents in the center and few closed-minded extremists with such low bounds of confidence that they play the role of a static group of attraction for the open-minded. Then, one can observe a central convergence of the open-minded agents, as well as convergence to both extremes or the drift of all central agents to one extreme.

2.4.4 Density-based bounded confidence models

We reformulate the homogeneous Hegselmann-Krause model 2.4.3 and homogeneous Deffuant-Weisbuch model 2.4.4 for a one-dimensional opinion space as density-based models with the same heuristics of repeated meetings and gossip. So, we switch from n agents with opinions in S to an idealized infinite population, which is divided into n opinion classes.

Let \underline{n} be a set of opinion classes and $p(0) \in \Delta^{n-1} \subset \mathbb{R}^n$ be an initial opinion distribution. A density based process is the interactive Markov chain (2.2) $p(t+1) = p(t)B(p(t))$ with the explicit definition of the transition matrix function $B(\cdot)$.

We need some preliminary definitions to define the transition matrix for the



2.4. The mathematical models

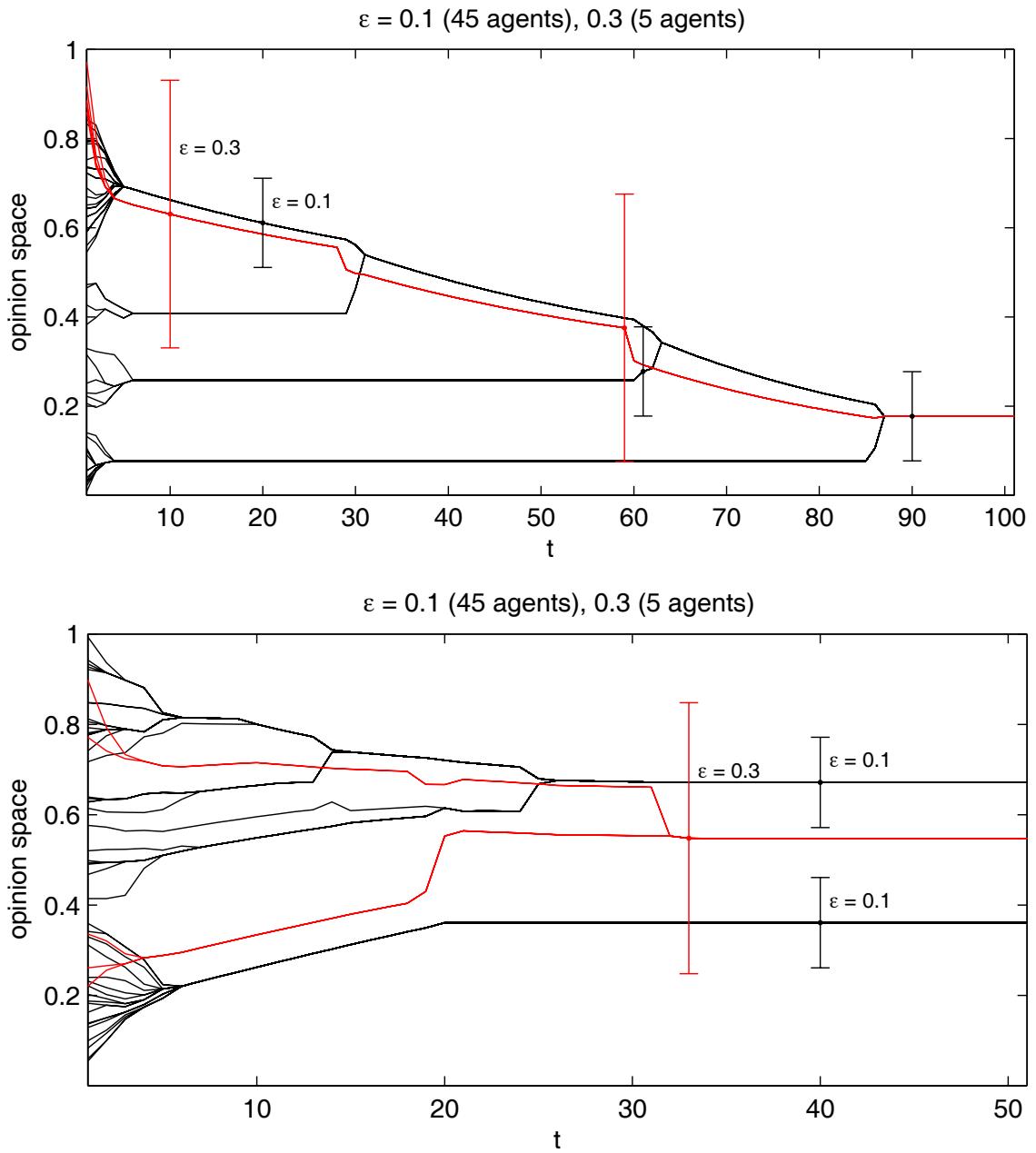
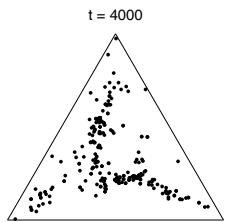
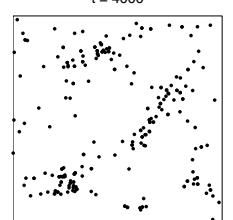
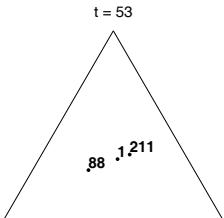


Figure 2.12: Two different possibilities what cluster formations may evolve when a small group of open-minded agents lives in a large group of closed minded agents. Consensus but with a severe drift to an extreme is possible as well as sitting between the chairs forever for the group of open-minded.

$t = 4000$





2. Modeling

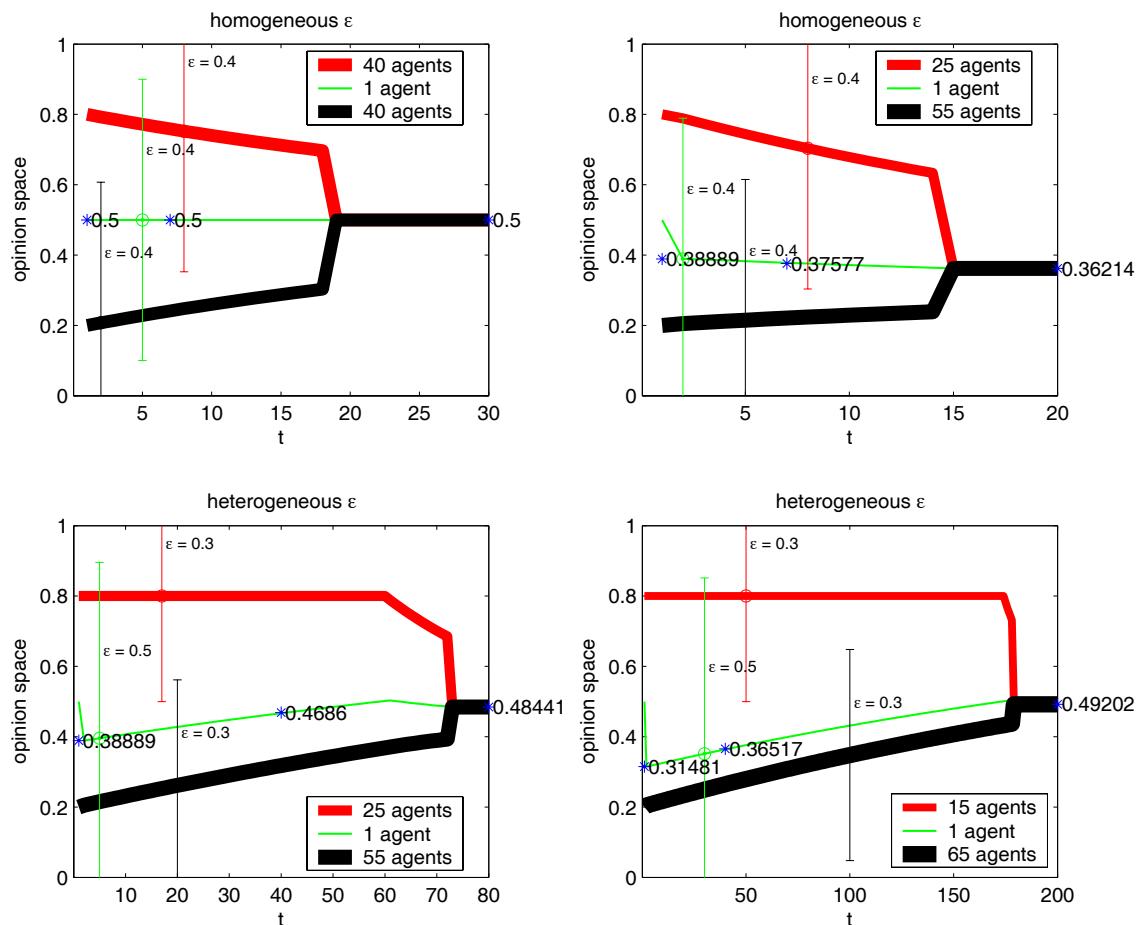
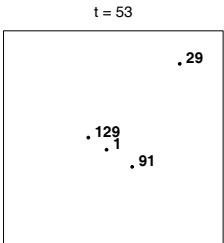
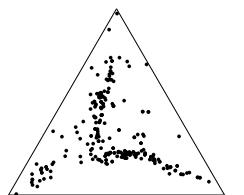


Figure 2.13: Under repeated meetings consensus often arises only due to a small central mediator group. This may also lead to significant drifting of the arithmetic mean of all opinions. In the homogeneous case there is only a small drift towards the biggest cluster. But in the heterogeneous case there can be a drift to the smaller cluster. This happens in random settings, too.





2.4. The mathematical models

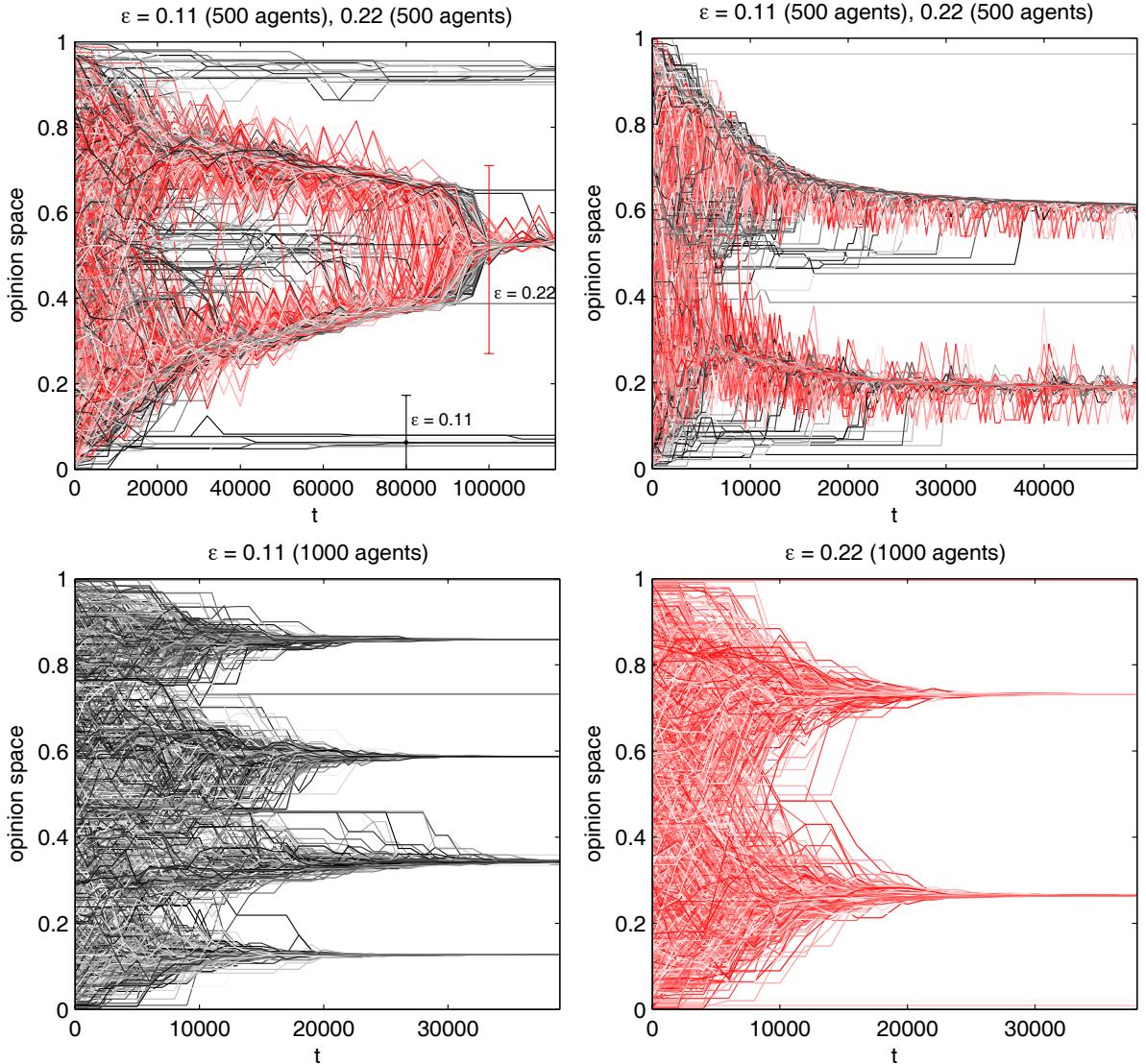
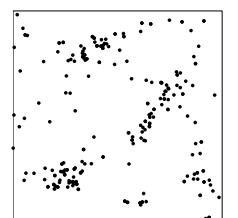
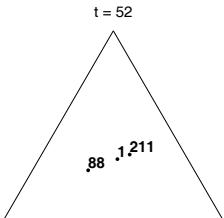


Figure 2.14: Two possible phenomena for gossip communication which come in when agents have heterogeneous bounds of confidence: Convergence of a vast majority to consensus (top left hand side), or a severe drift of opinions to one side (top right hand side). Interesting is especially that consensus is reached for $\varepsilon_1, \varepsilon_2$ which are to low to bring consensus under a homogeneous bound of confidence (bottom).





2. Modeling

interactive Markov chain with communication of repeated meetings like in the Hegselmann-Krause model.

Definition 2.4.5 (moments and barycenter in I). Let $I = \{i, +1, j\} \subset \underline{n}$ be a discrete interval and $p \in \Delta^{n-1}$ be an opinion distribution. We call

$$M_I^0(p) := \sum_{k \in I} p_k \text{ the } I\text{-mass (or 0th moment) of } p,$$

$$M_I^1(p) := \sum_{k \in I} kp_k \text{ the first } I\text{-moment of } p \text{ and}$$

$$M_I^{\text{bary}}(p) := \begin{cases} \frac{M_I^1(p)}{M_I^0(p)}, & \text{if } p_I \neq 0, \\ \frac{\max I + \min I}{2}, & \text{if } p_I = 0. \end{cases} \text{ the } I\text{-barycenter of } p.$$

Definition 2.4.6 (Hegselmann-Krause transition matrix). Let $p \in \Delta^{n-1}$ be an opinion distribution and $\epsilon \in \mathbb{N}$ be a discrete bound of confidence. For $i \in \underline{n}$ we abbreviate the ϵ -local mean as

$$M_i := M_{\{i-\epsilon, +1, i+\epsilon\}}^{\text{bary}}(p)$$

We define the *Hegselmann-Krause transition matrix* as

$$B_{ij}^{\text{HK}}(p, \epsilon) := \begin{cases} 1 & \text{if } j = M_i, \\ \lceil M_i \rceil - M_i & \text{if } j = \lfloor M_i \rfloor, j \neq M_i, \\ M_i - \lfloor M_i \rfloor & \text{if } j = \lceil M_i \rceil, j \neq M_i, \\ 0 & \text{otherwise.} \end{cases}$$

Each row of the transition matrix $B^{\text{HK}}(p, \epsilon)$ contains only one or two adjacent positive entries. The population with opinion i goes completely to the ϵ -local mean opinion if this is an integer. Otherwise they distribute to the two adjacent opinions. The fraction which goes to the lower (upper) opinion class depends on how close the ϵ -local mean lies to it. Thus, the heuristic of averaging all opinions in a local area is represented here.

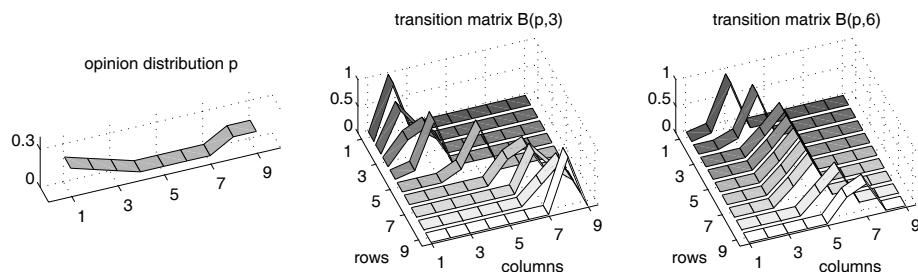
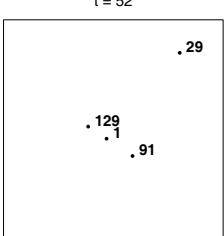
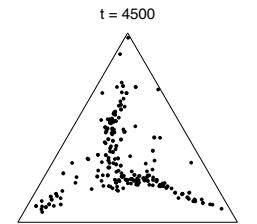


Figure 2.15: Visualizations of Hegselmann-Krause transition matrices for a given p with $n = 9, \epsilon = 3, 6$.

Figure 2.15 shows two examples. Each ribbon is one row of the transition matrix. Thus, the i -th ribbon gives the distribution vector how the mass p_i will distribute to the classes in the next step. The new population mass of class i computes as the scalar product of p and the i -th column (so cross the ribbons) of $B(p, \epsilon)$.





2.4. The mathematical models

Definition 2.4.7 (Deffuant-Weisbuch transition matrix). The *Deffuant-Weisbuch transition matrix* for an opinion distribution $p \in \Delta^{n-1}$, a discrete bound of confidence $\epsilon \in \mathbb{N}$ and a cautiousness parameter $0 < \mu \leq 0.5$ is defined by

$$B_{ij}^{\text{DW}}(p, \epsilon, \mu) := \begin{cases} \sum_{\{k \in \underline{n} \mid i + \mu(k-i) \in]j-1, j+1[\}} (1 - |i + \mu(k-i) - j|) \pi_k^i, & \text{if } i \neq j, \\ q_i, & \text{if } i = j. \end{cases}$$

with $q_i = 1 - \sum_{j \neq i, j=1}^n B_{ij}^{\text{DW}}(p, \epsilon, \mu)$ and

$$\pi_m^i := \begin{cases} p_m, & \text{if } |i - m| \leq \epsilon \\ 0, & \text{otherwise} \end{cases}$$

Remember that we defined $p_i = 0$ for all $i \notin \underline{n}$. For the most common setting $\mu = 0.5$ the definition simplifies to

$$B_{ij}^{\text{DW}}(p, \epsilon) := \begin{cases} \frac{\pi_{2j-i-1}^i}{2} + \pi_{2j-i}^i + \frac{\pi_{2j-i+1}^i}{2}, & \text{if } i \neq j, \\ q_i, & \text{if } i = j. \end{cases}$$

Nearly always in the following only the standard value $\mu = 0.5$ is studied. The general definition here is more a reference for future work.

We briefly describe how the agent-based gossip heuristics of the Deffuant-Weisbuch model 2.4.4 governs the transition matrix of the interactive Markov chain for the case $\mu = 0.5$. By the founding idea of the model an agent with opinion i moves to the new opinion j if he compromises with an agent with opinion $i + 2(j-i) = 2j-i$. The probability to communicate with an agent with opinion $2j-i$ is of course p_{2j-i} . Thus, the heuristic of random pairwise interaction is represented. The terms $\frac{\pi_{2j-i-1}^i}{2}, \frac{\pi_{2j-i+1}^i}{2}$ stand for the case when agents with opinion i communicate with agents with opinion j , but the distance $|i-j|$ is odd. In this case the population should go with probability $\frac{1}{2}$ to one of the two possible opinion classes $\lfloor \frac{i+j}{2} \rfloor, \lceil \frac{i+j}{2} \rceil$. For $\mu < 0.5$ it can also be checked with some care that the agent-based dynamics govern the transition matrix with similar arguments.

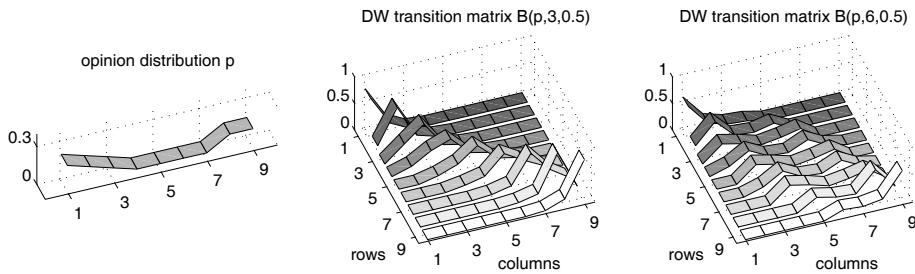
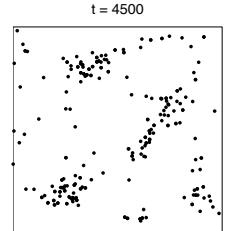
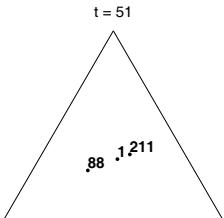


Figure 2.16: Visualizations of Deffuant-Weisbuch transition matrices for a given p with $n = 9, \epsilon = 3, 6$.

Figure 2.16 shows the same examples as figure 2.15 but for the Deffuant-Weisbuch transition matrix.

Now, we define coupled interactive Markov chains for populations of agents with heterogeneous bounds of confidence.





2. Modeling

Model Definition 2.4.8 ({DW,HK} Interactive Markov chain with heterogeneous bounds of confidence). Let n be opinion classes. Let $\epsilon_1, \dots, \epsilon_m \in \mathbb{N}$ be discrete bounds of confidence and $p^1(0), \dots, p^m(0)$ be positive vectors with $p(t) := \sum_{j=1}^m p^j(t) \in \Delta^{n-1}$. (p_i^j represents the fraction of agents which hold opinion i and have bounds of confidence ϵ_j .) Let $\text{CR} \in \{\text{HK}, \text{DW}\}$ determine the communication regime which governs the transition matrix. The {DW,HK} interactive Markov chain with heterogeneous bounds of confidence is the set of sequences $(p^1(t))_{t \in \mathbb{N}}, \dots, (p^m(t))_{t \in \mathbb{N}}$ recursively defined through the m coupled equations

$$\begin{aligned} p^1(t+1) &= p^1(t)B^{\text{CR}}(p(t), \epsilon_1) \\ &\vdots \\ p^m(t+1) &= p^m(t)B^{\text{CR}}(p(t), \epsilon_m). \end{aligned}$$

(For the DW transition matrix the cautiousness parameter μ has to be added in the argument of the transition matrix function.)

The interactive Markov chain under homogeneous bound of confidence appears in this definition as special case with $m = 1$.

2.4.5 Convergence for large numbers of agents and classes

Here we discuss the question of the 'equality' of the agent-based models and their density-based counter parts in the limit for large n .

Let us consider the agent-based Hegselmann-Krause model 2.4.3 with $n \in \mathbb{N}$ agents and opinion space $S = [0, 1]$ and let us consider the interactive Markov chain 2.4.8 with Hegselmann-Krause transition matrix 2.4.6 with n opinion classes. (Arguments hold analog for the Deffuant-Weisbuch model.)

We imagine that the opinion classes $1, \dots, n$ are representatives for an equidistant partition of the interval $[0, 1]$ in the way $i \leftrightarrow [\frac{i-1}{n}, \frac{i}{n}]$. Let us consider an initial opinion profile $x(0) \in [0, 1]^n$ and the initial opinion distribution $p(0) \in \Delta^{n-1}$. In the limit for large n the opinion distribution converges to a probability density on the opinion space $[0, 1]$. On the other hand the opinion profile $x(0) \in [0, 1]^n$ can be seen as a sample of a probability density function which can be better and better estimated in the limit for large n .

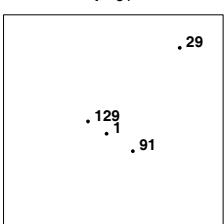
Due to this arguments we say that the agent- and the density-based model 'converge' to each other in the limit for large n .

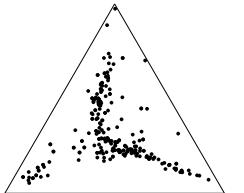
If the agents in the agent-based model have a homogeneous bound of confidence ε then the interactive Markov chain and the agent-based model 'converge' to each other when $\epsilon, n \rightarrow \infty$ but $\frac{\epsilon}{n} \rightarrow \varepsilon$. This is not a formal proof, just heuristic speculation which turn out to be plausible through simulation.

Under heterogeneous bounds of confidence similar convergence arguments are applicable.

2.4.6 Density-based examples and phenomena

Some examples should help to get a feeling about the density-based bounded confidence dynamics and show that the claim of convergence of the agent-based and the density-based model in the limit for large n in subsection 2.4.5 turns out to lead to similar behavior even not too large numbers of n .





2.4. The mathematical models

Like in the agent-based simulation we make the assumption that agents are distributed equally in the opinion space. The initial opinion distribution is thus $p_i(0) = \frac{1}{n}$.

First, we study the accuracy parameter n itself. Therefore, we choose ϵ such that $\epsilon = \frac{\varepsilon}{n} = 0.2$ which is an interesting value that has shown to lead sometimes to consensus and sometimes not. Figures 2.17 and 2.18 show the HK and the DW interactive Markov processes for $n = 10, 15, 20, 25$.

The first important thing to mention is that convergence seem to happen always in the density-based models, too.

Interesting is especially the HK model. Where we see convergence to consensus for low n , polarization for $n = 20$, but consensus again for $n = 25$. It turns out that n odd leads to different behavior than n even. This comes through the fact that for odd n the center consists of one class while for n even it consists of two adjacent classes. The mass in two classes is already splitted, while the mass in one class may not split, because it is the nature of the HK process that all agents in one class move in the same way. So, an odd n fosters consensus. This holds also for larger n (see [53] for a study). The effect is not yet known for the DW model and Figure 2.18 gives no hint, but it might be there.

In the following we will choose odd numbers of n because dynamics have shown to be closer to agent-based random settings in simulation. The odd n will sometimes lead to slight differences when we compare $\frac{\varepsilon}{n}$ in a density-based simulation (e.g. $\frac{200}{1001}$) to ε in an agent-based simulation (e.g. 0.2).

Next, we reproduce plurality, polarization and consensus as final distribution for different values of ϵ in the two density-based models in Figure 2.19. So, we resemble Figure 2.8 from the agent-based examples. In the agent-based example we had 50 agents, here we have 51 opinion classes. There are some slight differences to the agent-based setting. The HK model leads to consensus for $\frac{\varepsilon}{n} \approx 0.2$ in the density-based setup, which it did not in the agent-based simulation. But convergence is slow and via a metastable state. The DW model leads to consensus for $\frac{\varepsilon}{n} \approx 0.3$, which it does not in the agent-based example.

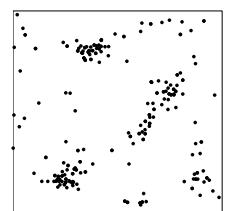
This is a hint that the density-based setup leads more easily to consensus than the agent-based setup. But on the other hand they ‘converge’ to each other with rising n . A simulation analysis of this hypothesis is done for the HK model in [55] which is also part of the dissertation. It delivers some evidence for the convergence and the trend that rising n fosters chances for consensus in the agent-based setting and weakens chances for consensus in the density-based setting.

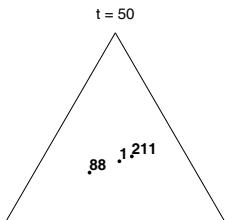
Figure 2.20 resembles the agent-based example in Figure 2.11. It is a more drastic example of an evolving metastable state.

Figure 2.21 shows a DW interactive Markov chain in detail. Interesting is especially that there evolve minor clusters at the extremes and also the central cluster is not visible without zooming. This shows that there is a structural reason for the frequently evolving outliers in gossip processes.

Next, we check if the features of drifting and sitting between the chairs which we discovered in agent-based simulation also occur in density-based simulation under heterogeneous bounds of confidence. Figure 2.22 gives an example where the HK interactive Markov chain leads to a sitting between the chairs situation for the open-minded agents.

The surprising agent-based example in Figure 2.14 where a heterogeneous gossip process leads to consensus although both bounds of confidence are low is





2. Modeling

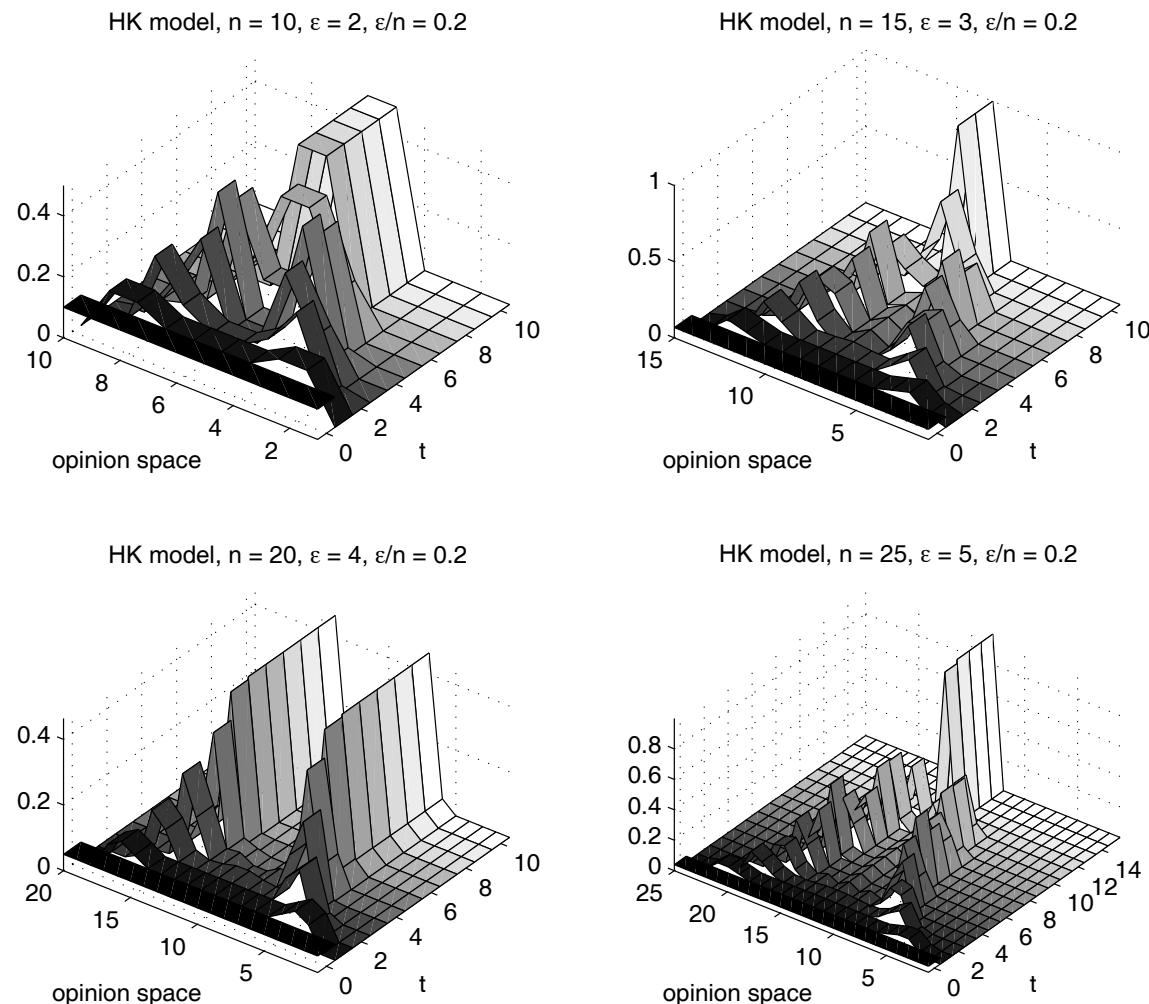
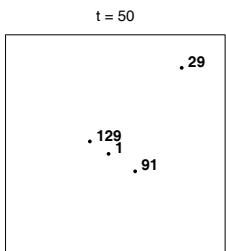
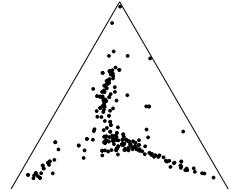


Figure 2.17: HK interactive Markov chains with uniform initial opinion distribution for $n = 10, 15, 20, 25$ and ϵ such that $\frac{\epsilon}{n} = 0.2$. Interesting is the difference in the number of evolving opinion peaks. This is a hint that odd n fosters consensus.





 2.4. The mathematical models

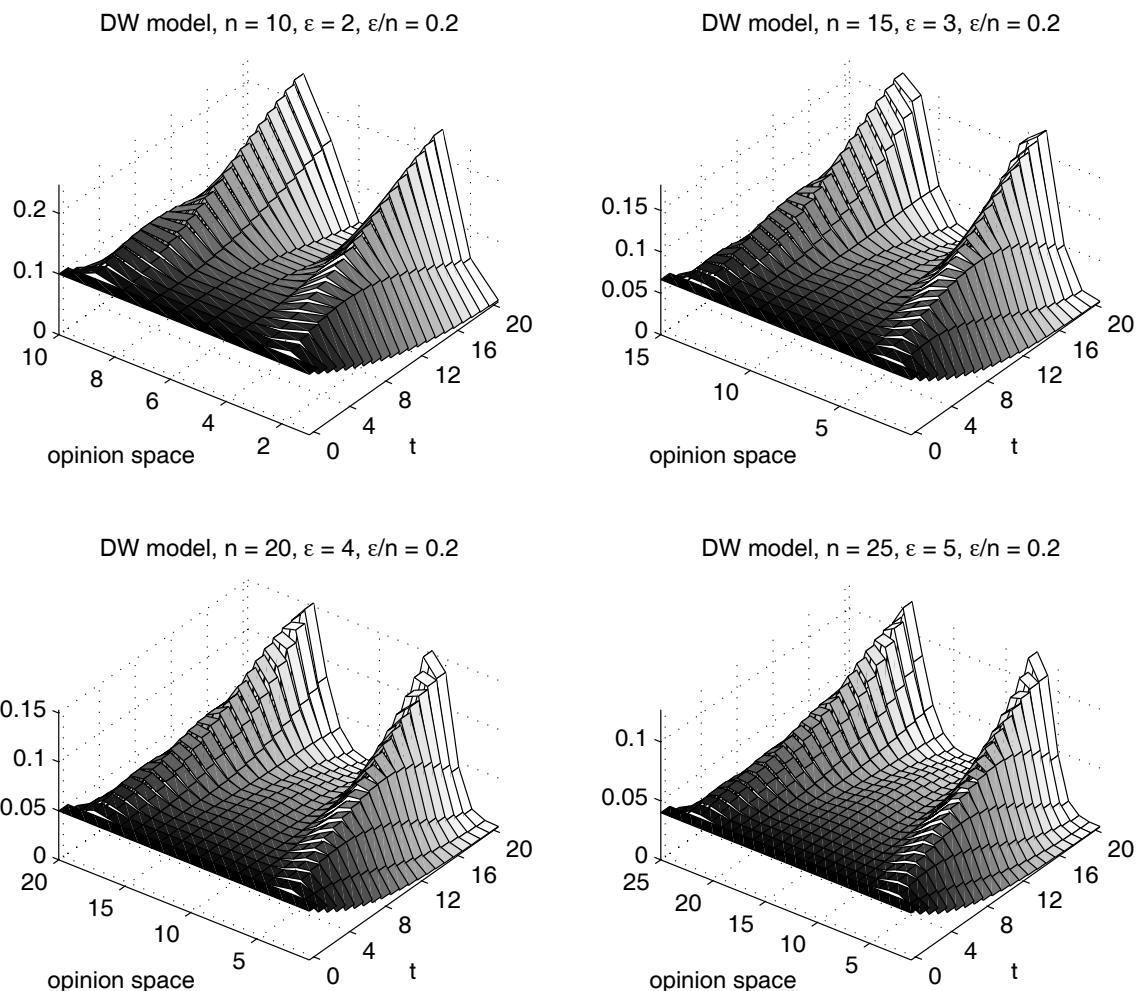
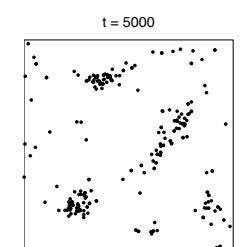
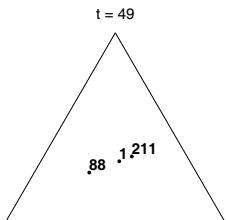


Figure 2.18: DW interactive Markov chains with uniform initial opinion distribution for $n = 10, 15, 20, 25$ and ε such that $\frac{\varepsilon}{n} = 0.2$.





2. Modeling

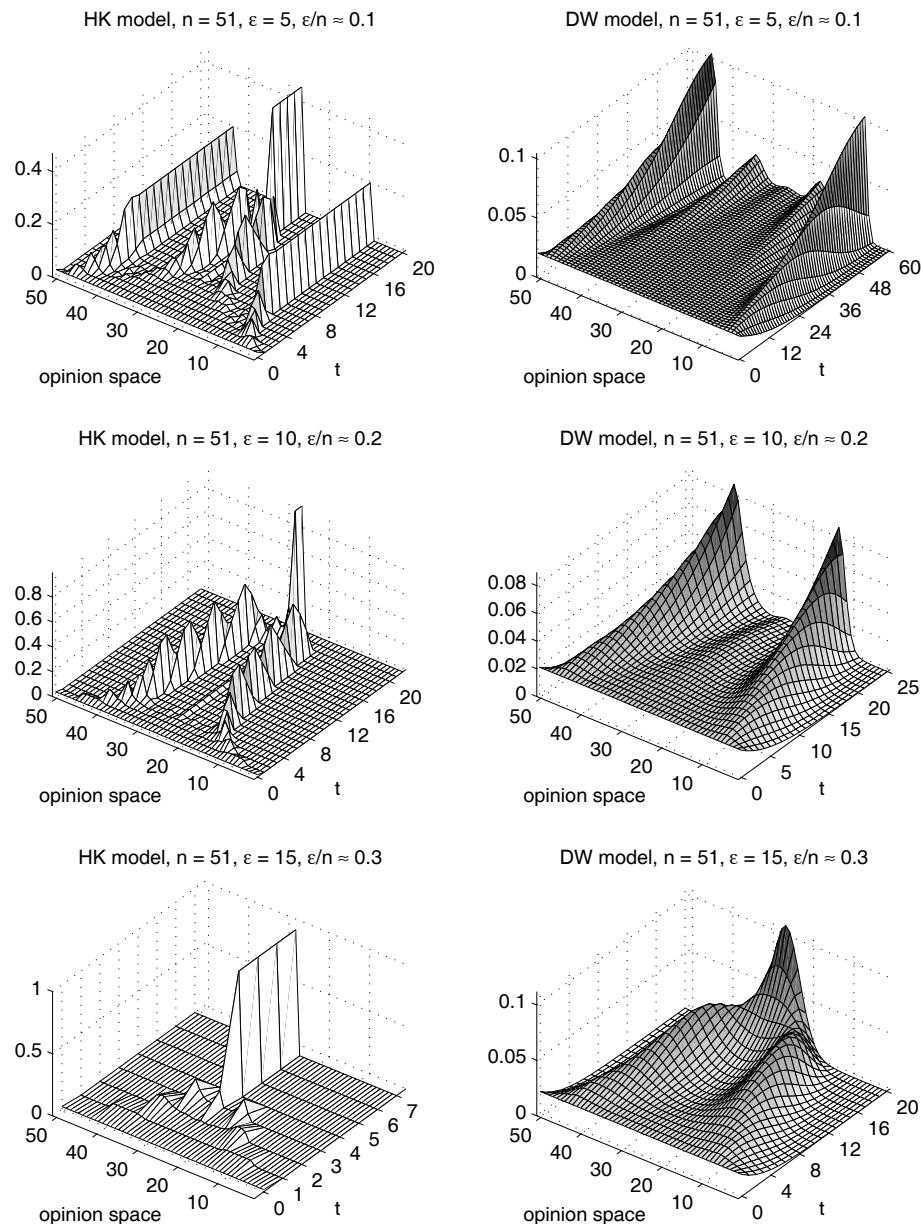
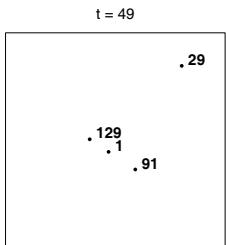


Figure 2.19: Plurality, Polarization and Consensus in the interactive Markov chains.



2.4. The mathematical models

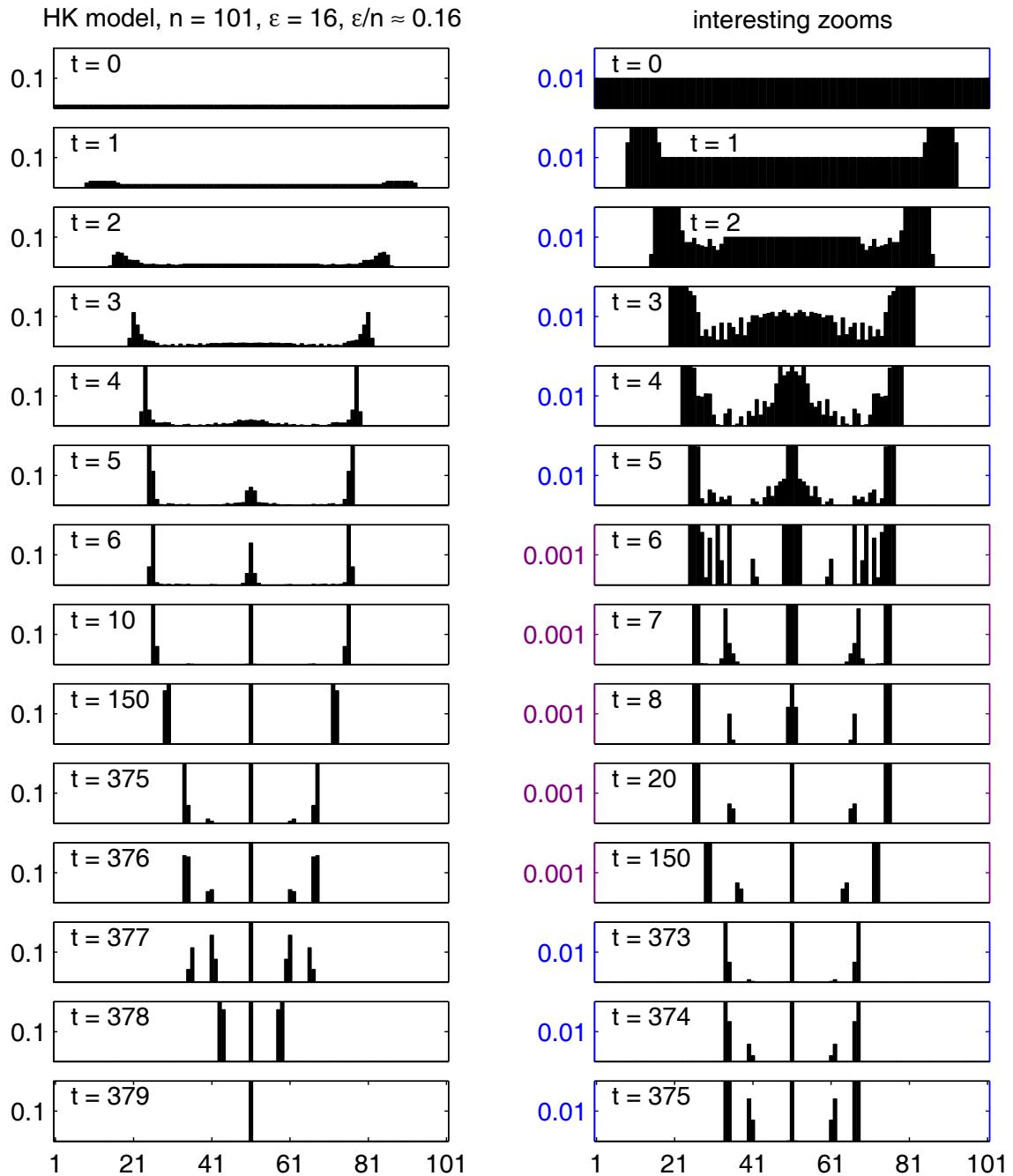
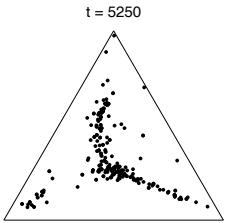
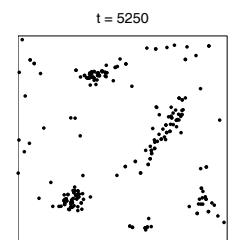
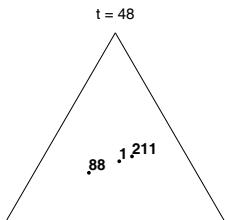


Figure 2.20: A drastic example for an evolving metastable state in a HK interactive Markov chain for interesting time steps.





2. Modeling

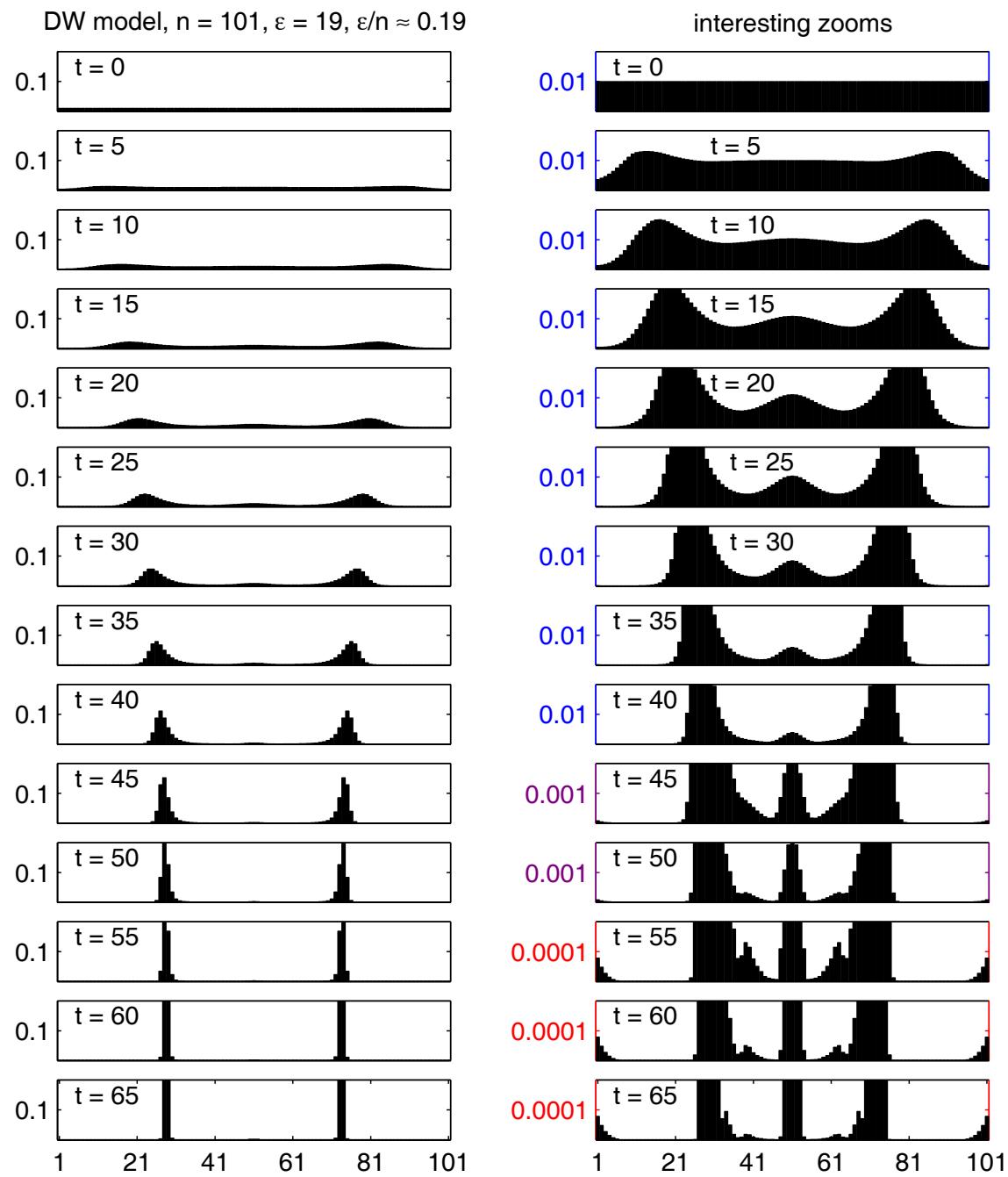
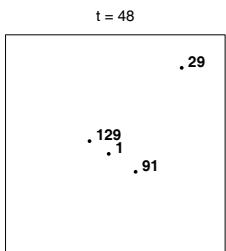
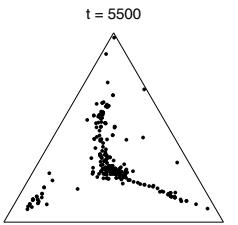
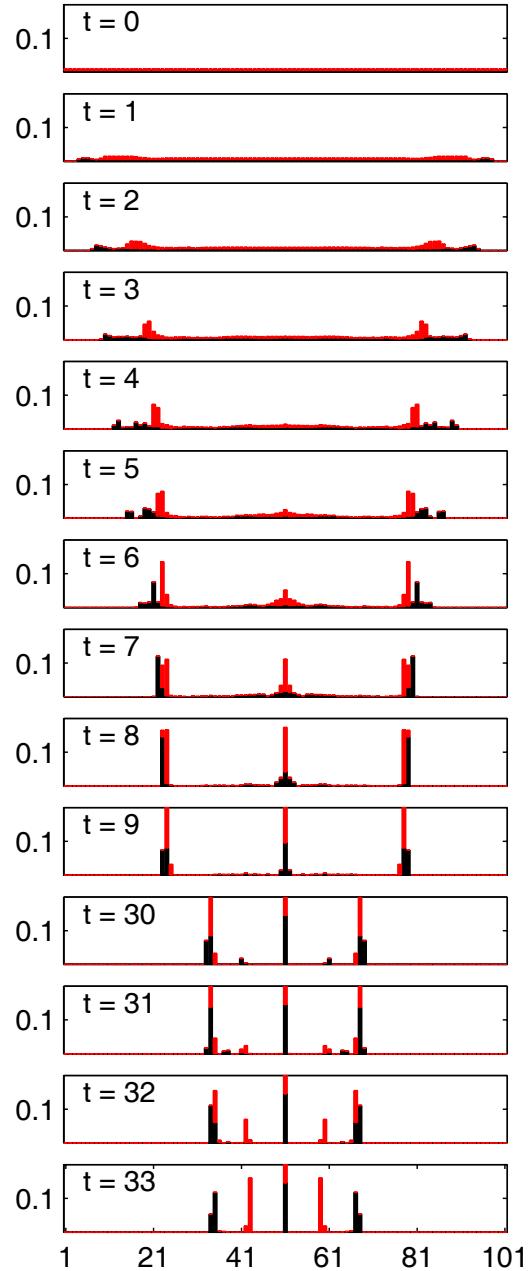


Figure 2.21: Evolution of minor clusters in the DW interactive Markov chain.





