Moderates and consensus formation in the Deffuant model

Ruan dos Santos Vieira Miranda

Thesis presented to the
Institute of Mathematics and Statistics
of the University of São Paulo
in partial fulfillment
of the requirements
for the degree of
Master of Science

Program: Probability and Statistics
Advisor: Prof. Dr. Fábio Machado

During the development of this work, the author received
financial support from CAPES – Finance code 001

São Paulo
November, 2024

Moderates and consensus formation in the Deffuant model

Ruan dos Santos Vieira Miranda

This is the original version of the thesis prepared by candidate Ruan dos Santos Vieira Miranda, as submitted to the Examining Committee.

Ficha catalográfica elaborada com dados inseridos pelo(a) autor(a)
Biblioteca Carlos Benjamin de Lyra
Instituto de Matemática e Estatística
Universidade de São Paulo

Miranda, Ruan dos Santos Vieira
Moderates and consensus formation in the Deffuant model / Ruan dos Santos Vieira Miranda; orientador, Fábio Prates Machado. - São Paulo, 2024.
79 p.: il.

Dissertação (Mestrado) - Programa de Pós-Graduação em Estatística / Instituto de Matemática e Estatística / Universidade de São Paulo.
Bibliografia
Versão original

1. Probabilidade. 2. Estatística. I. Prates Machado, Fábio. II. Título.

Bibliotecárias do Serviço de Informação e Biblioteca
Carlos Benjamin de Lyra do IME-USP, responsáveis pela
estrutura de catalogação da publicação de acordo com a AACR2:
Maria Lúcia Ribeiro CRB-8/2766; Stela do Nascimento Madruga CRB 8/7534.

Agradecimentos

A Deus, pela vida, por Seu Amor, Bondade e Misericórdia Infinitos e por todas as bênçãos e oportunidades que me são concedidas. A Nossa Senhora Maria Cristo, a Mãe de todos e todas nós, a Quem Jesus legou a humanidade no alto do Calvário[1], por sua Presença em nossas vidas, sua Proteção e Amor!

A Benjamin Teixeira de Aguiar (Mamin), amado Orientador Espiritual e Representante encarnado dos Cristos de Deus para a Terra (vide os Endossos Divinos[2]), por sua presença em nossas vidas, por nos auxiliar para que sigamos a nossa consciência, a nossa vocação e sejamos, verdadeiramente, felizes. Por ser Luz, em meio a tanta escuridão, e canalizar a Espiritualidade Sublime para as nossas vidas; e pela defesa da Feminilidade Sagrada, da Mãe Crística da humanidade, Maria Cristo!

Aos Guias Espirituais, Nossa Mãe e Mestra Espiritual Eugênia-Aspásia e Nosso Pai e Mestre Espiritual Matheus-Anacleto, que, graças à mediunidade-bênção de Mamin, esclarecem-nos e nos auxiliam para que amadureçamos e sejamos melhores servidores(as) do Bem.

Aos irmãos e irmãs em Ideal, minha família do Espírito, pelo acolhimento, fraternidade e amizade. Especialmente, Gabriel Brandão, Luidje Barboza e Vitor Zucchi pelas conversas, reflexões e resenhas; Ayslaine Cunha (Nane), pelo companheirismo e carinho; Isadora Guilherme (Isa), Jorge Fontes (Jorginho) e Luiza Prado (Lu), pelos rolês em São Paulo e acolhimento; Carlos Borbolla (Cadu) e Taís Fiscina de Oliveira (Taisinha), pelos esclarecimentos e auxílio nas sessões terapêuticas.

A toda minha família, especialmente a meu pai Raimundo Vieira, minha mãe Josy dos Santos; meu irmão Ramon Vieira e minhas irmãs Cileide dos Santos e Jaqueline dos Santos (*in memoriam*); minhas tias Fátima Vieira e Estela Vieira; meus tios Cícero Vieira, Ari Vieira e Carlos Vieira por todo suporte e acolhimento, sobretudo nesses dois anos em São Paulo.

[1] João, 19: 26-27.

[2] https://www.mariacristo.com.br/endossos/

Aos amigos e amigas da *Casa Maxixe*, onde morei por quase todo o período do mestrado, pela amizade e carinho, por todas as resenhas e por terem me apresentado a cultura paulista, paraense, capixaba e colombiana: Bruna Oliveira, Jonathan Castro, Lui Grenner, Manuel Espítia, Marcelo dos Santos e Valter Turri (Valtinho). Como também, ao nosso locador e amigo, que tornou tudo isso possível, Milton Francisconi.