HK model, $n = 101$, $\varepsilon = 6, 16$, $\varepsilon/n \approx 0.06, 0.16$



interesting zooms

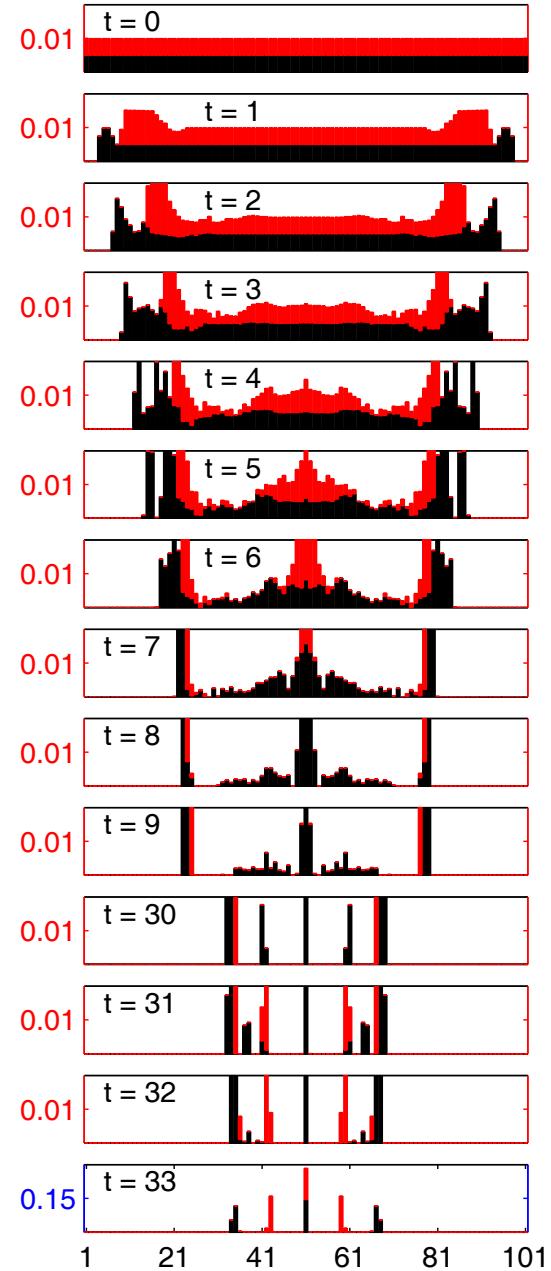
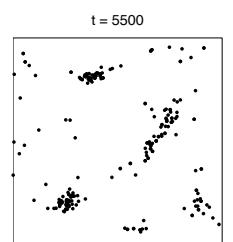
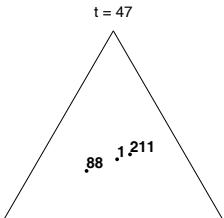


Figure 2.22: Sitting between the chairs in the HK model. Half of the population is open-minded, the other half is closed-minded.





2. Modeling

resembled for the DW interactive Markov chain in Figure 2.23.

Finally, we drop the uniform distribution and start with noisy initial opinion distributions to check if severe drifting to one side may also happen in density based settings under heterogeneous bounds of confidence. Figure 2.24 shows an example where a small population (20%) of open-minded agents pulls successively the whole population towards one side in a heterogeneous HK interactive Markov chain. This is quite similar to the agent-based example in Figure 2.12.

A similar drifting effect is shown in Figure 2.25 for the DW interactive Markov chain with heterogeneous bounds of confidence with half of the population closed-minded and the other half open-minded.

It should be emphasized that the random initial opinion distributions in Figures 2.24 and 2.25 are not styled to deliver a drift (as in the models of extremism propagation [3, 18, 20, 87]). Severe drifting is a generic feature which may happen under heterogeneous bounds of confidence when the initial condition is not totally uniform.

2.4.7 Continuous opinion dynamics and swarms

There is an appealing similarity of continuous opinion dynamics under bounded confidence and collective motion e.g. in animal swarms, flocks and herds.

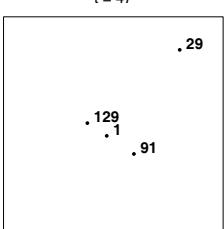
In individual-based biological models which try to simulate collective motion of animals [15, 66] different areas of perception around each individual are proposed. These areas are in analogy to an area of confidence in multidimensional continuous opinion dynamics. There are three areas around each individual. The *repulsion* area is the smallest area. The individual steers away from the barycenter of the individuals in its area of repulsion to avoid collisions. Larger is the *alignment* area. The individual adjusts its heading to the average heading of all individuals in the area of alignment. This is the closest analogy to continuous opinion dynamics. Even larger is the *attraction* area. Individuals try to steer towards the barycenter of the agents within their area of attraction to get not disconnected from the group. Often noise on the headings is added.

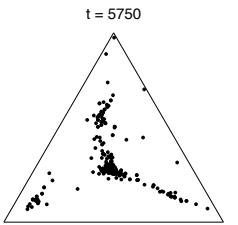
A group of individuals moving with this rules may lead to different attractive behavior, e.g. a swarm like bees where individuals remain cohesive but with no common direction and no movement of the whole swarm, collective motion like a flock of birds in a common direction or movement in a torus like a snake chasing their own tail which is sometimes observed in huge schools of fish. Which behavior is reached is mainly determined by the sizes of the areas of repulsion, alignment and attraction. Thus, some biologists and physicists started to analyze 'phase diagrams' of cohesive and collective motion [32, 84].

Aligning individuals try to average their headings under bounded perception. This is exactly the same heuristic as in the Hegselmann-Krause model of continuous opinion dynamics under bounded confidence. But also repulsion and attraction has to do with averaging because a barycenter has to be computed to find the attractive or repulsive point.

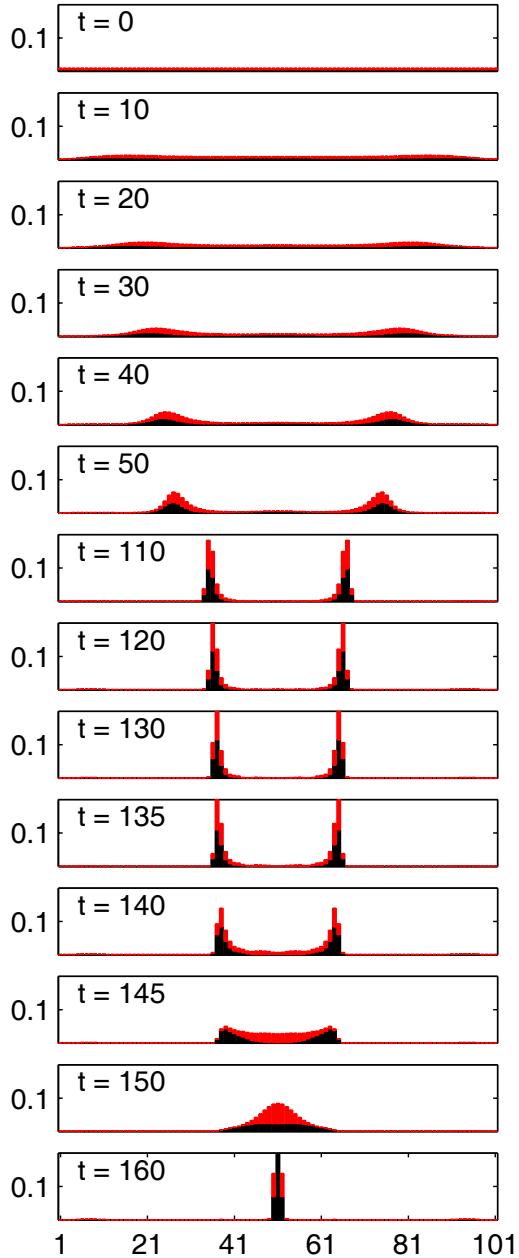
But there are two important issues in which swarm dynamics differ from continuous opinion dynamics as presented here.

First, the opinion for aligning individuals is their heading. For individuals moving in two-dimensional space this is an angle $\alpha \in [0, 2\pi]$ with 2π set equal to zero. So the opinion space has no borders and there are situations where averaging is not uniquely definable in a satisfactory manner. Consider two individuals –





DW model, $n = 101$, $\varepsilon = 11, 22$, $\varepsilon/n \approx 0.11, 0.22$



interesting zooms

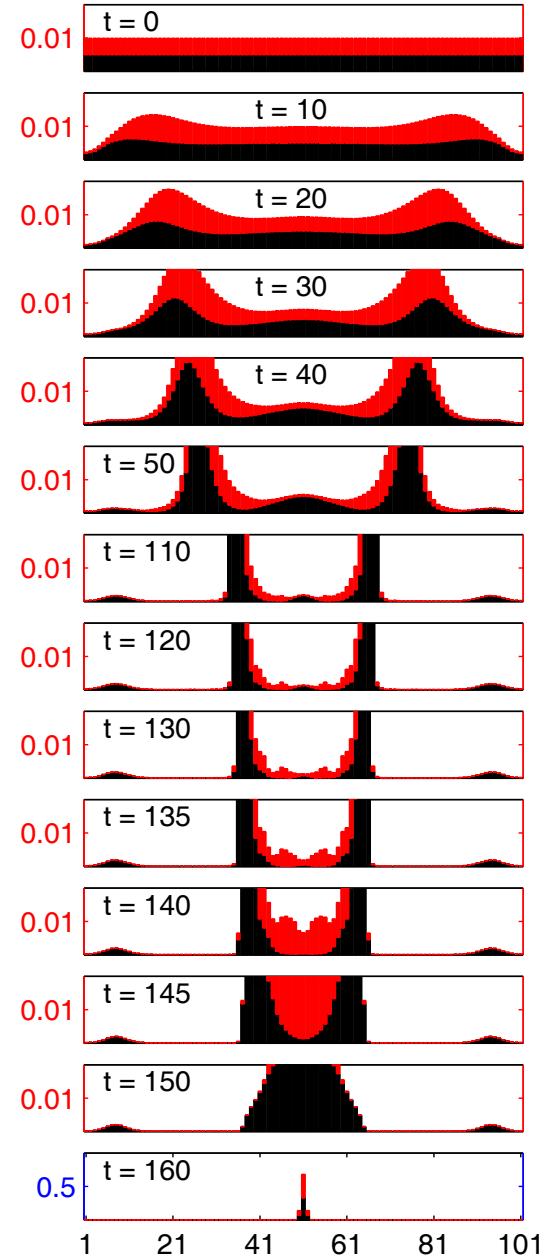
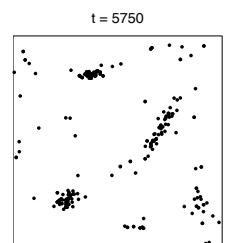
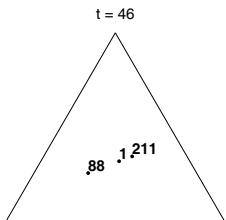


Figure 2.23: Heterogeneous agents reaching consensus, although both bounds of confidence are lower than necessary to produce consensus under a heterogeneous setting.





2. Modeling

HK model, $n = 101$, $\varepsilon = 10,30$, $\varepsilon/n \approx 0.1, 0.3$

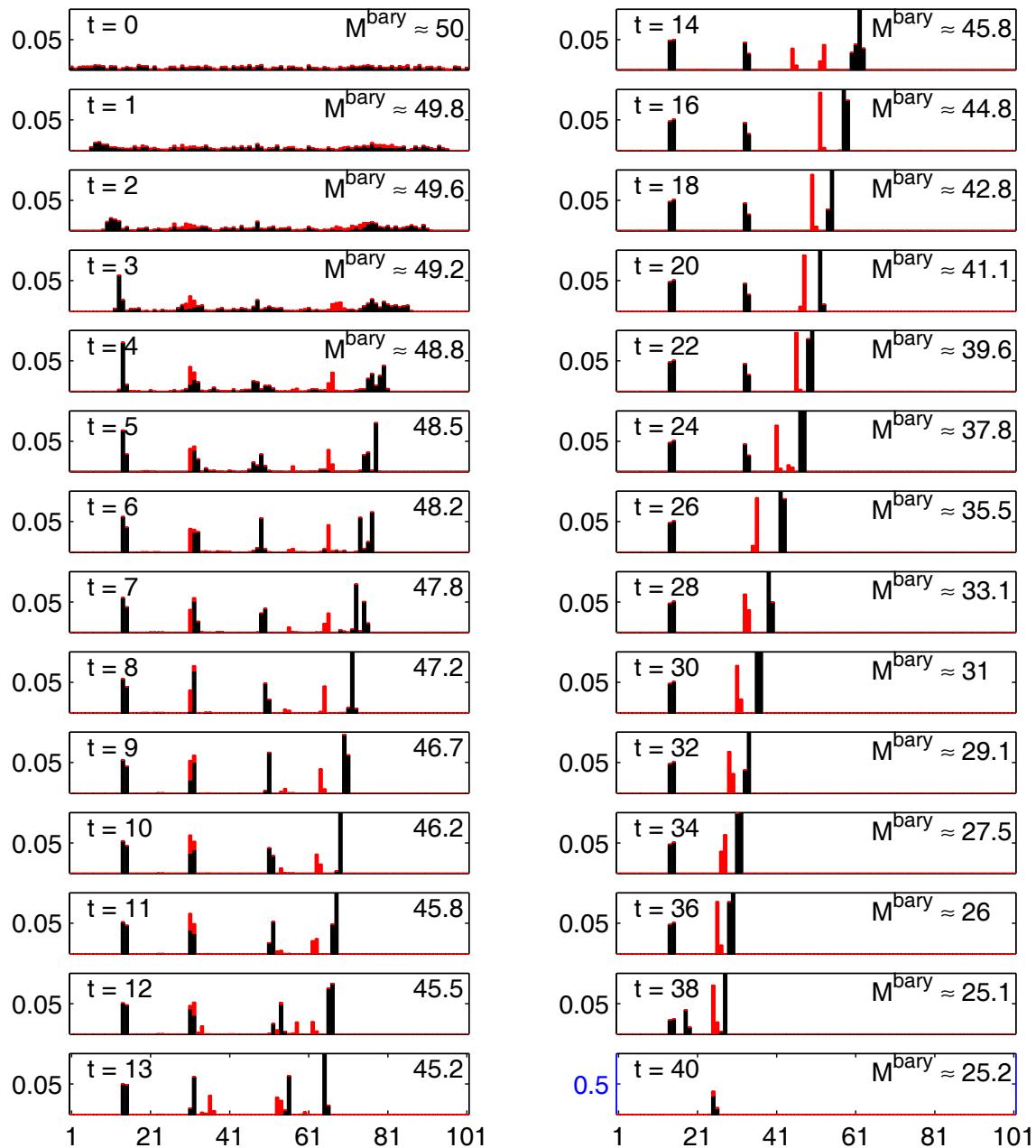
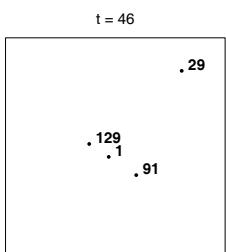


Figure 2.24: Drifting in the HK model. 20% open-minded and 80% closed-minded.



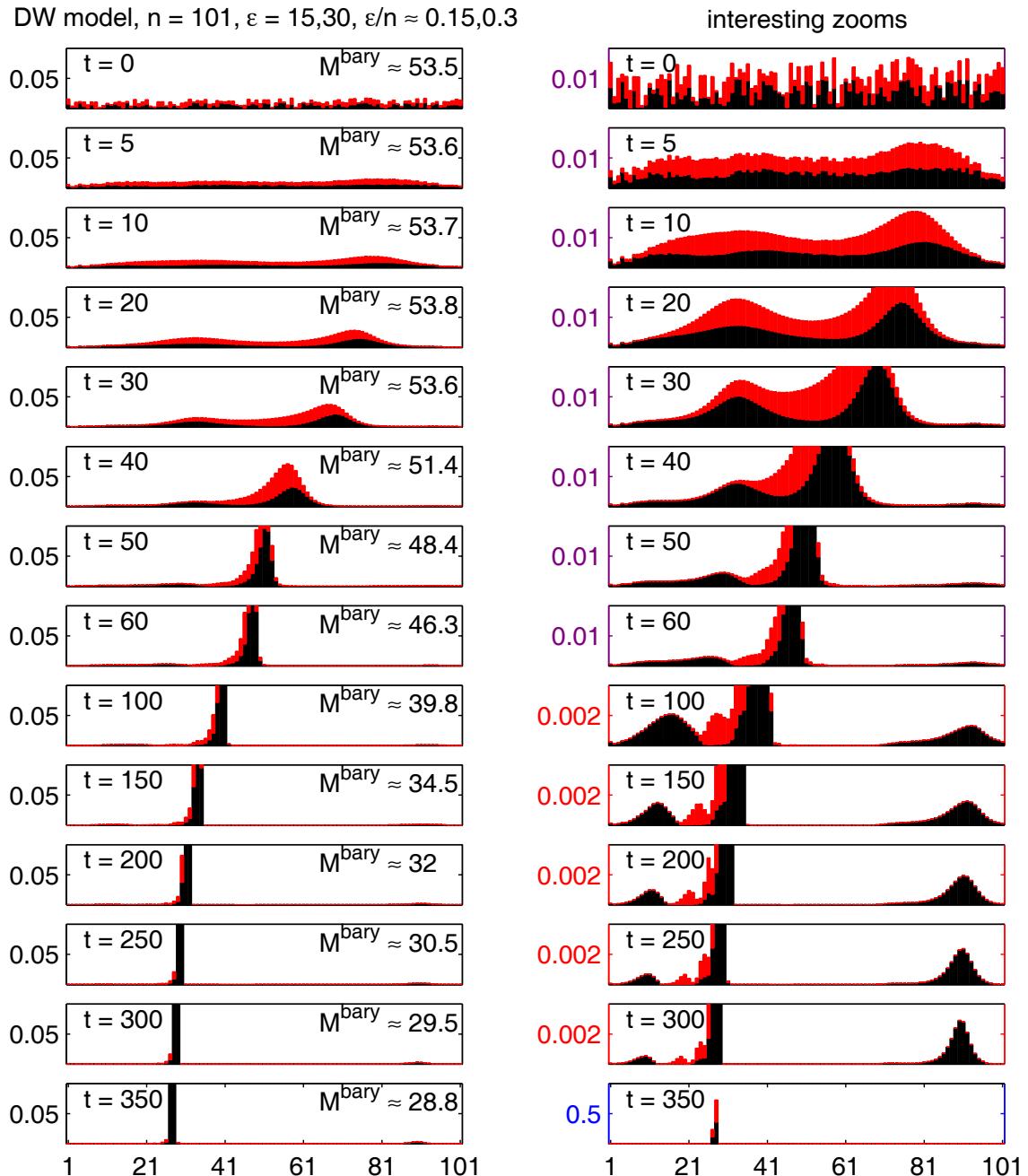
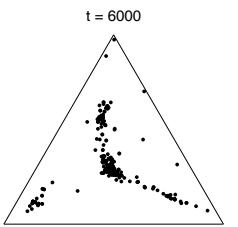
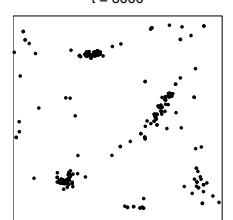
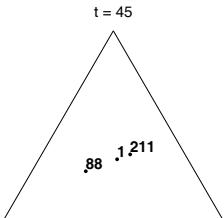


Figure 2.25: Drifting in the DW model. 50% open-minded and 50% closed-minded.

 $t = 6000$ 



2. Modeling

one moving east, one moving west. Then the compromise could either be north or south. An idea to circumvent this is to model not headings but velocity vectors. Then the individuals in the example would compromise by stopping (if they move at equal speed). But this is never done in the present swarm models because moving at a constant speed might be one important ingredient which produces such rich behavior like moving in a torus.

Second, individuals in a swarm do not have only an opinion (their heading) but also a position in space. Thus, there is one additional ingredient. Bounded perception is determined in the position space, while the averaging takes place in the heading space. This is in contrast to opinion dynamics where both are applied in the opinion space. The analysis of the interplay between dynamics in these two spaces is not well understood and a fruitful field for further research.

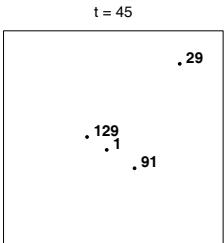
2.5 Some related models and some references

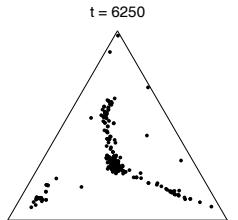
Additional to the references in the text, the basic HK and the DW model have been extended by the authors themselves [3, 20, 35, 38, 39, 86, 87] and others [23, 44, 45, 49, 81, 82, 83] including several papers of physicists in the International Journal of Modern Physics C [4, 24, 25, 26, 27, 42, 65, 75]. The rich clustering behavior of continuous opinion dynamics might have attracted the sociophysics community. In the meanwhile the number of publications on continuous opinions catches up with the work on discrete opinion dynamics inspired by physical spin-systems.

Kurz et al [49] used the Hegselmann-Krause model to define a game where political parties send campaigners out to catch voters. Jager and Amblard [44] extended the Deffuant-Weisbuch model by adding repulsive forces related from Social Judgment Theory. Salzarulo [68] presented a model of continuous opinion dynamics based on the Theory of Self-Categorization.

A slightly related old model is due to Hotelling [41]. Nowadays it is mostly cited as the ice-seller's problem and goes as follows: There are some ice-sellers uniformly distributed on two kilometers of promenade close to the beach. People on the beach naturally always go to the ice-seller which is closest by. So it pays for the ice-seller at the beginning of the promenade to move closer towards the center, because he will not lose customers but can gain some by stealing them from the next ice-seller. This produces an overall drift to the center. The model serves as an example that competition does not always produce an optimal welfare for the society, because when all ice-sellers have joined in the middle of the promenade the sum of the walk lengths for all people is larger than in the beginning.

Collignon and Al-Sadoon recently presented a work about the ‘Stochastic Consensus’ [13]. They argue that in great political organization such as the European Union there will always evolve a stochastic consensus about the most important issues but the question is, how long this will take and what organizational form could speed up this process.





Chapter 3

Mathematical Analysis

We bring mathematical machinery to work for answering the question of convergence in processes of continuous opinion dynamics, with convergence to consensus as a question of special interest. Section 3.1 analyzes iterated application of averaging maps which are very generally defined. Section 3.2 analyzes the matrix case and deals with infinite products of row-stochastic matrices. In section 3.3 results are applied for agent-based bounded confidence processes. Further on, their sets of fixed points are characterized. The set of fixed-points for density-based bounded confidence models is characterized in Section 3.4. Section 3.5 gives some references, which did not fit in the text. There is also the place where new contributions are emphasized and distinguished from cited work.

3.1 Averaging maps

In this section we give conditions for convergence to consensus when averaging maps are iteratively applied to an opinion profile by making use of compactness, continuity and convexity arguments. The results will be applied in the linear case. The next subsection is about one fixed averaging map. The following about a family of averaging maps with Theorem 3.1.8 as a core result of this thesis. We finish with a comparison to the theorem of Moreau [62] including an example where Moreau's Theorem is not applicable but Theorem 3.1.8 proves convergence to consensus.

3.1.1 The homogeneous case

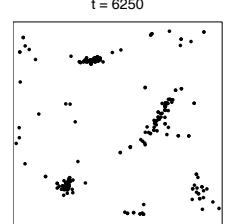
Let $S \subset \mathbb{R}^d$ be an appropriate opinion space. Let S^n be the state space and $f : S^n \rightarrow S^n$ be a continuous and proper averaging map. We recall from Subsection 2.3.5 that this means that for all $x \in S^n$ it either holds

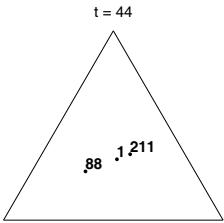
$$\text{conv}_{i \in \underline{n}}\{f_i(x)\} \subset \text{conv}_{i \in \underline{n}}\{x^i\} \quad (3.1)$$

for the convex hull of x or

$$\text{cube}_{i \in \underline{n}}\{f_i(x)\} \subset \text{cube}_{i \in \underline{n}}\{x^i\} \quad (3.2)$$

for the cube of x . In both cases proper means that for every x which is not a consensus it follows that subsets in (3.1) and (3.2) are strict. Sometimes it is





3. Mathematical Analysis

useful to look at the contraposition of the definition of proper: If equality holds in (3.1) or (3.2) this implies that x is a consensus.

Theorem 3.1.1 (Krause). *Let $x(0) \in S^n \subset (\mathbb{R}^d)^n$ with S being a compact and convex opinion space. Let f be a continuous and proper convex hull averaging map on S^n . Then the discrete dynamical system $x(t+1) = f(x(t))$ with $x(0) \in S^n$ converges to consensus. So, there is $\gamma(x(0)) \in S$ such that for all $i \in \underline{n}$ there is*

$$\lim_{t \rightarrow \infty} x^i(t) = \gamma(x(0))$$

Proof. Follows as a special case from theorem 3.1.8. Alternatively, see [48]. \square

Corollary 3.1.2. *Let f be a continuous averaging map on S^n and the k -times iterated function f^k be a proper averaging map. Then the discrete dynamical system converges to consensus.*

So an averaging map which is proper in some iteration is sufficient for convergence to consensus.

We give some examples that show, that the continuity of the averaging map is neither necessary nor neglectable.

Example 3.1.3. *Let $f : (\mathbb{R}_{\geq 0})^2 \rightarrow (\mathbb{R}_{\geq 0})^2$ with*

$$f(x^1, x^2) := \begin{cases} \left(\frac{3}{4}x^1 + \frac{1}{4}x^2, \frac{3}{4}x^1 + \frac{1}{4}x^2\right) & \text{if } x^1 + x^2 > 10, \\ \left((x^1)^{\frac{3}{4}}(x^2)^{\frac{1}{4}}, (x^1)^{\frac{1}{4}}(x^2)^{\frac{3}{4}}\right) & \text{otherwise.} \end{cases}$$

This averaging map converges to consensus but is not continuous. For $x(0) = (1, 9)$ it converges to $(3, 3)$ but for $x(0) = (1 + \varepsilon, 9)$ it converges to $(5 + \frac{\varepsilon}{2}, 5 + \frac{\varepsilon}{2})$. So continuity is not necessary for convergence.

But on the other hand continuity cannot be easily omitted because there are non-continuous averaging maps, which do not converge to consensus.

Example 3.1.4. *Let $f : (\mathbb{R})^3 \rightarrow (\mathbb{R})^3$ with*

$$f(x^1, x^2, x^3) := \begin{cases} (x^1, x^2, \frac{1}{2}x^3 + \frac{1}{2}\min\{x^1, x^2\}) & \text{if } x^3 < \min\{x^1, x^2\}, \\ (x^1, x^2, x^1) & \text{otherwise.} \end{cases}$$

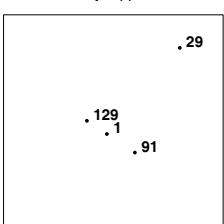
Starting with $x(0) = (2, 3, 1)$ the discrete dynamical system $x(t+1) = f(x(t))$ will converge to $(2, 3, 2)$, although it is a proper averaging map. But it is not continuous at all points where $x^3 = \min\{x^1, x^2\}$ and $x^1 \neq x^2$.

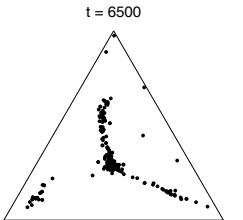
So, continuity of a proper averaging map is sufficient for reaching consensus but not necessary.

3.1.2 The inhomogeneous case

We extend the former result to a sequence of averaging maps f_t coming from a set $F \subset \{f : S^n \rightarrow S^n \mid f \text{ is a proper averaging map}\}$. So, in each time step each agent may apply another averaging map. F should contain either only convex hull averaging maps or only cube averaging maps.

To ensure convergence for a family of averaging maps some further definitions and assumptions must be made.





3.1. Averaging maps

Definition 3.1.5 ((uniformly) (equi)continuous). Let $(X, d_X), (Y, d_Y)$ be metric spaces, and F be a set of functions from X to Y . The set F is called *continuous* if

$$\forall \varepsilon > 0, f \in F, x \in X \exists \delta > 0 \forall \bar{x} \in X : d_X(x, \bar{x}) < \delta \Rightarrow d_Y(f(x), f(\bar{x})) < \varepsilon.$$

F is called *uniformly continuous* if

$$\forall \varepsilon > 0, f \in F \exists \delta > 0 \forall x \in X, \bar{x} \in X : d_X(x, \bar{x}) < \delta \Rightarrow d_Y(f(x), f(\bar{x})) < \varepsilon,$$

equicontinuous if

$$\forall \varepsilon > 0 \exists \delta > 0 \forall f \in F, x \in X, \bar{x} \in X : d_X(x, \bar{x}) < \delta \Rightarrow d_Y(f(x), f(\bar{x})) < \varepsilon,$$

uniformly equicontinuous if

$$\forall \varepsilon > 0 \exists \delta > 0 \forall f \in F, x \in X, \bar{x} \in X : d_X(x, \bar{x}) < \delta \Rightarrow d_Y(f(x), f(\bar{x})) < \varepsilon.$$

The notation with quantifiers emphasizes that the place of choice for δ is crucial. Note that if X is compact then (equi)continuity implies uniform (equi)continuity.

An equicontinuous family of functions is subject to the well-known Theorem of Arzelà-Ascoli, which we will use for the main theorem of this subsection.

Theorem 3.1.6 (Arzelà-Ascoli). Let (X, d_X) be a compact metric space and $(Y, \|\cdot\|)$ be a complete normed space. Let F be a subset of the continuous and bounded function from X to Y . It holds that F is relatively compact if and only if F is equicontinuous and for all $x \in X$ the set $\{f(x) | f \in F\} \subset Y$ is relatively compact.

Proof. See [1]. □

Notice that for subsets of \mathbb{R}^d relatively compact is equivalent to bounded. Uniform equicontinuity ensures, that a sequence $(f_t)_{t \in \mathbb{N}}$ in F can not converge pointwise to a noncontinuous function.

The next thing we need for the theorem of this section is the Hausdorff metric on the set of compact subsets of a metric space (X, d) . We define the distance of a point $x \in X$ and a nonempty compact set $C \subset X$ as

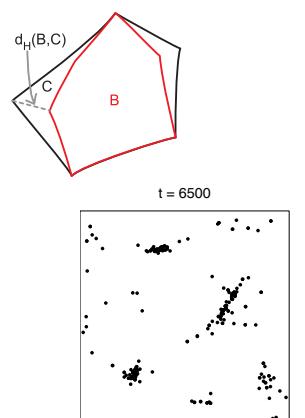
$$d(x, C) := \inf_{c \in C} d(x, c).$$

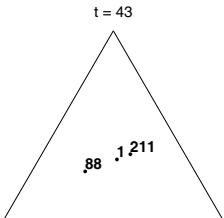
Let $B, C \subset X$ be nonempty and compact, then the *Hausdorff metric* is defined as

$$d_H(B, C) := \max\{\sup_{b \in B} d(b, C), \sup_{c \in C} d(c, B)\}.$$

The Hausdorff metric is the smallest ε such that the ε -neighborhood of B contains C and the ε -neighborhood of C contains B . It is easy to see that $d_H(B, C) = 0$ holds if and only if $B = C$. The Hausdorff metric defines a complete metric space on the set of compact subsets of X .

If $B \subset C$ (as in the case of the following proof) it holds $d_H(B, C) := \sup_{b \in B} d(b, C)$.





3. Mathematical Analysis

Definition 3.1.7 (equiproper). Let F be a family of proper (convex hull or cube) averaging maps on $S^n \subset (\mathbb{R}^d)^n$, with S an appropriate opinion space. F is called *equiproper* if for all $x \in S^n$ which are not a consensus there is $\delta(x) > 0$ such that for all $f \in F$

$$d_H(\text{conv}_{i \in \underline{n}}\{f_i(x)\}, \text{conv}_{i \in \underline{n}}\{x^i\}) > \delta(x) \quad (3.3)$$

if F is a family of convex hull averaging maps or

$$d_H(\text{cube}_{i \in \underline{n}}\{f_i(x)\}, \text{cube}_{i \in \underline{n}}\{x^i\}) > \delta(x) \quad (3.4)$$

if F is a family of cube averaging maps.

If F is equiproper this ensures that a sequence $(f_t)_{t \in \mathbb{N}}$ in F cannot converge to a non proper averaging map.

Theorem 3.1.8. Let F be a family of proper averaging maps on $S^n \subset (\mathbb{R}^d)^n$, with S an appropriate (especially compact) opinion space. Let F be uniformly equicontinuous and equiproper. Then it holds for any sequence $(f_t)_{t \in \mathbb{N}}$ with $f_t \in F$ and $x(0) \in S^n$ that the sequence $(x(t))_{t \in \mathbb{N}}$ recursively defined by

$$x(t+1) = f_t(x(t))$$

converges to a consensus. So, for all $i \in \underline{n}$ there is

$$\lim_{t \rightarrow \infty} x^i(t) = \gamma(x(0), (f_t)_{t \in \mathbb{N}}) =: \gamma.$$

Proof. The proof has to be done once for F being convex hull averaging maps and once for F being cube averaging maps. We do it first for cube averaging maps. In most arguments the convex hull case appears as special case. We will point this out afterwards. So, we assume that all $f \in F$ are proper cube averaging maps.

The proof goes in three steps while the middle step is the longest and goes in four substeps.

The idea of the three main steps is the following: We define

$$C(t) := \text{cube}_{i \in \underline{n}}\{x^i(t)\} \subset S$$

which is convex and compact. It holds $C(t+1) \subset C(t)$ because of the averaging property and $C := \bigcap_{t=0}^{\infty} C(t) \neq \emptyset$ because of compactness.

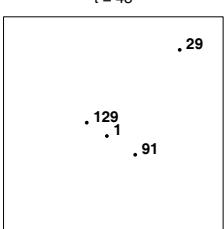
In the following we will show that $C(t)$ converges to a singleton in S and that for all $i \in \underline{n}$ the sequences $x^i(t)$ converge to it.

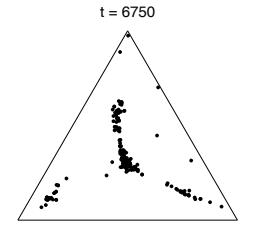
Because of compactness of $C(0)$ there is a subsequence t_s and

$$c := (c^1, \dots, c^n) \in C(0)^n$$

such that $\lim_{s \rightarrow \infty} x(t_s) = c$. (Compactness of $C(0)^n$ follows from Tychonoff's theorem [1] or directly via choosing subsequences one after the other.)

1. We show that $C = \text{cube}\{c^1, \dots, c^n\}$. To accept " \supset " see that for all $t_s \geq t$ there is $x^i(t_s) \in C(t)$ and thus $c^i \in C(t)$. This implies $c^i \in C$ because all the $C(t)$ are closed.





3.1. Averaging maps

To show " \subset " let $\varepsilon > 0$. Then there exists s_0 such that for all $i \in \underline{n}$ and $s \geq s_0$ it holds $\|x^i(t_s) - c^i\| < \varepsilon$. For $t \geq t_{s_0}$ and $x \in C$ it follows $x \in C(t)$. So, there exists componentwise convex coefficient vectors $a^1, \dots, a^n \in \mathbb{R}_{\geq 0}^d$ with $\sum_{i=1}^n a^i = \mathbf{1}$ such that $x = \sum_{i=1}^n a_i \bullet x^i(t_{s_0})$. Now we can conclude

$$\|x - \sum_{i=1}^n a^i \bullet c^i\| = \left\| \sum_{i=1}^n a^i \bullet (x^i(t_{s_0}) - c^i) \right\| \leq \sum_{i=1}^n \|x^i(t_{s_0}) - c^i\| = n\varepsilon.$$

It follows that $x \in \text{cube}\{c^1, \dots, c^n\}$ because $\text{cube}\{c^1, \dots, c^n\}$ is closed.

2. The next step is to show that c is a consensus, which means

$$c^1 = \dots = c^n. \quad (3.5)$$

F is uniformly equicontinuous and for all $x \in X$ it holds that $\{f(x) \mid f \in F\}$ is bounded (and thus relatively compact) because all the f are averaging maps. So, due to the theorem of Arzelà-Ascoli 3.1.6, F is relatively compact. Thus, there is a subsequence t_{s_r} such that $f_{t_{s_r}}$ converges uniformly to a continuous limit function g for $r \rightarrow \infty$.

We will show that g is a proper averaging map $\text{conv}_{i \in \underline{n}} g_i(c) = \text{conv}_{i \in \underline{n}} c^i$ which implies (3.5). This goes in four steps.

- (a) We show that g is a cube averaging map. So, we have to show that for all $i \in \underline{n}$ it must hold $g_i(x) \in \text{cube}_{i \in \underline{n}} x^i$. Let $\varepsilon > 0$, due to the pointwise convergence of $(f_{t_{s_r}})_i$ to g_i there is r_0 such that for all $r > r_0$ it holds

$$\|(f_{t_{s_r}})_i(x) - g_i(x)\| < \varepsilon.$$

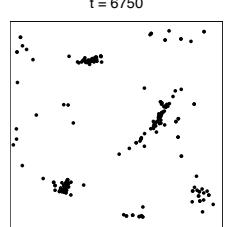
Due to $(f_{t_{s_r}})_i(x) \in \text{cube}_{i \in \underline{n}} x^i$ it follows that the maximal distance of $g_i(x)$ to $\text{cube}_{i \in \underline{n}} x^i$ is less than ε and thus it is in $\text{cube}_{i \in \underline{n}} x^i$ because $\text{cube}_{i \in \underline{n}} x^i$ is closed.

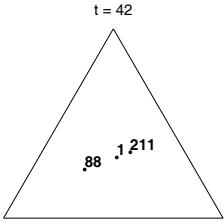
- (b) We show that g is proper. So, let $x \in S^n$ be not a consensus. We have to show that there is $y^* \in \text{cube}_{i \in \underline{n}} x^i$ but $y^* \notin \text{cube}_{i \in \underline{n}} g_i(x)$. We know that there is for each $r \in \mathbb{N}$ an $y(r) \in \text{cube}_{i \in \underline{n}} x^i$ with $y(r) \notin \text{cube}_{i \in \underline{n}} \{(f_{t_{s_r}})_i(x)\}$. According to the equiproper property it can be chosen such that the distance of $y(r)$ to $\text{cube}_{i \in \underline{n}} \{(f_{t_{s_r}})_i(x)\}$ is bigger than $\frac{\delta(x)}{2} > 0$ for all $r \in \mathbb{N}$. Further on, we know that the set difference $\text{cube}_{i \in \underline{n}} \{(f_{t_{s_r}})_i(x)\} \setminus \text{cube}_{i \in \underline{n}} x^i$ is non empty and bounded, thus there is yet another subsequence $t_{s_{r_q}}$ such that $y(t_{s_{r_q}})$ converges to an $y^* \in \text{cube}_{i \in \underline{n}} x^i$. Because of the construction it also holds $y^* \notin \text{cube}_{i \in \underline{n}} g_i(x)$.

- (c) We show that it holds $\lim_{r \rightarrow \infty} f_{t_{s_r}}(x_{t_{s_r}}) = g(c)$. We know that $f_{t_{s_r}} \rightarrow g$ uniformly and that $x^i(t_{s_r}) \rightarrow c$. Now we estimate

$$\|f_{t_{s_r}}(x(t_{s_r})) - g(c)\| \leq \|f_{t_{s_r}}(x(t_{s_r})) - f_{t_{s_r}}(c)\| + \|f_{t_{s_r}}(c) - g(c)\|$$

Both terms on the right hand side can be smaller than $\frac{\varepsilon}{2}$ for any ε and large enough r because of the equicontinuity of $f_{t_{s_r}}$ and the uniform convergence $f_{t_{s_r}} \rightarrow g$.





3. Mathematical Analysis

- (d) We show $\text{cube}_{i \in \underline{n}}\{g^i(c)\} = \text{cube}_{i \in \underline{n}}\{c^i\}$. " \subset " holds because g is a cube averaging map (see (a)). To show " \supset " let $x \in \text{cube}_{i \in \underline{n}}c^i$. Thus, for all r it holds $x \in C \subset C(t_{s_r} + 1)$ and thus there exist componentwise convex coefficient vectors with the componentwise convex combination $x = \sum_{i=1}^n a^i(r) \bullet x^i(t_{s_r} + 1)$. Now, $(a^1(r), \dots, a^n(r))_{r \in \mathbb{N}}$ is a sequence in the compact set of row-stochastic $d \times n$ matrices and thus there is a subsequence r_q such that $\lim_{q \rightarrow \infty} (a^1(r_q), \dots, a^n(r_q)) = (a^{1*} \dots a^{n*})$. Now due to (c) it holds,

$$x = \lim_{q \rightarrow \infty} x = \sum_{i=1}^n \lim_{q \rightarrow \infty} a^i(r_q) \bullet \lim_{q \rightarrow \infty} x^i(t_{s_{r_q}} + 1) = \sum_{i=1}^n a^i \bullet g^i(c).$$

Thus, $x \in \text{cube}_{i \in \underline{n}}\{g^i(c)\}$.

Now (d) implies that c is a consensus, because g is a proper averaging map (see (a) and (b)).

3. The last thing to show is that for all $i \in \underline{n}$ the sequence $(x^i(t))_{t \in \mathbb{N}}$ (and not only subsequences) converges to $\gamma := c^1 = \dots = c^n$ for $t \rightarrow \infty$. We know that for $\varepsilon > 0$ there is a r_0 such that for all $i \in \underline{n}$ it holds $\|x^i(t_{s_{r_0}}) - \gamma\| < \varepsilon$. Further on, for $t \geq t_{s_{r_0}}$ it holds $x^i(t) \in C(t) \subset C(t_{s_{r_0}})$. Thus, for all $i \in \underline{n}$ there are componentwise convex coefficient vectors with a componentwise convex combination $x^i(t) = \sum_{j=1}^n a^j \bullet x^j(t_{s_{r_0}})$. Now, we can conclude for all $t > t_{s_{r_0}}$

$$\|x^i(t) - \gamma\| = \left\| \sum_{j=1}^n a^j \bullet (x^j(t_{s_{r_0}}) - \gamma) \right\| \leq \sum_{j=1}^n \|x^j(t_{s_{r_0}}) - \gamma\| = n\varepsilon.$$

This proves the theorem for cube averaging maps.

For convex hull averaging maps all arguments hold analog with replacing 'cube' with 'conv', 'cube averaging map' with 'convex hull averaging map' and choosing componentwise convex coefficient vectors with each having all its entries equal. Then they are equivalent to ordinary convex coefficients. \square

Notice, that equiproper holds automatically if we have only one proper averaging map in the family F , which makes Theorem 3.1.1 a special case of Theorem 3.1.8. Some interesting corollaries can be deduced.

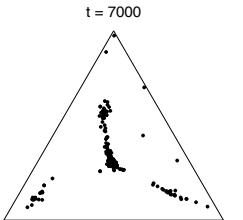
Corollary 3.1.9. *Let $F = \{f_1, \dots, f_m\}$ be a finite family of proper averaging maps on $S^n \subset (\mathbb{R}^d)^n$, with S an appropriate opinion space. Let F be uniformly continuous. Then it holds for a sequence $(f_t)_{t \in \mathbb{N}}$ with $f_t \in F$ and $x(0) \in S^n$ that the discrete dynamical system $x(t+1) = f_t(x(t))$ converges to consensus.*

Proof. First, a finite family of uniformly continuous functions is uniformly equicontinuous. It remains to show that F is equiproper. For $x \in S^n$ not a consensus it holds that for every $f \in F$ there is $\delta_f(x) > 0$ such that (3.3) (or (3.4)) is fulfilled because f is proper. We define $\delta(x) := \min_{f \in F} \delta_f(x)$. Now, $\delta(x) > 0$ because F is finite and (3.3) (or (3.4)) is fulfilled uniformly for all f by $\delta(x)$. Now, the theorem applies. \square

$t = 42$

$.29$

$.129$
 $.1$
 $.91$



3.1. Averaging maps

Corollary 3.1.10. Let F be a family of averaging maps on $S^n \subset (\mathbb{R}^d)^n$, with S an appropriate opinion space. Let F be uniformly equicontinuous and at least one element in F is proper and all proper elements of F are equiproper. Let $(f_t)_{t \in \mathbb{N}}$ be a sequence with $f_t \in F$ and t_s be a subsequence such that f_{t_s} is proper. Then, for $x(0) \in S^n$, the discrete dynamical system $x(t+1) = f_t(x(t))$ converges to consensus.

Proof. Apply the theorem for the sequence f_{t_s} with $s \rightarrow \infty$ and use arguments analog to step 3 of its proof to show that convergence to consensus holds also for f_t with $t \rightarrow \infty$. \square

Corollary 3.1.11. Let F be a family of averaging maps on $S^n \subset (\mathbb{R}^d)^n$, with S an appropriate opinion space. Let F be uniformly equicontinuous. Let $(f_t)_{t \in \mathbb{N}}$ be a sequence with $f_t \in F$ and $(t_s)_{s \in \mathbb{N}}$ be a subsequence (with $t_0 = 0$) such that the $f_s := f_{t_{s+1}-1} \circ f_{t_{s+1}-2} \circ \dots \circ f_{t_s+1} \circ f_{t_s}$ are equiproper. Then, for $x(0) \in S^n$, the discrete dynamical system $x(t+1) = f_t(x(t))$ converges to consensus.

Proof. It is easy to see that $\{f_s \mid s \in \underline{n}\}$ is an equicontinuous family of averaging maps. Now, apply the theorem for the sequence f_s with $s \rightarrow \infty$ and use arguments analog to step 3 of its proof to show that convergence to consensus holds also for f_t with $t \rightarrow \infty$. \square

We continue with some examples showing that the assumptions of the theorem are neither trivial nor sharp.

First ‘equi’ in equiproper cannot be omitted. The following example shows a sequence of proper averaging maps, which do not converge to consensus.

Example 3.1.12. Let $f_t : (\mathbb{R})^2 \rightarrow (\mathbb{R})^2$ with

$$f_t(x^1, x^2) := \left((1 - \frac{1}{4^t})x^1 + \frac{1}{4^t}x^2, \frac{1}{4^t}x^1 + (1 - \frac{1}{4^t})x^2 \right)$$

It is easy to see that for $t \geq 1$ and $x(1) = (0, 1)$ it holds that $x^1(t) < \frac{1}{3}$ and $x^2(t) > \frac{2}{3}$. Obviously, $\{f_t \mid t \in \mathbb{N}\}$ is not equiproper because f_t converges to the identity as $t \rightarrow \infty$.

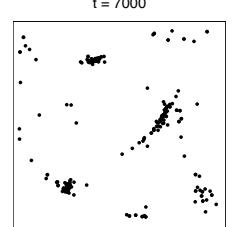
But on the other hand the next example shows a sequences of proper averaging maps which is not equiproper but converges to consensus for all initial values.

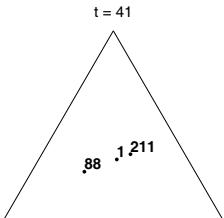
Example 3.1.13. Let $f_t : (\mathbb{R})^2 \rightarrow (\mathbb{R})^2$ with

$$f_t(x^1, x^2) := \left((1 - \frac{1}{t})x^1 + \frac{1}{t}x^2, x^2 \right)$$

This example is not equiproper, because f_t converges to the identity for $t \rightarrow \infty$. But for $t \geq 2$ and any $x(2) \in (\mathbb{R})^2$ the system $x(t+1) = f_t(x(t))$ has the solution $x(t) = (\frac{1}{t-1}x^1(2) + \frac{t-2}{t-1}x^2(2), x^2)$ and thus converges to consensus at $x^2(2)$.

So, there must be weaker conditions for reaching consensus than in the theorem. But compactness arguments in the proof at step 2b cannot easily be modified to a proof of convergence for this example. The limit function g is not a proper averaging map here. But nevertheless there is convergence to consensus.





3. Mathematical Analysis

At last we want to give an example on the equicontinuity of the family F . Of course, equicontinuity is not necessary. Example 3.1.3 gives a one-element family of not equicontinuous averaging maps which converge. Now, we show that a family of uniformly continuous proper averaging maps is not enough to ensure convergence. The example is inspired by bounded confidence.

Example 3.1.14 (Vanishing confidence). Let $f_t : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with

$$(f_t)_i(x) := \frac{\sum_{j=1}^n D_t(|x^i - x^j|)x^j}{\sum_{j=1}^n D_t(|x^i - x^j|)}$$

and $D_t : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$. Now, f_t is an averaging map for any choice of D_t . Further on, f_t is continuous if D_t is, and f_t is proper if D_t is strictly positive. The Hegselmann-Krause model 2.4.3 with homogeneous bound of confidence $\varepsilon > 0$ comes out for D_t being a noncontinuous cutoff function

$$D_t(y) = \begin{cases} 1 & \text{if } y \leq \varepsilon \\ 0 & \text{otherwise.} \end{cases}$$

We chose

$$D_t(y) := e^{-(\frac{y}{\varepsilon})^t}$$

as a sequence of functions which has the cutoff function as a limit function. So, D_t is continuous but $\{D_t \mid t \in \mathbb{N}\}$ is not equicontinuous.

Now, it holds for $x(0) = (0, 8)$, $\varepsilon = 1$ that the process $x(t) = f_t(x(t))$ does not converge to consensus although only proper averaging maps are involved. This holds because for $n = 2$ it holds that

$$|(f_t)_1(x^1, x^2) - (f_t)_2(x^1, x^2)| = \frac{1 - e^{|x^1 - x^2|^t}}{1 + e^{|x^1 - x^2|^t}} |x^1 - x^2|.$$

From this one can easily compute (with rough estimates) that for the specific initial profile $x(0) = (0, 8)$ it holds that for all $t \in \mathbb{N}$ it holds that $|x^1(t) - x^2(t)| \geq 8(1 - \frac{1}{4})(1 - \frac{1}{8})(1 - \frac{1}{16}) \dots \geq 4$.

But for other initial values and other ε convergence under vanishing confidence is possible, as numerical examples in Figure 3.1 show.

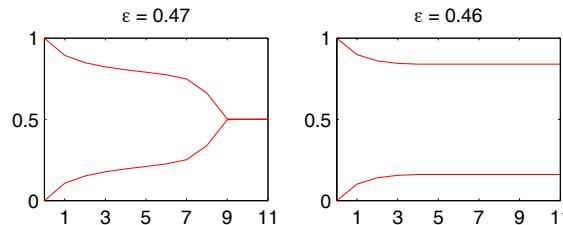
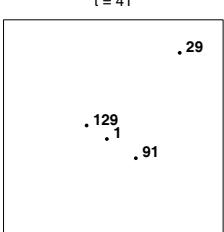
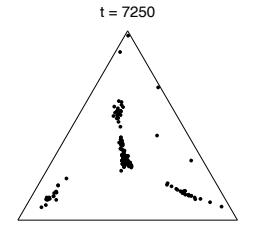


Figure 3.1: Vanishing confidence example with $x(0) = (0, 1)$ and $\varepsilon = 0.46, 0.47$.

3.1.3 Comparison with Moreau's theorem

Theorem 3.1.8 is similar to a theorem of Moreau [62, Theorem 2]. We cite it here to discuss similarities and differences. It incorporates changing communication networks into self-maps with averaging properties.





3.1. Averaging maps

Theorem 3.1.15 (Moreau). Let $(f_t)_{t \in \mathbb{N}}$ be a sequence of self maps on $S^n \subset (\mathbb{R}^d)^n$, with S convex and closed. Let $(N(t))_{t \in \mathbb{N}}$ be a communication regime where all networks have positive diagonals and where there is $T \in \mathbb{N}$ such that for all $t_0 \in \mathbb{N}$ the network $\text{inc} \left(\sum_{t=t_0}^{t_0+T} N(t) \right)$ has only one essential class¹.

Further on, it should exist for each network with positive diagonal N , each $x \in S^n$ and each agent $k \in \underline{n}$ a compact set $e_k(x, N)$ such that

1. For all $t \in \mathbb{N}$ it holds $(f_t)_k(x) \in e_k(x, N)$,
2. $e_k(x, N) \subset \text{ri conv}_{i \in \text{nb}(k, N)} \{x^i\}$,
3. $e_k(x, N)$ depends continuously on x (that means that the map $e_k : S^n \times \{0, 1\}^{n \times n} \rightarrow \mathcal{P}(S^n)$ is continuous with respect to the Hausdorff-metric on $\mathcal{P}(S^n)$ which is the set of all compact subsets of S^n .)

Then it holds for $x(0) \in S^n$ that the discrete dynamical system

$$x(t+1) = f_t(x(t))$$

converges to a consensus.

Notice that the relative interior of a singleton is the singleton itself. The theorem has been significantly modified in comparison with the original to fit it in our vocabulary. The most important one is that the original theorem is about "uniform global attractivity of the system with respect to the set of equilibrium solutions $x^1 = \dots = x^n = \text{constant}$ " which is equivalent to convergence to consensus for every $x(0) \in S^n$.

Items 1 and 2 in the assumptions of Theorem 3.1.15 gives something like a 'proper convex hull averaging map with respect to the actual network'. It is averaging due to conv, and proper due to ri (actually ri is a stronger assumption than proper). The continuity assumption in item 3 shows similarity to the assumption of equicontinuity in Theorem 3.1.8. Equiproper from Theorem 3.1.8 finds its analog in Theorem 3.1.15 in the fact that in item 1 it holds $(f_t)_k(x) \in e_k(x, N)$ and e_k is independent of t .

The next example is one where convergence to consensus is assured by Theorem 3.1.8 but not by Theorem 3.1.15.

Example 3.1.16. Let $g_1, g_2, g_3, g_4 : (\mathbb{R}^d)^3 \rightarrow \mathbb{R}^d$ with $g_1(x) := \max\{x^1, x^2, x^3\}$, $g_2(x) := \frac{1}{3}(x^1 + x^2 + x^3)$, $g_3(x) := \sqrt[3]{x^1 x^2 x^3}$ and $g_4(x) := \min\{x^1, x^2, x^3\}$ be general multidimensional means (all computations componentwise) and $f^{\sigma_1 \sigma_2 \sigma_3} : (\mathbb{R}^d)^3 \rightarrow (\mathbb{R}^d)^3$ with

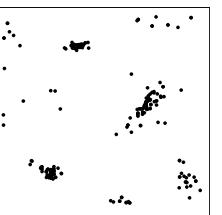
$$f^{\sigma_1 \sigma_2 \sigma_3} := (g_{\sigma_1}, g_{\sigma_2}, g_{\sigma_3})$$

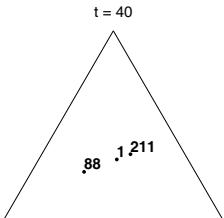
be averaging maps. Now it is easy to verify, that

$$F := \{f^{\sigma_1 \sigma_2 \sigma_3} \mid (\sigma_1, \sigma_2, \sigma_3) \in \{1, 2, 3, 4\}^3 \text{ but 1 and 4 not both in } (\sigma_1, \sigma_2, \sigma_3)\}$$

is an equicontinuous and equiproper (due to finiteness) set of cube averaging maps. Thus, for any sequence f_t with elements from F and $x(0) \in (\mathbb{R}^d)^3$ it holds that $x(t+1) = f_t(x(t))$ converges to consensus due to theorem 3.1.8. Theorem 3.1.15 is not applicable because item 2 does not hold for all elements of F .

¹See Section 3.2.1 for definitions of inc () and essential class.





3. Mathematical Analysis

Krause [48] shows another example where Moreau's Theorem fails because it needs a positive diagonal in each network which is not necessary.

3.2 Matrix-based analysis

In this section all matrices should be square unless stated otherwise.

Processes of opinion dynamics in the Models 2.4.1, 2.4.2, 2.4.3 and 2.4.4 can be computed by repeated multiplications of row-stochastic matrices from the left to an opinion profile. If we neglect the specific initial opinion profile we have an infinite product of row-stochastic matrices, where infinity is to the left (also called infinite backward product).

On the other hand a trajectory of the interactive Markov chain (2.2) (with arbitrary transition matrix) is also computed as a repeated multiplication of row-stochastic matrices from the right to an opinion distribution. So, there is an infinite product of row-stochastic matrices with infinity to the right (also called infinite forward product).

So, this section is about a sequence of row-stochastic matrices $(A(t))_{t \in \mathbb{N}}$ and the infinite accumulations $A(t, 0)$ (backward) and $A(0, t)$ (forward) for $t \rightarrow \infty$. Backward products appear in inhomogeneous processes of opinion pooling, because for an initial opinion profile $x(0)$ it holds

$$x(t) = A(t, 0)x(0).$$

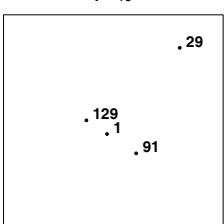
Forward products appear in inhomogeneous Markov processes, because for an initial distribution $p(0)$ it holds

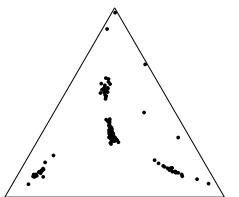
$$p(t) = p(0)A(0, t).$$

In the homogeneous case (where one has only one fixed matrix A) the analysis of Markov processes and processes of opinion pooling is done by analysis of the powers of A and thus quite similar. But we will point out that it divides into two worlds in the inhomogeneous case. Our focus is on the less studied backward products. Forward products were extensively studied in the theory of Markov chains. But unfortunately the interactive Markov chains with Hegselmann-Krause or Deffuant-Weisbuch transition matrix cannot be analyzed in classical inhomogeneous Markov theory because they are in no way ergodic. They are treated in Section 3.4 with other measures. Nevertheless, the common framework of infinite matrix products is appealing.

The analysis of infinite matrix products is a mathematical topic in its own interest. But infinite backward products also lead to some characteristic properties of convergence especially to consensus in opinion dynamic processes. Examples are the strong impact that self-confidence has on the stabilization of opinions or sufficient conditions about the connections (changing over time) which the agents need to reach consensus. Also results on acceptable hardening of positions or acceptable raising time to reach intercommunication which do not destroy the convergence to consensus are interesting in opinion dynamics.

There are different properties of row-stochastic matrices which bring different mathematical machineries to work. First, row-stochastic matrices are nonnegative, which makes the zero pattern of the matrix an interesting object to study and gives a connection to the Perron-Frobenius theory on nonnegative matrices.





3.2. Matrix-based analysis

Second, row-stochastic matrices are row allowable, which leads to interesting results on the evolution of the number of positive and zero entries in columns of the backward matrix accumulations.

Third, row-stochastic matrices have spectral radius one with one eigenvalue equal to one and eigenvector $\mathbf{1}$. This is the key to understand conditions for convergence to consensus in the autonomous case with one fixed confidence matrix.

Fourth, a row-stochastic matrix has max-norm equal to one. This ensures non increasing norms of growing products, which is a necessary condition for convergence.

Fifth, a multiplication with a row-stochastic matrix from the left to an opinion profile is a convex hull averaging map on the agents opinions. Further on, multiplication of a row-stochastic matrix from the left to another matrix is a convex hull averaging map on its rows. Thus, convergence results of the previous section are applicable.

We will give some results about matrices in a more general sense, where row-stochastic matrices are only special cases. This is done to open the scope for possible new results extending from these results, e.g. about a full understanding of converging matrix products (outlines in [34]) or the joint spectral radius. But the main line of this section remains to shed light on infinite products of row-stochastic matrices, especially the convergence properties of products infinite to the left.

3.2.1 The Gantmacher form for nonnegative matrices

Let A be a nonnegative $n \times n$ matrix.

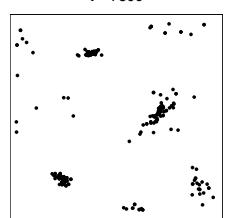
In the following we define a canonical form of a nonnegative matrix which sorts the positive and zero entries in a clear way by simultaneous row and column permutations. Simultaneous row and column permutation is equivalent to a similarity transformation with a permutation matrix. It conserves the fundamental structure (concerning e.g. the values of the entries, the norm, the spectrum and the structure of the eigenspaces). It might be useful for understanding this subsection to take frequent looks at the example at the end of the subsection.

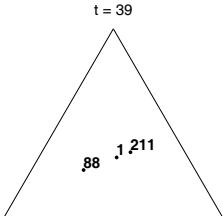
Classes of indices We define the *incidence matrix* of A as $\text{inc}(A)$ with $\text{inc}(A)_{ij} = 1$ if $a_{ij} > 0$ and zero otherwise. In this section every definition just matters in the *incidence sense*, which means that it is the same for all matrices with the same zero and positivity pattern. We say that A and another nonnegative matrix B are of the same *type* if they are equal in the incidence sense.

$$A \sim B \Leftrightarrow \text{inc}(A) = \text{inc}(B)$$

Consequently, A is *type-symmetric* if $A \sim A^T$. The following derivation of Gantmacher's canonical form of a matrix A is the same for all matrices of the same type.

For indices $i, j \in \underline{n}$ we say that there is a *path* from i to j if there is a sequence of indices $i = i_1, \dots, i_m = j$ such that for all $k \in \underline{m-1}$ it holds $a_{i_k, i_{k+1}} > 0$. We abbreviate this statement with $i \rightarrow j$. The *length* of a path is m .





3. Mathematical Analysis

We say $i, j \in \underline{n}$ communicate if $i \rightarrow j$ and $j \rightarrow i$, abbreviated $i \leftrightarrow j$. An index is self-communicating if there is a path $i \rightarrow i$. An index $i \in \underline{n}$ is called essential if for every $j \in \underline{n}$ with $i \rightarrow j$ it holds $j \rightarrow i$. An index is called inessential if it is not essential.

It is easy to see that for all self-communicating indices the relation ' \leftrightarrow ' is an equivalence relation. Thus, the set of self-communicating indices divides into disjoint equivalence classes of indices which we call *self-communicating classes*. All non self-communicating indices are inessential and we define each such index as an own non-self-communicating class. So, a *class of indices* is either a set of indices which all communicate and don't communicate with others or an isolated inessential index which does not communicate with itself (but may have paths to other indices). All together, the set of indices \underline{n} divides into h disjoint classes $\mathcal{I}_1, \dots, \mathcal{I}_h$.

The terms path, essential and inessential extend naturally to classes of indices. So, $\mathcal{J} \rightarrow \mathcal{I}$ means that there is a path from each index in \mathcal{J} to each index in \mathcal{I} . Thus, in this example for $\mathcal{J} \neq \mathcal{I}$ the class \mathcal{J} is an inessential class. \mathcal{I} may either be essential or inessential.

Let $0 \leq g \leq h$ be the number of essential classes (possibly zero). We sort the classes without loss of generality such that the first classes g in the set of classes $\mathcal{I}_1, \dots, \mathcal{I}_g, \mathcal{I}_{g+1}, \dots, \mathcal{I}_h$ are the essential classes. It can be shown that a row allowable matrix contains at least one essential class.

Further on, we sort the inessential classes $\mathcal{I}_{g+1}, \dots, \mathcal{I}_h$ without loss of generality such that there is no path $\mathcal{I}_i \rightarrow \mathcal{I}_j$ if $i < j$. This is possible due to the definition of a class.

A matrix is called *irreducible* if it has only one essential class ($h = 1$). It is called *reducible* if it is not irreducible.

If i is a self-communicating index, the *period* of i is the greatest common divisor of the lengths of all paths $i \rightarrow i$. It can be shown ([69]) that all indices in a self-communicating class have the same period. Thus, the period is naturally defined for self-communicating classes. For a self-communicating class \mathcal{I} we call $d(\mathcal{I})$ its period. For the non-self-communicating one-element classes we define the period to be one.

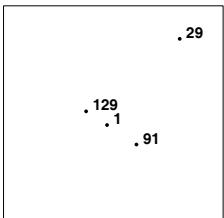
Further on, for all $i \in \underline{h}$ the class \mathcal{I}_i divides into $d(\mathcal{I}_i) \in \mathbb{N}$ disjoint subclasses $\mathcal{I}_{i,1}, \dots, \mathcal{I}_{i,d(\mathcal{I}_i)}$ such that for all $j \in d(\mathcal{I}) - 1$ and $k \in \mathcal{I}_{i,j}$ there is an index $l \in \mathcal{I}_{i,j+1}$ such that $a_{kl} > 0$ and for all $\overline{l} \notin \mathcal{I}_{i,j+1}$ it is $a_{kl} = 0$. And (to close the circle) for all $k \in \mathcal{I}_{i,d(\mathcal{I}_i)}$ there is $l \in \mathcal{I}_{i,1}$ such that $a_{kl} > 0$ and for all $l \notin \mathcal{I}_{i,1}$ it is $a_{kl} = 0$.

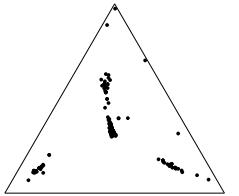
A matrix is called *primitive* if it has only one essential class with period one.

The Gantmacher form Let P be the $n \times n$ permutation matrix such that PA has the rows of A permuted in the order of the classes and subclasses

$$\underbrace{\mathcal{I}_{1,1}, \dots, \mathcal{I}_{1,d(\mathcal{I}_1)}, \dots, \mathcal{I}_{g,1}, \dots, \mathcal{I}_{g,d(\mathcal{I}_g)}, \dots, \mathcal{I}_{h,1}, \dots, \mathcal{I}_{h,d(\mathcal{I}_h)}}_{\mathcal{I}_1} \quad \dots \quad \underbrace{\mathcal{I}_{1,1}, \dots, \mathcal{I}_{g,1}, \dots, \mathcal{I}_{g,d(\mathcal{I}_g)}, \dots, \mathcal{I}_{h,1}, \dots, \mathcal{I}_{h,d(\mathcal{I}_h)}}_{\mathcal{I}_g} \quad \dots \quad \underbrace{\mathcal{I}_{1,1}, \dots, \mathcal{I}_{g,1}, \dots, \mathcal{I}_{g,d(\mathcal{I}_g)}, \dots, \mathcal{I}_{h,1}, \dots, \mathcal{I}_{h,d(\mathcal{I}_h)}}_{\mathcal{I}_h}.$$

Then PAP^T defines a simultaneous row and column permutation of the





3.2. Matrix-based analysis

entries in matrix A to the form

$$PAP^T = \left[\begin{array}{ccc|c} A_1 & & 0 & 0 \\ & \ddots & & \\ 0 & & A_g & \\ \hline A_{g+1,1} & \dots & A_{g+1,g} & A_{g+1} \\ \vdots & & \vdots & \vdots \\ A_{h,1} & \dots & A_{h,g} & A_{h,g+1} \dots A_h \end{array} \right] \quad (3.6)$$

where diagonal matrices A_1, \dots, A_h are square of sizes $\#\mathcal{I}_1, \dots, \#\mathcal{I}_h$ and A_1, \dots, A_g are irreducible and A_{g+1}, \dots, A_h either irreducible or zero and 1×1 . For each $i \in \{g+1, \dots, h\}$ with $A_i \neq 0$ it holds further on that at least one block $A_{i,1}, \dots, A_{i,i-1}$ must contain a positive entry. (In the case of a row allowable matrix there must be at least one block with at least one positive entry for every $i \in \{g+1, \dots, h\}$ because otherwise there must be a zero row.)

And for every $i \in \underline{h}$ and $d := d(\mathcal{I}_i)$ the diagonal block has the form

$$A_i = \left[\begin{array}{ccccc} 0 & Q_{12} & 0 & \dots & 0 \\ \vdots & 0 & Q_{23} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & Q_{d-1,d} \\ Q_{d,1} & 0 & \dots & \dots & 0 \end{array} \right] \quad (3.7)$$

where the d diagonal zero blocks are square of the sizes $\#\mathcal{I}_{i,1}, \dots, \#\mathcal{I}_{i,d(\mathcal{I}_i)}$ and the Q 's have at least one positive entry in each row and each column. (But notice that a matrix form (3.7) with at least one positive entry in each row and each column of each Q alone is not sufficient to define an irreducible matrix.)

This form is called *Gantmacher's canonical form of a nonnegative matrix* or *Gantmacher form*. The submatrices according to the classes are called *Gantmacher diagonal blocks* and *Gantmacher subdiagonal blocks* depending their position in the transformed matrix.

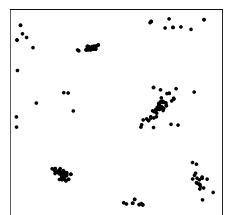
The Gantmacher form is not unique. Permutation of indices within subclasses does not destroy the structure, as well as permutation of whole essential classes. A permutation of two inessential classes \mathcal{I} and \mathcal{J} is possible, if there neither $\mathcal{I} \rightarrow \mathcal{J}$ nor $\mathcal{J} \rightarrow \mathcal{I}$.

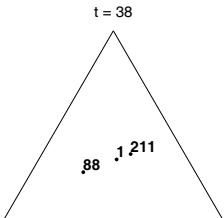
Remark. In the following we will often assume that a matrix is in Gantmacher without restriction of generality. This means that simultaneous row and column permutation will not bother the results. Often one should think of the same permutation for all involved matrices.

Gantmacher form of powers of a nonnegative matrix The concepts of paths and irreducible and primitive matrices have a strong relationship to the positivity of the entries in the powers of a nonnegative matrix.

Let $k \in \mathbb{N}$ and $i, j \in \underline{n}$ then $[(\text{inc}(A))^k]_{ij}$ gives the number of paths $i \rightarrow j$ of length k . This comes directly through the rules of matrix multiplication. Thus, there is a path $i \rightarrow j$ if and only if there is a $k \in \mathbb{N}$ such that $[A^k]_{ij} > 0$.

Due to that fact we can characterize 'irreducible' and 'primitive' in terms of powers of A .





3. Mathematical Analysis

The matrix A is irreducible if and only if for all $i, j \in \underline{n}$ there is $k \in \mathbb{N}$ such that $[A^k]_{ij} > 0$. The matrix A if and only if there is $k \in \mathbb{N}$ such that A^k is positive.

It is easy to see, that powers of A have the same Gantmacher form with the same permutation matrix P , due to the rules of matrix multiplication. For $k \in \mathbb{N}$ it holds that $[A^k]_{[\mathcal{I}_i, \mathcal{I}_i]} = A_i^k$. Thus, raising A to powers can be done separately in the diagonal blocks, but the subdiagonal blocks develop complex with rising k .

Further on it follows from the rules of matrix multiplication for a self-communicating class \mathcal{I}_i that

$$A_i^{d(\mathcal{I}_i)} = \begin{bmatrix} Q_{12} \dots Q_{d-1,d} Q_{d,1} & & & 0 \\ & Q_{23} \dots Q_{d,1} Q_{12} & & \\ & & \ddots & \\ 0 & & & Q_{d-1,d} Q_{12} \dots Q_{d-1,d} \end{bmatrix} \quad (3.8)$$

where each diagonal block is primitive. Irreducibility of each diagonal block is due to the fact that each index in a \mathcal{I}_i has a path to each other index in this class which length is dived by $d(\mathcal{I}_i)$. This includes self-paths. Due to the fact that $d(\mathcal{I}_i)$ is minimal there must be at least one positive entry on the diagonal of $A_i^{d(\mathcal{I}_i)}$ which implies primitivity (see [69, Lemma 1.1]) of A_i .

If we define d' as the least common multiple of $d(\mathcal{I}_1), \dots, d(\mathcal{I}_h)$ then $B := A^{d'}$ is transformed by P (the permutation to the Gantmacher form of (3.6)) to a *primitive Gantmacher form*

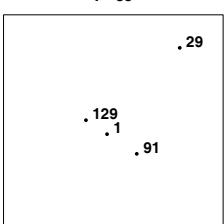
$$PBP^T = \left[\begin{array}{cc|c} B_1 & & 0 \\ \ddots & & \\ 0 & B_{g'} & \\ \hline B_{g'+1,1} & \dots & B_{g'+1,g'} & B_{g'+1} \\ \vdots & & \vdots & \vdots \\ B_{h',1} & \dots & B_{h',g'} & B_{h',g'+1} \dots B_{h'} \end{array} \right] \quad (3.9)$$

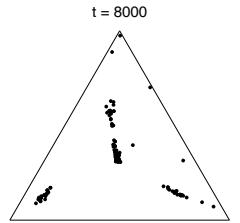
where $g' = d(\mathcal{I}_1) + \dots + d(\mathcal{I}_g)$, $h' = d(\mathcal{I}_1) + \dots + d(\mathcal{I}_h)$ and the diagonal matrices $B_1, \dots, B_{h'}$ are square of sizes $\#\mathcal{I}_{1,1}, \dots, \#\mathcal{I}_{1,d(\mathcal{I}_1)}, \dots, \#\mathcal{I}_{h,1}, \dots, \#\mathcal{I}_{h,d(\mathcal{I}_h)}$ and primitive or zero and 1×1 . Notice that there is a primitive Gantmacher form for every nonnegative matrix A because there is always a certain d' .

Graph theory and an example A graph $G(V, L)$ with a vertex set $V = \{1, \dots, n\}$ and an link set $L \subset V \times V$ can be represented by its adjacency matrix $A(G)$. The other way round, a nonnegative matrix A has an underlying graph $G(A)$ with its set of indices as vertices and a link (i, j) whenever $a_{ij} > 0$.

A graph is usually displayed as its vertices on a plane connected by arrows. Sometimes this representation is useful for interpretation.

We want to explain the terms defined in this section with an example matrix, its graph, and its Gantmacher form.





Let A be a nonnegative matrix, and we regard its incidence matrix to be

$$\begin{array}{l}
 1 \quad \left[\begin{array}{cccccccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \\
 . \quad (3.10)
 \end{array}$$

Figure 3.2 shows the graph of this incidence matrix. Starting from one vertex one can follow the arrows to find all vertices where it has paths to, where it gets paths from (by walking reverse on arrows) and which vertices it communicates with. Each class is colored with an own color and subclasses have different brightness if there is more than one subclass.

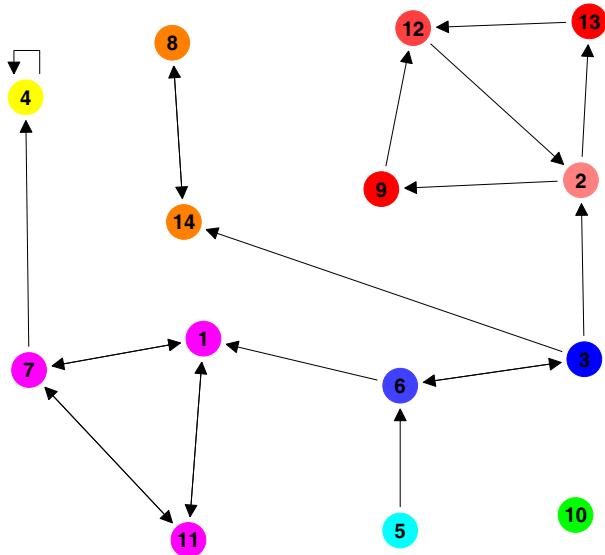
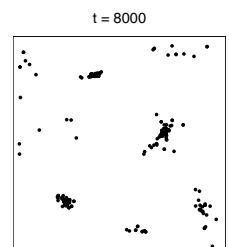
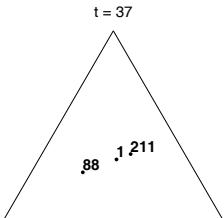


Figure 3.2: The graph of example matrix (3.10)





3. Mathematical Analysis

There are classes

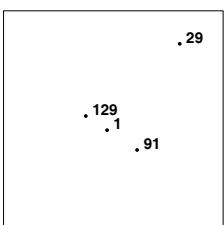
$$\begin{aligned}
 \mathcal{I}_1 &= \{4\} \\
 \mathcal{I}_2 &= \{8, 14\} \\
 \mathcal{I}_3 &= \{2, 9, 12, 13\} \text{ with subclasses } \mathcal{I}_{3,1} = \{9, 13\}, \mathcal{I}_{3,2} = \{12\}, \mathcal{I}_{3,3} = \{2\} \\
 \mathcal{I}_4 &= \{1, 7, 11\} \\
 \mathcal{I}_5 &= \{3, 6\} \text{ with subclasses } \mathcal{I}_{5,1} = \{3\}, \mathcal{I}_{5,2} = \{6\} \\
 \mathcal{I}_6 &= \{5\} \\
 \mathcal{I}_7 &= \{10\}
 \end{aligned}$$

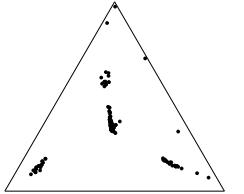
The example is chosen such that most cases appear. So \mathcal{I}_1 is a primitive essential class. \mathcal{I}_2 is essential and irreducible with period $d(\mathcal{I}_2) = 2$. \mathcal{I}_3 is essential and irreducible with period $d(\mathcal{I}_3) = 3$. \mathcal{I}_4 is inessential and primitive but with a zero diagonal. It has only paths $\mathcal{I}_4 \rightarrow \mathcal{I}_1$. \mathcal{I}_5 is inessential and irreducible with period $d(\mathcal{I}_5) = 2$ and has direct paths to $\mathcal{I}_2, \mathcal{I}_3, \mathcal{I}_4$ and an indirect path to \mathcal{I}_1 . \mathcal{I}_6 is inessential and not self-communicating. It has a direct path to \mathcal{I}_5 and thus indirect paths to all other classes except \mathcal{I}_7 . \mathcal{I}_7 is inessential and not self-communicating with no paths. Thus, it corresponds to an index with a zero row and zero column, so it is totally uninteresting.

A Gantmacher form of A looks

$$\begin{array}{c|ccccc|ccccc|c}
 & 4 & 8 & 14 & 9 & 13 & 12 & 2 & 1 & 7 & 11 & 3 & 6 & 5 & 10 & \\
 \hline
 4 & 1 & & & & & & & & & & & & & & 0 \\
 8 & 0 & 1 & & & & & & & & & & & & & 0 \\
 14 & 1 & 0 & & & & & & & & & & & & & \\
 \hline
 9 & & & 0 & 0 & 1 & 0 & & & & & & & & & \\
 13 & & & 0 & 0 & 1 & 0 & & & & & & & & & \\
 12 & & & 0 & 0 & 0 & 1 & & & & & & & & & \\
 2 & 0 & & 1 & 1 & 0 & 0 & & & & & & & & & \\
 \hline
 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & & & 0 \\
 7 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & & & \\
 11 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & & & & \\
 \hline
 3 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & & \\
 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \\
 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 \end{array} . \quad (3.11)$$

Because the least common multiple of 2 and 3 is 6 there is a primitive





3.2. Matrix-based analysis

Gantmacher form for A^6 . Its incidence matrix is

$$\begin{array}{c}
 \begin{array}{ccccccccc}
 4 & \boxed{1} & & & & & & & \\
 8 & & \boxed{1} & 0 & & & & & \\
 14 & & & \boxed{0} & 1 & & & & \\
 & & & & & & & & \\
 9 & & & & 1 & 1 & 0 & 0 & \\
 13 & & & & 1 & 1 & 0 & 0 & \\
 12 & & & & 0 & 0 & 1 & 0 & \\
 2 & & & & 0 & 0 & 0 & 1 & \\
 & & & & & & & & \\
 1 & 1 & & & & 0 & 1 & 1 & 1 & \\
 7 & 1 & & & & & 1 & 1 & 1 & \\
 11 & 1 & & & & & 1 & 1 & 1 & \\
 & & & & & & & & & \\
 3 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\
 6 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\
 5 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\
 10 & 0 & & & & & 0 & & & & 0 & \\
 & & & & & & & & & & & \\
 4 & 8 & 14 & 9 & 13 & 12 & 2 & 1 & 7 & 11 & 3 & 6 & 5 & 10
 \end{array} \\
 . \quad (3.12)
 \end{array}$$

Notice that the block $[3 \ 6 \mid 8 \ 14]$ is not fully filled with ones, because the classes \mathcal{I}_2 and \mathcal{I}_5 have the same period. Contrary the block $[3 \ 6 \mid 9 \ 13 \ 12 \ 2]$ has filled up with ones because the classes \mathcal{I}_3 and \mathcal{I}_5 have periods with no common divisors.

3.2.2 Powers of a row-stochastic matrix

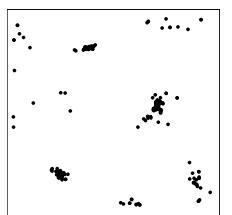
This subsection is to fully characterize the powers of an arbitrary row-stochastic matrix and discuss the implications for homogeneous Markov processes and processes of opinion pooling.

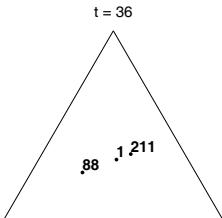
We start with the famous Perron-Frobenius-Theorem for nonnegative matrices and will get step by step to arbitrary row-stochastic matrices applying the primitive Gantmacher form.

Theorem 3.2.1 (Perron-Frobenius). *Let A be an irreducible $n \times n$ matrix with period $d \in \mathbb{N}$ and let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A sorted decreasing with their absolute value. Then it holds*

1. $|\lambda_1| = \dots = |\lambda_d| := \rho(A)$.
2. $\lambda_1, \dots, \lambda_d$ are simple roots of the characteristic equation.
3. $\lambda_1, \dots, \lambda_d$ are roots of the equation $\lambda^d - (\rho(A))^d = 0$. Thus for $j = 1, \dots, d$ it is $\lambda_j = \rho(A)e^{i\frac{j-1}{d}\pi}$ without restriction of generality. In particular λ_1 is real and positive.
4. Moreover, the whole spectrum $\lambda_1, \dots, \lambda_n$ as a cloud of points in the complex plane goes over into itself under a rotation of the plane by the angle $2\pi/d$.
5. To λ_1 there is a positive left and a positive right eigenvector.
6. There can be no other linearly independent nonnegative eigenvector of A .

Proof. See Gantmacher [30, XIII. 2 Theorem 2] for 1.–5. (The existence of a positive right and left eigenvector comes due to the fact that the transpose of an irreducible matrix is again irreducible). For 6. see [30, XIII. 2 Remark 3]. \square





3. Mathematical Analysis

The eigenvalue λ_1 is called the *Perron eigenvalue* and the corresponding left and right positive eigenvectors the *left or right Perron eigenvector*.

Obviously, at least the last n modulo d eigenvalues are zero, because they cannot have enough ‘rotating partners’.

Looking on reducible nonnegative matrices one finds that one needs specific conditions on the essential Gantmacher diagonal blocks such that they have a positive eigenvector. This is especially interesting for us because we know that a row-stochastic matrix has a positive eigenvector.

Theorem 3.2.2. *Let A be a nonnegative matrix in Gantmacher form. A has a positive eigenvector if and only if A_1, \dots, A_g have the same maximal eigenvalue and A_{g+1}, \dots, A_h have smaller maximal eigenvalues.*

Proof. See Gantmacher [30, XIII. §4 Theorem 6]. \square

For a primitive matrix ($d = 1$) theorem 3.2.1 says, that only the Perron eigenvalue has absolute value $\rho(A)$. The existence of this single dominant eigenvalue in a primitive nonnegative matrix leads to the following convergence theorem.

Theorem 3.2.3. *Let $A \in \mathbb{R}^{n \times n}$ be a primitive nonnegative matrix. Let the eigenvalues be sorted as $\rho(A) = \lambda_1 > |\lambda_2| \geq |\lambda_3| \geq \dots |\lambda_n|$ such that λ_2 has algebraic multiplicity m_2 which should be maximal among all eigenvalues of the same absolute value. It holds that*

$$A^t = \lambda_1^t w v + O(t^s |\lambda_2|^t)$$

where w is the right Perron eigenvector of A , v the left Perron eigenvector, normalized such that $vw = 1$ (as a scalar product) and $s := m_2 - 1$.

Proof. See Seneta [69, Theorem 1.2]. \square

Thus, for a primitive matrix A it holds $\lim_{t \rightarrow \infty} \rho(A)^{-t} A^t = vw$ the dyadic product of the right and the left Perron eigenvector with convergence geometrically fast.

For a primitive row-stochastic matrix A this simplifies to

$$\lim_{t \rightarrow \infty} A^t = \mathbf{1}v \quad (3.13)$$

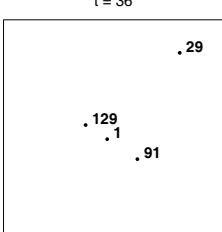
with v being the left Perron eigenvector and convergence geometrically fast.

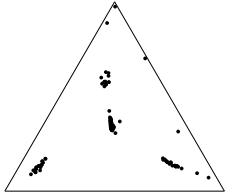
One can also easily see this with the Jordan form. Let J be the Jordan form of A and S be a matrix of eigenvectors and generalized eigenvectors of A in its columns such that

$$AS = SJ.$$

We assume without restriction of generality that the first Jordan block is $[1]$ and thus the first column of S is the right Perron vector (thus a multiple of $\mathbf{1}$). Now, it follows

$$A^t = SJ^t S^{-1} \xrightarrow{t \rightarrow \infty} S \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & 0 & \\ 0 & & & \end{bmatrix} S^{-1} = S_{[:,1]} (S^{-1})_{[1,:]}.$$





 3.2. Matrix-based analysis

And, it is easy to see that $S_{[:,1]}$ and $(S^{-1})_{[1,:]}$ are the right and left Perron eigenvector of A .

The next example shows that it is essential to use the Jordan form, because a primitive matrix is not necessarily diagonalizable.

Example 3.2.4.

$$AS = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = SJ$$

Zero is a double eigenvalue of A , but only the multiples of $[1, -1, -1]^T$ are eigenvectors.

The Jordan form is known to be unique up to permutations of the Jordan blocks. It is easy to see that for a nonnegative matrix the Jordan form can be arranged with respect to the Gantmacher form. Let A be a $n \times n$ nonnegative matrix in Gantmacher form (3.6) with classes $\mathcal{I}_1, \dots, \mathcal{I}_g, \dots, \mathcal{I}_h$. Then there exists a decomposition

$$A = SJS^{-1},$$

with

$$S = \left[\begin{array}{cc|c} S_1 & 0 & 0 \\ \ddots & & \\ 0 & S_g & \\ \hline * & \dots & * & S_{g+1} & 0 \\ \vdots & & \vdots & \ddots & \\ * & \dots & * & * & \dots & S_h \end{array} \right], J = \left[\begin{array}{cc|c} J_1 & 0 & 0 \\ \ddots & & \\ 0 & J_g & \\ \hline & & J_{g+1} & 0 \\ 0 & & 0 & \ddots \\ & & & J_h \end{array} \right]$$

with J_1, \dots, J_h containing only Jordan blocks and for all $i \in \underline{h}$ it holds for the i -th Gantmacher diagonal block that $A_i = S_i J_i S_i^{-1}$. We call this decomposition the *Jordan form with respect to the Gantmacher form*.

Now additionally, let A be row-stochastic. This implies that all g essential Gantmacher diagonal blocks have eigenvalue one which is simple (because Gantmacher blocks are irreducible). Further on, $\mathbf{1}$ is a right eigenvector in each Gantmacher diagonal block. Thus, for all $i \in \underline{g}$ the first column of S_i is $\mathbf{1}$. Additionally, we know that $\mathbf{1}$ is also a right eigenvector of A .

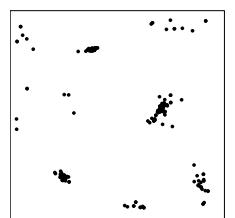
So, we know that the eigenspace to the eigenvalue 1 is

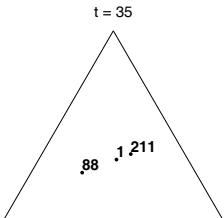
$$\text{eig}(A, 1) = \text{span} \left\{ \begin{bmatrix} \mathbf{1} \\ 0 \\ \vdots \\ 0 \\ * \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{1} \\ \vdots \\ 0 \\ * \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \mathbf{1} \end{bmatrix} \right\} \quad (3.14)$$

with the $*$ -parts summing up to $\mathbf{1}$.

If we assume further that the essential Gantmacher diagonal blocks of A are primitive, then they converge to consensus matrices.

We close this subsection with the full characterization of the powers of an arbitrary row-stochastic matrix up to possible cyclic behavior..





3. Mathematical Analysis

Theorem 3.2.5. Let $A \in \mathbb{R}^{n \times n}$ be a row-stochastic matrix in Gantmacher form. Let $d' \in \mathbb{N}$ be the exponent such that $A^{d'}$ is in primitive Gantmacher form

$$A^{d'} = \left[\begin{array}{ccc|c} A_1 & & 0 & 0 \\ & \ddots & & \\ 0 & & A_{g'} & \\ \hline R_1 & \dots & R_{g'} & Q \end{array} \right],$$

then it holds

$$\lim_{t \rightarrow \infty} (A^{d'})^t = \left[\begin{array}{ccc|c} \mathbf{1}v_1 & & 0 & \\ & \ddots & & \\ 0 & & \mathbf{1}v_{g'} & \\ \hline r_1 v_1 & \dots & r_{g'} v_{g'} & 0 \end{array} \right]$$

with v_i being the left Perron eigenvector of A_i with $\|v_i\|_1 = 1$,

$$r_i := (E - Q)^{-1} R_i \mathbf{1}$$

and convergence geometrically fast. It holds $\sum_{i=1}^{g'} r_i = \mathbf{1}$.

Proof. We conclude for the diagonal blocks and define for the off diagonal blocks that

$$(A^{d'})^t = \left[\begin{array}{ccc|c} A_1^t & & 0 & 0 \\ & \ddots & & \\ 0 & & A_{g'}^t & \\ \hline R_1^{(t)} & \dots & R_{g'}^{(t)} & Q^t \end{array} \right],$$

The convergence in the essential Gantmacher diagonal blocks thus follows from (3.13) and is geometrically fast.

Further on, $Q^t \xrightarrow{t \rightarrow \infty} 0$ geometrically fast, follows directly from the Jordan form with respect to the Gantmacher form, because all eigenvalues of Q are less than one (see Theorem 3.2.2).

Now the question is the convergence for $R_i^{(t)}$ for $t \rightarrow \infty$ and $i \in \underline{g'}$. It is easily checked that for all $i \in \underline{g'}$ it holds

$$R_i^{(t+1)} = \sum_{k=0}^t Q^k R_i A_i^{t-k}.$$

We define $M_i := A_i - \mathbf{1}v_i$, so we know that $M_i \xrightarrow{t \rightarrow \infty} 0$ geometrically fast. Then

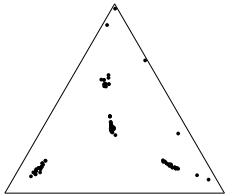
$$R_i^{(t+1)} = \left(\sum_{k=0}^t Q^k \right) R_i \mathbf{1}v_i + \sum_{k=0}^t Q^k R_i M_i^{t-k}. \quad (3.15)$$

From the geometric convergence speed we know that there are $c_Q, c_{M_i} \in \mathbb{R}_{>0}$ and $\delta_Q, \delta_{M_i} < 1$ such that Q^t is elementwise less than by $c_Q \delta_Q^t$ and M_i^t is elementwise less than by $c_{M_i} \delta_{M_i}^t$. So, we know that the right hand summand in (3.15) is less than $c_Q c_{M_i} \sum_{k=0}^t \delta_Q^t \delta_{M_i}^{t-k}$ and thus converges to zero geometrically fast.

t = 35

.29

.129
.1
.91



3.2. Matrix-based analysis

Further on, we know from the geometric convergence rate that

$$\sum_{k=0}^t Q^k = (E - Q)^{-1}.$$

It follows

$$\lim_{t \rightarrow \infty} R_i^{(t+1)} = (E - Q)^{-1} R_i \mathbf{1} v_i = r_i v_i.$$

Further on, it holds

$$\sum_{i=1}^{g'} r_i = \sum_{i=1}^{g'} (E - Q)^{-1} R_i \mathbf{1} = (E - Q)^{-1} \left(\sum_{i=1}^{g'} R_i \right) \mathbf{1} = (E - Q)^{-1} (E - Q) \mathbf{1} = \mathbf{1}.$$

□

Notice that in the case of only one primitive essential class in A the theorem implies that $r_1 = \mathbf{1}$ and thus A converges to a consensus matrix with zeros in all columns of the inessential indices.

Row-stochastic matrices with only one essential class which is primitive are called *regular*. It will turn out that for a row-stochastic matrix ‘regular’ is equivalent to convergence of the powers to a consensus matrix.

The term ‘regular’ is used in the sense of Seneta [69] and differs from the use of Gantmacher [30, §7]. For Gantmacher a regular matrix is a matrix where each essential class is primitive while a fully regular matrix has additionally only one essential class and is thus regular in the sense of Seneta.

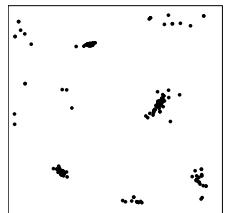
3.2.3 Homogeneous Markov processes and processes of opinion pooling

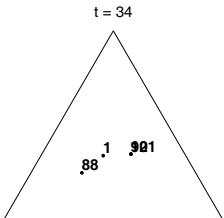
We conclude with a discussion what Theorem 3.2.5 implies for homogeneous Markov processes and processes of opinion pooling when they are determined by an arbitrary row-stochastic matrix. This will also help to understand what dynamics come in when matrices may change.

Markov process. Let us see what the result of Theorem 3.2.5 implies for a Markov process.

If we consider an initial distribution $p(0)$ and the process $p(t) = p(0)A^t$ then we know that we will reach a stable distribution p^* for $p(d't)$ with $t \rightarrow \infty$ and d' being the period determined by the primitive Gantmacher form. All the mass will get absorbed in the essential classes of the primitive Gantmacher form. In each class the mass will be distributed to the states in the proportion determined by the left Perron vector of the respective submatrix. The mass which is in the inessential classes in the beginning will be distributed during the process to the essential states. The proportions are determined by the r_i .