Aos amigos e amigas do *Lab 250*, pela receptividade, amizade, resenhas, estudos e partidas de futebol (*Pelab250*). Nomeadamente, Alex Nhacololo, Andrey Sarmento, Anny Rodrigues, Bruno de Assis, Danilo Silva, Davi Chaves, Eduardo Janotti, Enrico Weinstock, Felipe Toledo, Giovanni Pastori, Heitor de Assis, Hyeon Jeong (Gina), Leonardo Larrubia, Lucas Prates, Luigi Pavarini, Luís Henrique, Marília Sombra, Patrick Nasser, Pedro Reis, Rodolpho Jordan, Tiago Marques e Vinicius Soares.

Ao casal amigo de Salvador, que pude reencontrar na USP, por todas as resenhas e ajuda, Afonso Sant'Anna e Helena Beatriz.

A Fábio Prates, meu orientador acadêmico, pelo auxílio, aulas, paciência e todos ensinamentos durante o período do mestrado. Agradeço também por ter aprendido a regra *KISS*, *Keep It Super Simple* (adaptada), que me auxiliou, sobretudo no período da escrita da dissertação (risos).

Aos membros da banca, Dirk Erhard e Tertuliano Franco, pela paciência, revisão e sugestões de melhoria da dissertação, como também por toda ajuda durante os anos de minha graduação, na UFBA.

Aos(às) professores(as) da USP, por todo aprendizado e auxílio: Aline Duarte, Antonio Galves (*in memoriam*), Denise Botter, Elisabeti Kira (pela orientação com as monitorias e pelo exemplo no ensino), Florencia Leonardi, Luis Esteves, Miguel Abadi, Vera Giusti e Victor Fossaluza.

A Regiane Guimarães, pela solicitude, paciência e ajuda, sobretudo com as formalidades do mestrado; e a Josineide Pereira e Walquíria Fernandes, pelos cafés e risadas na copa do bloco A, no IME-USP.

À Universidade de São Paulo (USP), à Comissão da Pós-Graduação (CPG), pelo acolhimento, auxílio e suporte financeiro, sobretudo nos eventos acadêmicos; e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo financiamento que tornou possível a concretização deste mestrado.

A todos e todas que de alguma forma me auxiliaram, porque sei que nada construímos sozinhos(as) e que fazemos parte de uma rede de interdependência. Muito obrigado!

Salvador, Bahia, 15 de novembro de 2024

Resumo

Ruan dos Santos Vieira Miranda. **Moderados e formação de consenso no modelo Deffuant**. Dissertação (Mestrado). Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2024.

Entender as condições para a formação de consenso é um tópico fundamental nas ciências sociais, comportamentais, econômicas e em outras áreas. A dinâmica da opinião, particularmente no sentido matemático, fornece elucidações valiosas sobre esse fenômeno. Nesse sentido, modelo de Deffuant descreve a formação de consenso em que pares de agentes interagem somente se suas opiniões diferem em não mais do que um determinado limite θ, levando a uma aproximação de suas crenças. Tradicionalmente, esse modelo é definido em um grafo $G = (\mathbb{Z}, E)$, com $E = \{(x, x+1) : x \in \mathbb{Z}\}$ e opiniões iniciais uniformemente distribuídas no intervalo unitário. O valor crítico de θ para consenso nessa configuração é $1/2$. Em nossa pesquisa, estendemos o modelo para $\mathbb{Z}^2$ e consideramos opiniões em uma e duas dimensões. Além disso, exploramos várias distribuições iniciais para avaliar o papel dos moderados — agentes com visões dentro do espectro de opinião — além dos extremistas, na obtenção de consenso. Nosso objetivo é determinar a quantidade crítica de moderados necessária para a formação de consenso. Evidências numéricas sugerem que, em populações com alta tolerância à interação, mesmo um pequeno número de moderados pode impactar significativamente a velocidade de obtenção de consenso.

Palavras-chave: Dinâmica de opinião. Extremistas. Formação de consenso. Modelo de Deffuant. Moderados.

Abstract

Ruan dos Santos Vieira Miranda. **Moderates and consensus formation in the Deffuant model**. Thesis (Master's). Institute of Mathematics and Statistics, University of São Paulo, São Paulo, 2024.