If the Gantmacher form contains only one essential class which is primitive, the limit matrix is a consensus matrix (because $r_1 = \mathbf{1}$). The process $(p(t))_{t \in \mathbb{N}}$ is then called *ergodic*. Ergodic means that we reach the same distribution p^* for every initial distribution $p(0)$. In a probabilistic setting of a random walk with n states and transition probabilities a_{ij} this means that regardless in which state the system starts ($p(0)$ is a stochastic vector, with all the mass in one class), we





3. Mathematical Analysis

will end with fixed probabilities where the random walker is after a sufficiently long time. If the process is not ergodic, then the probabilities can also converge (if they are not cycling), but they will depend on the initial distribution. If the initial mass is in an essential class, it will stay there, if it is in an inessential class, it depends on the paths to essential classes to which of them the mass distributes. The magnitudes how much mass goes to each essential class is determined by the r_i and thus depends also on the paths in the inessential class itself.

So for the general setting of Theorem 3.2.5 we can say that the subprocesses related to a subclass of an essential class of the Gantmacher form of A is ergodic on time steps stretched by the period of the essential class.

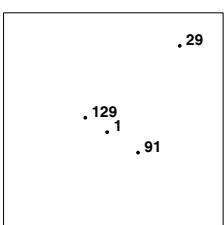
Process of opinion pooling. Let us see what the result of Theorem 3.2.5 implies for a process of opinion pooling.

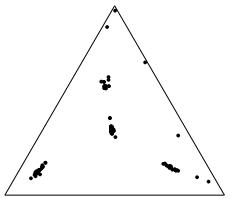
If we consider an initial opinion profile $x(0)$ and the process $x(t) = A^t x(0)$ we know that we will reach a stable opinion configuration x^* for $x(d't)$ with $t \rightarrow \infty$. All agents in the subclasses of the essential classes of the Gantmacher form will find internal consensus. But if the period of their essential class of the Gantmacher form is larger than one then the subgroups may regularly change their opinion with the other subclasses. The inessential agents form opinions as weighted averages of the values of internal consensus of the essential classes. How they distribute their weights between the essential classes is determined by the r_i 's.

We want to explain the link between classical Markov processes and processes of opinion pooling with the example of a group of experts which have individual real-valued opinions on a certain issue and fixed nonnegative recommendation weights for all their colleagues which sum up to one. So, one interpretation is now that A^t represents the recommendation matrix after several rounds of discussion between the agents. But there is an interpretation as a Markov chain. Let us consider a walker who wants to collect the best opinion on the issue. He starts with asking one expert, annotates his opinion and asks for a recommendation of another expert. He gets one with probabilities determined by the recommendation weights of the experts. So, the walker takes a random walk from expert to expert annotating opinions. After very long time he decides to compute the average of all collected opinions of the experts weighted by the proportion of visits he made to each expert. If the Markov process determined by A is ergodic the walker will come to the same result regardless with which expert he starts. If there are more than one essential groups of experts it depends on the first expert in which closed group of experts he ends. If the first expert is an inessential one he gets absorbed in one of the closed groups of experts with probabilities determined by the r_i 's. If one of the essential groups is not primitive the walker will compute his result as an average of all the opinions in the group determined by the weights and the numbers of agents in each subgroup. (But one might regard this situation as unrealistic, because in a non primitive essential class each agent has no recommendation weight for himself. This is a contradiction to the definition of an expert.)

Theorem 3.2.5 also sheds light on necessary conditions for reaching consensus in a homogeneous process of opinion pooling.

Theorem 3.2.6. *Let $x(0) \in S^n \subset (\mathbb{R}^d)^n$ be an initial opinion profile with S being an appropriate opinion space and let $A \in \mathbb{R}^{n \times n}$ be an arbitrary row-*





3.2. Matrix-based analysis

stochastic matrix. The homogeneous process of opinion pooling $(A^t x(0))_{t \in \mathbb{N}}$ converges to consensus if and only if

$$v_1 x^{\mathcal{I}_1}(0) = \dots = v_{g'} x^{\mathcal{I}_{g'}}(0). \quad (3.16)$$

With $v_1, \dots, v_{g'}$ being the right Perron eigenvectors of the diagonal matrices in the primitive Gantmacher form of $A^{d'}$ (with d' being the appropriate exponent such that $A^{d'}$ has only primitive diagonal blocks) and $\mathcal{I}_1, \dots, c_{g'}$ being all subclasses of the essential classes of A (which are all essential classes of $A^{d'}$).

Proof. Let $i \in g'$. With Theorem 3.2.5 it is easy to see that $x^{\mathcal{I}_i}(d't) \rightarrow v_i x^{\mathcal{I}_i}(0)$ componentwise. Thus, (3.16) must hold for convergence to consensus. Of course, the opinions in non primitive essential classes will cycle in steps not equal to $d't$ but this does not matter, if the opinions tend to consensus. \square

Another proof is also given by Berger [7].

3.2.4 Infinite products of row-stochastic matrices

Let us consider a sequence of row-stochastic matrices $(A(t))_{t \in \mathbb{N}}$. Now, we are interested in conditions which ensure convergence of the infinite backward product. So, the question is: Does $\lim_{t \rightarrow \infty} A(t, 0)$ exist? Especially of interest is convergence to a consensus matrix. So, we start with some facts about consensus matrices.

Lemma 3.2.7. *Let $A \in \mathbb{R}^{n \times n}$ be row stochastic. It holds that A is a consensus matrix if and only if for all row-stochastic matrices $B \in \mathbb{R}^{n \times n}$ it holds $BA = A$.*

Proof. If A is a consensus matrix this is equivalent to the fact that every convex combination of the rows of A is again the consensual row of A which is equivalent to the fact that for all row-stochastic B it holds for all $i, j \in \underline{n}$ that $\sum_{k=1}^n b_{ik} a_{kj} = a_{ij}$. \square

Lemma 3.2.8. *Let $A \in \mathbb{R}^{n \times n}$ be a consensus matrix. For all row-stochastic matrices $B \in \mathbb{R}^{n \times n}$ it holds that AB is a consensus matrix.*

Proof. Let B be a row-stochastic matrix. The i -th row of AB is a convex combination of the rows of B with the entries of the i -th row of A as convex coefficients. Because all rows in A are equal to each other each row in AB is the same convex combination of rows of B . Thus, AB is a consensus matrix, too. \square

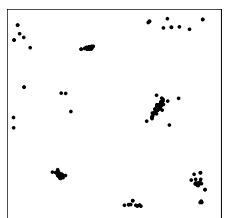
These two results lead us to two facts that show the fundamental theoretical difference between inhomogeneous processes of opinion pooling and inhomogeneous Markov processes.

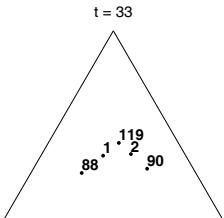
Let us consider for $t_0 \in \mathbb{N}$ that $A(t_0) := K$ is a consensus matrix. The following two facts can be easily checked with the two former lemmas. First it holds for the infinite backward product that

$$\lim_{t \rightarrow \infty} A(t, 0) = \dots A(t_0 + 1)KA(t_0 - 1) \dots A(1)A(0) = KA(t_0, 0). \quad (3.17)$$

Second it holds for all forward products with $t \geq t_0$ that

$$A(0, t) = A(0)A(1) \dots A(t_0 - 1)KA(t_0 + 1) \dots A(t) \quad (3.18)$$





3. Mathematical Analysis

is a consensus matrix. But it could be a different consensus matrix in each time step.

An infinite forward product of row stochastic matrices $A(0, t)$ which is getting closer and closer to the set of consensus matrices for $t \rightarrow \infty$ is called *weakly ergodic*. Formally this means that for $t \rightarrow \infty$ and all two rows $i, i' \in \underline{n}$

$$A(0, t)_{[i,:]} - A(0, t)_{[i',:]} \rightarrow 0. \quad (3.19)$$

The same for infinite backward products $A(t, 0)$.

For an inhomogeneous Markov process weak ergodicity of the infinite forward product implies that the process gets independent of the initial distribution $p(0)$. Thus, changes in the distribution appear only due to the inhomogeneity of the process.

Equation (3.18) shows that one consensus matrix in the sequence of transition matrices ensures weak ergodicity.

If a product converges to a fixed consensus matrix, this is called *strong ergodicity*.

From (3.17) we know that a process of opinion pooling leads to consensus in finite time if there is one consensus matrix in the sequence of confidence matrices. This is a sufficient condition for convergence to consensus in finite time. So (3.17) shows that one consensus matrix is enough to ensure strong ergodicity.

Ergodicity is thus an important concept for convergence to a consensus matrix, which we will point out in the next subsection.

3.2.5 Convergence to consensus matrices

It is easy to see that a row-stochastic matrix $A \in \mathbb{R}^{n \times n}$ which is seen as a map by multiplication from the left

$$A : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{n \times m}, B \mapsto AB$$

is a convex hull averaging map on the rows of the matrix B . This follows directly from A row-stochastic. So, it holds

$$\text{conv}_{i \in \underline{n}}(AB)_{[i,:]} \subset \text{conv}_{i \in \underline{n}}B_{[i,:]}.$$

This is especially of interest for $m = n$ and if A is seen as a self map on the row-stochastic matrices.

This subsection is to transform the results of section 3.1 into matrix language and use the concepts of ergodicity.

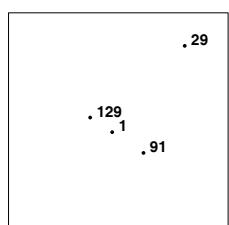
We define the *coefficient of ergodicity for row-stochastic matrices* as

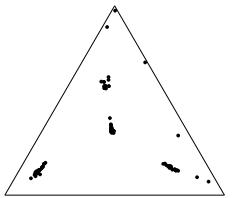
$$\tau_1(A) := \frac{1}{2} \max_{i, i' \in \underline{n}} \sum_{j=1}^n |a_{ij} - a_{i'j}| = \frac{1}{2} \max_{i, i' \in \underline{n}} \|A_{[i,:]} - A_{[i',:]}\|_1$$

It is equivalent to

$$\tau_1(A) = 1 - \min_{i, i' \in \underline{n}} \sum_{j=1}^n \min\{a_{ij}, a_{i'j}\}. \quad (3.20)$$

This follows from the fact that for any $a, b \in \mathbb{R}$ it holds $2 \min\{a, b\} = a + b - |a - b|$ and stochasticity.





 3.2. Matrix-based analysis

Proposition 3.2.9. *The coefficient of ergodicity for stochastic matrices τ_1 has the following properties.*

1. *It is continuous.*
2. *For all row-stochastic $A \in \mathbb{R}^{n \times n}$ it holds $0 \leq \tau_1(A) \leq 1$.*
3. *It is submultiplicative: For any row-stochastic $A, B \in \mathbb{R}^{n \times n}$ it holds*

$$\tau_1(AB) \leq \tau_1(A)\tau_1(B).$$

4. *It is proper in the sense that*

$$\tau_1(A) = 0 \Leftrightarrow A \text{ is a consensus matrix.}$$

Proof. Properties 1. and 4. " \Leftarrow " follow directly from the definition. 2. follows directly from (3.20). For submultiplicativity see the text below or Hartfiel [33, (1.7)]. Finally, 4 ' \Rightarrow ' follows by contraposition: If A is not a consensus matrix there must exist two rows i, i' which are not equal and thus $\tau_1(A) > 0$ by definition. \square

Coefficients of ergodicity can be defined more general. Seneta [70, Definition 4.6] defines a *coefficient of ergodicity* as a function τ from the set of row-stochastic matrices to the reals which fulfills properties 1., 2. and 4. of proposition 3.2.9.

This ensures that $\tau(A(0, t)) \rightarrow 0$ or $\tau(A(t, 0)) \rightarrow 0$ for $t \rightarrow \infty$ is equivalent to weak ergodicity [70, Lemma 4.1]. Thus, this holds for τ_1 , too.

Theorem 3.2.10. *For backward products of row-stochastic matrices weak and strong ergodicity are equivalent.*

Proof. One needs to show that weak ergodicity implies convergence. Consider a weakly ergodic infinite backward product of row-stochastic matrices $A(t, 0)$. Thus for $\varepsilon > 0$ there is t_0 such that for all rows $i, i' \in \underline{n}$ it holds

$$-\varepsilon \leq A(t, 0)_{[i,:]} - A(t, 0)_{[i',:]} \leq \varepsilon$$

By row-stochasticity of $A(t)$ it follows for arbitrary two rows $i, k \in \underline{n}$ that

$$A(t, 0)_{[i,:]} - \varepsilon \leq A(t+1, 0)_{[k,:]} \leq A(t, 0)_{[i,:]} + \varepsilon.$$

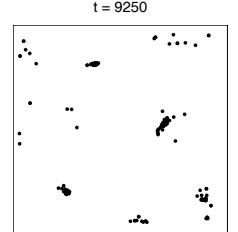
By induction for all $s \in \mathbb{N}$ it holds

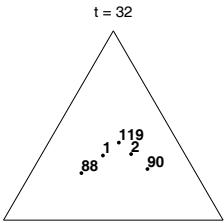
$$A(t, 0)_{[i,:]} - \varepsilon \leq A(t+s, 0)_{[k,:]} \leq A(t, 0)_{[i,:]} + \varepsilon.$$

This implies that $A(t, 0)_{[i,:]}$ is a Cauchy sequence for all rows $i \in \underline{n}$. \square

Thus, it is enough to show weak ergodicity of an infinite backward product of stochastic matrices to ensure convergence to a consensus matrix.

Another generalization of the coefficient of ergodicity for stochastic matrices is defined by Hartfiel [34]. A *contraction coefficient* is a real valued nonnegative function on a set of appropriate matrices (e.g. row-stochastic, but not restricted





3. Mathematical Analysis

to) which fulfills property 3 of proposition 3.2.9. Further on, he defines a *subspace contraction coefficient* τ_W for a subspace W in \mathbb{R}^n as

$$\tau_W(A) := \max_{p \in W, \|p\|=1} \|pA\|, \quad (3.21)$$

for any norm. Then, the coefficient of ergodicity for stochastic matrices τ_1 is a subspace contraction coefficient for $W = \text{span}\{\mathbf{1}\}^\perp$, the orthogonal complement of the subspace spanned by $\mathbf{1}$, as we will see: A basis of $\text{span}\{\mathbf{1}\}^\perp$ can be made of vectors having only two nonzero elements with equal absolute value but different signs. So, the vertices of the unit ball of the 1-norm in $\text{span}\{\mathbf{1}\}^\perp$ are all the vectors which have the two nonzero entries $\frac{1}{2}$ and $-\frac{1}{2}$. According to [34, Theorem 2.12] the vertices of the unit ball of a norm in W are enough to check in equation (3.21). Thus,

$$\tau_{\text{span}\{\mathbf{1}\}^\perp}(A) = \max_{i, i' \in \underline{n}} \left\| \frac{1}{2} A_{[i,:]} - \frac{1}{2} A_{[i',:]} \right\|_1 = \tau_1(A).$$

From (3.21) it follows that for any $p \in \mathbb{R}^n$ with $p\mathbf{1} = 0$ ($p \in \text{span}\{\mathbf{1}\}^\perp$) it holds

$$\|pA\|_1 = \tau_1(A) \|p\|_1.$$

This implies submultiplicativity of τ_1 because for row-stochastic $A, B \in \mathbb{R}^{n \times n}$ it holds

$$\|pAB\|_1 \leq \tau_1(B) \|pA\|_1 \leq \tau_1(A) \tau_1(B) \|p\|_1.$$

Submultiplicativity obviously extends to more than two row-stochastic matrices $A_1, \dots, A_m \in \mathbb{R}^{n \times n}$

$$\tau_1(A_1 \cdots A_m) \leq \tau_1(A_1) \cdots \tau_1(A_m).$$

Further on $\tau_1(A^t) \leq \tau_1(A)^t$. Thus, if $\tau_1(A) < 1$ then the homogeneous Markov process or homogeneous process of opinion pooling is strongly ergodic.

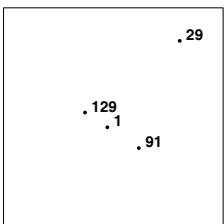
A row-stochastic matrix $A \in \mathbb{R}^{n \times n}$ is called *scrambling* if $\tau_1(A) < 1$. From (3.20) it follows directly that A is scrambling if and only if for all two rows $i, i' \in \underline{n}$ there is a column $j \in \underline{n}$ such that $a_{ij}, a_{i'j} > 0$. This characterization depends only on the zero pattern of A . In that sense the definition of scrambling extends to nonnegative matrices, which will be of interest in the next subsection. Another characterization of a scrambling matrix is that every two rows are not orthogonal.

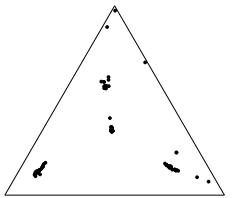
The next proposition shows that the coefficient of ergodicity for stochastic matrices also gives a measure how a matrix A shrinks the diameter of the convex hull of the rows of an arbitrary matrix by multiplication from the left.

Proposition 3.2.11. *Let $A \in \mathbb{R}^{n \times n}$ row-stochastic and $B \in \mathbb{R}^{n \times m}$. Then it holds*

$$\text{diam}(\text{conv}_{i \in \underline{n}}(AB)_{[i,:]}) \leq \tau_1(A) \text{diam}(\text{conv}_{i \in \underline{n}} B_{[i,:]}) .$$

Proof. For a $C = \text{conv}\{c_1, \dots, c_n\}$ it holds $\text{diam}(C) = \max_{i,j} \|c_i - c_j\|_2$. Thus,





3.2. Matrix-based analysis

it holds

$$\begin{aligned}
\text{diam}(\text{conv}_{i \in \underline{n}}(AB)_{[i,:]}) &= \max_{i,j \in \underline{n}} \left\| \sum_k a_{ik} B_{[k,:]} - \sum_{k \in \underline{n}} a_{jk} B_{[k,:]} \right\|_2 \\
&= \max_{i,j \in \underline{n}} \left\| \sum_{k \in \underline{n}} (a_{ik} - \min\{a_{ik}, a_{jk}\}) B_{[k,:]} - \sum_{k \in \underline{n}} (a_{jk} - \min\{a_{ik}, a_{jk}\}) B_{[k,:]} \right\|_2 \\
&= \max_{i,j \in \underline{n}} \left(r_{ij} \underbrace{\left\| \sum_{k \in \underline{n}} \frac{a_{ik} - \min\{a_{ik}, a_{jk}\}}{r_{ij}} B_{[k,:]} \right\|_2}_=: c - \underbrace{\sum_{k \in \underline{n}} \frac{a_{jk} - \min\{a_{ik}, a_{jk}\}}{r_{ij}} B_{[k,:]} \right\|_2}_=: d \right)
\end{aligned}$$

with $r_{ij} := 1 - \sum_{k \in \underline{n}} \min\{a_{ik}, a_{jk}\}$.

The summands c and d are convex combinations of the rows of B . The distance of two convex combinations of points in a convex set is by definition always lower than the diameter of the set. Thus it follows

$$\begin{aligned}
\max_{i,j} (r_{ij} \|c - d\|) &\leq (\max_{i,j} r_{ij}) \text{diam}(\text{conv}_{i \in \underline{n}} B_{[i,:]}) \\
&= \max_{i,j} \left(1 - \sum_k \min\{a_{ik}, a_{jk}\} \right) \text{diam}(\text{conv}_{i \in \underline{n}} B_{[i,:]}) \\
&= \left(1 - \min_{i,j} \sum_k \min\{a_{ik}, a_{jk}\} \right) \text{diam}(\text{conv}_{i \in \underline{n}} B_{[i,:]}) \\
&= \tau_1(A) \text{diam}(\text{conv}_{i \in \underline{n}} B_{[i,:]}) .
\end{aligned}$$

□

This proposition is sometimes called *shrinking lemma* and goes back to Krause [47]. Now we are able to translate the results of Section 3.1 to the matrix case.

Proposition 3.2.12. *Let $A \in \mathbb{R}^{n \times n}$ be a row-stochastic matrix. The map*

$$A : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{n \times m}, B \mapsto AB$$

is an averaging map on the rows of B . It is proper if and only if A is scrambling.

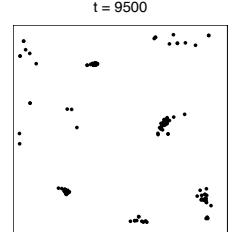
Proof. Let $B \in \mathbb{R}^{n \times n}$. The fact that A is an averaging map on the rows of B follows from A being row-stochastic and thus computing convex combinations of the rows of B .

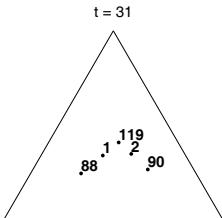
For the ‘only if’-part we consider that A is scrambling and that B is not a consensus. Then it holds by proposition 3.2.11 that

$$\text{diam}(\text{conv}_{i \in \underline{n}}(AB)_{[i,:]}) \leq \tau_1(A) \text{diam}(\text{conv}_{i \in \underline{n}} B_{[i,:]}) < \text{diam}(\text{conv}_{i \in \underline{n}} B_{[i,:]}) .$$

Thus, $\text{conv}_{i \in \underline{n}}(AB)_{[i,:]} \neq \text{conv}_{i \in \underline{n}} B_{[i,:]}$ and this implies that A is a proper averaging map.

We show the ‘if’-part by contraposition. If A is not scrambling, then there are $i, j \in \underline{n}$ such that for all $k \in \underline{n}$ either $a_{ik} = 0$ or $a_{jk} = 0$. Now, we define $B \in \mathbb{R}^{n \times m}$ such that $B_{[k,:]} = [1 0 \dots 0]$ if $a_{ik} = 0$ and $B_{[k,:]} = [0 \dots 0 1]$ otherwise. Then $\text{conv}_{i \in \underline{n}}(AB)_{[i,:]} = \text{conv}_{i \in \underline{n}} B_{[i,:]}$ but B is not a consensus matrix, thus A is not a proper averaging map. □





3. Mathematical Analysis

Now consider a set of $n \times n$ row-stochastic matrices Σ . If we regard Σ as a family of averaging maps as defined in Proposition 3.2.12, then it is clear that Σ is a set of proper averaging maps if and only if all the matrices in Σ are scrambling. The next proposition gives sufficient conditions that a set Σ is an equiproper set of averaging maps.

Proposition 3.2.13. *Let Σ be a set of $n \times n$ row-stochastic matrices regarded as averaging maps on the rows of $n \times m$ matrices by multiplication from the left. If there is $\delta > 0$ such that for all $A \in \Sigma$ it holds that $\tau_1(A) < 1 - \delta$ then Σ is an equiproper set of averaging maps.*

Proof. Let $\delta > 0$. Then there is $A \in \Sigma$ such that $\tau_1(A) < 1 - \delta$ and thus

$$\text{diam}(\text{conv}_{i \in \underline{n}}(AB)_{[i,:]}) < (1 - \delta)\text{diam}(\text{conv}_{i \in \underline{n}}B_{[i,:]})$$

Further on $\text{conv}_{i \in \underline{n}}B_{[i,:]} \subset \text{conv}_{i \in \underline{n}}(AB)_{[i,:]}$. In the worst case it must still hold that $d_H(\text{conv}_{i \in \underline{n}}(AB)_{[i,:]}, \text{conv}_{i \in \underline{n}}B_{[i,:]}) < \frac{\delta}{2}$, which proves that Σ is equiproper. \square

In the following we need another definition. For a nonnegative matrix A the *positive minimum* is defined as

$$\min^+ A := \min_{\{(i,j) \mid a_{ij} > 0\}} a_{ij}.$$

So, the minimum is taken over all positive entries of A .

Proposition 3.2.14. *Let A be a $n \times n$ row-stochastic matrix. If A is scrambling and $\min^+ A > \delta$ then $\tau_1(A) < 1 - \delta$.*

Proof. Follows from (3.20). \square

So, if Σ is a set of scrambling matrices which positive minimum is uniformly bounded from below by δ , then Proposition 3.2.13 applies and Σ is equiproper. But the converse of Proposition 3.2.14 need not be true. If Σ is an equiproper family of averaging maps this does not imply that all $A \in \Sigma$ are scrambling and $\min^+ A > \delta$, as the following example shows.

Example 3.2.15.

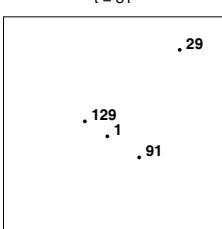
$$\Sigma := \left\{ \begin{bmatrix} 1 & 0 \\ 1 - \frac{1}{t} & \frac{1}{t} \end{bmatrix} \mid t \in \mathbb{N}_{\geq 2} \right\}$$

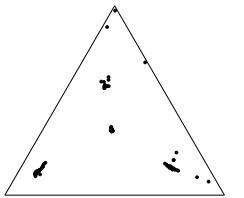
is an equiproper family of averaging maps (as defined in proposition 3.2.13), because for all $A \in \Sigma$ it holds $\tau_1(A) \geq \frac{1}{2}$, but there is no $\delta > 0$ such that for all $A \in \Sigma$ it holds $\min^+ A > \delta$. This is the case because the entry a_{22} which converges to zero is irrelevant for the scrambling property.

The description of proper averaging maps by scrambling plus a uniform positive minimum is easy to understand and easy to check, but not sharp. Nevertheless, a uniformly bounded positive minimum appears in applications (see Section 3.3). We will treat it again in Subsection 3.2.7.

The following equivalence remains as a conjecture.

Conjecture 3.2.16. *Let Σ be a set of $n \times n$ row-stochastic matrices regarded as averaging maps on the rows of $n \times m$ matrices by multiplication from the left. It holds that Σ is equiproper if and only if there is $\delta > 0$ such that for all $A \in \Sigma$ it holds $\tau_1(A) < 1 - \delta$.*





3.2. Matrix-based analysis

Nevertheless, the theorems about averaging maps in Section 3.1 are now widely applicable to the matrix case. We give the most general result.

Corollary 3.2.17. *Let Σ be a set of row-stochastic $n \times n$ matrices and let $(A(t))_{t \in \mathbb{N}}$ be a sequence of matrices in Σ . If there is a sequence of time steps $(t_s)_{s \in \mathbb{N}}$ and $0 < \delta < 1$ such that for all $s \in \mathbb{N}$ it holds that $\tau_1(A(t_{s+1}, t_s)) < 1 - \delta$ then there is a consensus matrix K such that*

$$\lim_{t \rightarrow \infty} A(t, 0) = K.$$

Proof. Proposition 3.2.13 makes Corollary 3.1.11 applicable. The composition of maps translates to the multiplication matrices. The initial opinion profile is analog to $A(0)$. \square

Until now the matrix formalism has not delivered anything more than the nonlinear case of averaging maps. If Conjecture 3.2.16 were true then averaging maps in the matrix case would be fully characterized, but until now it is not clear if the averaging map theorem has something more to say about the matrix case.

Now we present a theorem for the matrix case similar to Corollary 3.2.17 but more general. It makes direct use of the coefficient of ergodicity for stochastic matrices.

Theorem 3.2.18. *Let Σ be a set of row-stochastic $n \times n$ matrices and let $(A(t))_{t \in \mathbb{N}}$ be a sequence of matrices in Σ . If there is a sequence of time steps $(t_s)_{s \in \mathbb{N}}$ and a sequence of positive numbers $(\delta_s)_{s \in \mathbb{N}}$ with $\sum_{s=0}^{\infty} \delta_s = \infty$ such that for all $s \in \mathbb{N}$ it holds that $\tau_1(A(t_{s+1}, t_s)) < 1 - \delta_s$ then there is a consensus matrix K such that*

$$\lim_{t \rightarrow \infty} A(t, 0) = K.$$

Proof. We define $B(s) := A(t_{s+1}, t_s)$. Then, it holds due to (3.2.9 3.) and the definition of the coefficient of ergodicity that

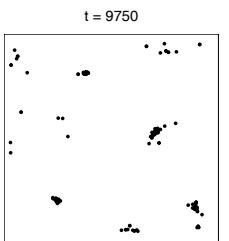
$$\lim_{s \rightarrow \infty} \tau(B(s, 0)) \leq \prod_{s=0}^{\infty} \tau(B(s)) = \prod_{s=0}^{\infty} (1 - \delta_s) \leq \prod_{s=0}^{\infty} e^{-\delta_s} = e^{-\sum_{s=0}^{\infty} \delta_s} = 0.$$

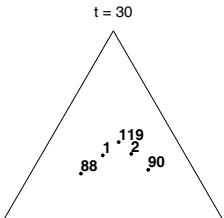
Thus, $B(s, 0)$ is weakly ergodic which implies strongly ergodic due to Theorem 3.2.10. Thus $A(t, 0)$ converges to consensus. \square

This result is more general than Corollary 3.2.17, because it allows $A(t)$ to converge to non-scrambling matrices but ensures convergence to consensus of $A(t, 0)$ if this does not happen too fast. The result goes back to Krause [47] but not in the context of coefficients of ergodicity.

Obviously, the assumption $\tau_1(A(t_{s+1}, t_s)) < 1 - \delta_s$ in Theorem 3.2.18 can be exchanged by $A(t_{s+1}, t_s)$ is scrambling and $\min^+ A(t_{s+1}, t_s) > \delta$. But this weakens the theorem.

The next subsection is about conditions on the zero patterns of individual matrices in the sequence $(A(t))_{t \in \mathbb{N}}$ which ensure that there exist accumulations such that Theorem 3.2.18 can be applied.





3. Mathematical Analysis

3.2.6 Convergence of the zero pattern

In this subsection we consider a sequence of row allowable matrices $(A(t))_{t \in \mathbb{N}}$. It is easy to see that products of row allowable matrices are row allowable. Here, we are interested in convergence of the zero patterns in $A(t, 0)$, so on convergence in the incidence sense. The assumption of row allowability ensures that every sequence of row-stochastic matrices is a special case. Especially we are interested in conditions on the zero patterns of the matrices in $(A(t))_{t \in \mathbb{N}}$ such that there is a sequence of time steps $(t_s)_{s \in \mathbb{N}}$ such that $A(t_{s+1}, t_s)$ has certain properties such that theorems of the former subsection can be applied.

Some basic results

First, two facts on row allowable matrices.

Lemma 3.2.19. *A row allowable matrix has at least one essential class of indices.*

Proof. See Seneta [69, Lemma 1.1]. □

Proposition 3.2.20. *Let A, B be row allowable matrices. Then it holds:*

1. *If column $j \in \underline{n}$ in A is positive, then column j in BA is positive.*
2. *If column $j \in \underline{n}$ in A is zero, then column j in BA is zero.*

Proof. Obvious with the rules of matrix multiplication. □

This proposition shows that if a product of row allowable matrices infinite to the left $A(t, 0)$ has a positive (respectively zero) column for some $t \in \mathbb{N}$ then these column remains positive (respectively zero) for further time steps. Unfortunately, one cannot say anything about the nonzero and non positive columns.

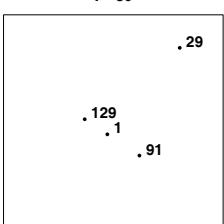
Second, we state a lemma which shows that one can use the concept of a majorizing or minorizing sequence for infinite products of nonnegative matrices in the incidence sense.

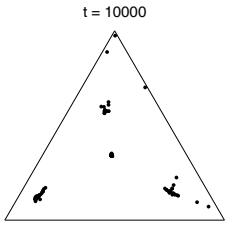
Lemma 3.2.21. *Let $(A(t))_{t \in \mathbb{N}}$ and $(B(t))_{t \in \mathbb{N}}$ be sequences of nonnegative matrices. If it holds for all $t \in \mathbb{N}$ that $\text{inc}(A(t)) \leq \text{inc}(B(t))$ then it holds for all $s, t \in \mathbb{N}$ that $\text{inc}(A(s, t)) \leq \text{inc}(B(s, t))$.*

Proof. Follows directly from the rules of matrix multiplication. □

Conditions for scrambling sub-accumulations

The central question in this subsection is about the conditions on the zero patterns of the matrices in $(A(t))_{t \in \mathbb{N}}$ such that there is a sequence of time steps $(t_s)_{s \in \mathbb{N}}$ such that $A(t_{s+1}, t_s)$ is scrambling. The sequence $(A(t_{s+1}, t_s))_{s \in \mathbb{N}}$ is called a sequence of *sub-accumulations*. The question is important because scrambling sub-accumulations are crucial for convergence to a consensus matrix in theorem 3.2.18. The *length of an accumulation* $A(s, t)$ is $|s - t|$. Often it is of interest if $(t_s)_{s \in \mathbb{N}}$ can be chosen such that the lengths of the sub-accumulations are bounded.





In the incidence sense it is especially of interest that A is scrambling if and only if for all two rows $i, i' \in \underline{n}$ there is a column $j \in \underline{n}$ such that $a_{ij}, a_{i'j} > 0$.

We recall that a matrix A is regular if it has only one essential class which is primitive. So, regular is determined only by the zero pattern, too. We know from Theorem 3.2.5 that A^t converges to a consensus matrix if and only if A is regular. So, it is of interest if this property is also important in inhomogeneous products of row-stochastic matrices.

Proposition 3.2.22. *A scrambling matrix is regular.*

Proof. Consider A to be not regular. So due to Lemma 3.2.19, it either has an essential class which is periodic with $d \geq 2$ or it has at least two essential classes, both implies that A has two rows which have in every column no joint positive entries due to the Gantmacher form structure. Thus, A is not scrambling. \square

For $(A(t))_{t \in \mathbb{N}}$ we define the *set of important zero patterns* $\mathcal{G}((A(t))_{t \in \mathbb{N}})$ as the set of all incidence matrices that appear infinitely often as zero patterns in $(A(t))_{t \in \mathbb{N}}$. Clearly, $\mathcal{G}((A(t))_{t \in \mathbb{N}})$ is not empty and finite.

Obviously, there is a time step t_0 such that for all $t \geq t_0$ it holds that $\text{inc}(A(t)) \in \mathcal{G}((A(t))_{t \in \mathbb{N}})$.

Theorem 3.2.23 (Sufficient conditions for scrambling sub-accumulations). *Let $(A(t))_{t \in \mathbb{N}}$ be a sequence of nonnegative matrices and their set of important zero patterns be $\mathcal{G} := \mathcal{G}((A(t))_{t \in \mathbb{N}})$. There exists a sequence of time steps $(t_s)_{s \in \mathbb{N}}$ such that for all $s \in \mathbb{N}$ the accumulation $A(t_{s+1}, t_s)$ is scrambling if either*

1. *there is $B \in \mathcal{G}$ which is scrambling, or*
2. *all finite products of matrices in \mathcal{G} are regular.*

In case 1. the length of the sub-accumulations $t_{s+1} - t_s$ is bounded by the maximal distance of two scrambling matrices in $(A(t))_{t \in \mathbb{N}}$ if it exists. In case 2. the length of the sub-accumulations $t_{s+1} - t_s$ is bounded by $T + 1$ with T being the number of all regular zero patterns of size $n \times n$.

Proof. To prove 1. it is enough to prove that for a scrambling matrix A and a row allowable matrix B it holds that AB and BA is scrambling. The proof is simple, see Seneta [70, Theorem 4.11 and Lemma 4.6].

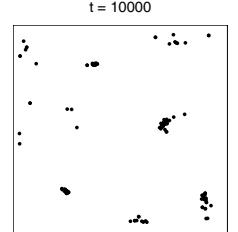
To prove 2. we first show that if B is a regular matrix and $AB \sim A$, then A is scrambling. $AB \sim A$ implies $AB^2 \sim AB \sim A$ and thus for all $t \in \mathbb{N}$ it holds $AB^t \sim A$. For t large, B^t has a positive column and is thus scrambling. As stated above this implies that AB^t is scrambling which is of the same type as A .

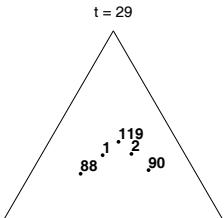
Now, let us consider an arbitrary accumulation of length $T + 1$. We relabel it $A_1 A_2 \dots A_{T+1}$. By the definition of T there exist $r, s \in \mathbb{N}$ with $0 < r < s \leq T + 1$ such that

$$A_1 \dots A_r \sim A_1 \dots A_s.$$

This implies that $A_1 \dots A_r$ is scrambling and thus the whole accumulation. \square

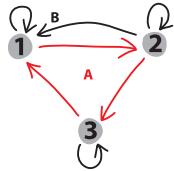
The idea of this proof is due to Wolfowitz [89, Lemma 4]. There is an additional sufficient condition for scrambling sub-accumulations in the next paragraph on matrices with positive diagonals. The next example shows that both conditions in the theorem are not necessary.





3. Mathematical Analysis

Example 3.2.24. We define two matrices which both are not regular and thus also not scrambling.



$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then it holds

$$BAB = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

which is scrambling. Further on $(AB)^t$ and $(BA)^t$ is positive for $t \geq 4$, while $(AAAB)^t$ never gets positive for all $t \in \mathbb{N}$. So, there are scrambling (and positive) accumulations as desired for Theorem 3.2.23 e.g. in the infinite product $\dots BABA$. But there are none in the infinite product $\dots AAABAAAB$.

The next example shows that it is not enough to assume that all matrices in $\mathcal{G}((A(t))_{t \in \mathbb{N}})$ are regular in Theorem 3.2.23.

Example 3.2.25. We define two regular matrices

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

A is primitive and B has only the essential class $\{1, 2\}$ which is primitive. But

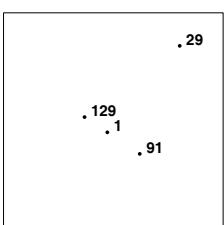
$$\text{inc}(AB) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

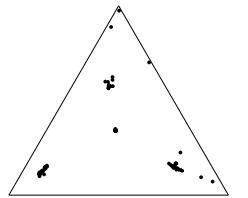
has two essential classes $\{1\}$ and $\{2\}$. So, the infinite product $\dots ABABAB$ cannot be divided into scrambling accumulations.

Theorem 3.2.26 (Necessary conditions for scrambling sub-accumulations). Let $(A(t))_{t \in \mathbb{N}}$ be a sequence of row allowable matrices and let $\mathcal{G} := \mathcal{G}((A(t))_{t \in \mathbb{N}})$ be their set of important zero patterns. If there exists a sequence of time steps $(t_s)_{s \in \mathbb{N}}$ such that for all $s \in \mathbb{N}$ the accumulation $A(t_{s+1}, t_s)$ is scrambling then $\text{inc}(\sum_{B \in \mathcal{G}} B)$ must be regular.

Proof. By contraposition. If $\text{inc}(\sum_{B \in \mathcal{G}} B)$ is not regular then it must have (due to lemma 3.2.19) either a periodic essential class with period $d \geq 2$ or has at least two essential classes. This must hold for each matrix in $(A(t))_{t \in \mathbb{N}}$. Let A^* be a matrix with zero pattern that has the same Gantmacher structure as $\text{inc}(\sum_{B \in \mathcal{G}} B)$ but with $\text{inc}(A(t)) \leq \text{inc}(A^*)$ for all $t \in \mathbb{N}$. So the sequence $\dots A^* A^* A^* A^*$ is a majorant in the incidence sense for $\dots A(3)A(2)A(1)A(0)$. But every power of A^* is either periodic with period greater than 2 or has more than one essential classes. Both is in contradiction to scrambling. \square

The next example shows that the conditions in Theorem 3.2.26 are not sufficient.





 3.2. Matrix-based analysis

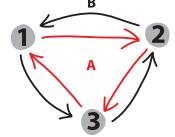
Example 3.2.27. We define two permutation matrices

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Then it holds

$$\text{inc}(A + B) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

which is primitive and thus also regular. But every infinite product of A 's and B 's will always remain only a permutation matrix.

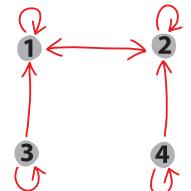


The term $\text{inc}(\sum_{B \in G} B)$ appears similar Moreau's Theorem 3.1.15. There $\text{inc}(\sum_{B \in G} B)$ being regular together with a positive diagonal in all network matrices is essential for convergence. The role of the positive diagonal is important. We will study it in the following.

Seneta [69] defines as G_1 the set of all regular matrices and G_3 the set of all scrambling matrices. While a scrambling matrix in a product ensures that the whole product is scrambling, this does not hold for regular matrices. So, he defines the other way round as G_2 the set of row-stochastic regular matrices that preserve the regular property when multiplied to an arbitrary regular matrix. The matrices in G_2 need not be scrambling as the following example shows.

Example 3.2.28 (Seneta).

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix},$$



is not scrambling but ensures that for every regular B it holds that AB and BA are regular.

The matrix in the example has a positive diagonal which is crucial for being in G_2 . The next paragraph is about nonnegative matrices with positive diagonals which have also some other good properties which may lead to convergence of infinite backward accumulations of row-stochastic matrices.

Matrices with positive diagonals

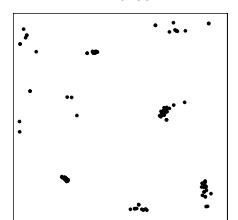
Here, we assume that all matrices in $(A(t))_{t \in \mathbb{N}}$ have positive diagonals. The main effect is, that a positive diagonal conserves positive entries and prevents cyclic behavior. It has thus a strong impact on convergence of the zero pattern.

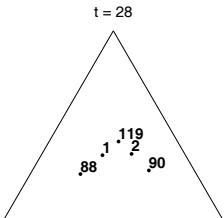
The source of this impact is the following lemma.

Lemma 3.2.29. Let A and B be nonnegative matrices and let A have a positive diagonal. Then $\text{inc}(BA) \geq \text{inc}(B)$ and $\text{inc}(AB) \geq \text{inc}(B)$.

Proof. We decompose A into the sum of a diagonal matrix and a matrix with all off diagonal entries of A .

$$A := A_{\text{diag}} + A_{\text{offdiag}}$$





3. Mathematical Analysis

Then,

$$\text{inc}(BA) = \text{inc}(B(A_{\text{diag}} + A_{\text{offdiag}})) \geq \text{inc}(BA_{\text{diag}}) = \text{inc}(B).$$

For AB analog. \square

The lemma makes another sufficient condition for scrambling sub-accumulations additional to the ones in Theorem 3.2.23 possible.

Theorem 3.2.30 (Sufficient conditions for scrambling sub-accumulations (continued)). Let $(A(t))_{t \in \mathbb{N}}$ be a sequence of nonnegative matrices and let $\mathcal{G} := \mathcal{G}((A(t))_{t \in \mathbb{N}})$ be their set of important zero patterns. There exists a sequence of time steps $(t_s)_{s \in \mathbb{N}}$ such that for all $s \in \mathbb{N}$ the accumulation $A(t_{s+1}, t_s)$ is scrambling if

- 3. all matrices in \mathcal{G} are regular and have positive diagonals.

The length of the sub-accumulations $t_{s+1} - t_s$ is bounded by $T + 1$ with T being the number of all regular zero patterns with positive diagonals of size $n \times n$.

Proof. By lemma 3.2.29 it holds that all products of matrices in \mathcal{G} are regular. Thus condition 2. of theorem 3.2.23 is fulfilled. \square

But positive diagonals can also play a role for convergence in general when matrices are not regular. The first version of this Theorem is in Lorenz [52, Abschnitt 2.2].

Theorem 3.2.31. Let $(A(t))_{t \in \mathbb{N}}$ be a sequence of nonnegative matrices with positive diagonals. Then there exists a sequence of natural numbers $(t_s)_{s \in \mathbb{N}}$ such that for all $s \in \mathbb{N}$ it holds

$$A(t_{s+1}, t_s) \sim A(t_1, t_0). \quad (3.22)$$

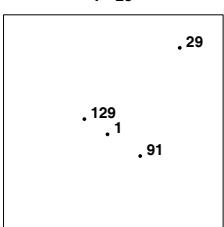
Thus, $A(t_{s+1}, t_s)$ can be brought to the same Gantmacher form for all $s \in \mathbb{N}$. Further on, all Gantmacher diagonal blocks are positive and all non diagonal Gantmacher-Blocks are either positive or zero.

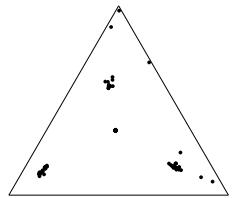
The same holds for forward accumulations.

Proof. The proof works with a double monotonic argument on the positivity of entries: Due to Lemma 3.2.29 more and more (or exactly the same) positive entries appear in $A(t, 0)$ monotonously increasing with rising t . Due to finiteness of the number of entries we reach a maximum at a time step t_0^* . We cut $A(t_0^*, 0)$ off and find t_1^* when $A(t, t_0^*)$ reaches maximal positivity again with rising t . We go on like this and get the sequence $(A(t_{s+1}^*, t_s^*))_{s \in \mathbb{N}}$. Obviously, less and less (or exactly the same) positive entries appear monotonously decreasing with rising s and we reach a minimum at k . We relabel $t_j := t_{k+j}^*$ and thus have the desired sequence $(t_s)_{s \in \mathbb{N}}$ with $A(t_{s+1}, t_s)$ having the same zero-pattern.

So the product of sub-accumulations $(A(t_{s+1}, t_s))_{s \in \mathbb{N}}$ behaves in the incidence sense as the powers of $A(t_1, t_0)$. Positivity of Gantmacher blocks follows for all blocks $A(t_{i+1}, t_i)_{[\mathcal{J}, \mathcal{I}]}$ where we have a path $\mathcal{J} \rightarrow \mathcal{I}$. If we have such a path, then there is a path from each index in \mathcal{J} to each index in \mathcal{I} and thus every entry must be positive in a long enough accumulation. Thus, the block has to be positive already, otherwise $(t_i)_{i \in \mathcal{I}}$ is chosen wrong.

To prove the result for forward accumulations, we can use the same arguments. \square





3.2. Matrix-based analysis

The positive diagonal in each matrix in $(A(t))_{t \in \mathbb{N}}$ thus delivers a sequence of sub-accumulations with the same Gantmacher structure and primitive Gantmacher diagonal blocks. Unfortunately, the length of the sub-accumulations $t_{s+1} - t_s$ need not be bounded.

Corollary 3.2.32. *Let $\mathcal{G} := \mathcal{G}((A(t))_{t \in \mathbb{N}})$ be the set of important zero patterns for the sequence of matrices $(A(t))_{t \in \mathbb{N}}$ which all have positive diagonals. Then there is an exponent $k \in \mathbb{N}$ such that*

$$\text{inc} \left(\sum_{B \in \mathcal{G}} B \right)^k \sim A(t_1, t_0)$$

with t_0, t_1 defined in Theorem 3.2.31.

3.2.7 Convergence

Here, we come back to the central question of this section. What are conditions for convergence of the infinite backward accumulation $A(t, 0)$ for a sequence of row-stochastic matrices $(A(t))_{t \in \mathbb{N}}$? First, we put the results of the three former subsections together.

Putting results together

In Subsection 3.2.4 we saw that one consensus matrix in $(A(t))_{t \in \mathbb{N}}$ is enough to ensure convergence in finite time to this consensus matrix.

In Subsection 3.2.5 we saw that the existence of a sequence of sub-accumulations $A(t_{s+1}, t_s)$ with a coefficient of ergodicity less than $(1 - \delta_s)$ with $\sum_{s=1}^{\infty} \delta_s = \infty$ is sufficient to ensure convergence to a consensus matrix. One easy to check sufficient but not necessary condition is that all sub-accumulations are scrambling and have a positive minimum $\min^+ A(t_{s+1}, t_s) \geq \delta_s$. So, the positive minimum may converge to zero, but not too fast.

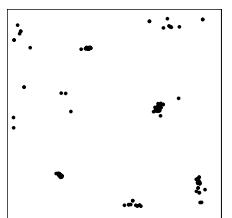
Further on, we know that the diameter of the convex hull of the rows in $A(t, 0)$ can only shrink if multiplied from the left with a scrambling matrix. Thus, these conditions are close to necessary conditions besides convergence in finite time or very specific assumptions on the matrices.

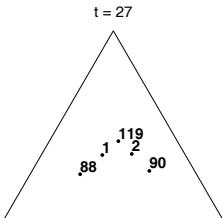
In Subsection 3.2.6 we gave some sufficient and some necessary conditions for the existence of scrambling sub-accumulations. There we studied the zero patterns of the matrices in the set of important zero patterns of $(A(t))_{t \in \mathbb{N}}$ and came to the conclusion that their existence can heavily depend on the specific order of matrices. Three necessary conditions are given in the Theorems 3.2.23 and 3.2.30. But, it is still necessary to assume $\tau_1(A(t_{s+1}, t_s)) \leq 1 - \delta_s$ to ensure convergence to a consensus matrix. Thus, assumptions on the zero patterns of individual matrices in $(A(t))_{t \in \mathbb{N}}$ are not enough to ensure convergence.

Further on, we saw that a positive diagonal in every matrix of $(A(t))_{t \in \mathbb{N}}$ leads to the existence of a sequence of sub-accumulation with the same zero pattern, positive Gantmacher diagonal blocks and Gantmacher subdiagonal blocks which are either positive or zero.

So, the following theorem can be derived.

Theorem 3.2.33. *Let $(A(t))_{t \in \mathbb{N}}$ be a sequence of row-stochastic matrices with positive diagonals, $(t_s)_{s \in \mathbb{N}}$ be the sequence of time steps defined by Theorem*





3. Mathematical Analysis

3.2.31, $\mathcal{I}_1, \dots, \mathcal{I}_g$ be the essential classes and \mathcal{J} be the union of all inessential classes of $A(t_1, t_0)$.

If for all $s \in \mathbb{N}$ it holds $\min^+(A(t_{s+1}, t_s)) \geq \delta_s$ and $\sum_{s=1}^{\infty} \delta_s = \infty$, then

$$\lim_{t \rightarrow \infty} A(t, 0) = \left[\begin{array}{ccc|c} K_1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & K_g & 0 \\ \hline & \text{not converging} & & 0 \end{array} \right] A(t_0, 0)$$

where K_1, \dots, K_g are consensus matrices. (All matrices must be transformed simultaneously to the Gantmacher form.)

Proof. Due to Theorem 3.2.31 the Gantmacher diagonal blocks $A(t_{s+1}, t_s)$ are all positive and thus scrambling. For each block $\mathcal{I} \in \{\mathcal{I}_1, \dots, \mathcal{I}_g\}$ it holds further on that $\min^+(A(t_{s+1}, t_s)_{[\mathcal{I}, \mathcal{I}]}) \geq \delta_s$. Thus, convergence to a consensus matrix in the essential classes is proved by Theorem 3.2.18.

It remains to show that the diagonal block of all inessential classes converges to zero. We define $\|\cdot\|$ as the maximum-row-sum-norm for nonnegative matrices. It holds $\|A(t_{s+1}, t_s)_{[\mathcal{J}, \mathcal{J}]}\| \leq (1 - \delta_s)$ because there is at least one positive subdiagonal Gantmacher block for each inessential class. Thus, it holds

$$\|A_{[\mathcal{J}, \mathcal{J}]}(\infty, t_0)\| \leq \prod_{s=1}^{\infty} \|A_{[\mathcal{J}, \mathcal{J}]}(t_{s+1}, t_s)\| \leq \prod_{s=1}^{\infty} (1 - \delta_s) \leq \prod_{s=1}^{\infty} e^{-\delta_s} = e^{-\sum_{s=1}^{\infty} \delta_s}.$$

This proves that $\lim_{t \rightarrow \infty} A_{[\mathcal{J}, \mathcal{J}]}(t, 0) = 0$. \square

Notice that the formulation with \min^+ is crucial for the convergence to zero in the inessential blocks.

In Subsection 3.2.9 we give an explanation what may happen in the subdiagonal Gantmacher blocks which do not converge.

Conditions for $\min^+ A(t_{s+1}, t_s) \geq \delta_s$

If we can ensure scrambling sub-accumulations in an arbitrary infinite product $A(t, 0)$ by Theorem 3.2.23 or 3.2.30 we still have to assume something more to ensure convergence to a consensus matrix. A sufficient condition is

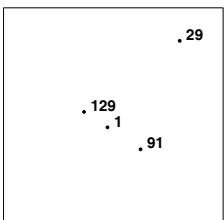
$$\min^+ A(t_{s+1}, t_s) \geq \delta_s \text{ for } \sum_{s=1}^{\infty} \delta_s = \infty. \quad (3.23)$$

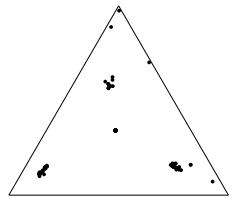
The same assumption appears in Theorem 3.2.33. It would be desirable to have assumptions on the individual matrices which ensure (3.23).

First, we want to emphasize that $\min^+ A(t)$ may go to zero with $t \rightarrow \infty$. There is already the trivial Example 3.2.15 but here we give an example where also $\tau_1(A(t))$ goes to one.

Example 3.2.34.

$$A(t) := \left\{ \begin{bmatrix} 1 & 0 \\ \frac{1}{t} & 1 - \frac{1}{t} \end{bmatrix} \mid t \in \mathbb{N}_{\geq 2} \right\}$$





3.2. Matrix-based analysis

for $t \in \mathbb{N}_{\geq 1}$ is a sequence of scrambling matrices, and for $\delta_t := \frac{1}{t}$ it holds that $\tau_1(A(t)) = 1 - \frac{1}{t}$ and $\sum_{t=1}^{\infty} \delta_t = \infty$. Thus, $A(t, 0)$ converges to the consensus matrix $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$. But $\min^+ A(t) = \frac{1}{t}$ goes to zero. This example is similar to Example 3.1.13.

But often one knows that there is $\delta > 0$ such that for all $t \in \mathbb{N}$ it holds $\min^+ A(t) \geq \delta$. So, now we will study a sequence of row-stochastic matrices for which this holds.

The positive minimum is *supermultiplicative* due to the rules of matrix multiplication. For a set of row-stochastic matrices A_0, \dots, A_t it holds

$$\min^+(A_t \cdots A_0) \geq \min^+ A_t \cdots \min^+ A_0. \quad (3.24)$$

Thus, for a sequence of scrambling sub-accumulations $(A(t_{s+1}, t_s))_{s \in \mathbb{N}}$, the lengths of the sub-accumulations is of interest for convergence results. We will finish this subsection with three theorems on convergence. But we need two preliminary lemmas.

Lemma 3.2.35. *Let $0 < \delta < 1$ and $T \in \mathbb{R}_{>0}$ then*

$$\sum_{s=3}^{\infty} \delta^{T \log(\log(s))} = \infty. \quad (3.25)$$

Proof. We can use the integral test for the series $\sum_{s=1}^{\infty} \delta^{T \log(\log(s))}$ because $f(x) := \delta^{T \log(\log(x))}$ is positive and monotonously decreasing on $[3, \infty[$.

With substitution $y = \log(\log(x))$ (thus $dx = e^{(y+e^y)} dy$) it holds

$$\begin{aligned} \int_3^{\infty} \delta^{T \log(\log(x))} dx &= \int_1^{\infty} e^{T \log(\delta) \log(\log(x))} dx = \int_1^{\infty} e^{T \log(\delta) y} e^{y+e^y} dy \\ &= \int_1^{\infty} e^{Ty(\log(\delta)+1)+e^y} dy \end{aligned}$$

The integral diverges because $Ty(\log(\delta) + 1) + e^y \rightarrow \infty$ as $y \rightarrow \infty$. \square

Lemma 3.2.36. *Let $(A(t))_{t \in \mathbb{N}} \in \mathbb{R}_{\geq 0}^{n \times n}$ be a series of row-stochastic matrices which have positive diagonals, are type-symmetric and there is $\delta > 0$ such that for all $t \in \mathbb{N}$ it holds $\min^+ A(t) > \delta$. Then it holds for every two time steps $t_0 < t_1$ that the lowest positive entry of $A(t_0, t_1)$ is greater than δ^{n^2-n+2} .*

Proof. See Lorenz [54, Proposition 4]. \square

Theorem 3.2.37 (Convergence to a consensus matrix). *Let $(A(t))_{t \in \mathbb{N}}$ be a sequence of row-stochastic matrices and $\delta > 0$ be such that for all $t \in \mathbb{N}$ it holds $\min^+ A(t) \geq \delta$. Let $\mathcal{G} := \mathcal{G}((A(t))_{t \in \mathbb{N}})$ be their set of important zero patterns. There exists a consensus matrix K such that*

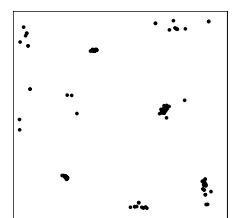
$$\lim_{t \rightarrow \infty} A(t, 0) = K$$

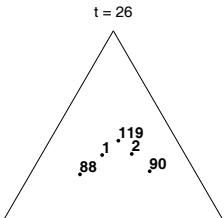
if either

1. there exists a sequence of time steps $(t_s)_{s \in \mathbb{N}}$ and $T \in \mathbb{N}$ such that

$$t_{s+1} - t_s \leq T \log(\log(s+2))$$

and such that for all $s \in \mathbb{N}$ the accumulation $A(t_{s+1}, t_s)$ is scrambling, or





3. Mathematical Analysis

2. there is $B \in \mathcal{G}$ which is scrambling and there is $T \in \mathbb{N}$ such that distance of a scrambling matrix to the next does not grow faster than $T \log(\log(t))$, or
3. all finite products of matrices in \mathcal{G} are regular, or
4. all matrices in \mathcal{G} are regular and have a positive diagonal.

Proof. 1. follows from Theorem 3.2.18 with Lemma 3.2.35 and (3.24) ensuring the diverging infinite sum. 2. follows from Theorems 3.2.23 and 3.2.18, Lemma 3.2.35 and (3.24). 3. follows from Theorems 3.2.23 and 3.2.18 with the boundedness of the length of the sub-accumulations and supermultiplicativity of the positive minimum (3.24) ensuring that there exist sub-accumulations such that for all $s \in \mathbb{N}$ it holds that $\min^+ A(t_{s+1}, t_s) \geq \delta^{T+1}$, with T as defined in Theorem 3.2.23. 4. follows analog from Theorems 3.2.30 and 3.2.18 with (3.24). \square

The length of the sub-accumulation is called length of the *intercommunication intervals*. In [8] (which goes partly back to Tsitsiklis [79]) bounded intercommunication intervals reaching connectivity and a positive diagonal were demanded to ensure convergence to consensus. We have shown here that a positive diagonal is not necessary and that a slight increasing of the length of intercommunication intervals is possible without destroying convergence to consensus. The increase can be as fast as $\log(\log(s))$. If the increase is as fast as $\log(s)$ then one can derive analog to lemma 3.2.35 that for $0 < \delta < 1$ and $T \in \mathbb{N}$ that

$$\sum_{s=1}^{\infty} \delta^{T \log(s)} < \infty \iff \delta < e^{-1} \approx 0.3679.$$

Thus, the positive minimum in each matrix has to be greater than e^{-1} to ensure divergence which can only hold if each row contains not more than two positive entries (because $e^{-1} > \frac{1}{3}$). So, a growing of the intercommunication interval with $\log(s)$ could be too fast.

We continue with results on matrices with positive diagonals.

Theorem 3.2.38 (Convergence positive diagonal matrices). Let $(A(t))_{t \in \mathbb{N}}$ be a sequence of row-stochastic matrices with positive diagonals and let $\delta > 0$ be such that for all $t \in \mathbb{N}$ it holds $\min^+ A(t) \geq \delta$. Let $(t_s)_{s \in \mathbb{N}}$ be the sequence of time steps defined by Theorem 3.2.31, $\mathcal{I}_1, \dots, \mathcal{I}_g$ be the essential classes and \mathcal{J} be the union of all inessential classes of $A(t_1, t_0)$.

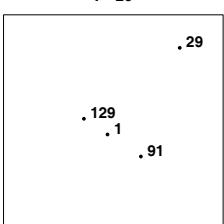
If there exists $T \in \mathbb{N}$ such that for all $s \in \mathbb{N}$ it holds

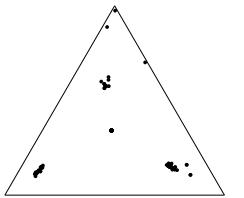
$$t_{s+1} - t_s \leq T \log(\log(s + 2)),$$

then

$$\lim_{t \rightarrow \infty} A(t, 0) = \left[\begin{array}{ccc|c} K_1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & K_g & 0 \\ \hline & \text{not converging} & & 0 \end{array} \right] A(t_0, 0)$$

where K_1, \dots, K_g are consensus matrices. (All matrices must be transformed simultaneously to the Gantmacher form.)





3.2. Matrix-based analysis

Proof. Follows from Theorem 3.2.33 with Lemma 3.2.35 and (3.24) ensuring the conditions. \square

Finally, in the type-symmetric case we can drop the assumption on the length of the sub-accumulations due to Lemma 3.2.36.

Theorem 3.2.39 (Convergence type-symmetric pos. diag. matrices). *Let $(A(t))_{t \in \mathbb{N}}$ be a sequence of row-stochastic and type-symmetric matrices with positive diagonals and let $\delta > 0$ be such that for all $t \in \mathbb{N}$ it holds $\min^+ A(t) \geq \delta$. Let $(t_s)_{s \in \mathbb{N}}$ be the sequence of time steps defined by Theorem 3.2.31, $\mathcal{I}_1, \dots, \mathcal{I}_g$ be the essential classes of $A(t_1, t_0)$.*

$$\lim_{t \rightarrow \infty} A(t, 0) = \begin{bmatrix} K_1 & & 0 \\ & \ddots & \\ 0 & & K_g \end{bmatrix} A(t_0, 0),$$

where K_1, \dots, K_g are consensus matrices. (All matrices must be transformed simultaneously to the Gantmacher form.)

Proof. Obviously, there do not exist inessential classes due to type-symmetry. Lemma 3.2.36 ensures that every accumulation $\min^+ A(t_{s+1}, t_s) \geq \delta^{n^2-n+2}$ and thus Theorem 3.2.18 is applicable which delivers convergence of the essential Gantmacher diagonal blocks. \square

3.2.8 Trying to apply the joint spectral radius

The convergence behavior of the powers of a row-stochastic matrix is determined by the left and right eigenspaces to the eigenvalue one. So, it is natural to try to extend this to inhomogeneous matrix products. It is possible to same extent which we will point out by extending an idea of Theys and Blondel [8, 78] which leads to an application of the joint spectral radius. The joint spectral radius of a set of matrices represents the maximal growth rate of products from this set. Unfortunately, some bad properties of the joint spectral radius make fruitful results difficult.

The joint spectral radius

There are two convergence results for the spectral radius of a matrix A :

$$\rho(A) = \lim_{t \rightarrow \infty} (\rho(A^t))^{\frac{1}{t}}, \quad \rho(A) = \lim_{t \rightarrow \infty} \|A^t\|^{\frac{1}{t}}.$$

The second one holds for any matrix norm.

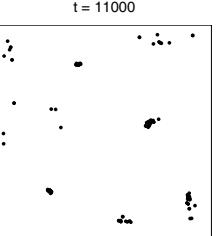
Both concepts have been generalized to a set of matrices Σ (see e.g. Daubechies and Lagarias [16]). Let

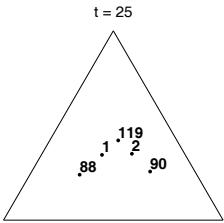
$$\rho_t(\Sigma) := \sup \left\{ \rho(A_{i_t} \dots A_{i_1})^{\frac{1}{t}} \mid A_{i_j} \in \Sigma \text{ for all } j \in \underline{t} \right\},$$

$$\hat{\rho}_t(\Sigma, \|\cdot\|) := \sup \left\{ \|A_{i_t} \dots A_{i_1}\|^{\frac{1}{t}} \mid A_{i_j} \in \Sigma \text{ for all } j \in \underline{t} \right\}.$$

Then we define the *generalized spectral radius* ρ and the *joint spectral radius* $\hat{\rho}$ as

$$\rho(\Sigma) := \limsup_{t \rightarrow \infty} (\rho_t(\Sigma))^{\frac{1}{t}}, \quad \hat{\rho}(\Sigma, \|\cdot\|) := \limsup_{t \rightarrow \infty} (\hat{\rho}_t(\Sigma, \|\cdot\|))^{\frac{1}{t}}$$





3. Mathematical Analysis

We can write $\hat{\rho}(\Sigma, \|\cdot\|) = \rho(\Sigma)$ because it does not depend on the norm, due to the equivalence of matrix norms. If Σ is entrywise bounded, then it holds

$$\rho_k(\Sigma)^{\frac{1}{k}} \leq \rho(\Sigma) = \hat{\rho}(\Sigma) \leq \hat{\rho}(\Sigma)^{\frac{1}{k}}.$$

(See [6] and [34] for the equality of ρ and $\hat{\rho}$.)

If Σ is a bounded, compact and $\rho(\Sigma) < 1$ then it follows that all infinite products with matrices from Σ converge to zero, see [34].

A joint projection on the orthogonal complement of $\text{eig}(\Sigma, 1)$

We regard a sequence of row-stochastic matrices $(A(t))_{t \in \mathbb{N}}$ and define $t_0 \in \mathbb{N}$ as the time step until every matrix in $(A(t))_{t \geq t_0}$ has a zero pattern of the set of important zero patterns $\mathcal{G}((A(t))_{t \geq t_0})$.

We know that $G := \text{inc}(\sum_{B \in \mathcal{G}} B)$ is a zero pattern which is greater than every zero pattern in $(A(t))_{t \in \mathbb{N}}$. We perform simultaneous row and column permutations jointly on every matrix in $(A(t))_{t \in \mathbb{N}}$ such that $\text{inc}(\sum_{B \in \mathcal{G}} B)$ is in Gantmacher form.

By row-stochasticity it holds for all $t \in \mathbb{N}$ that $\rho(A(t)) = 1$ and thus $\hat{\rho}(\Sigma) = 1$, too. But we can do a joint transformation of all matrices in Σ which leads us to a situation where the joint spectral radius is more interesting.

We know by row-stochasticity that for all $t \geq t_0$ there is to each essential class \mathcal{I}_i in G an eigenvalue one to the eigenvector $\mathbf{1}$ in $A_i(t)$. Thus, we have a common subspace of the 1-eigenspaces for every $t \geq t_0$ at least

$$\text{eig}(A(t), 1) \supset \text{span} \left\{ \begin{bmatrix} \mathbf{1} \\ 0 \\ \vdots \\ 0 \\ * \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{1} \\ \vdots \\ 0 \\ * \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{1} \\ * \end{bmatrix} \right\} \quad (3.26)$$

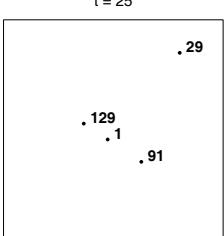
with the $*$ -parts summing up to $\mathbf{1}$.

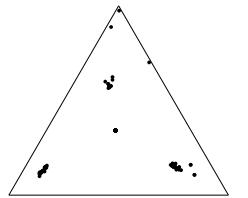
Let \mathcal{I}_k be the k -th essential class of G . According to an idea outlined in [8, 78] we can make a transformation $P_k A(t)_{\mathcal{I}_k, \mathcal{I}_k} P_k^T =: A'(t)_k$ such that the spectrum stays the same but with eigenvalue one removed. This P_k is $\#\mathcal{I}_k - 1 \times \#\mathcal{I}_k$ and its rows build an orthogonal basis of the orthogonal complement of $\text{span}\{\mathbf{1}\}$. (This can be normalized vectors with two nonzero entries which have the same absolute value and different signs.)

Thus, $A'(t)_k$ is square with dimension $\#\mathcal{I}_m - 1 \times \#\mathcal{I}_m - 1$. To see that the spectrum of $A'(t)_k$ is the spectrum of $A(t)_k$ without one consider an eigenvalue $\lambda \neq 1$ and one of its eigenvectors x . Then $y := P_k x$ is not zero and an eigenvector of $A'(t)_k$ for the eigenvalue λ . ($A(t)_k x = \lambda x \Rightarrow P_k A(t)_k P_k^T y = \lambda y \Rightarrow A'(t)_k y = \lambda y$.)

Now, we can generalize this transformation idea to our setting. Let $g \in \mathbb{N}$ be the number of essential classes and $h \in \mathbb{N}$ the number of self-communicating classes of G . We define the $(n - g) \times n$ matrix

$$P := \begin{bmatrix} P_1 & & 0 \\ & \ddots & \\ 0 & & P_g & E \end{bmatrix}$$





3.2. Matrix-based analysis

Notice that the blocks P_1, \dots, P_g are not square and thus not diagonal. Now, it holds $A'(t) :=$

$$PA(t)P^T = \begin{bmatrix} A'(t)_1 & & & & 0 \\ & \ddots & & & \\ & & A'(t)_g & & \\ & 0 & & & \\ A(t)_{g+1,1}P_1^T & \dots & A(t)_{g+1,g}P_g^T & A(t)_{g+1} & \\ \vdots & & \vdots & & \ddots \\ A(t)_{h,1}P_1^T & \dots & A(t)_{h,g}P_g^T & A(t)_{p,g+1} & \dots & A(t)_h \end{bmatrix} \quad (3.27)$$

Now we can study the joint spectral radius of $\Sigma' := \{PA(t)P^T \mid t \in \mathbb{N}\}$. Unfortunately, the matrices in Σ' are neither row-stochastic nor nonnegative anymore.

With some additional assumptions one can say a little bit more. If all $A(t)$ have positive diagonals then Theorem 3.2.31 delivers a sequence of time steps $(t_s)_{s \in \mathbb{N}}$ such that $B(s) := A(t_{s+1}, t_s)$ have the same zero pattern, positive Gantmacher diagonal blocks and Gantmacher subdiagonal blocks either positive or zero. We do the transformation 3.27 and get the matrices on the diagonal of $B'(s)$ as $B'(s)_1, \dots, B'(s)_g$.

Now, it holds for the spectral radii that $\rho(B(s)_1') < 1, \dots, \rho(B(s)_g') < 1$, due to the fact that $B(s)_1, \dots, B(s)_g$ were positive and thus had no other eigenvalues of absolute value one. Further on, the spectral radii of $B(s)_{g+1}, \dots, B(s)_p$ are less than one because for $k \in \{g+1, \dots, h\}$ it holds $\rho(B(i)_k) \leq \|B(i)_k\| < 1$. The second inequality holds due to the fact that all row sums in $B(s)_k$ are less than one. ($\|B\| := \max_i \sum_j |b_{ij}|$ in this case.)

Thus, it holds $\rho(B'(s)) < 1$ for all $t \in \mathbb{N}$. But unfortunately even this does not imply $\hat{\rho}(\Sigma') < 1$, see [9] and the following examples.

Example 3.2.40. Let $\Sigma := \{A, B\}$ with

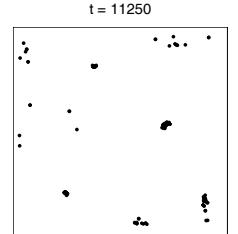
$$A = \begin{bmatrix} 0 & \frac{1}{8} \\ 2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 2 \\ \frac{1}{8} & 0 \end{bmatrix}.$$

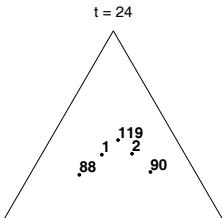
It holds $\rho(A) = \rho(B) = \frac{1}{2}$, but $\rho(AB) = 4$. These kind of examples is not restricted to irreducible matrices, which one can easily check with the same matrices but small fluctuations in the entries.

3.2.9 Inhomogeneous processes of opinion pooling

We conclude with a discussion what the results in Subsections 3.2.4 – 3.2.7 imply for processes of opinion pooling defined in Model 2.4.1. Regarding an initial opinion profile $x(0)$ the process $A(t, 0)x(0)$ converges to consensus under the conditions on the sequence of confidence matrices stated in Theorems 3.2.18 or 3.2.37. The main assumption is the scrambling property on sub-accumulations. So each two agents should find in a certain time interval (at least indirectly) a common third agent which they both give a positive weight. If these weights do not go too fast to zero during the process, this ensures convergence to consensus.

The process converges, but not necessary to consensus, under the conditions of Theorem 3.2.39. There it must hold in every time step that (i) agents have





3. Mathematical Analysis

a little bit of self-confidence (positive diagonal), (ii) confidence is mutual (type-symmetric confidence matrix) and (iii) both former properties do not fade away.

The process converges partially for agents in essential classes to internal consensus in each essential group under the conditions of Theorems 3.2.33 or 3.2.38. There it must hold in every time step that (i) agents have a little bit of self-confidence (positive diagonal) and (ii) either the positive minimum on inter-communication intervals does not fade away too fast or the positive minimum in each confidence matrix is uniformly bounded from below and the length of intercommunication intervals does not grow faster than $\log(\log(t))$.

We give an example what the inessential agents may do in cases where partial convergence as in Theorems 3.2.33 or 3.2.38 appears.

Example 3.2.41. Let $x(0) := [0 \ 0.3 \ 0.7 \ 1 \ 0 \ 1]^T$ and $A(t)$ be arbitrary row-stochastic matrices of the form

$$\begin{bmatrix} * & * \\ * & * \\ & * & * \\ & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{bmatrix}$$

with positive entries at all * such that their positive minimum is uniformly greater than a small $\delta > 0$. Then the essential groups of agents $\{1, 2\}$ and $\{3, 4\}$ will quickly find internal consensus (as it follows from Theorem 3.2.38). But the inessential agents $\{5, 6\}$ continue forever in building convex combinations of the opinions of the essential agents. Figure 3.3 gives an example.

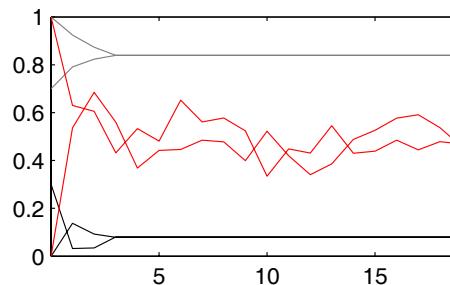


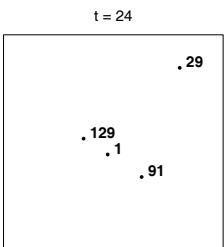
Figure 3.3: Hopping of inessential agents as described in Example 3.2.41.

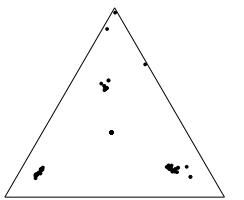
Finally, we want to point out, how growing intercommunication intervals can prevent consensus, even if there are scrambling sub-accumulations and the positive minimum of all matrices is uniformly bounded from below.

Example 3.2.42 (Moreau). Let $x(0) := [0 \ 1 \ 1]^T$ and

$$A_1 := \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_2 := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A_3 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}, \quad A_4 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$





3.3. Agent-based bounded confidence models

The sequence of confidence matrices is then defined as B_1, B_2, B_3, \dots with

$$B_s = \underbrace{A_1, \dots, A_1}_{2s \text{ times}}, \underbrace{A_2, A_3, \dots, A_3}_{2s+1 \text{ times}}, A_4.$$

Now, the backward accumulations of all matrices in B_s are scrambling. But $x(t) := A(t, 0)x(0)$ does not converge, as Figure 3.4 shows. The length of the accumulation of the matrices in B_s is $4s + 3$ and thus growing faster than $\log(\log(s))$. So this example shows that the assumption on slow growing inter-communication intervals is not neglectable.

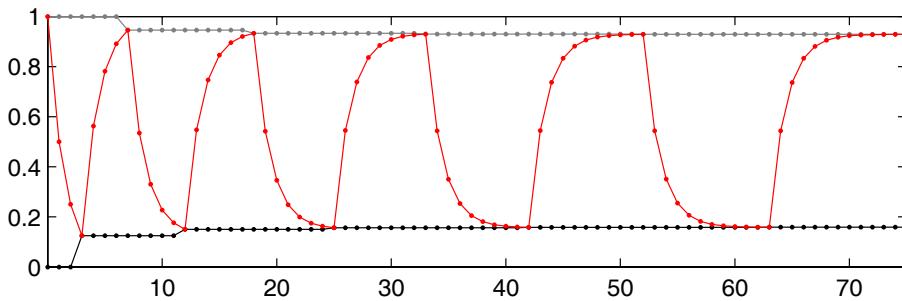


Figure 3.4: A process of opinion pooling as defined in Example 3.2.42.

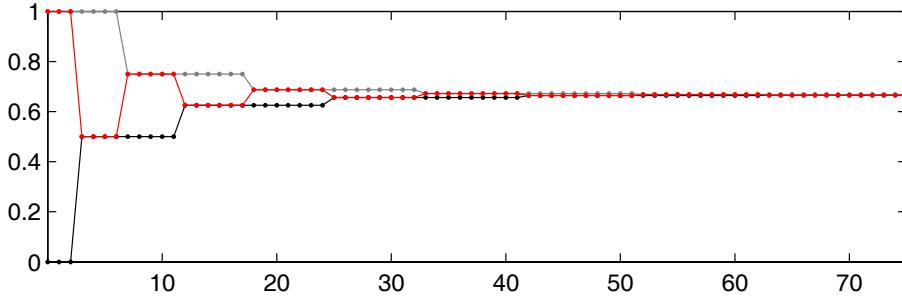


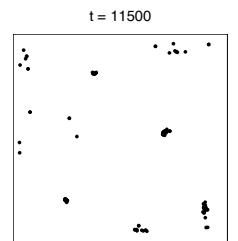
Figure 3.5: A process of opinion pooling as defined in Example 3.2.42 with A_1, A_3 replaced by the unit matrix.

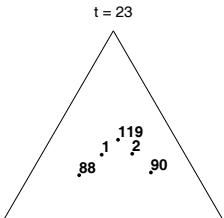
Interesting is that replacing A_1 and A_3 by the unit matrix, so reducing communication in the sequence, will lead to convergence, as figure 3.5 shows.

The example goes back to Moreau [62]. He provides a proof that there can not be convergence to consensus.

3.3 Agent-based bounded confidence models

We derive some facts about the agent-based bounded confidence models of Hegselmann-Krause and Deffuant-Weisbuch. We restrict mostly to the homogeneous case.





3. Mathematical Analysis

3.3.1 Convergence

Theorem 3.3.1. Let $\varepsilon > 0$ be a bound of confidence, p be a p -norm parameter, $n, d \in \mathbb{N}$ be the agent and the opinion dimension. Then

1. the repeated meeting process in the homogeneous Hegselmann-Krause model 2.4.3 converges in finite time.
2. the gossip process in the homogeneous Deffuant-Weisbuch model 2.4.4 with cautiousness parameter $0 < \mu \leq 0.5$ converges.

Proof. All confidence matrices which may evolve during the process fulfill the conditions of Theorem 3.2.39. Type-symmetry is given due to the homogeneous bound of confidence and it is clear from the definition of a confidence matrix A that $\min^+ A \geq \frac{1}{n}$.

Thus, there exist only essential classes of indices and the corresponding agents converge to consensus within their class.

It remains to show the finiteness of convergence time in the repeated meeting process. There is a step in time when all agents in a class are closer than ε to each other. Due to the definition of the confidence matrix these agents will find consensus in the next time step. This holds for all classes, thus the process converges in finite time. \square

For the HK model there is a proof of the same result by induction on the numbers of agents which does not use matrix techniques in Sieveking [72]. The earlier proof of Dittmer [23] works only for $d = 1$.

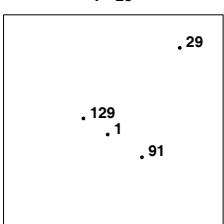
Conjecture 3.3.2. The repeated meeting and the gossip process under heterogeneous bounds of confidence converge.

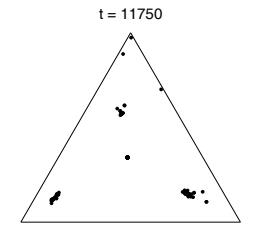
The confidence matrices in these processes fulfill the assumptions of Theorem 3.2.38 except for the bound on the intercommunication intervals $t_{s+1} - t_s$. Further on, the agents corresponding to inessential classes need not converge even under bounded intercommunication intervals. A counter-example which does not converge even in essential classes could perhaps be constructed with unbounded and fast growing intercommunication intervals (like in example 3.2.42). But this seems very implausible under the bounded confidence assumption. A counter-example which does not converge for agents corresponding to inessential agents seems implausible for the same reason. But a formal proof of convergence is still lacking.

As seen in Theorem 3.3.1 the repeated meeting process under homogeneous bound of confidence converges in finite time. This does neither hold for heterogeneous bounds nor for the gossip process.

The gossip process converges only for very specific profiles and choices of communication partners in finite time. Usually, convergence time is infinite. One can check this for three agents with opinions $0, 1, 1$ and $\varepsilon = 1, \mu = 0.5$. In this example consensus must be reached at $\frac{2}{3}$ (see following Proposition 3.3.3) but this is obviously not possible for every sequence of communication pairs.

In the repeated meeting process with heterogeneous bounds of confidence convergence time may be infinite. An example is in Figure 2.12 where the red agents which ‘sit between the chairs’ in the lower subfigure do not converge in finite time, although it looks like. The reason is that there remains an accumulation for each red agent a positive weight on its initial opinion although the





3.3. Agent-based bounded confidence models

red agents converge to a value only computed as a convex combination of the opinions of the black agents.

3.3.2 Mean conservation in homogeneous gossip processes

As we have seen in the figure 2.13 the arithmetic mean of all opinions may vary even under homogeneous bounds of confidence. Under heterogeneous bounds very drastic drifting of the mean opinion can occur for repeated meetings as well as for gossip. But the gossip process under a homogeneous bound of confidence conserves the arithmetic mean.

Proposition 3.3.3. *Let $x(0) \in S^n \subset (\mathbb{R}^d)^n$ be an opinion profile in an appropriate opinion space S . Let $\varepsilon > 0$ be a bound of confidence and $(x(t))_{t \in \mathbb{N}}$ be the gossip process. For every $t \in \mathbb{N}$ it holds that*

$$\frac{1}{n} \sum_{i=1}^n x^i(t) = \frac{1}{n} \sum_{i=1}^n x^i(0).$$

Proof. We consider the gossip step from t to $t+1$ with randomly chosen agents i, j which interact. Then $\frac{1}{n} \sum_{k=1}^n x^k(t+1) = \frac{1}{n}((1-\mu)x^i(t) + \mu x^j(t) + (1-\mu)x^j(t) + \mu x^i(t) + \sum_{k \in \underline{n} \setminus \{i,j\}} x^k(t)) = \frac{1}{n} \sum_{k=1}^n x^k(t+1)$. The proposition holds by induction over t . \square

3.3.3 Scaling and translation invariance

Let $x(0) \in (\mathbb{R}^d)^n$ be an opinion profile, $\varepsilon_1, \dots, \varepsilon_n > 0$ be bounds of confidence, $(N(t))_{t \in \mathbb{N}}$ be a communication regime and $(x(t))_{t \in \mathbb{N}}$ be the evolving general bounded confidence process.

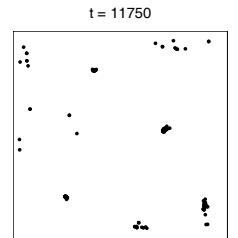
Then we define for $\lambda \in \mathbb{R}_{>0}$ and a consensus $c \in (\mathbb{R}^d)^n$ that $x'(0) := \lambda x(0) + c$ and $\varepsilon'_1 := \lambda \varepsilon_1, \dots, \varepsilon'_n := \lambda \varepsilon_n$ and compute the general bounded confidence process $(x'(t))_{t \in \mathbb{N}}$. It holds for all $t \in \mathbb{N}$ that $x'(t) := \lambda x(t) + c$. Both holds due scaling and translation invariance of weighted arithmetic means and due to the fact that areas of confidence and distances of opinions when scaled by λ leave the bounded confidence network invariant.

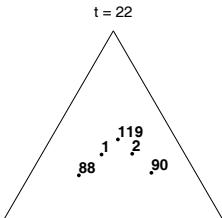
In conclusion, this shows that the relative relation between initial opinion profile and the bounds of confidence is of interest. For S a one-dimensional interval it is appropriate to study $[0, 1]$ and adjust ε .

3.3.4 The impact of the communication regime

Figure 2.10 demonstrates that the specific choice of communication partners in a gossip process can manipulate the outcome significantly. The question how big the impact of the communication regime may be is mathematically studied by Lorenz and Urbig in [61]. The paper is part of the dissertation. Section 3 of the paper deals with the questions: Given an opinion profile $x(0)$,

- how low may ε be in a homogeneous gossip process such that consensus is possible for at least one specific communication regime?
- how high may ε be such that preventing consensus is possible with at least one communication regime even if we switch back to random communication in the end?





3. Mathematical Analysis

On the first question there are fair approximations. The second question is completely answered. The ε -range where either consensus or dissensus may be enforced by the communication order is usually really huge.

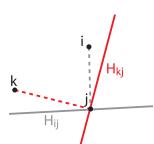
3.3.5 The set of fixed points

In this subsection we consider a homogeneous bound of confidence $\varepsilon > 0$, a p -norm parameter, a cautiousness parameter $0 < \mu \leq 0.5$, a number of agents $n \in \mathbb{N}$ and a number of opinion issues $d \in \mathbb{N}$. We call $x^* \in (\mathbb{R}^d)^n$ a *fixed point of the HK model* if $A(x^*, \varepsilon)x^* = x^*$. We call $x^* \in (\mathbb{R}^d)^n$ a *fixed point of the DW model* if for all choices $i, j \in \underline{n}$ the profile x^* does not change if agents i and j communicate. Further, $F^{\text{HK}} \subset (\mathbb{R}^d)^n$ and $F^{\text{DW}} \subset (\mathbb{R}^d)^n$ are the sets of fixed points of the corresponding models.

In the following we describe these sets and show that they are equal. Surprisingly, the proof for the HK model is not totally trivial. It relies on the finiteness of the number of agents. The following lemma will be helpful. Beforehand we define for an opinion profile $x \in (\mathbb{R}^d)^n$ and two agents $i, j \in \underline{n}$ as $H_{ij} \subset \mathbb{R}^d$ the hyperplane that is orthogonal to $x^j - x^i$ which goes through x^j . Furthermore, $H_{ij}^+ \subset \mathbb{R}^d$ is the closed half-space defined by H_{ij} which does not contain x^i .

Lemma 3.3.4. *Let $x^* \in (\mathbb{R}^d)^n$ be a fixed point of the homogeneous HK model with bound of confidence $\varepsilon > 0$. Let there be $i, j \in \underline{n}$ with $x^i \neq x^j$ such that $i \in I_\varepsilon(j, x)$ and let there be $k \in \underline{n}$ such that $x^k \notin H_{ij}^+$. Then there exists $m \in I_\varepsilon(j, x)$ different from i, j, k such that $x^m \in H_{kj}^+$.*

Proof. We abbreviate $x := x^*$. Due to x being a fixed point it must hold that



So, x^j is the barycenter of all the opinions in of agents in the confidence set $I_\varepsilon(j, x)$. By definition $i \in I_\varepsilon(j, x)$. Due to the fact that $k \notin H_{ij}^+$ and that H_{ij}^+ is closed, the angle between $x^i - x^j$ and $x^k - x^j$ is less than $\frac{\pi}{2}$ and thus $i \notin H_{kj}^+$.

There must be at least one more agent in $I_\varepsilon(j, x)$ besides i and j , because otherwise x^j is not the barycenter of x^i and x^j . If all these other agents were not in H_{jk}^+ then x^j would be an extreme point² of the convex hull of the opinions of the agents in $I_\varepsilon(j, x)$. Thus, there must be $m \in I_\varepsilon(j, x)$ such that $x^m \in H_{jk}^+$ and $m \neq j$. \square

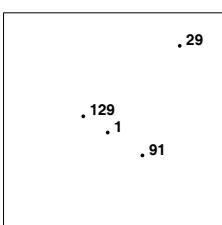
Proposition 3.3.5. *Let $\varepsilon > 0$ be a bound of confidence and $p \in \mathbb{N} \cup \infty$ a norm parameter which defines the homogeneous HK model and the DW model (with $0 < \mu \leq 0.5$) on the opinion space $(\mathbb{R}^d)^n$. It holds that*

$$F^{\text{HK}} = F^{\text{DW}} = \{x \in (\mathbb{R}^d)^n \mid \forall i, j \in \underline{n} : \|x^i - x^j\|_p > \varepsilon \text{ or } x^i = x^j\}. \quad (3.28)$$

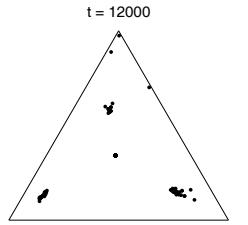
Proof. If x is in the set as described in (3.28) each two agents either reached consensus or are too far away from each other to interact. Thus, x is a fixed point in the DW and in the HK model.

Let x be not in the set as described in (3.28) then there are $i, j \in \underline{n}$ such that $\|x^i - x^j\|_p \leq \varepsilon$ and $x^i \neq x^j$.

²See Rockafellar [67] for convex analysis.



3.3. Agent-based bounded confidence models



Then x can not be a fixed point of the DW model, because if i, j are chosen as communication partners both agents will move towards each other.

It remains to show that x cannot be a fixed point of the HK model. We assume that x is a fixed point of the HK model and derive a contradiction.

Due to Lemma 3.3.4 there exists $m_0 \in I_\varepsilon(j, x)$ with $x^{m_0} \in H_{ij}^+$ with

$$\|x^i - x^{m_0}\|_2 > \|x^i - x^j\|_2$$

(we set k in the lemma equal to i). Now we apply the lemma again for $j \in I_\varepsilon(m_0, x)$. Then obviously $i \notin H_{jm_0}^+$, and thus there is $m_1 \in H_{im_0}^+$ such that $\|x^i - x^{m_1}\|_2 > \|x^i - x^{m_0}\|_2$. We can conclude like this to derive a sequence of agents m_0, m_1, m_2, \dots such that $\|x^i - x^{m_0}\|_2 < \|x^i - x^{m_1}\|_2 < \|x^i - x^{m_2}\|_2 < \dots$. This is a contradiction to the finiteness of the number of agents. \square

Figure 3.6 gives impressions how the set of fixed points F^{HK} and F^{DW} looks for the opinion space $[0, 1] \subset \mathbb{R}$ (so $d = 1$), $n = 2, 3$ and $\varepsilon = 0.3$. The figures show the complete state space $[0, 1]^n$ in contrast to Figure 2.5 where we show only opinion spaces but multidimensional.

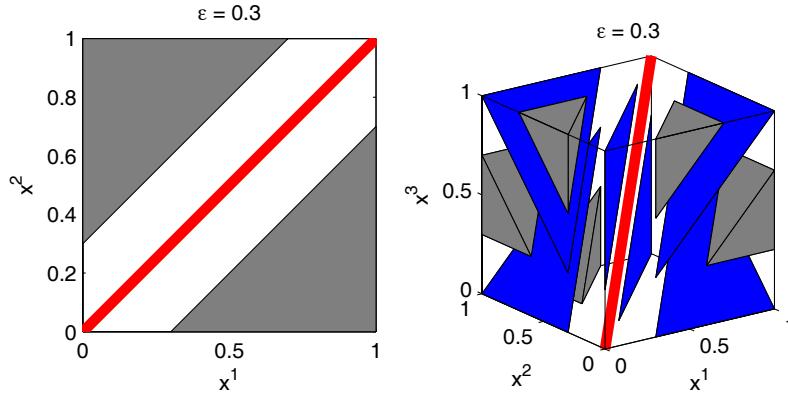
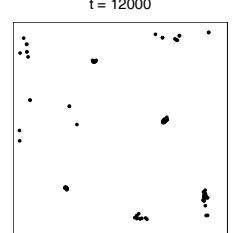
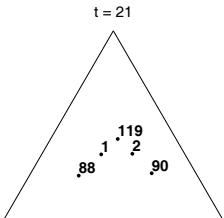


Figure 3.6: The set of fixed points F^{HK} and F^{DW} for the opinion space $[0, 1] \subset \mathbb{R}$ (so $d = 1$), $n = 2, 3$ and $\varepsilon = 0.3$. The red line represents all consensus points. The blue patches all points where two agents found consensus, while the other is far enough away. The gray regions represent all fixed points where each agent has an individual opinion. The 'invisible' space are thus all points where dynamics happen.

For higher n (but still $d = 1$) one can imagine this set like: Take the whole state space and remove successively points. First take all subspaces where numbers in two dimensions must be equal and remove the closed ε -region around this subspaces from the whole space but keep the subspaces itself. Then take from every of these subspaces all subspaces where either a third number must be equal to the former two, or two other numbers must be equal and remove their ε -region but keep the subspaces them self. Continuing like this spans a lattice of subspaces of the same kind as the lattice of partitions into subsets of the set $\{1, 2, \dots, n\}$. The number of subspaces to treat is much bigger than d it scales with the Bell numbers [1].





3. Mathematical Analysis

3.4 Density-based bounded confidence models

We will describe the set of fixed points of the interactive Markov chains

$$p(t+1) = p(t)B^{\text{CR}}(p(t), \epsilon) \quad (3.29)$$

defined in Model 2.4.8 with DW and HK transition matrix from Definitions 2.4.6 and 2.4.7. We restrict the analysis of the DW transition matrix case to $\mu = 0.5$ to keep the fight with indices at an acceptable level. We will show that the sets of fixed points are equal for both communication heuristics.

Further on, we give a Lyapunov-function for the interactive Markov chain with DW transition matrix which rules out cycles. Convergence to fixed point remains as conjectures for the DW as well as for the HK transition matrix.

For both interactive Markov chains it is useful to look at their difference equation, because it can play the role of a discrete master equation (see [31]), which displays gain and loss terms for the mass changes in one class at one time step.

Let

$$\Delta p := pB^{\text{CR}}(p, \epsilon) - p = p(B^{\text{CR}}(p, \epsilon) - E), \quad (3.30)$$

then the interactive Markov chain is a trajectory of the equation

$$p(t+1) = p(t) + \Delta p(t) = p(t) + p(t)(B^{\text{CR}}(p(t), \epsilon) - E).$$

An opinion distribution p^* is a *fixed point* of the *interactive Markov chain* (3.29) if $p^* = p^*B^{\text{CR}}(p^*, \epsilon)$. Obviously, this is equivalent to $\Delta p^* = 0$.

3.4.1 The Deffuant-Weisbuch model

We take a look at the difference Δp in detail. Simply calculating equation (3.30) with $B^{\text{CR}}(p, \epsilon) := B^{\text{DW}}(p, \epsilon)$ leads to the the following explanatory difference equation for all $k \in \underline{n}$.

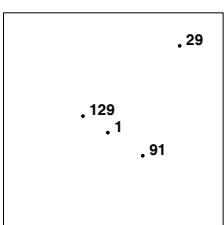
$$\Delta p_k = \underbrace{\sum_{\frac{i+j}{2}=k, 2 \leq |i-j| \leq \epsilon} p_i p_j}_{\text{fraction joining state } k} + \underbrace{\frac{1}{2} \sum_{\frac{i+j}{2}=k \pm \frac{1}{2}, 2 \leq |i-j| \leq \epsilon} p_i p_j}_{\text{fraction leaving state } k} - \underbrace{p_k \sum_{2 \leq |j-k| \leq \epsilon} p_j}_{\text{fraction leaving state } k} \quad (3.31)$$

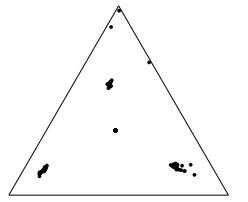
(The first two sums go over all $(i, j) \in \underline{n} \times \underline{n}$, the third over $j \in \underline{n}$, under restriction of the equations below.) This is analog to a master equation determining the fraction leaving a state and the fraction joining a state, but discrete in state and time.

Theorem 3.4.1. *An opinion distribution $p \in \Delta^{n-1} \subset \mathbb{R}^n$ is a fixed point of the interactive Markov chain (3.29) with DW transition matrix and discrete bound of confidence $\epsilon \in \mathbb{N}$ if and only if it holds for all $k \in \underline{n}$ that*

$$p_k > 0 \Rightarrow p_m = 0 \text{ for all } m \in \{k - \epsilon, \pm 1, k - 2, k + 2, \pm 1, k + \epsilon\} \cap \underline{n}. \quad (3.32)$$

Proof. For the ‘if’-part let us assume that p is a fixed point and show that (3.32) holds. If p is a fixed point it holds $\Delta p_k = 0$ for all $k \in \underline{n}$.





3.4. Density-based bounded confidence models

Let $k \in \underline{n}$ be such that $p_k > 0$. For an indirect proof let us assume that there is $m_0 \in \{k - \epsilon, \dots, k - 2, k + 2, \dots, k + \epsilon\} \cap \underline{n}$ such that $p_{m_0} > 0$ and find a contradiction.

We can conclude from $\Delta p_k = 0$ and equation (3.31) that it holds

$$\underbrace{\sum_{j=0}^{p_k} p_j}_{>0 \text{ because it contains } p_{m_0}} = \sum_{\frac{i+j}{2}=k, 2 \leq |i-j| \leq \epsilon} p_i p_j + \frac{1}{2} \sum_{\frac{i+j}{2}=k \pm \frac{1}{2}, 2 \leq |i-j| \leq \epsilon} p_i p_j$$

Thus, on the right hand side one addend $p_{m_1} p_{n_1}$ must be positive. A careful look at the summation index sets will help us to conclude further. If we assume without loss of generality $m_1 < n_1$ then we can conclude $m_1 < k$.