Understanding the conditions for consensus formation is a pivotal topic in social, behavioral, economic, and other sciences. Opinion dynamics, particularly in the mathematical sense, provides valuable insights into this phenomenon. The Deffuant model describes consensus formation where pairs of agents interact only if their opinions differ by no more than a given threshold, leading to an approximation of their beliefs. Traditionally, this model is defined on a graph $G = (\mathbb{Z}, E)$ with $E = \{(x, x+1) : x \in \mathbb{Z}\}$ and initial opinions uniformly distributed on the unit interval. The critical threshold value for consensus in this setup is $1/2$. In our research, we extend the model to $\mathbb{Z}^2$ and consider opinions in one and two dimensions. We explore various initial distributions to assess the role of moderates — agents with views inside the opinion spectrum —, in addition to the extremists, in achieving consensus. We aim to determine the critical quantity of moderates necessary for consensus formation. Numerical evidence suggests that in populations with a high tolerance to interact, even a small number of moderates can significantly impact the speed of achieving consensus.

Keywords: Consensus formation. Deffuant model. Extremists. Moderates. Opinion dynamics.

List of abbreviations

IPS	Interacting Particle Systems
IME	Institute of Mathematics and Statistics
USP	University of São Paulo
SAD	Sharing a Drink

List of symbols

$\{\eta_t(y)\}_{y\in\mathbb{Z}}$	Opinion profile at time t on the Deffuant model
$\{\gamma_t(y)\}_{y\in\mathbb{Z}}$	Opinion profile at time t on the multivariate Deffuant model
$\{\xi_t(y)\}_{y\in\mathbb{Z}}$	Opinion profile at time t on the SAD model
$c_{i\to j}$	Rate of change from state i to state j on IPS models
o_t	Configuration of an IPS model at time t
$G(V,E)$	Graph G represented with its set of vertices V and set of edges E
μ	Convergence parameter on the Deffuant model
θ	Tolerance parameter on the Deffuant model
τ	Tolerance parameter on the multivariate Deffuant model
S	State space for IPS models

List of Figures

List of Tables

Contents

Introduction

Understanding crowd behavior has been a great challenge throughout history. Social, behavioral, and theoretical sciences have long explored the influence of individuals on collective behaviors. Interacting Particle Systems (IPS) have emerged as a pivotal framework for modeling and exploring interactions within societies, specially in the mathematical sense.

This thesis focuses on the opinion dynamics, specifically the Deffuant model. Despite the extensive research on this topic, the role of moderates — individuals whose opinions lie within the opinion space — in influencing consensus formation still has open questions. This study investigates how moderates impact opinion convergence in the Deffuant model on various structures, including the complete graph, as well as the integer and square lattices.

The main objective of this study is to analyze the influence of moderates on consensus dynamics in bounded confidence models. By running the Deffuant model on different social structures, we seek to understand the role of moderates in consensus formation. Our main goal is to answer the following question: is there a critical number of moderates that set the population progressively towards consensus?

This work is divided into three chapters. The first chapter provides an introduction to Interacting Particle Systems (IPS), introducing a brief historical overview and two models: the Ising model for magnetism, and the Voter Model for opinion dynamics. The second chapter focuses on theoretical results, by presenting and demonstrating the pivotal results on convergence of the traditional Deffuant model. Furthermore, it presents varied results from recent papers on the topic. Finally, the third chapter explores the role of moderates through simulations, applying the model to different social networks. In addition, it tests the validity of the results from Chapter Two, by analyzing the critical quantity of moderates in the population to achieve a consensus.

Chapter 1

Interacting Particle Systems

Formalized in the 1970s through the independent works of Frank Spitzer and R. L. Dobrušin, Interacting Particle Systems (IPS) have become a significant area of mathematics, particularly in the study of consensus formation. In this chapter, we briefly explore the history and applications of IPS, presenting two introductory examples along with key results.

1.1 Opinion dynamics models

Interacting Particle Systems (IPS) are mathematical models used to describe the behavior of a collection of particles (or agents) on a spatial configuration over time. The evolution of the system is governed by stochastic rules related to the theory of Markov processes.