We can conclude from $\Delta p_{m_1}^* = 0$ and equation (3.31) that

$$\underbrace{\sum_{j=0}^{p_{m_1}^*} p_j^*}_{>0 \text{ because it contains } p_{n_1}^*} = \sum_{\frac{i+j}{2}=m_1, 2 \leq |i-j| \leq \epsilon} p_i p_j + \frac{1}{2} \sum_{\frac{i+j}{2}=m_1 \pm \frac{1}{2}, 2 \leq |i-j| \leq \epsilon} p_i p_j$$

Thus on the right hand side one addend $p_{m_2} p_{n_2}$ must be positive again and there is $m_2 < m_1 < k$.

We conclude by induction until we reach an index $m_z < 1$ for which p_{m_z} must be positive – a contradiction.

To prove the ‘only if’-part we assume that for all $k \in \underline{n}$ it holds (3.32). We have to check that $\Delta p_k = 0$ in equation (3.31) for all $k \in \underline{n}$. We see that every addend in each equation is of the form $p_i p_j$ with $2 \leq |j - l| \leq \epsilon$ and $\frac{i+j}{2} \in \{k, k \pm \frac{1}{2}\}$. From (3.32) we know that in every case either p_i or p_j are zero. \square

Thus, it is possible for at most two positive opinion classes to lie adjacent to each other. The structure of the set of fixed points is thus: all opinion classes with positive mass lie in adjacent pairs or isolated. Pairs and isolated classes must have a distance greater than ϵ to each other. In an adjacent pair of classes in a fixed point there are no further restrictions on the proportion of agents in the two classes. So, fixed points lie in certain lines in the simplex Δ^{n-1} .

Theorem 3.4.2. *For every $p(0) \in \Delta^{n-1} \subset \mathbb{R}^n$ the interactive DW Markov chain $(p(t))_{t \in \mathbb{N}_0}$ can not be periodic.*

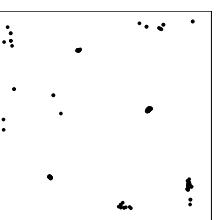
Proof. We define a Lyapunov function $L : S_n \rightarrow \mathbb{R}$ which is continuous and strictly decreasing on $(p(t))_{t \in \mathbb{N}_0}$ for every initial distribution $p(0)$ as long as we do not reach a fixed point rules. Let

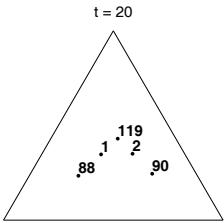
$$L(p) := \sum_{i=1}^n 2^i p_i.$$

Now we have to show that for every p which is not a fixed point it holds that $L(p) > L(pB(p, \epsilon))$.

Because of the linearity of L we can transform the inequality such that we have to show

$$0 > L(pB(p, \epsilon) - p) = L(\Delta(p)).$$





3. Mathematical Analysis

Due to (3.31) it holds

$$\begin{aligned}
 L(\Delta p) &= \sum_{k \in \underline{n}} 2^k \left(\sum_{\frac{i+j}{2} = k, 2 \leq |i-j| \leq \epsilon} p_i p_j + \frac{1}{2} \sum_{\frac{i+j}{2} = k \pm \frac{1}{2}, 2 \leq |i-j| \leq \epsilon} p_i p_j \right. \\
 &\quad \left. - p_k \sum_{2 \leq |j-k| \leq \epsilon} p_j \right) \\
 &= \sum_{2 \leq |i-j| \leq \epsilon} (2^{\lfloor \frac{i+j}{2} \rfloor} + 2^{\lceil \frac{i+j}{2} \rceil}) p_i p_j + \sum_{2 \leq |i-j| \leq \epsilon} (2^i + 2^j) p_i p_j \\
 &= \sum_{2 \leq |i-j| \leq \epsilon} (2^{\lfloor \frac{i+j}{2} \rfloor} + 2^{\lceil \frac{i+j}{2} \rceil} - 2^i - 2^j) p_i p_j
 \end{aligned}$$

It holds $(2^{\lfloor \frac{i+j}{2} \rfloor} + 2^{\lceil \frac{i+j}{2} \rceil} - 2^i - 2^j) < 0$ for all i, j with $|i-j| \geq 2$ and thus it holds $L(\Delta p_i) < 0$.

Due to the existence of the Lyapunov function it holds that $(p(t))_{t \in \mathbb{N}_0}$ can not have cycles. Because if we consider that there is a period $T \in \mathbb{N}$ such that $p(t) = p(t+T)$ then the sum $\sum_{s=t}^{t+T-1} L(\delta p(s))$ would be negative, but on the other hand it also holds

$$\sum_{s=t}^{t+T} L(\delta p(s)) = \sum_{s=t}^{t+T} L(p(s+1) - p(s)) = \sum_{s=t}^{t+T} L(p(s+1)) - L(p(s)) = 0.$$

Thus there is a contradiction to a periodic solution. \square

If one would define a Lyapunov function which is zero on every fixed point one might prove convergence to a fixed point.

Conjecture 3.4.3. *For every $p(0) \in \Delta^{n-1} \subset \mathbb{R}^n$ the interactive DW Markov chain $(p(t))_{t \in \mathbb{N}_0}$ converges to a fixed point.*

There is evidence from simulation for this conjecture.

3.4.2 The Hegselmann-Krause model

Here we show that the fixed points of the interactive Markov chain (3.29) with HK transition matrix (defined in Definition 2.4.6) are the same as for the DW transition matrix.

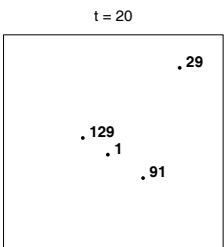
We start with a lemma on the I -barycenters.

Lemma 3.4.4. *Let $p \in \Delta^{n-1} \subset \mathbb{R}^n$ be an opinion distribution and discrete intervals $I_0 = \{i_0, \dots, j_0\} \subset \underline{n}$ and $I_1 = \{i_1, \dots, j_1\} \subset \underline{n}$. It holds*

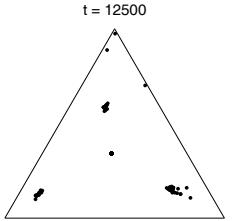
(a) $i_0 \leq i_1$ and $j_0 \leq j_1 \implies M_{I_0}^{\text{bary}}(p) \leq M_{I_1}^{\text{bary}}(p)$,

(b) if $i_0 \leq i_1$ and $j_0 \leq j_1$

$$M_{I_0}^{\text{bary}}(p) < M_{I_1}^{\text{bary}}(p) \iff \exists m \in (I_0 \cup I_1) \setminus (I_0 \cap I_1) \text{ with } p_m > 0.$$



3.4. Density-based bounded confidence models



Proof. In a first step we assume $p_{I_0} \neq 0$ and $p_{I_1} \neq 0$. Thus there is $m_0 \in I_0$ with $p_{m_0} > 0$ and one $m_1 \in I_1$ with $p_{m_1} > 0$ thus the following equation is well defined:

$$\begin{aligned} M_{I_0}^{\text{bary}}(p) &= \frac{M_{I_0}^1(p)}{M_{I_0}^0(p)} = \frac{M_{I_0}^1(p)M_{I_1}^0(p)}{M_{I_0}^0(p)M_{I_1}^1(p)} M_{I_1}^{\text{bary}}(p) \\ &= \frac{\sum_{(m_0, m_1) \in I_0 \times I_1} m_0 p_{m_0} p_{m_1}}{\sum_{(m_0, m_1) \in I_0 \times I_1} m_1 p_{m_0} p_{m_1}} M_{I_1}^{\text{bary}}(p) \end{aligned}$$

To prove (a) we have to show that the fraction in equation (3.33) is less or equal than one.

We compare the summands in the numerator and the denominator. If $m_0, m_1 \in I_0 \cap I_1$ then the summands $m_0 p_{m_0} p_{m_1}$ and $m_1 p_{m_0} p_{m_1}$ appear in both. In all other combination of indices it holds either $(m_0, m_1) \in (I_0 \setminus I_1) \times I_1$ or $(m_0, m_1) \in I_0 \times (I_1 \setminus I_0)$. Due to $i_0 \leq i_1$ and $j_0 \leq j_1$ it holds $m_0 < m_1$ and thus the numerator is less or equal to the denominator and the fraction is less or equal to one.

To prove (b) we have to show that fraction in (3.33) is strictly less than one. This holds if there is a pair $(m_0, m_1) \in (I_0 \setminus I_1) \times I_1$ or $(m_0, m_1) \in I_0 \times (I_1 \setminus I_0)$ for which $p_{m_0} > 0$ and $p_{m_1} > 0$. This is obviously the case due to the claim in (b) and the assumption $p_{I_0} \neq 0$ and $p_{I_1} \neq 0$.

At least we have to check the case, where $p_{I_0} = 0$ or $p_{I_1} = 0$. The same steps as in equation (3.33) lead with definition 2.4.5 to the equations

$$\begin{aligned} p_{I_0} \neq 0, p_{I_1} = 0 \Rightarrow M_{I_0}^{\text{bc}}(p) &= \frac{2 \sum_{m \in I_0 \setminus I_1} mp_m}{(j_1 + i_1) \sum_{m \in I_0 \setminus I_1} p_m} M_{I_1}^{\text{bc}}(p) \\ p_{I_0} = 0, p_{I_1} \neq 0 \Rightarrow M_{I_0}^{\text{bc}}(p) &= \frac{(j_0 + i_0) \sum_{m \in I_1 \setminus I_0} p_m}{2 \sum_{m \in I_1 \setminus I_0} mp_m} M_{I_1}^{\text{bc}}(p) \\ p_{I_0} = 0, p_{I_1} = 0 \Rightarrow M_{I_0}^{\text{bc}}(p) &= \frac{j_0 + i_0}{j_1 + i_1} M_{I_1}^{\text{bc}}(p) \end{aligned}$$

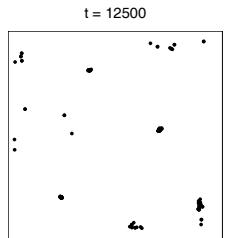
(We can chose the summation index sets $I_0 \setminus I_1$ instead of I_0 in the upper equation, because all summands with indices out of $I_0 \cap I_1$ are obviously zero. Analog for the middle equation.) For all three equations we can conclude like above to get (a) and (b). \square

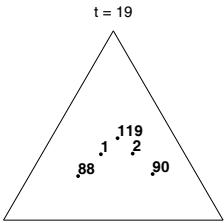
For an opinion distribution $p \in \Delta^{n-1} \subset \mathbb{R}^n$ and a discrete bound of confidence $\epsilon \in \underline{n}$ we recall the abbreviation $M_i := M_{\{i-\epsilon, \dots, i+\epsilon\}}^{\text{bary}}(p)$. Due to Lemma 3.4.4 it holds

$$M_1 \leq M_2 \leq \dots \leq M_n. \quad (3.33)$$

Analog to the former subsection we reformulate (3.30), which leads to the following explanatory difference equation for all $k \in \underline{n}$ (again in analogy to a master equation).

$$\Delta p_k = \underbrace{\sum_{j \in I_k^{\lceil \cdot \rceil}} (M_j - \lfloor M_j \rfloor) p_j + \sum_{j \in I_k} p_j + \sum_{j \in I_k^{\lfloor \cdot \rfloor}} (\lceil M_j \rceil - M_j) p_j}_{\text{fraction joining } k} - \underbrace{p_k}_{\text{fraction leaving } k} \quad (3.34)$$





3. Mathematical Analysis

with

$$I_k^{[\cdot]} := \{j \in \underline{n} \mid M_j \neq k = \lceil M_j \rceil \text{ and } p_k > 0\},$$

$$I_k^* := \{j \in \underline{n} \mid k = M_j \text{ and } p_k > 0\} \text{ and}$$

$$I_k^{[\cdot,\cdot]} := \{j \in \underline{n} \mid M_j \neq k = \lfloor M_j \rfloor \text{ and } p_j k > 0\}.$$

It is easy to see with (3.33) that the sets $I_i^{[\cdot]}, I_i^*$ and $I_i^{[\cdot,\cdot]}$ are all discrete intervals, that they are pairwise disjoint and that their union

$$I_i := I_i^{[\cdot]} \cup I_i^* \cup I_i^{[\cdot,\cdot]}$$

is a discrete interval, too. We know also that the coefficients $(M_j - \lfloor M_j \rfloor)$ and $(\lceil M_j \rceil - M_j)$ in (3.34) are always positive and strictly less than one by definition.

The following proposition shows that an opinion class with positive mass has a local barycenter which is less than one class away and that the adjacent class has positive mass too and a local barycenter between the two classes.

Proposition 3.4.5. *Let $p \in \Delta^{n-1} \subset \mathbb{R}^n$ be a fixed point of the interactive HK Markov chain (3.29) with HK transition matrix and let $p_i > 0$ then it holds*

either $M_i = i$ and $I_i = \{i\}$

or $i < M_i \leq M_{i+1} < i + 1, p_{i+1} > 0$ and $I_i = \{i, i + 1\} = I_{i+1}$,

or $i - 1 < M_{i-1} \leq M_i < i, p_{i-1} > 0$ and $I_i = \{i, i - 1\} = I_{i-1}$

Proof. We define $p = [p_1 \dots p_n]$.

In a first step we will show that $i - 1 < M_i < i + 1$. Let us assume for an indirect proof that $M_i \geq i + 1$.

The fact that p is a fixed point implies $\Delta p = 0$ and thus we can derive from equation (3.34) that

$$p_i = \sum_{j \in I_i^{[\cdot]}} (M_j - \lfloor M_j \rfloor) p_j + \sum_{j \in I_i^*} p_j + \sum_{j \in I_i^{[\cdot,\cdot]}} (\lceil M_j \rceil - M_j) p_j \quad (3.35)$$

Due to $M_i \geq i + 1$ it holds that $i \notin I_i$ (the union of all index sets) and due to Lemma 3.4.4 it holds for $j \in I_i$ that $j \leq i - 1$. Let $i_1 := \max I_i$. Thus it is clear that $i_1 < i, p_{i_1} > 0$ and $i_1 < M_{i_1}$.

We conclude further with equation (3.34) that

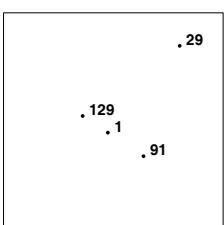
$$p_{i_1} = \sum_{j \in I_{i_1}^{[\cdot]}} (M_j - \lfloor M_j \rfloor) p_j + \sum_{j \in I_{i_1}^*} p_j + \sum_{j \in I_{i_1}^{[\cdot,\cdot]}} (\lceil M_j \rceil - M_j) p_j \quad (3.36)$$

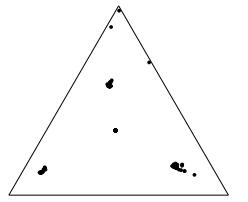
It may $i_1 \in I_{i_1}^{[\cdot]}$ but it holds $\max I_{i_1} \leq i_1$ and due to $(M_j - \lfloor M_j \rfloor) < 1$ it holds that there must exist $i_2 := \max I_{i_1} \setminus \{i_1\}$ with $p_{i_2} > 0$ and $i_2 < M_{i_2}$.

We derive by induction further on the existence of a decreasing chain of indices $i > i_1 > i_2 > \dots$ with $p_i > 0, p_{i_1} > 0, p_{i_2} > 0, \dots$. Thus there must be $z < 1$ with $p_z > 0$, a contradiction, thus $M_i < i + 1$.

If we assume $M_i \leq i - 1$ we can derive analog that there must be $z > n$ with $p_z > 0$. Thus we know $i - 1 < M_i < i + 1$.

In the second step we show $M_i > i \Rightarrow M_{i+1} < i + 1, p_{i+1} > 0$. It is clear by Lemma 3.4.4 that $M_{i+1} \geq M_i$, lets assume $M_{i+1} \geq i + 1$. Then we find (looking





 3.4. Density-based bounded confidence models

at equation (3.35)) that $i \in I_i^{\lfloor \cdot \rfloor}$ and $i+1 \notin I_i$ thus we can conclude in the same way as after equation (3.36) that there exist $z < 1$ with $p_z > 0$. Thus it follows by this contradiction that $M_{i+1} < i+1$. Analog we derive $M_i < i \Rightarrow M_{i-1} > i-1$

In the third step we show that $M_i > i$ implies $I_i = \{i, i+1\} = I_{i+1}$ and $p_{i+1} > 0$. From equation $\Delta p = 0$ and equation (3.34) we can derive the two equations

$$\begin{aligned} p_i &= (\lceil M_i \rceil - M_i)p_i + (\lceil M_{i+1} \rceil - M_{i+1})p_{i+1} + \sum_{j \in I_i \setminus \{i, i+1\}} \text{positive terms} \\ p_{i+1} &= (M_i - \lfloor M_i \rfloor)p_i + (M_{i+1} - \lfloor M_{i+1} \rfloor)p_{i+1} + \sum_{j \in I_{i+1} \setminus \{i, i+1\}} \text{positive terms} \end{aligned}$$

If we add both equations we get by calculation

$$0 = \sum_{j \in I_i \setminus \{i, i+1\}} \text{positive terms} = \sum_{j \in I_{i+1} \setminus \{i, i+1\}} \text{positive terms}$$

and thus $I_i \setminus \{i, i+1\}$ and $I_{i+1} \setminus \{i, i+1\}$ must be empty. And due $i+1 \in I_i$ it holds $p_{i+1} > 0$.

Analog, we prove that $M_i < i$ implies $I_i = \{i-1, i\} = I_{i-1}$ and $p_{i-1} > 0$. \square

So, for the fixed point p and $p_i > 0$ we know that either $I_i = \{i\}$ or $I_i = \{i, i+1\}$ with $p_{i+1} > 0$ or $I_i = \{i-1, i\}$ with $p_{i-1} > 0$. We define two new discrete intervals

$$I_i^{-\epsilon} := \{(\min I_i) - \epsilon, \dots, (\min I_i) - 1\},$$

$$I_i^{+\epsilon} := \{(\max I_i) + 1, \dots, (\max I_i) + \epsilon\}.$$

The discrete interval $I_i^{-\epsilon} \cup I_i \cup I_i^{+\epsilon}$ is the interval which contains all the classes where the imaginary agents in the classes of I_i interact with. The next proposition shows that the class(es) in $I_i^{-\epsilon}$ and $I_i^{+\epsilon}$ can only both contain mass or both contain no mass.

Proposition 3.4.6. *Let $p \in \Delta^{n-1} \subset \mathbb{R}^n$ be a fixed point of the interactive HK Markov chain (3.29) with HK transition matrix and let $p_i > 0$ then it holds*

$$p_{I_i^{-\epsilon}} = 0 \Leftrightarrow p_{I_i^{+\epsilon}} = 0.$$

Proof. First we consider $I_i = \{i, i+1\}$. Thus, due to Proposition 3.4.5 it holds $i < M_i \leq M_{i+1} < i+1$. It holds $\Delta p = 0$ because p is a fixed point. From (3.34) we can thus derive

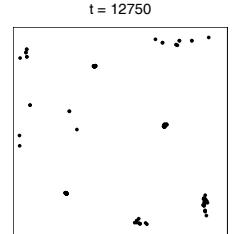
$$p_i = (\lceil M_i \rceil - M_i)p_i + (\lceil M_{i+1} \rceil - M_{i+1})p_{i+1}$$

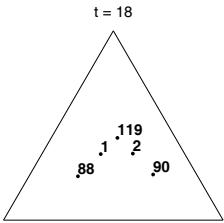
With $\lceil M_i \rceil = \lceil M_{i+1} \rceil = i+1$ it follows

$$p_i = ((i+1) - M_i)p_i + ((i+1) - M_{i+1})p_{i+1}.$$

This can be transformed to

$$M_i p_i + M_{i+1} p_{i+1} = i p_i + (i+1) p_{i+1} \tag{3.37}$$





3. Mathematical Analysis

Now, we assume for an indirect proof that $p_{I_i^{-\epsilon}} = 0$ and $p_{I_i^{+\epsilon}} \neq 0$ and derive a contradiction. Due to this assumption it holds $M_i = M_{\{i, \pm 1, i+\epsilon\}}^{\text{bary}}$ and $M_{i+1} = M_{\{i, \pm 1, i+1+\epsilon\}}^{\text{bary}}$. Then it follows from lemma 3.4.4 that $M_{\{i, i+1\}}^{\text{bary}} < M_{i+1}$ and $M_{\{i, i+1\}}^{\text{bary}} \leq M_i$. Now, we conclude from (3.37) that

$$M_{\{i, i+1\}}^{\text{bary}} p_i + M_{\{i, i+1\}}^{\text{bary}} p_{i+1} < i p_i + (i+1) p_{i+1}.$$

Both sides divided by the positive term $(p_i + p_{i+1})$ delivers

$$M_{\{i, i+1\}}^{\text{bary}} < \frac{i p_i + (i+1) p_{i+1}}{p_i + p_{i+1}} = M_{\{i, i+1\}}^{\text{bary}}.$$

A similar contradiction can be derived for the assumption $p_{I_i^{-\epsilon}} \neq 0$ and $p_{I_i^{+\epsilon}} = 0$. This proves $p_{I_i^{-\epsilon}} = 0 \Leftrightarrow p_{I_i^{+\epsilon}} = 0$.

For $I_i = \{i-1, i\}$ arguments are the same after renumbering $i \rightarrow i-1$.

For $I_i = \{i\}$ it holds $M_i = i$. Again, we assume for an indirect proof that $p_{I_i^{-\epsilon}} = 0$ and $p_{I_i^{+\epsilon}} > 0$ and derive a contradiction:

$$M_i = M_{\{i, \pm 1, i+\epsilon\}}^{\text{bary}} < M_{\{i\}}^{\text{bary}} = i = M_i.$$

□

Now, we show that the set of fixed points of the interactive Markov chain with HK transition matrix is the same as for the DW transition matrix.

Theorem 3.4.7. *An opinion distribution $p \in \Delta^{n-1} \subset \mathbb{R}^n$ is a fixed point of the interactive Markov chain (3.29) with HK transition matrix and discrete bound of confidence $\epsilon \in \mathbb{N}$ if and only if it holds for all $k \in \underline{n}$ that*

$$p_k > 0 \Rightarrow p_m = 0 \text{ for all } m \in \{k - \epsilon, \dots, k - 2, k + 2, \dots, k + \epsilon\} \cap \underline{n}. \quad (3.38)$$

Proof. For the ‘if’-part let us assume that p is a fixed point and show that (3.38) holds. For an indirect proof we assume that there are $i, j \in \underline{n}$ such that $i < j$, $2 \leq |i - j| \leq \epsilon$ and $p_i, p_j > 0$ and find a contradiction.

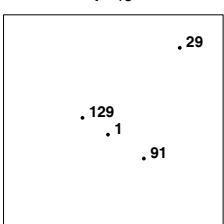
From Proposition 3.4.5 we know that I_i and I_j are disjoint. From Proposition 3.4.6 we know that there must exist $m_0 \in \mathbb{N}$ such that $m_0 < i$, $|i - m_0| \leq \epsilon$, $p_{m_0} > 0$ and I_{m_0} and I_i are disjoint. Comparing m_0 and i we know with the same arguments that there must exist $m_1 \in \mathbb{N}$ with $m_1 < m_0$, $|m_0 - m_1| \leq \epsilon$, $p_{m_1} > 0$ and I_{m_1} and I_{m_0} are disjoint. By induction we can construct a sequence of natural numbers $m_0 > m_1 > m_2 > \dots$ with $p_{m_0}, p_{m_1}, p_{m_2}, \dots > 0$. Thus there must exist $z \in \mathbb{N}$ such that $m_z < 1$ and $p_{m_z} > 0$, which is a contradiction.

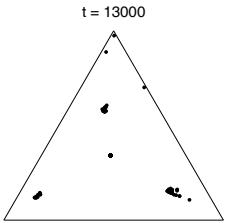
To prove the ‘only if’-part we assume that for all $k \in \underline{n}$ it holds (3.38). We have to check that $\Delta p_i = 0$ in (3.34) for all $i \in \underline{n}$. We see that every addend in each equation is of the form $p_i p_j$ with $2 \leq |j - l| \leq \epsilon$ and $\frac{i+j}{2} \in \{k, k \pm \frac{1}{2}\}$. From (3.38) we know that in every case either p_i or p_j are zero. □

The convergence to a fixed point remains as a conjecture.

Conjecture 3.4.8. *For every $p(0) \in \Delta^{n-1} \subset \mathbb{R}^n$ the interactive HK Markov chain $(p(t))_{t \in \mathbb{N}_0}$ converges to a fixed point. Convergence occurs in finite time.*

There is strong evidence from simulation for the conjecture, but a prove seems difficult, because convergence time can be arbitrary long due to the existence of metastable states like in figure 2.20.





3.5 References and relations

The work of Section 3.1 is inspired by preprint of Krause [48] and the work of Moreau [62]. The extension to a family of averaging maps is a new result. It is joint work with Dirk A. Lorenz. Moreau based his work on the new framework of set-valued Lyapunov-functions which are able to describe stability and attractivity for sets of fixed points where fixed points are not isolated, but in contrast to other set stability notions it is demanded that a process converges to one of the fixed points (and does not e.g. drifts a closer and closer along the stability set). Basically, one can see the proof the main theorem also as a kind of set-valued Lyapunov technique. The convex hull (or the cube) is shrinking along the trajectory and the proper property ensures that we must reach a consensus.

These stabilization and attraction results are especially of interest in the analysis of consensus protocols of autonomous agents – a field which has been studied by Tsitsiklis [79] in the context of distributed computing in the eighties and has recently got discussed a lot while searching for conditions about successful collective motion of unmanned aerial vehicles [43, 64]. These studies also have to deal with the differences of consensus and swarm dynamics mentioned in Subsection 2.4.7: A successful algorithm which leads a group to consensus relies on the existence of borders in the opinion space, like the convex hull here. The opinions of swarming particle are angles. These problems were circumvented by [43] by making the unrealistic assumption that one can average angles in $[0, 2\pi]$ in the same way as opinions. This leads to the problem that angles ε and $2\pi - \varepsilon$ are averaged to π although they are very close together for small ε and should in this case be averaged to zero. Therefore, in [40, 62] the application of consensus algorithms by repeated averaging to the swarming case are restricted to swarms where the initial headings of all agents do not spread more than π .

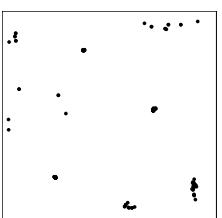
The results in Subsections 3.2.1 to 3.2.7 resemble known results with references to Gantmacher [29, 30], Seneta [69, 70], Hartfiel [33, 34], Wolfowitz [89] and Krause [47] with the aim of a characterization of infinite backward products of row-stochastic matrices. Thereby, we give in Theorem 3.2.5 a description about the convergence of the powers of an arbitrary row-stochastic matrix. The convergence results in Subsections 3.2.5 to 3.2.7 are mostly new. Collected results or strong relations to former results are mentioned there.

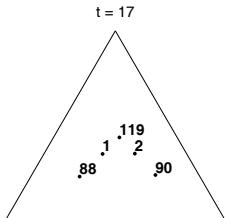
A tight and quick characterization of all sequences of row-stochastic matrices which backward accumulations converge seems difficult. This is especially underpinned by the examples in Subsections 3.2.6. Nevertheless, there are (non generalizable) results which characterize conditions for converging sequences of row-stochastic 2×2 matrices shown by Sieveking [73].

Especially Theorem 3.2.38 about the allowed growing of intercommunication intervals is a generalization of the formerly know theorems about bounded intercommunication intervals [40, 52, 62].

Another fruitful line of research to proceed is the determination of necessary conditions for convergence to a consensus matrix. The result in this thesis – Theorem 3.2.26 – is just a starting point in this direction.

In Subsection 3.2.8 we tried to apply results on the joint spectral radius to the problem via a joint transformation which goes back to Blondel and Theys [8, 78]. This transformation is familiar with the subspace contraction coefficients outline in 3.2.5 which goes back to Hartfiel [34]. But we got stuck (like probably many others) with the problem that it is not possible to decide on the



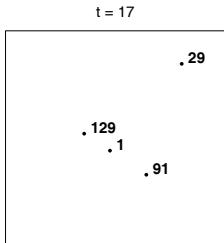


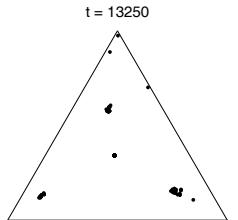
3. Mathematical Analysis

boundedness of all products of a pair of matrices [9]. Nevertheless, this line of research is worth to deepen especially in generalizing the results to other than row-stochastic matrices. The theory about sets of matrices which are LCP (have the left convergence property) founded in Daubechies and Lagarias [16] (with the name RCP) should be of help there. Hartfiel [34] extracted from their work the basic property that matrices which appear infinitely often in a converging inhomogeneous matrix product must have the same eigenspace for the eigenvalue one. This gives rise to the question if row-stochastic matrices can serve as a kind of prototype for converging inhomogeneous matrix products. (Neumann and Schneider [63] can be another source of inspiration.)

The work in Section 3.3 is to some extent already in the Diploma thesis [52] but presented here in English and put in the broader context. The description of the set of fixed points is new.

The work in Section 3.4 is new. The idea for the formulation of density-based dynamics was inspired by Ben-Naim et al. [5]. The setup as interactive Markov chains (as well as the term) is inspired by Conlisk [14]. The derivation of the difference equation is in pretty good analogy to the method of determining a master equation of a particle system in statistical physics [31]. But it is rarely done in discrete time and space although this has obvious advantages for simulation. It remains to show that the chains always converge to a fixed point. Perhaps this is possible with a Lyapunov method. Another task is to describe the class of all processes that converge to such clustered fixed point.





Chapter 4

Simulation

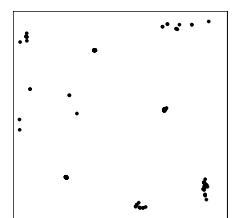
Simulation can help to study dynamical behavior when mathematical analysis got stuck. In Chapter 2 we saw already a couple of examples for the agent-based as well as for the density-based bounded confidence models. They all lead to a characteristic pattern of clusters. The aim here is to study this more systematic. Our basic assumption is that agents are initially uniformly distributed in the opinion space. Section 4.1 summarizes the knowledge about stabilization and defines clustering which is the basis for further simulation results. The density-based models are employed for this because one simulation stands for a whole collection of agent-based runs. Then we study in Section 4.2 the bifurcation of the clustering pattern with respect to the bound of confidence and compare the results to the results of Ben-Naim and others [5] for the DW model obtained with a differential equation. In Section 4.3 we analyze populations where agents may have two different bounds of confidence $\varepsilon_1, \varepsilon_2$ and derive extended phase diagrams to see how the consensus transition is controlled by $(\varepsilon_1, \varepsilon_2)$. Section 4.4 gives a guideline for reading the accompanying papers, which are also part of the dissertation.

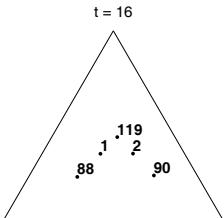
4.1 Stabilization and clustering in time evolution

In the following we will take it as given that every repeated meeting process and every gossip process converges – for the agent-based formulation as well as for the density-based formulation – for a homogeneous bound of confidence as well as for heterogeneous bounds of confidence.

There is strong evidence from simulation that this must be true. Additionally, there are proofs for the repeated meeting process and the gossip process for a homogeneous bound of confidence (Theorem 3.3.1). For the density-based DW model with a homogeneous bound of confidence and cautiousness parameter $\mu = 0.5$ we know at least that it can not have cycles by Theorem 3.4.2.

So, we assume that every agent-based process of continuous opinion dynamics in the framework of the HK model 2.4.3 and the DW model 2.4.4 converges to a *limit opinion profile* $x^* \in S^n \subset \mathbb{R}^n$. It depends on the initial profile $x(0)$, the bounds of confidence $\varepsilon_1, \dots, \varepsilon_n$ and for the DW model also on the specific choice of the communication regime and the cautiousness parameter μ .





4. Simulation

Analog, we assume that for every interactive Markov chain in the framework of model 2.4.8 there is a *limit opinion distribution* $p^* \in \Delta^{n-1}$ where the process converges to. It depends on the initial distributions $p^1(0), \dots, p^m(0)$, the bounds of confidence $\epsilon_1, \dots, \epsilon_m$ and the communication regime {HK, DW}. For DW-communication it also depends on the cautiousness parameter μ . If we make some assumptions on the initial distributions or the bounds of confidence the number of opinion classes n or the number of different bounds of confidence can also be of interest.

Further on, we assume that limit profiles and limit distributions are usually fixed point (besides rare examples e.g. for the heterogeneous agent-based HK model). The fixed points are e.g. characterized in Proposition 3.3.5 for the homogeneous agent-based models and by Theorems 3.4.1 and 3.4.7 for the interactive Markov chains under homogeneous bounds of confidence. A limit profile of an agent-based model is characterized by the clustering of agents in groups which have internal consensus. The limit distribution of a density-based model is characterized by the clustering of the population mass in few opinion classes.

Crucial to the simulation analysis is thus the definition of a *cluster*.

Definition 4.1.1 (Cluster in an opinion profile). Consider an opinion space $S \subset \mathbb{R}^d$ convex, agents $\{1, \dots, n\}$, an opinion profile $x \in S^n$, bounds of confidence $\epsilon_1, \dots, \epsilon_n \in \mathbb{R}_{>0}$ and a norm $\|\cdot\|$.

A set $\mathcal{I} \subset \underline{n}$ is a *consensus cluster* if $x^\mathcal{I}$ is a consensus and for all $i \in \mathcal{I}, j \in \underline{n} \setminus \mathcal{I}$ it holds $x^i \neq x^j$.

Let $N^{\text{BC}}(x, \epsilon_1, \dots, \epsilon_n)$ be the bounded confidence network matrix.

The set \mathcal{I} is a *network cluster* if $N_{[\mathcal{I}, \mathcal{I}]}^{\text{BC}}(x, \epsilon_1, \dots, \epsilon_n) = 1$ and for all $\mathcal{J} \supset \mathcal{I}$ (with $\mathcal{I} \neq \mathcal{J}$) it holds $N_{[\mathcal{J}, \mathcal{J}]}^{\text{BC}}(x, \epsilon_1, \dots, \epsilon_n)$ is not a matrix of ones.

A consensus or network cluster \mathcal{I} is *essential* if $N_{[\mathcal{I}, \underline{n} \setminus \mathcal{I}]}^{\text{BC}} = 0$. A cluster is *inessential* if it is not essential.

An essential consensus or network cluster \mathcal{I} is *isolated* if $N_{[\underline{n} \setminus \mathcal{I}, \mathcal{I}]}^{\text{BC}} = 0$.

The *mass* of a network or consensus cluster \mathcal{I} is $\#\mathcal{I}/n$.

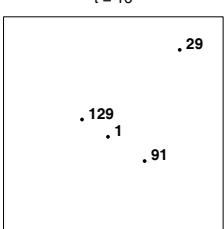
The *location* of a consensus cluster \mathcal{I} is the opinion of its agents. The location of a network cluster is the arithmetic mean of the opinions of the agents in the cluster $\frac{1}{\#\mathcal{I}} \sum_{i \in \mathcal{I}} x^i$.

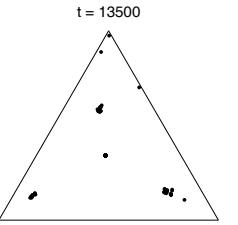
A consensus cluster is always a subset (or equal) to a network cluster. For an agent in an essential network cluster it holds that all agents outside the cluster are out of his actual area of confidence while all agents in the cluster are within his actual area of confidence. For all agents in an isolated network cluster it holds further on, that all agents outside the cluster have none of them in their actual area of confidence.

It is important to notice that a cluster need not stay a cluster for all the time. This holds e.g. for $d \geq 2$ or under heterogeneous bounds of confidence.

The definition of a network cluster is important because convergence time can be infinite. The homogeneous HK model converges in finite time thus no stopping criteria is needed for running simulation. The only danger (besides floating point errors) is to run into a metastable state with very long convergence time.

Things get difficult in the homogeneous DW model were convergence time is infinite. But the computation of the exact limit profile is possible after a finite number of time steps. For this one repeatedly checks the existence of





4.2. Bifurcation of clusters in the evolution of the bound of confidence

network clusters. If there are only isolated network clusters and there are no agents left who are not member of a network cluster, then we know that each network cluster will remain an isolated network cluster for ever. It is ensured by Theorem 3.2.39 that we will reach such a situation. For such a network cluster which stays a network cluster forever we can compute the limit opinion of all its agents as the arithmetic mean of all opinions because of Proposition 3.3.3. The only problem which may evolve is that it may take very long until we reach this situation. An example is a situation with three agents which are isolated from every body else but with only one agent connected to both others. So this is not a network cluster and it may split if the central agent communicates with one of the others. But the probability that the central agent and one of the two others is chosen in random pairwise communication is $\frac{2}{n^2}$. For thousand agents this is $\frac{1}{500.000}$.¹

It follows a cluster definition for density-based models.

Definition 4.1.2 (Cluster in an opinion distribution). Let $p \in \Delta^{n-1}$ be an opinion distribution. A discrete interval $\mathcal{I} = \{i, \pm 1, j\} \subset \underline{n}$ is a *cluster* if $p_{\mathcal{I}} > 0$ and $p_{i-1}, p_{j+1} = 0$. For a discrete bound of confidence $\epsilon \in \mathbb{N}$ a the cluster is isolated if $p_{i-\epsilon, \pm 1, i-1}, p_{j+1, \pm 1, j+\epsilon} = 0$. The cluster is *contracted* if $j - i \leq \epsilon$.

Let $\delta > 0$ be a (small) *level of precision*. Then \mathcal{I} is *cluster with precision δ* if \mathcal{I} is a cluster in the distribution equal to $\bar{p}/\|\bar{p}\|_1$ with \bar{p} be equal to p but with all entries of p which are below δ set to zero.

The *mass* of a cluster \mathcal{I} is its \mathcal{I} -mass $M_{\mathcal{I}}^0(p)$.

The *location* of a cluster \mathcal{I} is its \mathcal{I} -barycenter $M_{\mathcal{I}}^{\text{bary}}(p)$.

4.2 Bifurcation of clusters in the evolution of the bound of confidence

In a first step we analyze continuous opinion dynamics under bounded confidence where all agents have a homogeneous bound of confidence $\varepsilon > 0$.

Let us assume that agent's opinions are initially uniformly distributed in a one-dimensional opinion space. The opinion space should be bounded and convex, thus it is an interval. Due to scaling and translation invariance (see Subsection 3.3.3) in the agent-based model it is enough to study the interval $[0, 1]$.

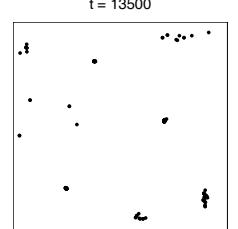
Under the assumption of uniformly distributed initial opinions there is a typical clustering in the limit profile or the limit distribution for each value of ε . Of interest is

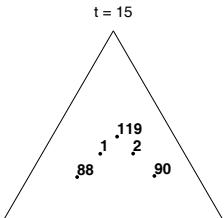
- the number,
- the mass, and
- the location in the opinion space

of clusters in the limit profile or limit distribution.

In the following we study bifurcations of the clustering patterns in the limit profiles and limit distributions with respect to variations of the bound of confidence ε .

¹This is also the reason for the surprising fact that convergence to the limit opinion within a small cluster lasts longer than convergence within a large cluster. An example is in Figure 2.10 (bottom left hand side).





4. Simulation

4.2.1 Aggregated data of multiple agent-based simulation runs

The inventors of the two models, Hegselmann and Krause respectively Deffuant, Weisbuch and others, have analyzed their models via agent-based simulation: They fixed the number of agents $n \in \mathbb{N}$ (not too low), defined a set of ε -steps and picked a sufficiently large number of random initial profiles ($s \in \mathbb{N}$ called the sample size) and then computed the limit profiles for each value of ε . The sets of limit profiles for the specific values of ε can then be analyzed and visualized.

Hegselmann and Krause [37] fixed the parameters $n = 625, s = 50$. Then they divided the opinion space $[0, 1]$ into the subintervals $[\frac{i-1}{n}, \frac{i}{n}]$ for $i = 1, \dots, 100$ and counted for each $\varepsilon = 0, +0.01, 0.4$ and each subinterval the number of agents which end with an opinion within this subinterval over all limit profiles in the sample. This gives a figure which we call *added limit profiles*. We reproduced this kind of visualization with $i = 1, \dots, 25, n = 200$ and $s = 250$ in Figure 4.1 (upper row). So we see a histogram of agents opinions after stabilization.

In this kind of visualization one can see that there are attractive regions for the limit opinions which change with ε . Especially one sees the consensus mountain for large ε and two hills which stand for a high probability that there is a cluster close to the extremes. How close to the extremes is determined by the bound of confidence. The larger ε is the more does each hill moves towards the center and grows. Before they join to the consensus mountain the hills stand for a group of agents polarized into two clusters.

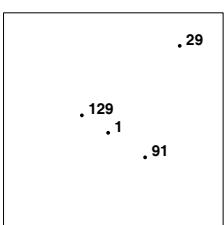
The problem with this visualization is that one do not see how many clusters do really evolve. Especially it is not clear what happens for low ε . The figure is noisy there such that one does not see any attractive regions anymore.

Deffuant et al [19] fixed $n = 1000$ and $s = 250$. They counted the number of clusters in each limit profile. Then they calculate for each ε and each number of clusters the fraction of runs which leads to these number of clusters. We give these kind of diagrams in Figure 4.1 for our setting $n = 200, s = 250$ and $\varepsilon = 0, +0.01, 0.4$. One clearly sees that there are attractive ε -phases for each number of clusters. The problem with this is that one counts every cluster even if it is only one isolated outlier. As we will see, these outliers or minor clusters appear for structural reasons. If one is only interested in the big clusters one has to define a rule which cluster should be regarded as a minor cluster. We show that this is very relevant by the dotted lines in the figure. Further on, the diagram contains no information about locations of clusters.

One reason for the problems to get a complete view about attractive states of the dynamics is that we only have the aggregated data of random settings and unfortunately the noisy situation does not wash out, at least not for sample sizes below $s = 1000$. Nevertheless, there is some evidence that there exist ε -phases for which the processes lead to a characteristic clustering pattern and that this pattern is already attractive for medial number of agents about 200 to 1000. Therefore, we will use here an approach with the homogeneous interactive Markov chains.

4.2.2 The basis of bifurcation diagrams

As outlined in Subsection 2.4.5 on initial opinion distribution $p(0) \in \Delta^{n-1}$ with a huge enough number of opinion classes can capture the dynamics of a



4.2. Bifurcation of clusters in the evolution of the bound of confidence

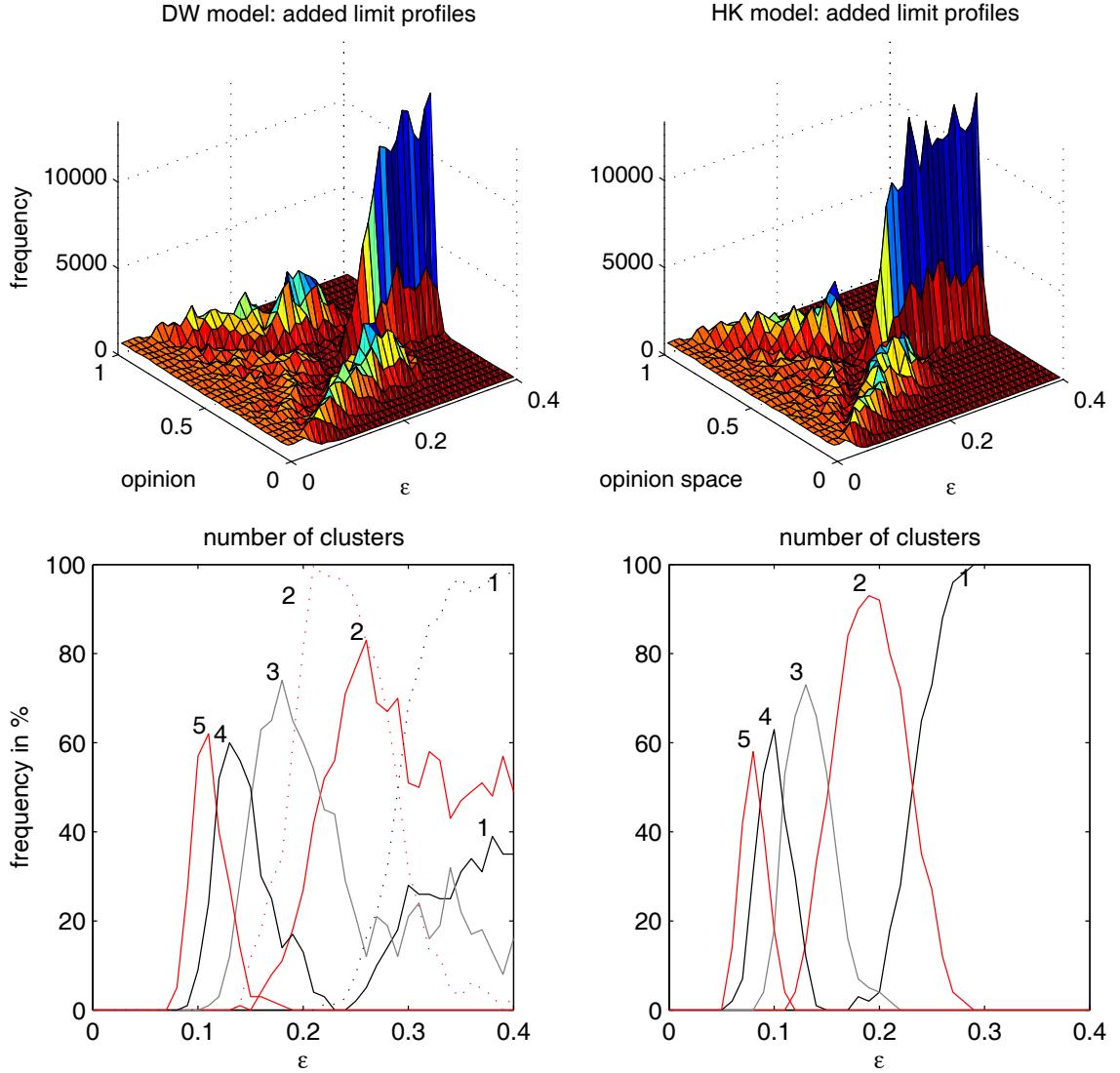
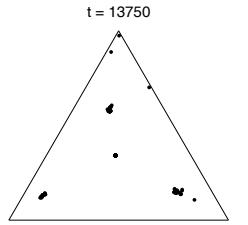
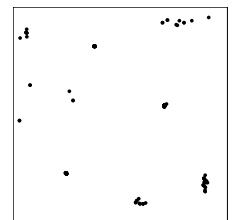
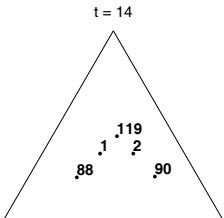


Figure 4.1: Data for a set of 200 initial opinion profiles in $x(0) \in [0, 1]^{200}$. The opinions in each initial opinion profile are random and uniformly distributed. For each initial profile the limit profile is computed for the HK and the DW model and $\epsilon = 0, +0.01, 0.4$. Then added limit profiles (top) and frequencies of the number of evolving clusters (bottom) are shown (for the dotted lines only clusters with more than five agents were counted).

$t = 13750$





4. Simulation

collection of agent-based simulation runs with a huge enough number of agents when the random initial opinions are chosen with probabilities defined by the opinion distribution $p(0)$. So, a collection of initial opinion profiles with random and uniformly distributed initial opinions is captured by the initial opinion distribution $p_i(0) = \frac{1}{n}$ with $i = 1, \dots, n$. In the following we will deal with $n = 201, 1001$.

We have pointed out the ‘equality in the limit for large n ’ for the agent-based and the density-based approach. But this is just a heuristic analogy in construction, there is no formal proof. But there is also evidence from example simulation runs that the two approaches capture the same phenomena already for not too large n . To check this one can compare the Figures 2.11 vs. 2.20, 2.12 vs. 2.24 and 2.22, and 2.14 vs. 2.23.