The first formalization of this field traces back to the 1970s, with the independent works of Frank Spitzer and R. L. Dobrušin. The former, in the United States, with the paper titled "Interaction of Markov Processes" published in the Advances in Mathematics journal.(Spitzer, 1970) The second, in the Soviet Union, with "Markov processes with a large number of locally interacting components: Existence of a limit process and its ergodicity". (Dobrušin, 1971) The original motivation for IPS arose from statistical mechanics, particularly models related to physical systems.

Consider an undirected graph $G = (V, E)$, which consists of a set of vertices V representing agents or particles, and a set of edges E, representing connections between them. At any given time, each particle has a state, which could be represented by a number, letter or color from a set S. In general, the configuration of the system at time t is described by a function:

$$o_t : V = \mathbb{Z}^d \to S \text{ where } o_t(x) = \text{state of vertex } x \text{ at time } t. \tag{1.1}$$

The dynamics of the system evolve as a continuous-time Markov chain, meaning that the future evolution depends only on the current state, not on the past of the system. Typically, the dynamics are driven by a Poisson clock assigned to each edge $e \in E$. Interactions between particles (or agents) can occur only if there is an edge connecting them. The neighborhood of a vertex x, denoted N_x, is defined as:

$$N_x = \{y \in V : (x, y) \in E\}, \quad \text{for all } x \in V. \tag{1.2}$$

The transition rate at witch vertex x changes from state $i \in S$ to state $j \in S$, denoted by $c_{i\to j}$, only depends on the configuration of the neighborhood of x at time t. Mathematically, the transition rates are defined by:

$$c_{i\to j}(x, o_t) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\mathbb{P}(o_{t+\varepsilon}(x) = j | o_t(x) = i, o_t) = c_{i\to j}(x, o_t(y) : y \in N_x) \tag{1.3}$$

This local dependence with the Markov property implies that the dynamics are governed by local transition rates in both space and time. Commonly, (1.3) are finite.

For infinite graphs, defining the evolution of the process becomes more complex due to the non-existence of the "next" time step. However, in this paper, we focus on finite or locally finite graphs (with bounded degree) and finite sets of states and transition rates, where the process is well-defined. For further details, see Liggett, 1985.

To better understand IPS, let us define the Ising model, introduced by the physicist Wilhelm Lenz, (Lenz, 1920) who gave it as a problem to his student Ernst Ising. This is a famous model for magnetism. Consider a Markov process with state space $\{+1, -1\}^{\mathbb{Z}^d}$, where $+1$ and -1 can be interpreted as the orientation (spin) of the magnets dispersed over $\mathbb{Z}^d$. For each site $x \in \mathbb{Z}^d$ there is a configuration $o(x) \in \{+1, -1\}$ such that its energy is given by the Hamiltonian equation:

$$H(o) = -\sum_{x\sim y} J_{xy}o(x)o(y) - \mu \sum_{y} h(y)o(y), \tag{1.4}$$

where the first sum is over adjacent sites, J_{xy} is a coupling constant (often taken as 1 for neighbors and 0 otherwise), μ the magnetic moment, and $h(y)$ the external magnetic field (usually, for simplification, taken as zero).

The probability of a given configuration o is given by the Boltzmann distribution:

$$\mathbb{P}_\beta(o) = \frac{1}{Z(\beta)}e^{-\beta H(o)}, \tag{1.5}$$

where $\beta \geq 0$ represents the reciprocal of the temperature of the system and $Z(\beta)$ is the normalization function given by:

$$Z(\beta) := \sum_{o\in\Omega} \exp\{-\beta H(o)\}, \tag{1.6}$$

where Ω is the set of all possible configurations.

From 1.4, it is clear that the energy is higher when the spin of site x is different from most of its neighbors, which makes the rate of change (1.5) higher. In other words, the system "favors" configurations in which the spins are aligned with one another. In the physical context, this is called ferromagnetism.

At $\beta = 0$, hypothetical case of infinity temperature, H has no effect in the model, and the spins $o(x)$ are independent two-state Markov chains. In this case, the system has a unique invariant measure. As $\beta > 0$ increases (representing lower temperatures), the system tends towards low energy configurations, and interesting behaviors arise, particularly in higher dimensions.