As Figure 2.17 shows it is important if one chooses an odd or an even number of opinion classes. This is because for a symmetric initial distribution dynamics are very much influenced by the evolution of the mass in the central class(es). If there are two central classes (n even) then the mass may split easier. We chose here the odd alternative because it seems closer to the agent-based behavior in simulation.

It is usual to study parameters which rise on axes of a diagram. Here, ε will rise on the abscissa. Bifurcations of a cluster into more clusters mostly occurs when lowering ε . Thus, the presented diagrams may better be called *reverse bifurcation diagrams*.

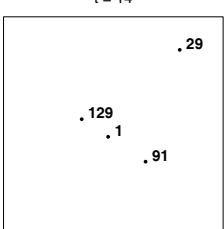
4.2.3 Bifurcations in the DW model

In Figure 4.2 we see the opinion distributions of the homogeneous DW interactive Markov chain which start with a uniform initial opinion distribution after 500 time steps for a ‘continuum’ of ε . On the ordinate there is the opinion space $[0, 1]$ divided into $n = 1001$ opinion classes and on the abscissa there is ε in discrete steps of $\varepsilon/1001$ with $\varepsilon = 50, \dots, 300$. For each point $(i, \varepsilon) \in \underline{n} \times \{50, \dots, 300\}$ we draw a rectangle with the color in a gray scale which stands for the proportion of agents with white means no agents and black all agents. The scale is not linear and emphasizes even very low proportions with a certain amount of gray to make them visible.

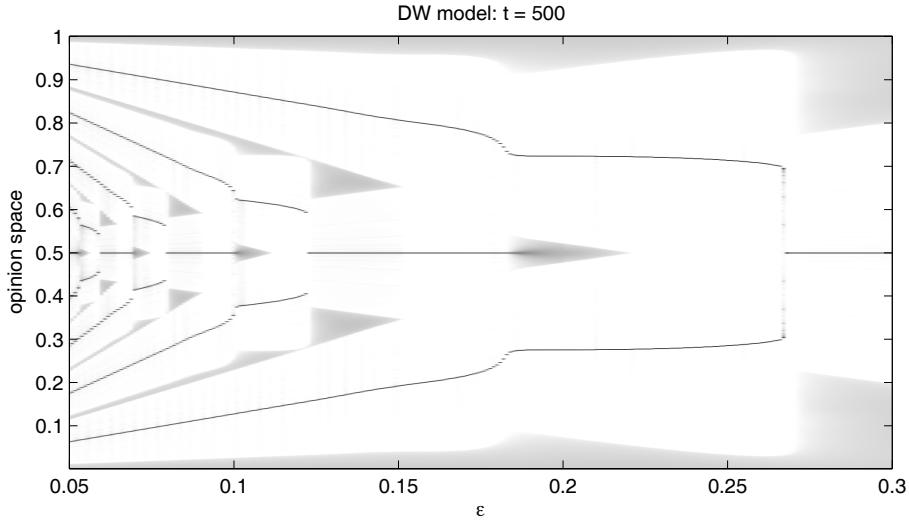
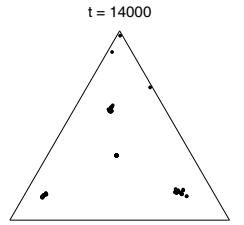
The figure shows that the location of the final clusters evolves in lines which bifurcate at certain points in time. Further on, it shows that minor clusters appear structurally nearly always between the major clusters and at the extremes. While the major clusters are nearly converged to a situation where mass has contracted into two classes² the minor clusters still spread across several clusters, although (or better because) they have a significantly lower proportion of the agent population. The effect of lower convergence in minor clusters is also observed in agent-based simulations, e.g. in Figure 2.10 (bottom right hand side).

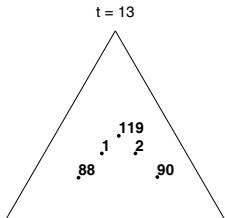
Obviously, the proportion in many classes converges to zero but on the other hand no class will reach zero in finite time. The clear calculation of clusters works only with the definition of a level of precision.

²The two classes are not really visible in the figure due to high resolution and small size. We know from Subsection 3.4 that in fixed points of the model adjacent positive classes may occur.



4.2. Bifurcation of clusters in the evolution of the bound of confidence





4. Simulation

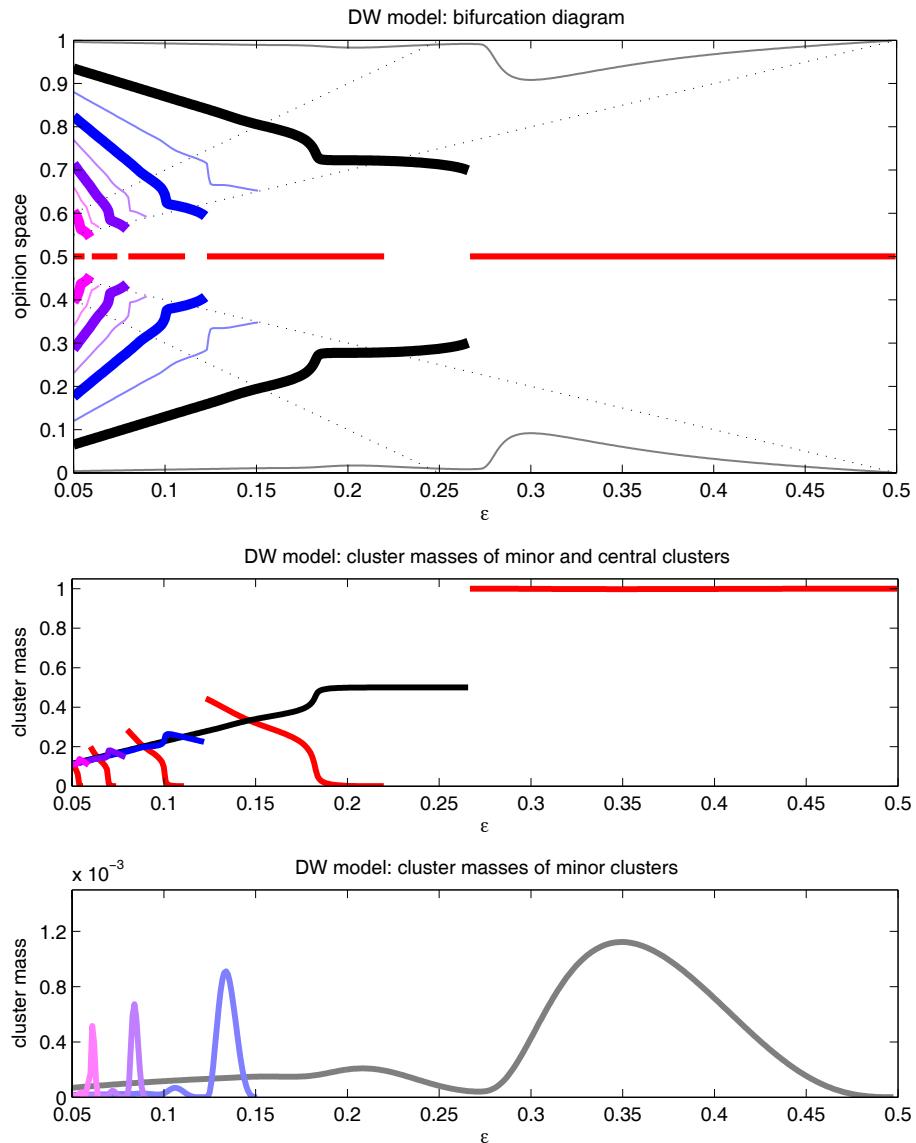
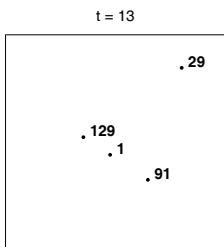
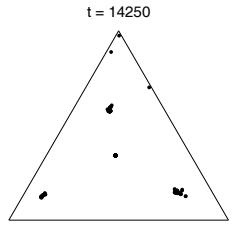


Figure 4.3: Exact location and masses of clusters with level of precision $\delta = 10^{-9}$ in the DW model when cluster formation has stabilized, computed with interactive Markov chains with $[0, 1]$ divided into 1001 opinion classes and $\varepsilon = \frac{50}{1001}, \dots, \frac{500}{1001}$ and uniform initial distribution. The dotted lines stands for the ε -interval around the central cluster.





The phase ‘polarization with marginal center’ and ‘polarization with equipollent center’ are not divided by a change in the number of clusters, but by the ε -value where the mass of the central cluster jumps quite quick from a neglectable mass to an amount comparable to the major clusters.

So, transitions from one ε -phase to another are of four types. If we go down with ε the transition of ε -phases occur as

1. Nucleation of two minor clusters symmetric and with distance ε to the center.
2. Nucleation of two major clusters symmetric and with distances less than ε to the center and vanishing of the central cluster.
3. Nucleation of a marginal central cluster.
4. Sudden increase in the mass of the central cluster and sudden drift outwards in the location of the two major clusters.

These pattern repeats $1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, \dots$ with $\varepsilon \rightarrow 0$ until the accuracy is not fine enough.

4.2.4 Bifurcations in the HK model

The same simulation setup as in the former Subsection 4.2.3 for the DW interactive Markov chain model can be applied to the HK interactive Markov chain. The problem of determining a converged state is not present here, because convergence happens in finite time. An analysis of convergence time is also of interest because we know from examples that convergence time can be very long when metastable states evolve. And, indeed the last Markov chain converged at $t = 21431$.

Another important issue is numerical instability of the process. Computing it as it is sometimes leads to nonsymmetric opinion distributions, which are theoretically not possible. To circumvent this we make the profile p symmetric again by $p=(\text{flip}(p)+p)/2$ after each step.

Figure 4.4 shows the diagram of all the HK interactive Markov chains for interesting time steps.

One can see how consensus evolves very slowly in the ε -interval $[0.19, 0.21]$ and very very slowly for some values in the ε -interval $[0.15, 0.175]$.

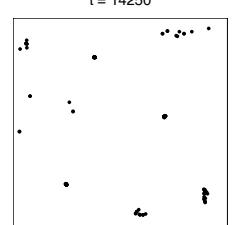
Figure 4.5 gives a better overview with colored lines of clusters, cluster masses and convergence time.

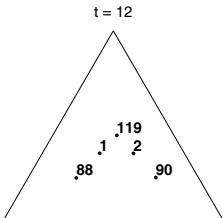
We give the ε -phases in Table 4.2.

ε -region	phase name	clusters
$0.19 < \varepsilon$	total consensus	central
$0.172 < \varepsilon < 0.19$	polarization	marginal central, 2 major
$0.151 < \varepsilon < 0.172$	consensus strikes back	instable cons./pol.
$0.1 < \varepsilon < 0.151$	polarization with center	equipollent center, 2 major
$0.098 < \varepsilon < 0.1$	4 major clusters	marginal central, 4 major

Table 4.2: ε -phases for the homogeneous HK interactive Markov chain with uniform initial distribution.

The phase ‘polarization with center’ is similar to the phase ‘total consensus’ if one removes the first outer major clusters. Since then the same ε -phases repeat





4. Simulation

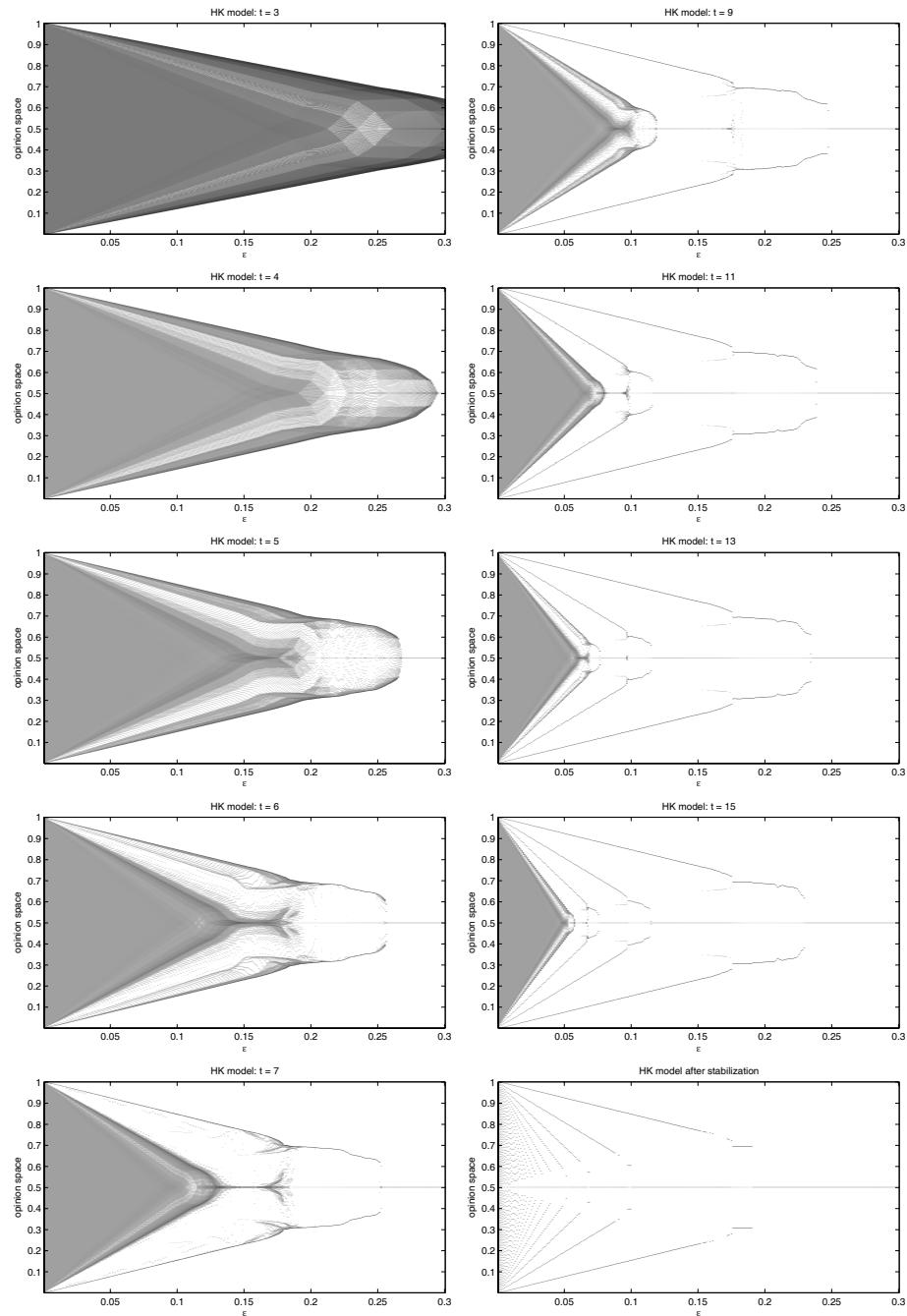
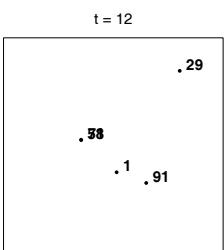
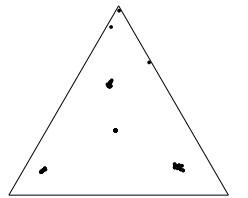


Figure 4.4: Opinion distribution of homogeneous HK interactive Markov chains starting with uniform initial distribution for opinion space $[0, 1]$ divided into 1001 opinion classes and $\varepsilon = \frac{50}{1001}, \dots, \frac{300}{1001}$ for time steps 3, 4, 5, 6, 7, 9, 11, 13, 15, and after convergence at time step 21431.





4.2. Bifurcation of clusters in the evolution of the bound of confidence

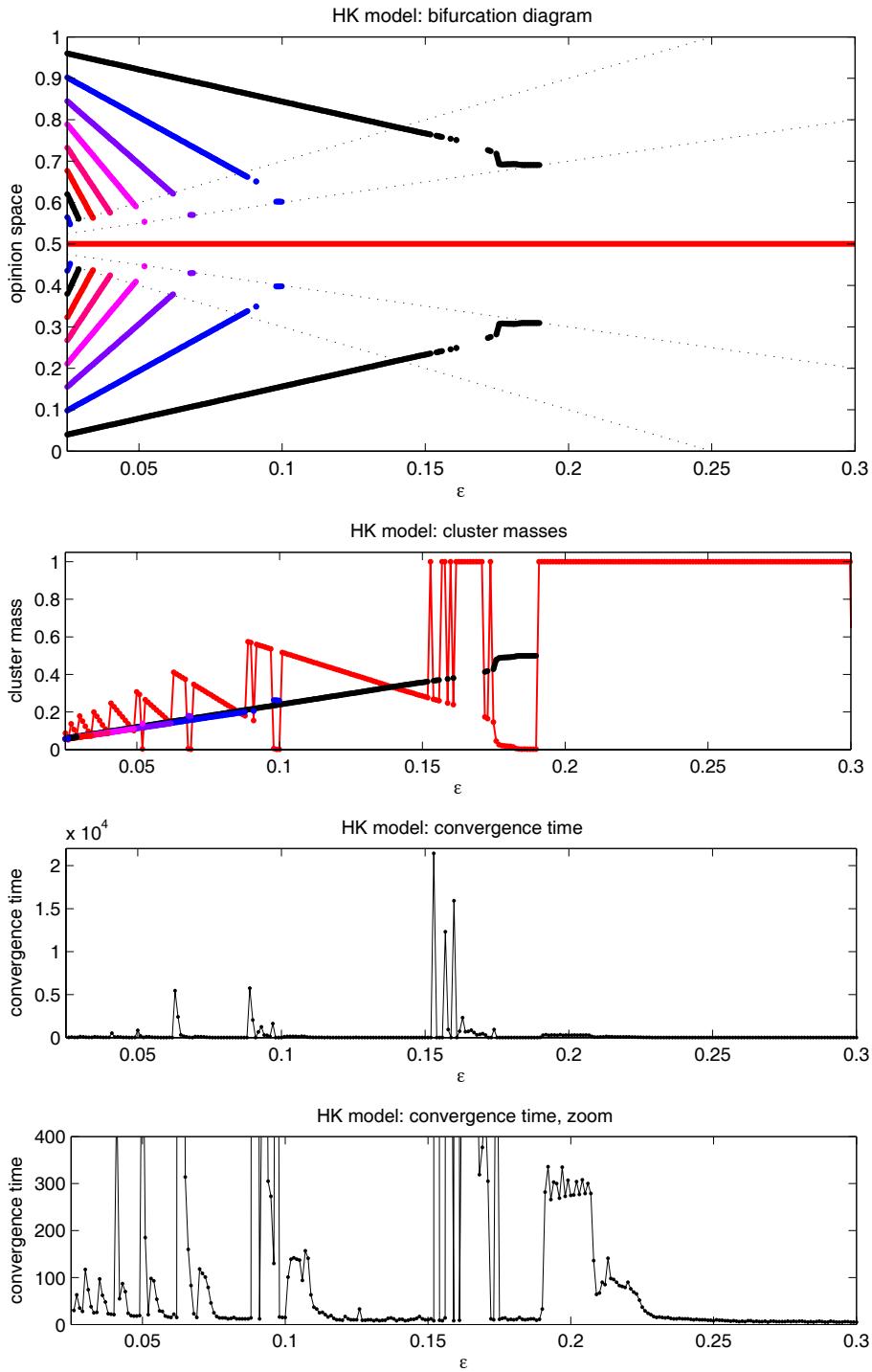
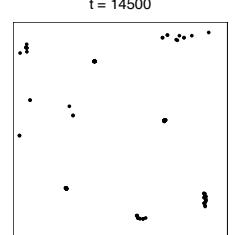
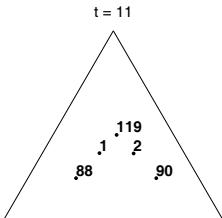


Figure 4.5: Exact location and masses of clusters in the HK model after stabilization, computed with interactive Markov chains with $[0, 1]$ divided into 1001 opinion classes and $\varepsilon = \frac{50}{1001}, \dots, \frac{500}{1001}$ and uniform initial distribution. The dotted lines stands for the ε -interval around the central cluster.





4. Simulation

with always two majors added on increasingly shorter ε -phases.

So, transitions from one ε -phase to another are of three types. If we go down with ε the transition of ε -phases occur as

2. Nucleation of two clusters symmetric and with distances ε to the center and nearly vanishing of the central cluster.
3. Unstable vanishing of the two clusters to the benefit of the central cluster.
4. Stabilization of the existence of two clusters and an equipollent central cluster.

These pattern repeats 2, 3, 4, 2, 3, 4, 2, 3, 4, ... with $\varepsilon \rightarrow 0$ until the accuracy is not fine enough. The numbers of the types begin with 2 because then transitions are analog to transitions in the DW model. The HK model does not have minor clusters, thus transition 1 for the DW model does not have a counterpart in the HK transitions.

4.2.5 Comparison between the DW and the HK model

In Figure 4.6 we see the bifurcation diagrams for the DW and the HK interactive Markov chain as in Figures 4.3 and 4.5 together.

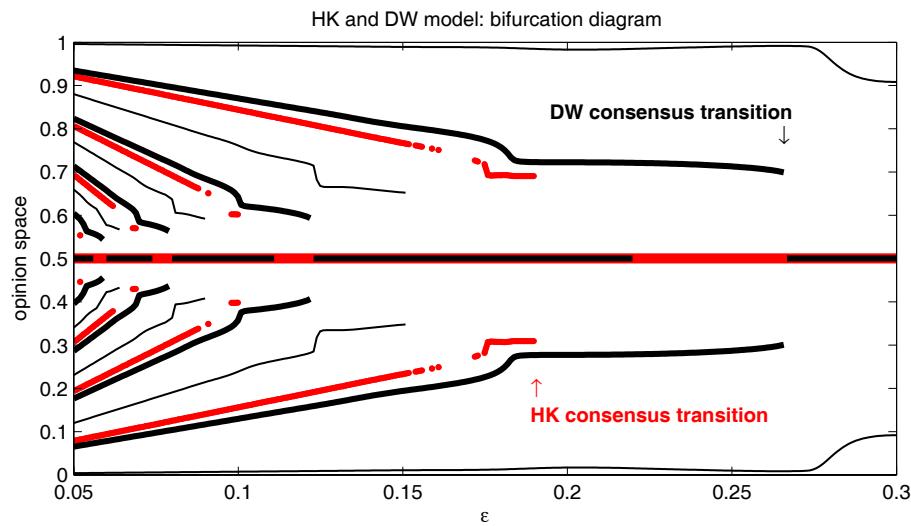
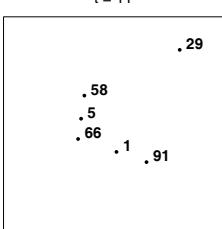
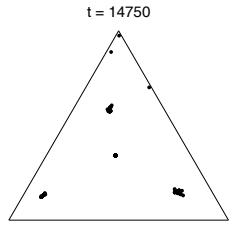


Figure 4.6: Bifurcation diagrams for the DW and HK interactive Markov chains of Figures 4.3 and 4.5 together. Emphasized is the critical ε -value for the consensus transition.

For each major cluster in the HK model there is an adjacent major cluster in the DW model. In general, the HK major clusters lie closer to the center than its adjacent major cluster in the DW model. Minor clusters exist only in the DW model, although there would also be space between the major clusters in the HK model.

The bifurcation points where two new major clusters nucleate (and the central cluster is reduced to zero or to a marginal size) appear for much lower ε in the HK model compared to the DW model. This comes mainly due to the slow convergence of evolving metastable states with two major clusters and a small central cluster.





Further on, there is a surprising ‘consensus strikes back’ phase where the metastable state contains a small central and two minute clusters which connect the central cluster to the two major clusters. The ‘consensus strikes back’ phase is unstable. So, sometimes convergence to consensus works, sometimes it fails. One may speculate that it will work more often when the number of classes is raised but probably at the cost of even worse convergence times. The ε -value when the ‘consensus strikes back’ phase ends coincides with the ε -value where minor clusters are born between the central and the major clusters in the DW model.

It is not visible in the figure but, there is also a very short ε -phase where convergence in the DW model also happens via a metastable state. For ε slightly above 0.266 convergence to consensus also happens very slowly. This is visible in Figure 4.2, where one can see a non converged situation at the transition from ‘consensus and minors’ to ‘polarization’.

4.2.6 Is there a universal scale for transition points?

For this subsection we want to change the scaling of the bifurcation diagrams, such that they are comparable to the results in Ben-Naim et al. [5]. They fixed the bound of confidence to one and then ran simulations for opinion spaces $[-y, y]$ with $y \in \mathbb{R}$ growing.

We can transform our diagrams to this form by mapping the opinion spaces depending on ε from $[0, 1]$ to $\frac{1}{\varepsilon}[-\frac{1}{2}, \frac{1}{2}] =: [-y, y]$ and mapping ε to $\frac{1}{2\varepsilon} =: y$. The total mass of the population in the setup of Ben-Naim et al. is the length of the opinion space interval $[-y, y]$. So cluster masses from our simulation have to be transformed, too.

The transformed diagrams are shown in Figure 4.7 for the DW model and 4.8 for the HK model.

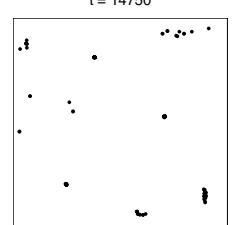
The advantage of this transformation is that the lines of clusters evolve into straight lines with rising y and slope 1 respectively -1.

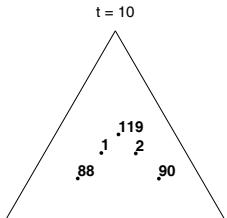
Further on, the distances of transition points of the same type seems more or less constant. We give the transition points for the transitions of the types defined in Subsections 4.2.3 and 4.2.4 of the transformed data in Table 4.3.

DW				HK		
type 1	type 2	type 3	type 4	type 2	type 3	type 4
1.0000	1.8816	2.2750	2.7500	2.6340	2.8600	3.2928
3.3146	4.1025	4.5090	5.0050	5.0050	5.1070	5.6875
5.5611	6.3364	6.7635	7.2536	7.2536	7.3603	8.0726
7.8203	8.4831	8.9375	9.4434	9.2650	9.2650	10.2143

Table 4.3: The y -values of transition points for transitions of the types defined in Subsections 4.2.3 and 4.2.4 of data transformed as described in Subsection 4.2.6.

The data in the lower rows gets inaccurate because data points get sparse for large y due to the scaling $y = \frac{1}{\varepsilon}$. Further on, the data for the DW model varies a little bit compared with the data of Ben-Naim et al. (even in the upper rows where the data of this work is more accurate). This may result from the different framework. Here, we compute the interactive Markov chain which are





4. Simulation

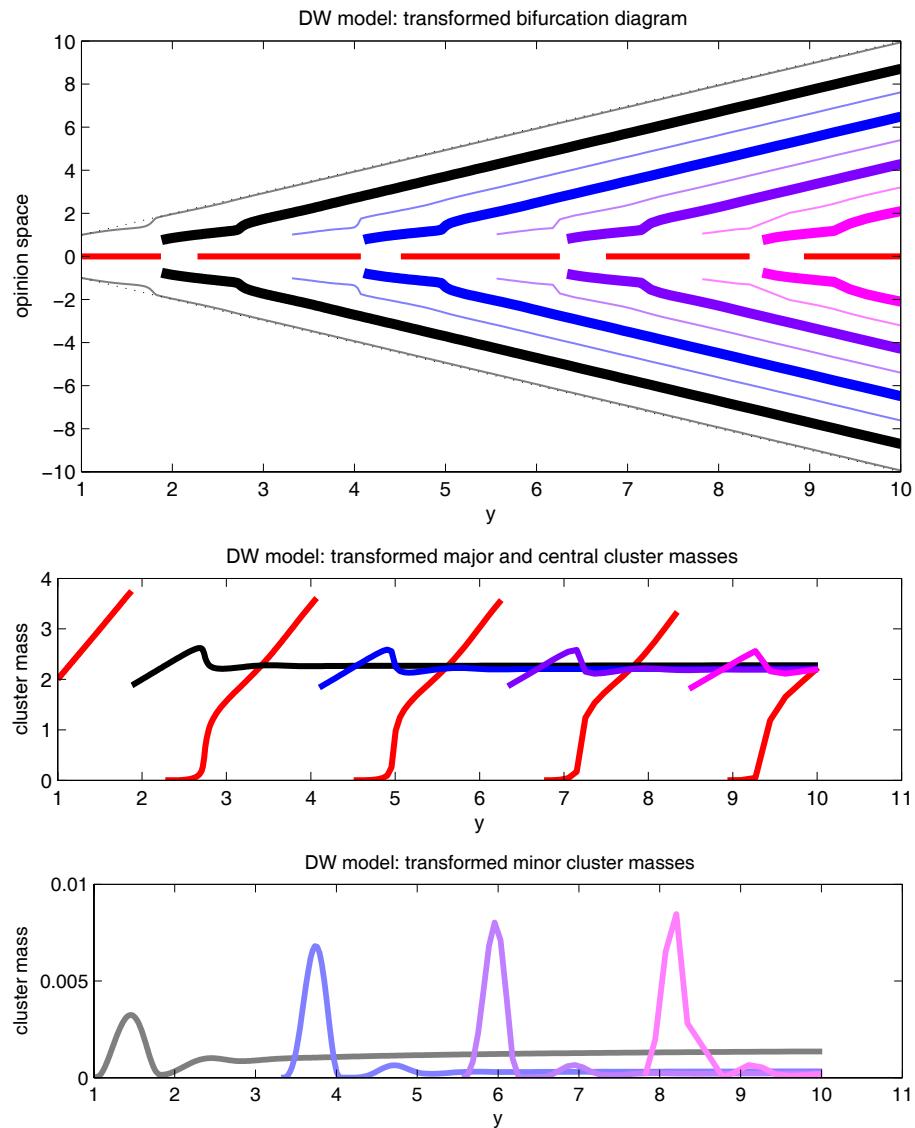
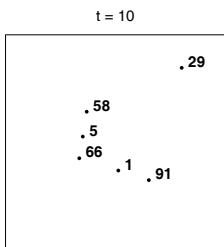


Figure 4.7: DW Diagrams from Figure 4.3 with opinion space $[0, 1]$ transformed to $\frac{1}{\varepsilon}[-\frac{1}{2}, \frac{1}{2}] =: [-y, y]$ and ε to $\frac{1}{2\varepsilon} =: y$. Comparable to the model of Ben-Naim et al. [5].



4.2. Bifurcation of clusters in the evolution of the bound of confidence

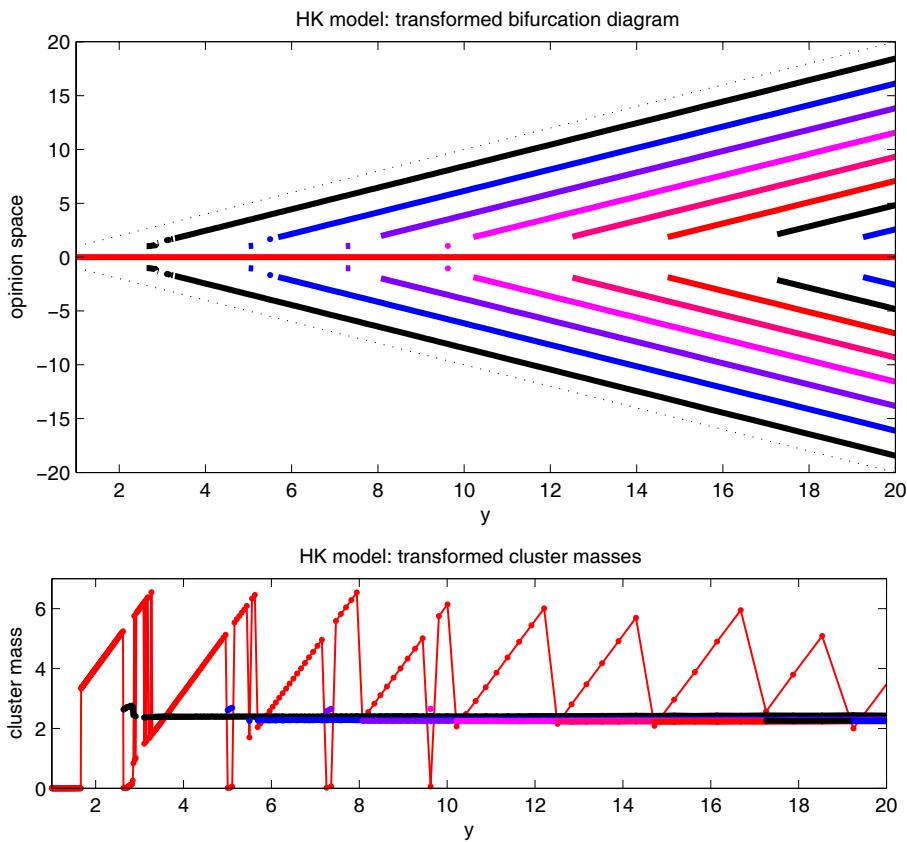
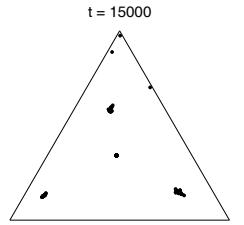
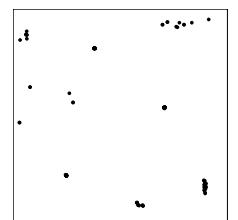
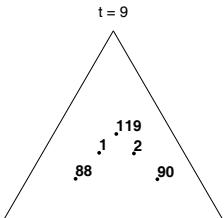


Figure 4.8: HK diagrams from Figure 4.3 with opinion space $[0, 1]$ transformed to $\frac{1}{\varepsilon}[-\frac{1}{2}, \frac{1}{2}] =: [-y, y]$ and ε to $\frac{1}{2\varepsilon} =: y$. Comparable to the diagrams in Ben-Naim et al. [5].

$t = 15000$





4. Simulation

discrete in time and state. In Ben-Naim et al. they made a forth order Adams-Bashforth computation of a differential equation with another discretization of the opinion space.

Ben-Naim et al. claimed in [5] that in the DW model the distances of transition points of the same type converge to a constant value of about 2.155. This coincides with the even rougher ‘ $1/2\varepsilon$ ’-rule which should give the number of evolving (major) clusters in the DW model as stated in [19].

If one extrapolates such a value form the data in this work one reaches at a values also slightly larger than two. Mostly there is a tendency that the differences of a transition point to the next one of the same type is getting a little less than in the step before. This might converge to a fixed rate. But this is a tendency with exceptions. So, it seems not so clear how to extrapolate.

Basically, the data for the HK model which is presented here for the first time looks similar than in the DW model.

It remains an open question if the distances of transition points of the same type really converge or if they begin to fluctuate for larger values of y (lower values of ε). Evidence for the convergence is given by the very regular looking figures here, either for the DW as well as for the HK model.

But there is also evidence for the claim that the HK model will lead to fluctuations for much larger y (much lower ε). Check Lorenz [55] (which is also part of the dissertation) for the evidence. There – in Figure 8 – one sees simulations for one sided dynamics. As we know, dynamics start at the extremes of the opinion space under uniform distributions. Thus, one can study dynamics only on one side by making the opinion space successively longer such that there is no influence from the dynamics coming from the other side. In Figure 8 in [55] one sees that the first twenty clusters evolve quite regular with low fluctuations in their differences but then big fluctuations evolve. This happens for two very different levels of accuracy, which underpins the evidence.

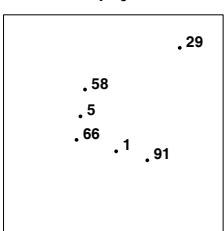
A brief explanation for the fluctuations when we regard one sided dynamics as in Figure 8 of [55]: The formation of a cluster produces some low fluctuations in upper classes these fluctuations propagate quicker than the clustering to upper classes and at some time they have impact on the so far more or less regular clustering.

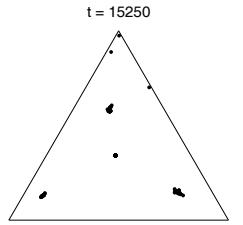
It is an open problem if this occurs only due to floating point errors or to low levels of accuracy or for structural reasons of dynamics. Further on, it is open if this might happen in the DW model, too.

Another source of inspiration to this question is the diploma thesis of Dittmer [22]. He studies the one sided agent-based HK model with equidistant opinion profile $x(0) = [1 \ 2 \ 3 \ 4 \ \dots]^T$ and $\varepsilon = 1$. In this setting he computes the evolving clusters and finds numerically up to the 155th cluster that the distance gets stable and every cluster contains three agents, which is evidence for stabilization of the distances of clusters again.

4.3 Heterogeneous bounds of confidence

In the following we study the impact of heterogeneity of the bounds of confidence on the chances for consensus. This study is only a very first step in analysis because we will only deal with two levels of heterogeneous bounds of confidence $\varepsilon_1, \varepsilon_2 > 0$. Further on, we assume that half of the population has bound of





confidence ε_1 and is uniformly distributed, the same for the other half which has bound of confidence ε_2 .

We study this in the framework of HK and DW interactive Markov chains with heterogeneous bounds of confidence with opinion space $[0, 1]$ divided into $n = 201$ opinion classes and $\varepsilon_1, \varepsilon_2 = \frac{10}{201}, \dots, \frac{70}{201} \approx 0.05, \dots, 0.35$. We ran the interactive Markov chains for each communication regime and each $(\varepsilon_1, \varepsilon_2)$ combination.

The heterogeneous case leads to a much richer variety of stabilized profiles e.g. due to sitting between the chairs situations or slow drifting of a group of closed-minded by a group of open-minded.

Therefore, we study in a first approach the consensus transition. We want to determine the $(\varepsilon_1, \varepsilon_2)$ -region where the dynamics lead to a consensus of a vast majority of agents. The right measure for this is the *mass of the biggest cluster*. This mass can only get larger than 50% if the biggest cluster is the central cluster for reasons of conserved symmetry. Further on, the mass in a central cluster will contract into one opinion class due to the setting with an odd number of opinion classes.

We color the plane of all $(\varepsilon_1, \varepsilon_2)$ points with the mass of the biggest cluster after stabilization. We call this diagram an *extended phase diagram of the consensus transitions*. It is called ‘extended’, because it does not only determine the borders where transitions happen but the exact values of the mass of the biggest cluster. So, regions of certain degrees of consensus are also visible as well as a variety of continuous and abrupt transitions.

4.3.1 Extended phase diagram of the heterogeneous HK model

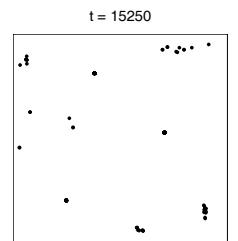
For the HK model the simulation were run until nothing happened anymore. Structurally, the HK chain need not converge in finite time, but it reached a stable state all the time in simulation. Probably because the accuracy of the computer could not capture changes anymore.

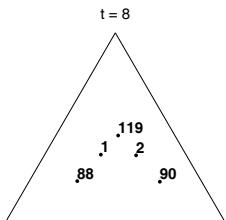
Numerical problems evolve when distributions got unsymmetric for the reason of floating point errors. These problems were circumvented by making the opinion distribution symmetric again after each iteration with $p=(\text{flip}(p)+p)/2$.

Clusters were determined according to definition 4.1.2. Sometimes the maximal distance to the next cluster is lower than than $\max\{\varepsilon_1, \varepsilon_2\}$ because if a ‘sitting between the chairs’ situation occurs, cluster formation has stabilized but not all clusters are disconnected.

Figure 4.9 shows the extended phase diagram for the HK interactive Markov chain.

Let us first take a look on the diagonal of the diagram. The diagonal represents the homogeneous situation $\varepsilon_1 = \varepsilon_2$. Thus, it resembles the maximum of the cluster masses which are shown in Figure 4.5. We see the polarization into two classes which each class capturing half of the population at about $(0.19, 0.19)$ and the ‘consensus strikes back’ phase at about $(0.17, 0.17)$. If we go down further on the diagonal the mass of the biggest cluster is only about one third of the population but increases with lowering $(\varepsilon, \varepsilon)$ until a value of nearly 60% of the population. This is the growing central cluster which even grows to a majority before bifurcating, the highest values is reached for such low value as $(0.09, 0.09)$.





4. Simulation

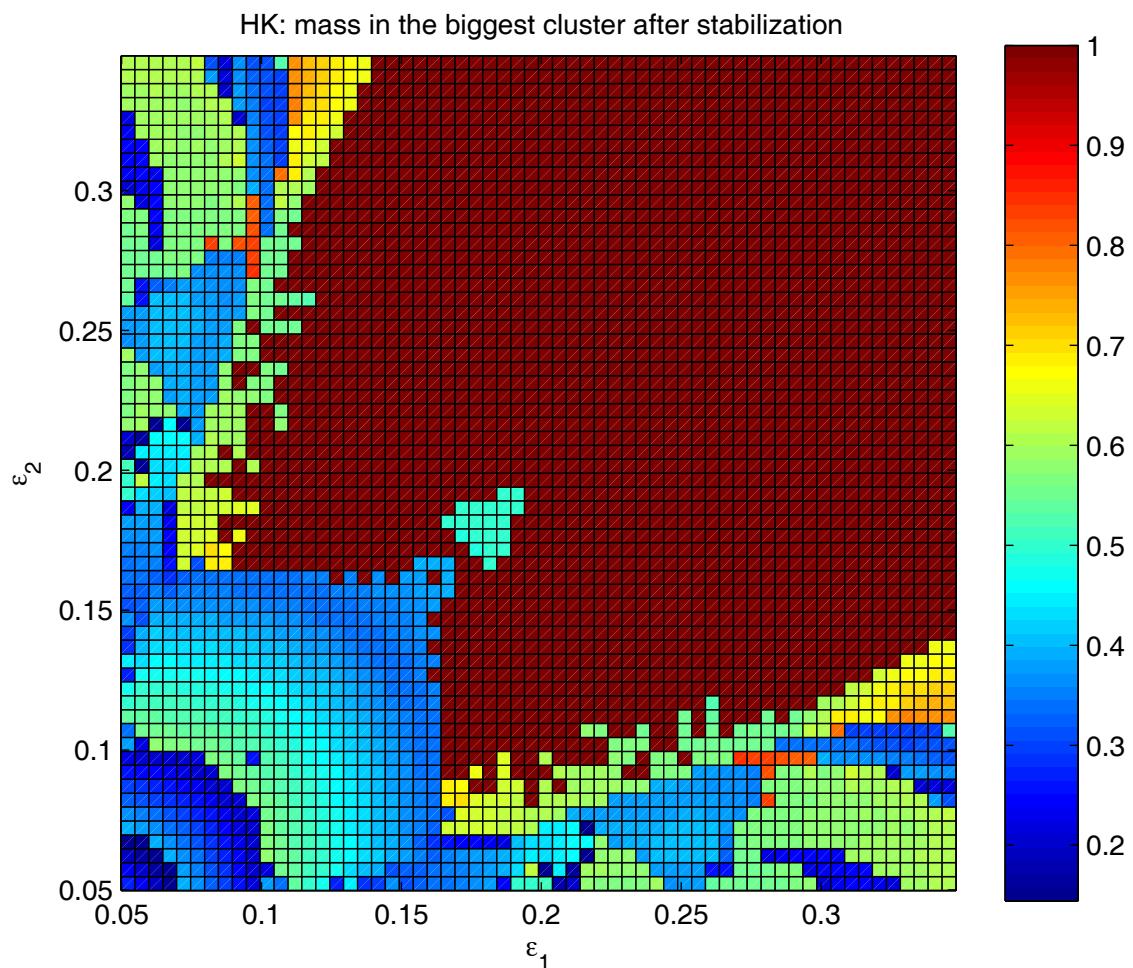
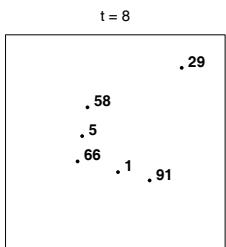
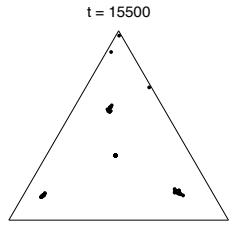


Figure 4.9: Mass of the biggest cluster after stabilization for HK interactive Markov chain with population divided into equal proportions with bounds of confidence ε_1 and ε_2 . Both subpopulation were initially uniformly distributed in the opinion space $[0, 1]$ divided into 201 opinion classes.





But the real surprise comes when we look on the heterogeneous situations beside the diagonal. We observe that even for low differences in the two bounds of confidence the polarization into two 50% clusters vanishes. Consensus is e.g. possible with $\varepsilon_1 = 0.17$ and $\varepsilon_2 \in [0.1, 0.15]$ which is both lower than the consensus transition value from the homogeneous case 0.19. But on the other hand if ε_1 goes up, ε_2 cannot be as low as in the example before, so if $\varepsilon_1 = 0.35$ then ε_2 must be larger than 0.14 to lead to consensus. Probably the chances for sitting between the chairs situations rise with a larger spread between ε_1 and ε_2 , while for a lower spread drifting to the center is more likely to occur. We also expect that the symmetric situation has also a strong impact on the surprising positive impact that heterogeneity has on the chances for consensus in the HK model.

We also observe that the bounds between the consensus and the non consensus regions are often fuzzy. This coincides with the observation of the unstable ‘consensus strikes back’ phase in the bifurcation diagram of the homogeneous model.

4.3.2 Extended phase diagram of the heterogeneous DW model

For the heterogeneous DW model stopping criteria are more complicated because it has a rich variety of types of convergence which are not fully classified and understood until now. Further on, convergence can last very long and it is difficult to decide whether convergence will lead to another drastic change once or not. There are three figures to give an idea what happens.

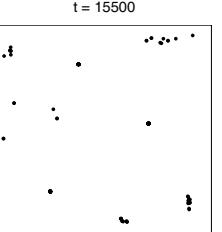
For Figure 4.10 we compute clusters with precision level 10^{-4} on $p^1(t) + p^2(t)$. We then join clusters which are closer than $\max\{\varepsilon_1, \varepsilon_2\}$ and stopped dynamics if the length of these clusters was below $\max\{\varepsilon_1, \varepsilon_2\}$. The figure also shows a drastic positive impact of heterogeneity on the chances for a consensus of a vast majority. The impact looks even more drastic as for the HK model.

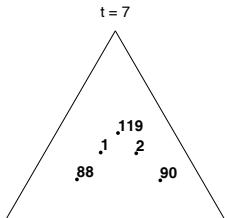
It is important to notice that the diagram might look very different for different levels of precision. In the Figure 4.10 most points are converged except for some in the upper left und bottom right corner (which points these are is visible in Figure 4.12). So for unconverged states it might happen that clusters may split and lead to less mass in the biggest cluster. But for converged points it might be the case that dynamics would go on when classes with mass less than δ were not neglected. This might lead then to a union of clusters and produce a larger mass of the biggest cluster.

But on the other hand if one thinks about a population of agents whose bounds of confidence and initial opinions are more or less determined by the opinion distribution in these simulation then one needs a really huge number of agents such that one could expect enough agents in these minor clusters such that this very low dynamics really occur in agent-based settings.

So, for some pairs $(\varepsilon_1, \varepsilon_2)$ where consensus is reached this might be possible only with millions of agents in agent-based simulation and additionally after very long time.

Figure 4.11 gives a better view what happens after some not too long time for $t = 200$. Because clusters can not be determined at this level we simply plot the maximal proportion of population over all opinion classes. One can see that especially in the region around $(\varepsilon_1, \varepsilon_2) = (0.11, 0.22)$ the maximum has already exceeded 50%. So, in this region consensus will appear after a reasonable amount





4. Simulation

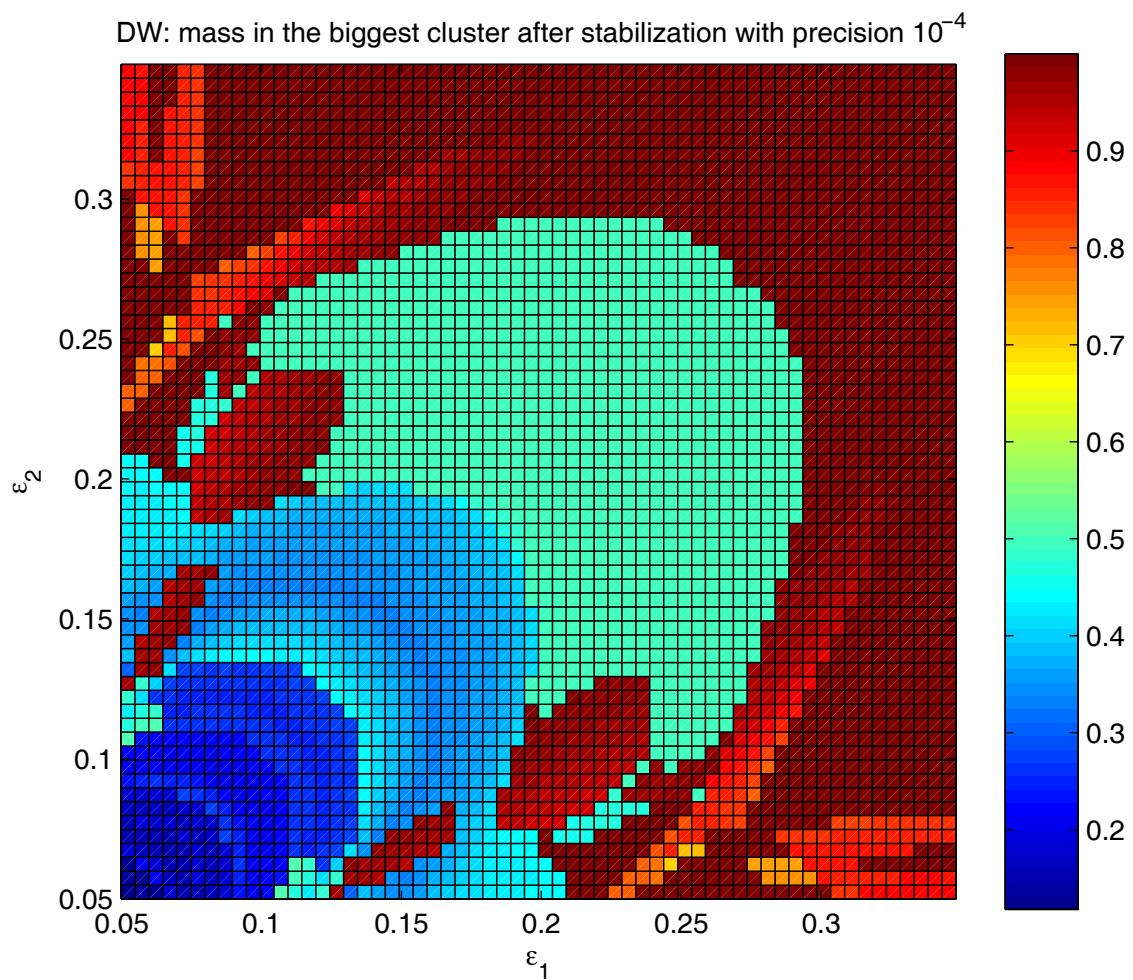
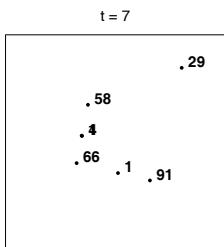
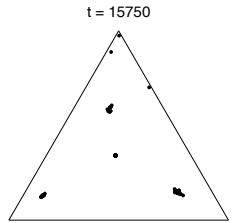


Figure 4.10: Mass of the biggest cluster after stabilization (with level of precision $\delta = 10^{-4}$ for DW interactive Markov chain with population divided into equal proportions with bounds of confidence ε_1 and ε_2 . Both subpopulation were initially uniformly distributed in the opinion space $[0, 1]$ divided into 201 opinion classes.



4.3. Heterogeneous bounds of confidence



of time. Surprisingly ε_1 and ε_2 are both much less than the value for consensus transition in the homogeneous case which is 0.266. And indeed in Figure 2.14 we see an example where this type of convergence to consensus happens with 1000 agents. So, this example is not hand-picked but there is a generic reason for this type of convergence. So, effects of heterogeneity can be more drastic as stated by Weisbuch et al [88]. They only claimed that the dynamics of the higher ε will govern the evolution of clusters in the long run. Here we see that two heterogeneous low ε can even work together and reach consensus which is impossible for them if they were homogeneous.

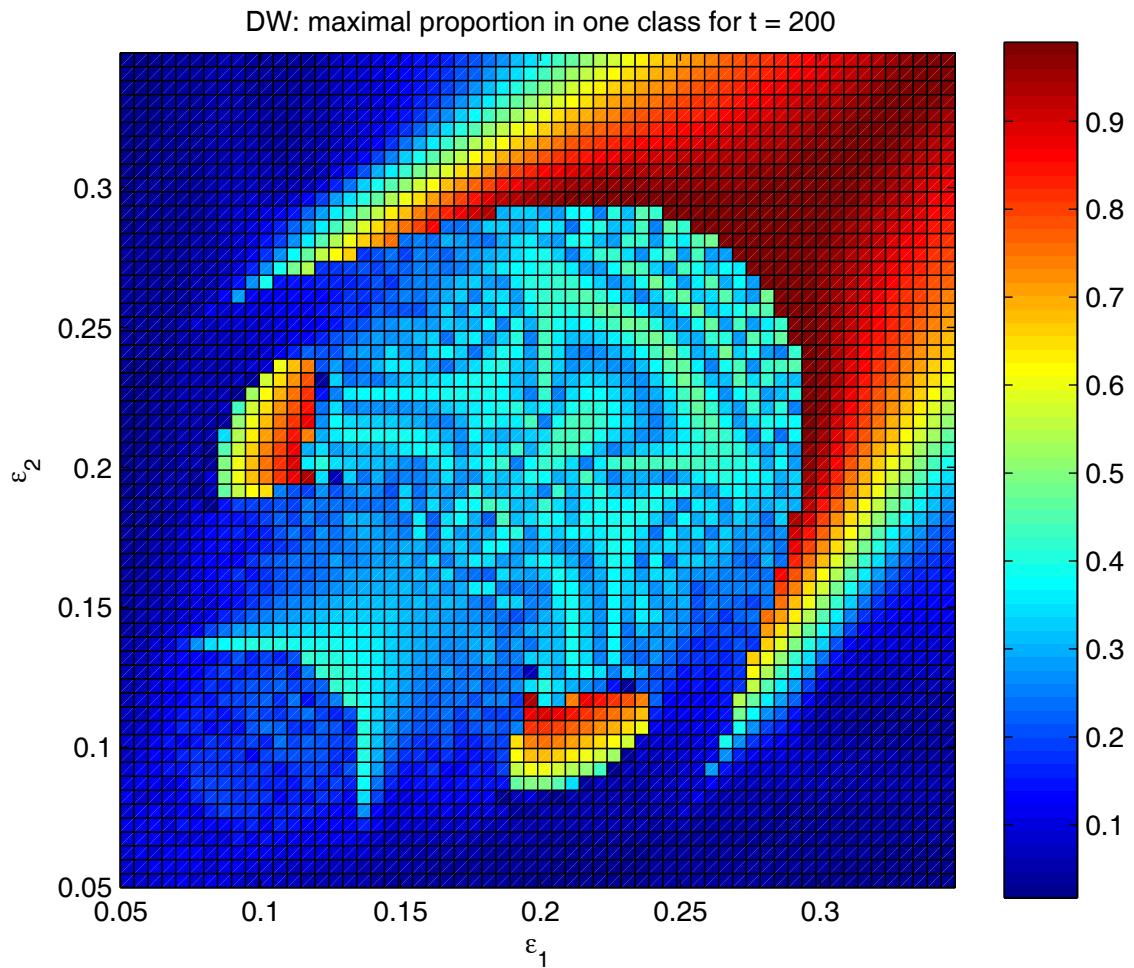
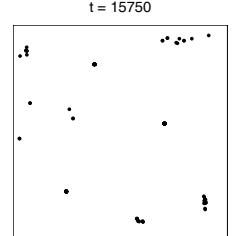
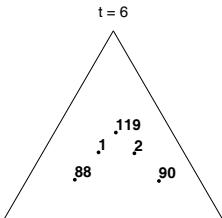


Figure 4.11: Mass of the class with the most mass for $t = 200$ for DW interactive Markov chain with population divided into equal proportions with bounds of confidence ε_1 and ε_2 . Both subpopulation were initially uniformly distributed in the opinion space $[0, 1]$ divided into 201 opinion classes.

At last, Figure 4.12 shows the time steps when the central class contains more than 50% of the mass. So, this is another measure for convergence to consensus in a reasonable time. The color axis has been adjusted from zero to thousand.





4. Simulation

For all white points the central cluster has not exceeded 50% either because of convergence to polarization or plurality (around the diagonal) or because of too low convergence (corners). We emphasize that the dark red stands for a huge time interval, e.g. the converged states in the upper left and bottom right corner converged after 45.000 time steps! Further on, we emphasize that there is one example for long convergence time on the diagonal, so for the homogeneous case. This is an example for the very short ε -phase where convergence to consensus in the DW is via a metastable polarized state. We know that this ε -phase is really huge for the HK model.

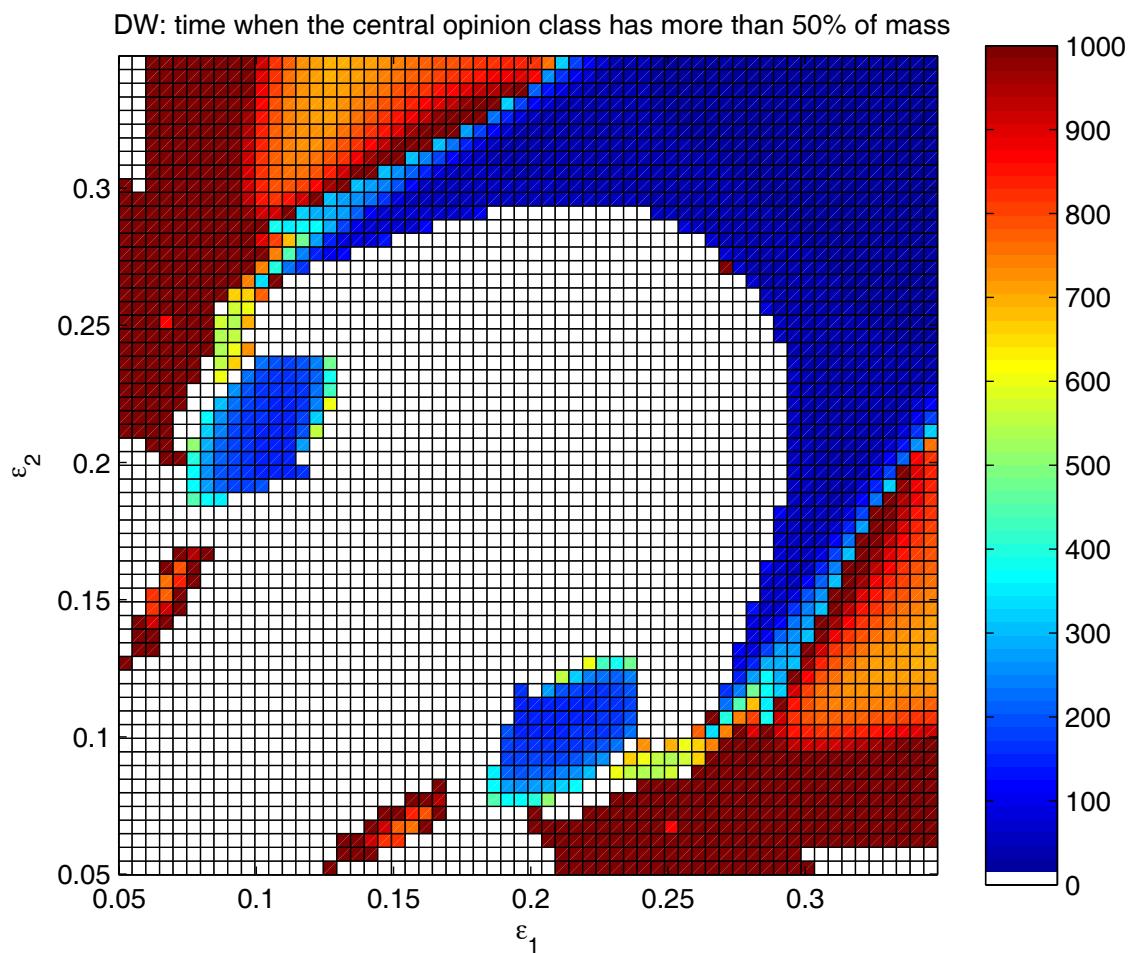
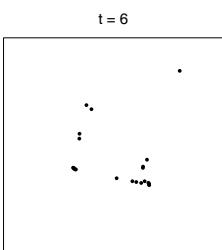
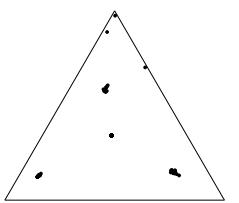


Figure 4.12: Time when the mass in the central class exceeds 50% for DW interactive Markov chain with population divided into equal proportions with bounds of confidence ε_1 and ε_2 . Both subpopulation were initially uniformly distributed in the opinion space $[0, 1]$ divided into 201 opinion classes.





 4.4. Results in accompanying papers

4.4 Results in accompanying papers

Here, short summaries are given for the simulation results in the accompanying papers which are also part of the dissertation [55, 56, 58, 61]. This should give the reader a guideline for neglecting redundant content in the papers which is also covered by this thesis.

4.4.1 Multidimensional opinions

Papers [56, 58] deal with agent-based opinion dynamics of multidimensional opinions. The question was if multidimensionality may foster the chances for consensus. In both papers simulations have been done with 200 agents.

In [56] the opinion spaces are simplices Δ^{n-1} for $n = 2, 3, 4, 5, 6, 7, 8$. So, n stands for the length of an opinion vector. Attention! In this paper n determines the number of opinion issues (here d) and $m = 200$ the number of agents. The norm parameter p is set to $p = 2$ there and simulations were carried out for the DW as well as for the HK model.

The opinion issues in an opinion coming from the simplex have to be positive and sum up to one. This is interpreted as a *budget constraint*. There is a fixed amount of money which has to be distributed to n departments. The main question is than if a larger number of departments has an impact on the chances for consensus.

The paper can be read with a brief look on Section 1, omitting Section 2 which only defines models and shows examples and the bifurcation diagrams. Simulation descriptions and results are then to be read in Section 3 including some mathematical reasons for the results. Interpretations are then in the conclusions (Section 4). The main message in a nutshell is:

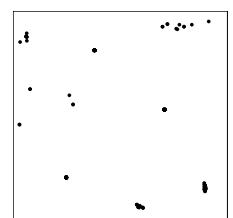
Consensus is fostered by raising the number of departments if the opinions are under budget constraints. One drawback is that the number of outliers raises, too. This happens drastically in the DW model but also in the HK model. Further on, the lowering of the consensus transition when raising the number of departments slows down. There seems to be a lower limit where the ε -value of the consensus transition can not fall below. Thus, raising the number of departments ad infinitum does not make sense also from this point of view.

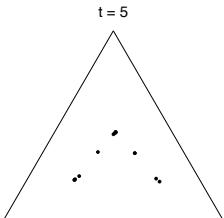
Paper [58] deals mainly with the same problem but has a broader and diverse parameter space. We distinguish there the opinion spaces the simplex Δ^d and the cube $\square^n := [0, 1]^n$ (has nothing to do with cube defined in Chapter 3) for $d = 1, 2, 3$. For the simplex opinion issues are regarded to be under budget constraints, for the cube they are regarded to be independent. Further on, two norm parameters $p = 1, \infty$ were distinguished. Agents which use the 1-norm are regarded as compensator and the ones with the ∞ -norm as noncompensators. Finally, we distinguish the two communication regimes HK and DW.

The paper can be read with a brief look on Sections 1, 2 and 3. Simulation descriptions and results are then to be read in Section 4 including interpretations followed by a summary and outlook in Section 5. The main message in a nutshell is:

Consensus can be fostered by raising the number of opinion issues if they are under budget constraints, while consensus is weakened if issues are not under

t = 16000





4. Simulation

budget constraints. Further on, some slight effects are shown that compensating agents might foster consensus a little bit better in this setting than noncompensating agents.

4.4.2 Communication strategies

The paper [61] is joint work with Diemo Urbig. The paper concerns only the agent-based DW model and the one-dimensional opinion space $[0, 1]$. The basic questions was on the impact of the choice of communication partners. The mathematical part of the paper has already been described in Section 3.3.4.

Further on, there is a simulation part where we implemented the communication strategies ‘balancing’ and ‘curious’. The idea is that agents remember their last positive communication (when they changed their opinion) and especially if their partner was an agent with a lower or an upper opinion. Then curious agents want to find a communication partner of the same side again, while balancing agents want to find a communication partner from the other side. They did this by refusing to compromise if the other agent is close enough but comes from the wrong side. But agents refuse only until a frustration maximum f_{\max} is reached. If it is reached they are open to all agents within their area of confidence again.

To get the simulation results one should briefly read Section 1, skip Sections 2 and 3, read Section 4 and the discussion in Section 5 as well as Appendix B for the details of the balancing and curious strategy implementation. The main message in a nutshell is:

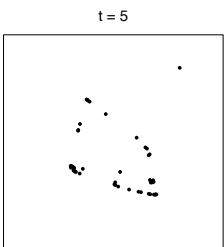
Balancing agents foster consensus in general. But there is an impact of the cautiousness parameter μ (which has neglectable impact in the standard model). The positive effect of balancing agents on reaching consensus gets even better if agents are more cautious (have lower μ). For curious agents the situation is less clear. Curious agents which are cautious need a very high frustration maximum such that consensus is fostered for low frustration maximum chances for consensus are even weakened. For the standard cautiousness value $\mu = 0.5$ dynamics of cautious and balancing agents are equivalent.

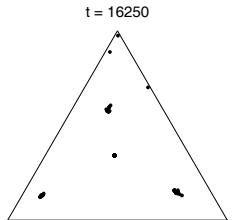
4.4.3 Theoretic questions

The simulations in paper [55] deals with more theoretic questions of the HK model. We cited already some results throughout in the thesis.

One question is the comparison between the agent-based and the density-based HK model. The tool of one sided dynamics which simplify the dynamics is introduced there. A pretty good analogy of the agent-based and density-based HK model even quantitatively is been observed in simulation. This underpins the speculations of convergence of both approaches in the limits for large n in Subsection 2.4.5.

Another question is Hegselmann’s conjecture that for each ε one finds a number of agents such that agents with a uniform distribution in the opinion space must find consensus in HK dynamics. We give evidence that this is probably wrong.





Chapter 5

Conclusion

This conclusion is to briefly sketch possible real-world implications for the design of processes of continuous opinion dynamics (Section 5.1), to summarize the scientific achievements of this dissertation from the point of view of the author (Section 5.2), and to provide a list of open problems for further research (Section 5.3).

5.1 How to foster or dilute consensus

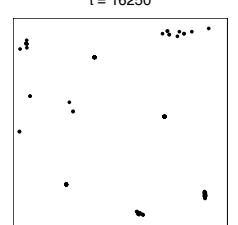
First, we regard a general process of opinion pooling where agents perform repeated averaging of their opinions.

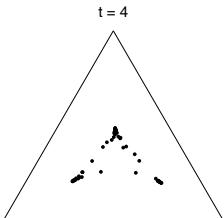
The core question in the early works [11, 21, 50] beginning with DeGroot was on mathematical conditions for reaching consensus. These conditions are, colloquially speaking, that there has to be a somehow connected confidence structure between the agents.

If we regard a fixed confidence structure, agents have to be connected in a way that there is only one group of agents which is internally connected such that everyone takes everyone else's opinion into account at least indirectly. All agents which are not recognized by this group must have an (at least indirect) connection to this core group. The consensus will then evolve in the core group with the other agents following.

It turns out then that even under very heavy changing of the confidence structure over time convergence to consensus is also possible. Important is that a possible indirect confidence structure as described above evolves frequently over certain intercommunication intervals in time. Therefore, the order of different confidence structures in the process can be crucial to produce convergence to consensus or not. If consensus is not achieved there are basically only three possibilities. Either agents get disconnected or they begin an endless alternation of opinions in a way 'we say what you say, you say what she says, she says what we say and so on forever', or agents reach the desired connectivity but on increasingly lower level, such that convergence does not reach consensus in a way 'we approach step by step but we will never be close enough to shake hands'. The phenomenon of endless alternation is ruled out when every agent got at least a bit of self-confidence and has thus a strong impact on stabilization.

Now, we make the additional bounded confidence assumption. So, agents





5. Conclusion

restrict their confidence to other agents dynamically with respect to their distance in opinion. This leads to convergence to diverse patterns of opinion clusters. Cluster formation depends then on the initial opinions, the agent's bounds of confidence and the communication regime. Some of these ingredients may be manipulable by the organizer or initiator of opinion dynamics. We give the following list of advices how to foster reaching consensus by collecting results from this thesis and the accompanying publications.

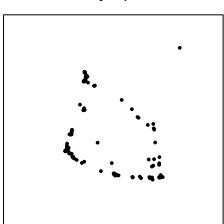
1. Appealing to the agents to raise their bounds of confidence is in most cases a good idea. But it can also lead to negative effects by fostering a clear polarization and diluting the chances for the evolution of small mediator groups.
2. Do not rely only on gossip, under repeated meetings chances for consensus are much higher, but often paid with very long time until consensus is reached.
3. If you have more issues to discuss, then put them under budget constraints.
4. If there are different independent issues to consider, try to reduce the number of issues.
5. If there are different issues under budget constraints to consider, don't be shy to add one more. But there is a limit when there is no effect anymore and one has to be aware of the growing danger that agents get isolated while heavily promoting one single issue.
6. If there are different issues to discuss than probably agents that judgment opinion distances with compensation between the issues should be fostered instead of non-compensators.
7. In a gossip environment it would help if agents act balancing by trying to chose communication partners alternating from different sides.
8. In a gossip environment it can also be a good idea to appeal to agents to be curious by trying to find communication partner with opinions even more to the direction of where their former partner came from. But don't forget to appeal additionally to them to be not cautious when being curious because curious agents that foster consensus usually have to make a journey cross the whole opinion space and back to the center.
9. In most cases it would be positive if agents have diverse bounds of confidence. It could even pay to appeal to some agents to get more closed minded. But be aware that a consensus reached in mixed societies mixed of closed-minded and open-minded agents may be surprisingly far from the initial average opinion, due to possible drifting of clusters.

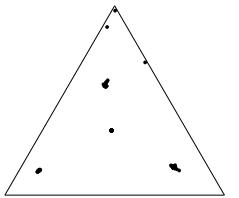
If dissensus is the aim, then just act against all this advices.

5.2 Achievements

This dissertation studies dynamics of repeated averaging under bounded confidence. It seems to be the first larger study of systems with these two driving forces. It was mainly inspired by the bounded confidence models of opinion dynamics of Hegselmann-Krause and Deffuant-Weisbuch.

The contribution regarding modeling lies on the one hand in the compilation of various applications where results may play a role from opinion dynamics to distributed computing and collective motion. On the other hand the agent-based





5.3. Open problems

and density based modeling tools have been defined to help in analysis on the mathematical and the simulation level. These tools are general averaging maps, infinite backward products of row-stochastic matrices and interactive Markov chains which are related to a master equation approach.

Regarding mathematics, we reached a new theorem on convergence to consensus under the repeated application of very generally defined averaging maps. In matrix theory we deliver a concise description of the powers of an arbitrary row-stochastic matrix by emphasizing the existence of a primitive Gantmacher form in some power. Then a bunch of theorems about convergence of infinite backward products of row-stochastic matrices is derived. Some of these theorems are new, some are collected from various works. This seems to be the collection of theorems which are known until now. The difficulties of finding a unifying characterization of converging sequences are pointed out with several examples. Further on, we derived the sets of fixed points of the interactive Markov chains which are defined under the heuristics of repeated averaging under bounded confidence. They seem to be representatives of a very interesting class of systems which converge to fixed points but where fixed points are not isolated and some fixed points seem to be more attractive than others. The model dynamics show interesting interplay between the speeds of contracting and disconnecting. Further on, drifting phenomena occur when processes with different parameters are coupled.

Regarding simulation, bifurcation diagrams and extended phase diagrams are delivered which give an new overview on attractive states of continuous opinion dynamics under bounded confidence.

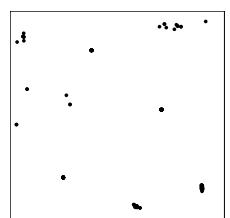
The last contribution lies in the provision of a list of implications for opinion dynamics in the real world and a list of open problems to encourage further research.

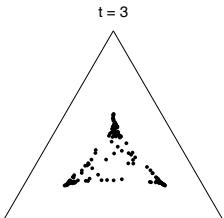
5.3 Open problems

Here is a list of interesting open problems.

Modeling:

1. Model strategic play into continuous opinion dynamics. How to incorporate the fact that agents may additionally act strategically to foster or dilute consensus or to drive the evolving consensual value into their desired direction?
2. Model the impact of voting and decision systems on opinion dynamics.
3. Define a general class of density-based systems on the heuristics of averaging and bounded confidence that captures the dynamics of the DW and the HK model.
4. Define a general framework of continuous opinion dynamics and swarms.





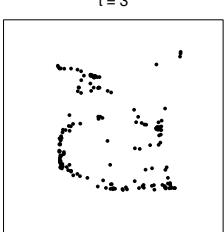
5. Conclusion

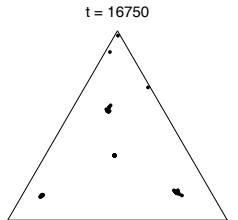
Mathematics:

1. Prove or disprove Conjecture 3.2.16, that a set of row-stochastic matrices Σ which is equiproper, does imply that $\tau_1(A)$ is uniformly bounded from below for all $A \in \Sigma$.
2. Characterize all converging sequences of row-stochastic matrices which backward accumulation converges.
3. Prove or disprove that the HK and the DW interactive Markov chains converge to one of their fixed points. Further on, derive which fixed point will be approached.
4. How to characterize the set of matrices $\{0, 1\}$ -matrices that leave a regular matrix regular after multiplication (Senetas class G_2), especially the matrices which do not have a positive diagonal?
5. Is there a universal distance of transition points of the same type evolving in the DW and the HK model as proposed by Ben-Naim et al [5] with evidence but where there is also evidence for the contrary [55]?

Simulation:

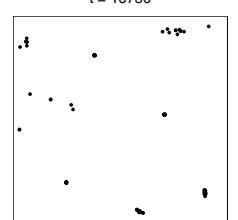
1. Study in detail the role of the accuracy of interactive Markov chains n .
2. Study coupled interactive Markov chains with heterogeneous bounds of confidence in detail. What happens when more than two bounds of confidence are possible?
3. Study nonuniform initial distributions in interactive Markov chains. Does this reduce the phenomena of outliers for the DW model and metastable states for the HK model?
4. How does the danger of drift under heterogeneous bounds of confidence depends on initial conditions and the configuration of the bounds of confidence?

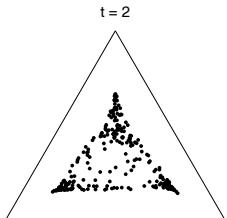




Bibliography

- [1] There is an entry in the English Wikipedia for this mathematical topic. so, one need not expect major changes of the core content.
- [2] Giovanni Allegretti and Carsten Herzberg. Participatory Budgets in Europe Between Efficiency and Growing Local Democracy. *TNI and the Centre for Democratic Policy-Making, TNI Briefing Series*, (2004/5), 2004.
- [3] Frédéric Amblard and Guillaume Deffuant. The Role of Network Topology on Extremism Propagation with the Relative Agreement Opinion Dynamics. *Physica A*, 343:725–738, 2004.
- [4] Patrick Assmann. Monte Carlo Simulation of Deffuant Opinion Dynamics with Quality Differences. *International Journal of Modern Physics C*, 15(10):1439–1447, 2004.
- [5] Eli Ben-Naim, Sidney Redner, and Paul L. Krapivsky. Bifurcation and Patterns in Compromise Processes. *Physica D*, 183:190–204, 2003.
- [6] Marc A. Berger and Yang Wang. Bounded Semigroups of Matrices. *Linear Algebra and its Application*, 166:21–27, 1992.
- [7] Roger L. Berger. A Necessary and Sufficient Condition for Reaching Consensus Using DeGroot’s Method. *Journal of the American Statistical Association*, 76:415–418, 1981.
- [8] Vincent D. Blondel, Julien M. Hendrickx, Alex Olshevsky, and John N. Tsitsiklis. Convergence in Multiagent Coordination, Consensus, and Flocking. In *Proceedings of the 44th IEEE Conference on Decision and Control, Seville, Spain*, 2005.
- [9] Vincent D. Blondel and John N. Tsitsiklis. The Boundedness of All Products of a Pair of Matrices is Undecidable. *Systems & Control Letters*, 41:135–140, 2000.
- [10] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. New York : Oxford University Press, 1999.
- [11] Samprit Chatterjee and Eugene Seneta. Towards Consensus: Some Convergence Theorems on Repeated Averaging. *J. Appl. Prob.*, 14:159–164, 1977.

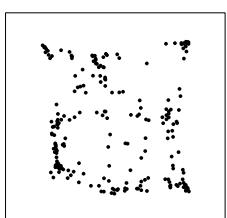


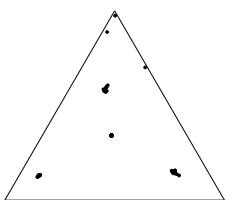


BIBLIOGRAPHY

- [12] Kay-Yut Chen, Leslie R. Fine, and Bernardo A. Huberman. Eliminating Public Knowledge Biases in Information-Aggregation Mechanisms. *Management Science*, 50(7):983–994, 2004.
- [13] Stefan Collignon and Majid Al-Sadoon. A Theory of Stochastic Consensus. *Preprint*, 2006.
- [14] John Conlisk. Interactive Markov Chains. *Journal of Mathematical Sociology*, 4:157–185, 1976.
- [15] Iain D. Couzin, Jens Krause, Richard James, Graeme D. Ruxton, and Nigel R. Franks. Collective Memory and Spatial Sorting in Animal Groups. *Journal of theoretical Biology*, 218:1–11, 2002.
- [16] Ingrid Daubechies and Jeffrey C. Lagarias. Sets of Matrices All Infinite Products of Which Converge. *Linear Algebra and its Application*, 161:227–263, 1992.
- [17] Boaventura de Sousa Santos. Participatory Budgeting in Porto Alegre: Toward a Redistributive Democracy. *Politics & Society*, 26(4):461–510, 1998.
- [18] Guillaume Deffuant. Comparing Extremism Propagation Patterns in Continuous Opinion Models. *Journal of Artificial Societies and Social Simulation*, 9(3), 2006.
- [19] Guillaume Deffuant, David Neau, Frédéric. Amblard, and Gérard Weisbuch. Mixing Beliefs among Interacting Agents. *Advances in Complex Systems*, 3:87–98, 2000.
- [20] Guillaume Deffuant, David Neau, Frédéric. Amblard, and Gérard Weisbuch. How Can Extremism Prevail? A Study Based on the Relative Agreement Interaction Model. *Journal of Artificial Societies and Social Simulation*, 5(4), 2002.
- [21] Morris H. DeGroot. Reaching a Consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.
- [22] Jan Christian Dittmer. *Diskrete Nichtlineare Modelle der Konsensbildung*. Diploma thesis, Universität Bremen, April 2000.
- [23] Jan Christian Dittmer. Consensus Formation Under Bounded Confidence. *Nonlinear Analysis*, 47:4615–4621, 2001.
- [24] Santo Fortunato. The Krause-Hegselmann Consensus Model with Discrete Opinions. *International Journal of Modern Physics C*.
- [25] Santo Fortunato. Universality of the Threshold for Complete Consensus for the Opinion Dynamics of Deffuant et al. *International Journal of Modern Physics C*, 15(9):1301–1307, 2004.
- [26] Santo Fortunato. On the Consensus Threshold for the Opinion Dynamics of KrauseHegselmann. *International Journal of Modern Physics C*, 16(2):259–270, 2005.

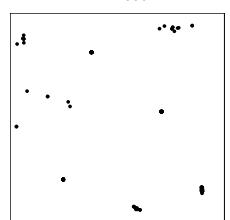
t = 2

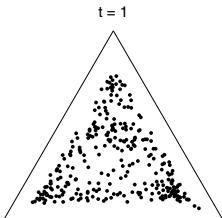




BIBLIOGRAPHY

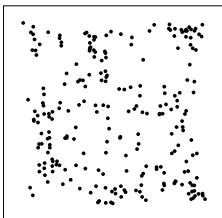
- [27] Santo Fortunato, Vito Latora, Alessandro Pluchino, and Andrea Rapisarda. Vector Opinion Dynamics in a Bounded Confidence Consensus Model. *International Journal of Modern Physics C*, 16(10):1535–1551, 2005.
- [28] Francis Galton. Vox Populi. *Nature*, 75(1949):450–451, 1907.
- [29] Felix R. Gantmacher. *The Theory of Matrices Vol. 1*. AMS Chelsea Publishing, 1959.
- [30] Felix R. Gantmacher. *The Theory of Matrices Vol. 2*. AMS Chelsea Publishing, 1959.
- [31] Daniel T. Gillespie. *Markov Processes - An Introduction for the Physical Scientist*. Academic Press, Inc., 1992.
- [32] Guillaume Grégoire and Hugues Chaté. Onset of Collective and Cohesive Motion. *Physical Review Letters*, 92(2), 2004.
- [33] Darald J. Hartfiel. *Markov Set-Chains*, volume 1695 of *Lecture notes in mathematics*. Springer, 1998.
- [34] Darald J. Hartfiel. *Nonhomogeneous Matrix Products*. World Scientific Publishing Co., 2002.
- [35] R. Hegselmann. *Laws and Models in Science*, chapter Opinion dynamics: Insights by radically simplifying models, page 129. 2004.
- [36] Rainer Hegselmann. Communication in Seminar ‘Effizienz Dezentraler Strukturen’, Universität Bayreuth, 11 2006.
- [37] Rainer Hegselmann and Ulrich Krause. Opinion Dynamics and Bounded Confidence, Models, Analysis and Simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.
- [38] Rainer Hegselmann and Ulrich Krause. Opinion Dynamics Driven by Various Ways of Averaging. *Computational Economics*, 25(4):381–405, 2004.
- [39] Rainer Hegselmann and Ulrich Krause. Truth and cognitive division of labour: First steps towards a computer aided social epistemology. *Journal of Artificial Societies and Social Simulation*, 9(3), 2006.
- [40] Julien M. Hendrickx and Vincent D. Blondel. Convergence of Different Linear and Non-Linear Vicsek Models. *CESAME research report 2005.57*, 2005.
- [41] Harold Hotelling. Stability in Competition. *The Economic Journal*, 39(153):41–57, 1929.
- [42] Dirk Jacobmeier. Multidimensional Consensus Model on a Barabási-Albert Network. *International Journal of Modern Physics C*, 16(4):633–646, 2004.
- [43] Ali Jadbabaie, Jie Lin, and A. Stephen Morse. Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor rules. *IEEE Transactions on Automatic Control*, 48(6), 2003.

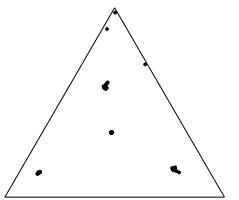




BIBLIOGRAPHY

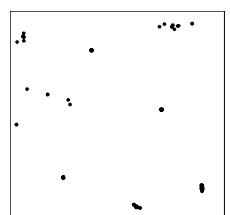
- [44] Wander Jager and Frédéric Amblard. Uniformity, Bipolarization and Pluriformity Captured as Generic Stylized Behavior with an Agent-Based Simulation Model of Attitude Change. *Computational & Mathematical Organization Theory*, 10:295–303, 2004.
- [45] Wander Jager and Frédéric Amblard. Multiple Attitude Dynamics in Large Populations. *Proceedings of The Agent 2005 Conference on: Generative Social Processes, Models, and Mechanisms, Argonne National Laboratory, The University of Chicago*, 2005.
- [46] Ulrich Krause. A Discrete Nonlinear and Non-Autonomous Model of Consensus Formation. In S. Elyadi, G. Ladas, J. Popenda, and J. Rakowski, editors, *Communications in Difference Equations*, pages 227–236. Gordon and Breach Pub., Amsterdam, 2000.
- [47] Ulrich Krause. Positive particle interaction. In A. De Santis L. Benvenuti and L. Farina, editors, *Positive Systems*, volume 294 of *Lecture Notes in Control and Information Sciences*, pages 199–206. Springer-Verlag Berlin Heidelberg, 2004.
- [48] Ulrich Krause. How to Compromise if You Must. *preprint*, 2006.
- [49] Sascha Kurz and Martin Schymalla. Note for the Seminar ‘Effizienz Dezentraler Strukturen’, Universitat Bayreuth, 12 2006. in German <http://www.wm.uni-bayreuth.de/index.php?id=parteienspiel>.
- [50] Keith Lehrer and Carl Wagner. *Rational Consensus in Science and Society*. D. Reidel Publishing Company, Dordrecht, Holland, 1981.
- [51] Harold A. Linstone and Murray Turoff. *The Delphi Method: Techniques and Applications*. freely available on the web, earlier print version from 1975 edition, 2002.
- [52] Jan Lorenz. Mehrdimensionale Meinungsdynamik bei Wechselndem Vertrauen. Master’s thesis, University of Bremen, 2003. Find it at www.janlo.de.
- [53] Jan Lorenz. Continuous Opinion Dynamics: Insights through Interactive Markov Chains;. *Proceedings of IASTED Conference "Modelling, Simulation and Optimization MSO 2005" August 29-31, Oranjestad, Aruba*, 2005.
- [54] Jan Lorenz. A Stabilization Theorem for Dynamics of Continuous Opinions. *Physica A*, 355(1):217–223, 2005.
- [55] Jan Lorenz. Consensus Strikes Back in the Hegselmann-Krause Model of Continuous Opinion Dynamics under Bounded Confidence. *Journal of Artificial Societies and Social Simulation*, 9(1), 2006.
- [56] Jan Lorenz. Convergence of Products of Stochastic Matrices with Positive Diagonals and the Opinion Dynamics Background. In *Positive Systems (Proceedings of the second Multidisciplinary International Symposium on Positive Systems: Theory and Applications (POSTA 06) Grenoble, France, Aug. 30-31, Sept. 1, 2006)*, volume 341 of *Lecture Notes in Control and Information Sciences*, pages 209–216, 2006.

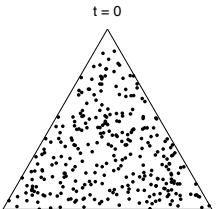




BIBLIOGRAPHY

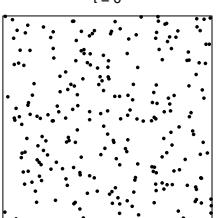
- [57] Jan Lorenz. Continuous Opinion Dynamics of Multidimensional Allocation Problems under Bounded Confidence: More Dimensions Lead to Better Chances for Consensus. *accepted at European Journal of Economic and Social Systems*, 2007.
- [58] Jan Lorenz. *Managing Complexity*, Editor Dirk Helbing, chapter Foster- ing Consensus in Multidimensional Continuous Opinion Dynamics under bounded Confidence. Springer, 2007.
- [59] Jan Lorenz and Frederico Menino B. de Oliveira. Designing Participatory Budgeting: Mathematics of Opinion Dynamics and Aggregation. *Policy Paper of the Brazilian-German Project of the Friedrich-Ebert Foundation*, 2005. Available at www.janlo.de and www.brasilalemanha.org.
- [60] Jan Lorenz and undergraduate students of psychology in the third semester. Meinungsdynamik-Experiment, SS 2006. available on request from post@janlo.de.
- [61] Jan Lorenz and Diemo Urbig. About the Power to Enforce and Prevent Consensus. *accepted at Advances in Complex Systems*, 2007.
- [62] Luc Moreau. Stability of Multiagent Systems with Time-Dependent Communication Links. *IEEE Transactions on Automatic Control*, 50(2), 2005.
- [63] Michael Neumann and Hans Schneider. The Convergence of General Products of Matrices and the Weak Ergodicity of Markov Chains. *Linear Algebra and its Applications*, 287(1-3):307–314, 1999.
- [64] Reza Olafati-Saber and Richard. M. Murray. Flocking with Obstacle Avoidance: Cooperation with Limited Communication in Mobile Networks. *Proc. of the 42nd IEEE Conference on Decision and Control*, 2:20222028, 2003.
- [65] Alessandro Pluchino, Vito Latora, and Andrea Rapisarda. Compromise and Synchronization in Opinion Dynamics. *European Physical Journal B*, 50:169–176, 2006.
- [66] Craig W. Reynolds. Flocks, Herds and Schools: A Distributed Behavioral Model. *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 25 – 34, 1987.
- [67] R. Tyrell Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics. Princeton University Press, New Jersey, 1970.
- [68] Laurent Salzarulo. A Continuous Opinion Dynamics Model Based on the Principle of Meta-Contrast. *Journal of Artificial Societies and Social Simulation*, 9(1), 2006.
- [69] Eugene Seneta. *Non-Negative Matrices*. George Allen & Unwin Ltd., London, 1973.
- [70] Eugene Seneta. *Non-Negative Matrices and Markov Chains. Second Edition*. Springer Verlag, 1981.





BIBLIOGRAPHY

- [71] M. Sherif and C.I. Hovland. *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. Yale University Press, New Haven, CT, 1961.
- [72] Malte Sieveking. Condensing. In *Positive Systems (Proceedings of the second Multidisciplinary International Symposium on Positive Systems: Theory and Applications (POSTA 06) Grenoble, France, Aug. 30-31, Sept. 1, 2006)*, volume 341 of *Lecture Notes in Control and Information Sciences*, pages 247–254, 2006.
- [73] Malte Sieveking. Converging Products. *preprint*, 2006.
- [74] Dietrich Stauffer. Sociophysics Simulations II: Opinion Dynamics. *preprint*, 2005.
- [75] Dietrich Stauffer and Hildegard Meyer-Ortmanns. Simulation of Consensus Model of Deffuant et al on a Barabási-Albert Network. *International Journal of Modern Physics C*, 15(2):241–246, 2004.
- [76] Mervyn Stone. The Opinion Pool. *The Annals of Mathematical Statistics*, 32(4):1339–1342, 1961.
- [77] James Surowiecki. *The Wisdom of Crowds: Why the Many are Smarter than the Few*. Doubleday, New York, 2004.
- [78] Jacques Theys. *Joint Spectral Radius: Theory and Approximation*. PhD thesis, Université Catholique de Louvain, 2005.
- [79] John N. Tsitsiklis. *Problems in Decentralized Decision Making and Computation*. PhD thesis, Massachusetts Institute of Technology, 1984.
- [80] John C. Turner, Michael A. Hogg, Penelope J. Oakes, Stephen D. Reicher, and Margaret S. Wetherell, editors. *Rediscovering the Social Group : A Self-Categorization Theory*. Blackwell, Oxford, 1987.
- [81] Diemo Urbig. Attitude Dynamics with Limited Verbalisation Capabilities. *Journal of Artificial Societies and Social Simulation*, 6(1), 2003.
- [82] Diemo Urbig and Jan Lorenz. Communication Regimes in Opinion Dynamics: Changing the Number of Communicating Agents. In *Proceedings of the Second Conference of the European Social Simulation Association (ESSA)*, September 2004. download at www.janlo.de or www.diemo.de.
- [83] Diemo Urbig and Robin Malitz. Dynamics of Structured Attitudes and Opinions. *Proceedings of the Third Conference of the European Social Simulation Association (ESSA)*, September 5-9, Koblenz, German, 2005.
- [84] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofers Shochet. Novel Type of Phase Transition in a System of Self-Driven Particles. *Physical Review Letters*, 75(6):1226–1229, 1995.
- [85] Bauke Visser and Otto H. Swank. On Committees of Experts. *Quarterly Journal of Economics*, 122(1):337–372, 2007.



BIBLIOGRAPHY

- [86] G. Weisbuch. Bounded Confidence and Social Networks. *European Physical Journal B*, 38:339–343, 2004.
- [87] Gérard Weisbuch, Guillaume Deffuant, and Frédéric Amblard. Persuasion Dynamics. *Physica A*, 353:555–575, 2005.
- [88] Gérard Weisbuch, Guillaume Deffuant, Frédéric Amblard, and Jean-Pierre Nadal. Meet, Discuss and Segregate! *Complexity*, 7(3):55–63, 2002.
- [89] Jacob Wolfowitz. Products of Indecomposable, Aperiodic, Stochastic Matrices. *Proceedings of the American Mathematical Society*, 15:733–737, 1963.

Index

- I -barycenter, 40
 I -mass, 40
•, 12
 a -mean, 24
 d -dimensional interval, 11
 f -mean, 24

added limit profiles, 114
adjacency matrix, 12
affine hull, 11
agent, 19
agent dimension, 20
agent-based, 8, 19
agent-based process, 21
alignment, 50
appropriate opinion space, 20
area of confidence, 22
arithmetic mean, 24
attraction, 50
averaging map, 25

backward accumulation, 12
bound of confidence, 22
bounded confidence, 7, 22
bounded confidence matrix, 27
bounded confidence network, 27
bounded confidence process, 29
budget constraint, 133
budget constraints, 20

cautiousness parameter, 30
class of indices, 66
cluster, 112
cluster in opinion distribution, 113
cluster location, 112, 113
cluster mass, 112, 113
cluster with precision δ , 113
coefficient of ergodicity, 79
coefficient of ergodicity for row-stochastic matrices, 78
communicate, 66

communication regime, 27
compensating, 22
componentwise convex coefficient vectors, 25
componentwise convex combination, 25
confidence matrix, 26
confidence set, 23
connected, 27
consensus, 20
consensus cluster, 112
consensus matrix, 27
continuous, 57
continuous opinion dynamics, 7
contracted, 113
convex, 11
convex coefficients, 11
convex combination, 11
convex hull, 11
convex hull averaging map, 25
cube, 25
cube averaging map, 25
cutoff function, 62

Delphi effect, 15
density based, 19
discrete bound of confidence, 23, 26
discrete dynamical system, 21
discrete interval, 11
DW, 29
dynamic, 21

equicontinuous, 57
equiproper, 58
ergodic, 75
essential, 66
essential cluster, 112
extended phase diagram of the consensus transitions, 127

first I -moment, 40
fixed point of the DW model, 100

- fixed point of the HK model, 100
fixed point of the interactive Markov chain, 102
forward accumulation, 12

Gantmacher diagonal blocks, 67
Gantmacher form, 67
Gantmacher subdiagonal blocks, 67
Gantmacher's canonical form of a non-negative matrix, 67
general convex hull mean, 25
general cube mean, 25
general mean, 23–25
generalized spectral radius, 93
geometric mean, 24
gossip process, 30

Hadamard product, 12
harmonic mean, 24
Hausdorff metric, 57
heterogeneous gossip process, 30
heterogeneous repeated meeting process, 29
HK, 29
homogeneous bound of confidence, 22
homogeneous gossip process, 30
homogeneous repeated meeting process, 29

incidence matrix, 65
incidence sense, 65
independent of time, 21
inessential, 66
inessential cluster, 112
interactive Markov chain, 22
intercommunication intervals, 92
irreducible, 66
isolated cluster, 112
italics, 9

joint spectral radius, 93
Jordan form with respect to the Gantmacher form, 73

length of a path, 65
length of an accumulation, 84
level of precision, 113
limit opinion distribution, 112
limit opinion profile, 111
linear, 21

mass of the biggest cluster, 127
mathematical model, 13
meeting, 30
metastable, 33
model, 13
Monte-Carlo analysis, 14

neighbors, 12
network, 26
network cluster, 112
non-compensating, 22

opinion, 20
opinion classes, 20
opinion dimension, 20
opinion distribution, 20
opinion issues, 20
opinion pool, 28
opinion profile, 20
opinion spaces, 20

partial abstract means, 24
path, 65
period, 66
permutation matrix, 12
Perron eigenvalue, 72
Perron eigenvector, 72
positive minimum, 82
power mean, 24
primitive, 66
process of opinion pooling, 28
proper, 26

reducible, 66
regular, 75
relative interior, 12
relatively compact, 11
repeated averaging, 7
repeated meeting process, 29
repulsion, 50
reverse bifurcation diagrams, 116
row allowable, 12
row-stochastic, 12
row-stochastified, 12

sandwich inequality, 23
scrambling, 80
self-communicating, 66
self-communicating classes, 66
shrinking lemma, 81
simplex, 11

INDEX

solution, 21
spectral radius, 12
state, 19
state space, 19
state-dependent, 21
stochastic, 11
strong ergodicity, 78
sub-accumulations, 84
submultiplicative, 79
subspace contraction coefficient, 80
supermultiplicative, 91
support, 26

time-dependent, 21
trajectory, 21
transition matrix, 22
type, 65
type-symmetric, 65

uniformly continuous, 57
uniformly equicontinuous, 57
unit simplex, 11

weakly ergodic, 78
weighted arithmetic mean, 24
weighted geometric mean, 24