Since the formalization of IPS, many other models arose. This study will primarily focus on opinion dynamics models. This refer to IPS models that represent the spread of opinions or preferences among agents. The state space can be discrete, such as $\{-1, +1\}$, which represents two opposite opinions, or continuous, taking values on $[0, 1]$. We will be interested, specifically, on models with bounded confidence, where agents can only interact with neighbors whose opinions lie within a certain spectrum of compatibility or proximity. For a wide review on dynamic models, see the joint work of Claudio Castellano, Santo Fortunato, and Vittorio Loreto "Statistical physics of social dynamics".(Castellano *et al.*, 2009)

1.1.1 The Voter Model

Independently introduced by Clifford and Sudbury, 1973 and Holley and Liggett, 1975, the voter model is a type of spin system characterized by transition rates $c(x, o)$ given by:

$$c(x, o) = \begin{cases} \sum_y p(x, y)o(y) & \text{if } o(x) = 0, \\ \sum_y p(x, y)[1 - o(y)] & \text{if } o(x) = 1, \end{cases} \tag{1.7}$$

where $p(x, y) \geq 0$ for all $x, y \in S$ represents the transition probabilities of an irreducible Markov chain, and

$$\sum_y p(x, y) = 1 \text{ for all } x \in S. \tag{1.8}$$

Here, S denotes the set of sites in the model.

The voter model simulates individuals on a lattice (or graph) where each site x is occupied by an opinion $o(x)$ which can be either 0 or 1. According to a Poisson process with rate 1, each individual re-evaluates their opinion by choosing one of its neighbors with probability $p(x, y)$ and mimicking their opinion. This process can also be viewed as a branching process where the change in opinion represents the replacement of offspring.

In the specific case of the lattice $\mathbb{Z}^d$, where each site interacts with its $2d$ nearest neighbors, the transition probabilities are $p(x, y) = \frac{1}{2d}$ for $y \in N_x$, where N_x denotes the set of neighbors of x. The transition rates can then be written as:

$$c_{0\to1}(x, o_t) = \frac{1}{2d} \sum_{y \in N_x} o_t(y) \text{ and } c_{1\to0}(x, o_t) = \frac{1}{2d} \sum_{y \in N_x} (1 - o_t(y)). \tag{1.9}$$

Here, o_t denotes the opinion configuration at time t. The voter model is a classical opinion dynamics model used to study how opinions spread over a network. Key questions include whether the model converges to a consensus state and under what conditions this occurs. According to the transition rates given in equation (1.9), it is evident that the configurations where all sites have the same opinion, i.e., $o \equiv 0$ or $o \equiv 1$, are invariant measures of the

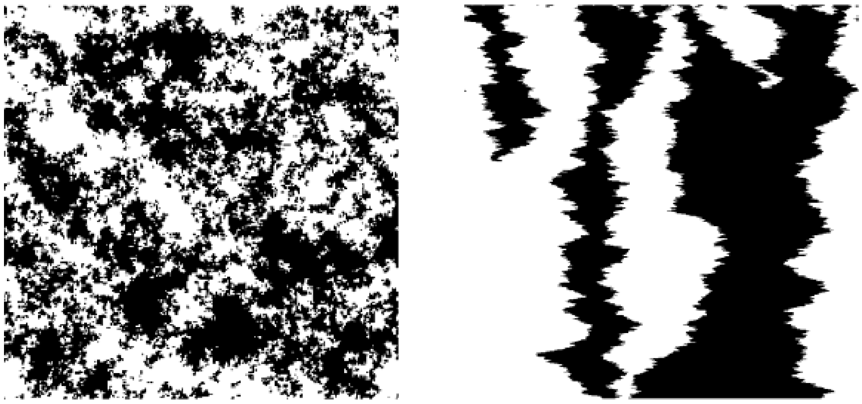

Figure 1.1: *Snapshot at time* 1000 *of the voter model on a* 300×300 *lattice with periodic boundary conditions (left) and realization of the voter model from time* 0 *to time* 10,000 *on the one-dimensional torus with* 600 *vertices.* ***Source****: Lanchier, 2017.*

model.

An interesting tool for exploring the dynamics of the voter model is the Harris graphical representation, commonly used for analyzing similar stochastic models such as the contact process. Harris, 1974 This graphical approach allows for a simultaneous view of both time and space structures.

In the Harris graphical representation, each site in the lattice is associated with a Poisson process with rate 1. Time evolves vertically in this representation. The arrows $\leftarrow$ and $\rightarrow$ are used to represent the updating mechanism of the voter model, indicating changes in opinion. Since there are only two possible opinions (0 and 1), each vertical line in the representation — whether solid or dashed — corresponds to one of these opinions. This allows for a clear visualization of how opinions evolve over time and space. To better understand what is the dual of the process, let us present some basic definitons.

Definition 1.1.1 (Lanchier, 2017)**.** *There is a path* $(z, s) \rightarrow (x, t)$ *if there exist*

$$s = s_0 < s_1 < \cdots < s_{n+1} = t \quad \textit{and} \quad z = z_0, x_1, x_2, \ldots, x_n = x \in \mathbb{Z}^d$$

such that the following two conditions hold:

- *for* $i = 1, 2, \ldots, n$*, there is an arrow* $x_{i-1} \rightarrow x_i$ *at time* s_i*; and*
- *for* $i = 0, 1, \ldots, n$*, there is no arrow pointing at* $\{x_i\} \times (s_i, s_{i+1}]$.

If it holds, there is also a dual path $(x, t) \rightarrow (z, s)$, which represents the process backward in time, following the arrows of the graphical representation in the opposite way. Below, Figure 1.2 illustrates the Harris graphical representation of the voter model and its dual. For each $A \subset \mathbb{Z}^d$ the *dual process* starting at (A, t) is then defined as:

$$\hat{o}_s(x, t) = \left\{ z \in \mathbb{Z}^d \ : \ \text{there is a dual path } (x, t) \rightarrow (z, t - s) \text{ for some } x \in A \right\} \tag{1.10}$$

for all $o \leq s \leq t$. In the voter model, dual processes follow the ancestor. In particular, when $A = \{x\}$, we have the dual relationship.:

$$o_t(x) = o_{t-s}(\hat{o}_s(x, t)) \text{ for all } 0 \leq s \leq t. \tag{1.11}$$

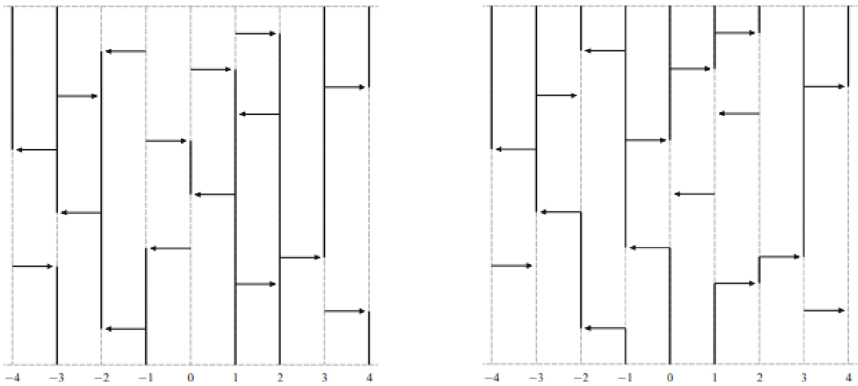

Figure 1.2: *Graphical representation of the voter model and its dual process.* ***Source***: *LANCHIER, 2017.*

The arrows in the graphical representation follow a Poisson process with intensity one and can originate from any of the $2d$ neighbors with equal likelihood. As a result, the dual process evolves based on a continuous-time symmetric random walk at a rate of one. Additionally, as shown in Figure 1.2, the dual processes coalesce when they meet, making them coalescing random walks. From this, many properties from random walk theory arise.

Clustering vs coexistence

Assume that the model starts with a product measure $\rho \in (0, 1)$. Then,
Theorem 1.1.1 (HOLLEY and LIGGETT, 1975). *Starting with a density ρ of type 1 individuals,*

- ***Consensus:*** $o_t \xrightarrow{d} (1-\rho)\delta_0 + \rho\delta_1$ *when $d \leq 2$, where δ_i is the measure that concentrates on the configuration in which all sites are of type i.*
- ***Coexistence:*** $o_t \xrightarrow{d} o_\infty$ *when $d > 2$, where $P(o_\infty(x) = 1) = \rho$.*

In other words, at a large t, with probability closer to one, in dimensions less than or equal to 2, any region is either occupied by only type 1 individuals with probability ρ; or occupied by type 0 individuals, with probability $1 - \rho$. Therefore, consensus is achieved. Conversely, for higher dimensions, coexistence occurs. For further results, see LANCHIER, 2017.

An extension of the voter model was presented in 2003 by F. Vazquez, P. L. Krapivsky, and S. Redner (VAZQUEZ *et al.*, 2003). Their model introduces three types of opinion profiles: leftist, centrist, and rightist. A key aspect of this model is that leftists and rightists, being highly incompatible, do not interact directly, whereas centrists can interact with both. The authors concluded that the barrier between extremists significantly impacts the dynamics, particularly by slowing the convergence process.

This is one type of bounded confidence opinion model, which is the main focus of this paper. In the next chapter, we will explore this concept with another example: the Deffuant model.

Chapter 2

Deffuant Model

Originally presented in 2000 by Guillaume Deffuant, David Neau, Frederic Amblard, and Gerard Weisbuch (Deffuant *et al.*, 2000), the Deffuant Model features two key parameters: $\theta > 0$ and $\mu \in (0, 1/2]$. The first parameter, known as the confidence threshold, determines whether two agents will interact based on the difference in their opinions. The second parameter, susceptibility to change, dictates how much agents influence each other during interactions. Besides these, another distinctive characteristic of the Deffuant Model is its use of non-binary opinions, where agents have continuous opinions in the range $[0, 1]$. This contrasts with previous models that typically used binary opinions. In this chapter, we describe the model across different spatial and opinion configurations and present the most significant results related to consensus formation.

2.1 Model Description

Let $G = (V, E)$ be a connected undirected graph with no loops, either finite or infinite, with a bounded degree. The Deffuant Model (Deffuant *et al.*, 2000) is a stochastic model for opinion dynamics, where interactions occur pairwise. The model features two important parameters: $\theta > 0$, the confidence threshold, and $\mu \in (0, 1/2]$, which represents the susceptibility to compromise.

The model dynamics are as follows: at time $t = 0$ each agent $u \in V$ receives a random opinion value, independent of all others, uniformly distributed in $[0, 1]$. Each edge $e \in E$ is assigned an independent unit rate Poisson process, which determines the times of interactions.

Let $\eta_t(u)$ be the opinion of agent u at time t. When a Poisson clock rings at an epoch t at edge $e = \langle u, v \rangle$, such that $\eta_{t^-} := \lim_{s \to t} \eta_s(u) = a$ and $\eta_{t^-} := \lim_{s \to t} \eta_s(v) = b$, the opinion update occurs as follows:

$$\eta_t(u) = \begin{cases} a + \mu(b - a) \text{ if } |a - b| \leq \theta, \\ a \text{ otherwise.} \end{cases} \quad \text{and} \quad \eta_t(v) = \begin{cases} b + \mu(a - b) \text{ if } |a - b| \leq \theta, \\ b \text{ otherwise.} \end{cases} \tag{2.1}$$

The process is well-defined for a finite set of edges E, since there will almost surely be no two Poisson events occurring simultaneously. For the case of an infinite graph with bounded degree, the techniques to prove this are well-known in the general theory of Interacting Particle Systems and can be found in (Liggett, 1985).

The model describes a pattern of social interactions where each agent's exposure is limited by the confidence threshold θ, meaning interactions occur only within a certain opinion spectrum. This reflects people's tendency to listen to similar opinions (homophily). In addition, the parameter μ dictates the agents' willingness to compromise.

In the literature, many variations of the Deffuant model are explored, for new properties arise with the change of its features. The most popular variations are presented following (see Bernardo, 2016 for a review on the subject in Portuguese):

- Discrete or continuous time;
- Unidimensional or multidimensional opinion space;
- Finite or infinite number of agents;
- Agent or density-based model;
- Homogeneous or non-homogeneous model (see (H.-L. Li, 2021) for a study with parameters varying over time for example).

One of the most common questions about social dynamics models is the circumstances under which consensus is achieved. In the following sections, we will delve into the traditional Deffuant model. It assumes $G = (\mathbb{Z}, E)$, where $E = \{< v, v+1 >:\ v \in \mathbb{Z}\}$ with initial opinions independent and identically uniformly distributed in $[0, 1]$. We will present the critical value for θ that shows the population behavior change towards agreement. Additionally, we will explore an extension to the square lattice, where opinions will assume bidimensional values.

2.2 Consensus formation on $\mathbb{Z}$

When first introduced by its authors, a critical value for the confidence parameter in the Deffuant Model was conjectured. Based on numerical evidence, it was found that when the confidence parameter θ_c was equal to $1/2$, there was a change in the global behavior, regardless of the value of the convergence parameter μ. Above this value, consensus was established, while below it, coexistence occurred. Figures 2.1 and 2.2 show the time chart of opinions. One time unit corresponds to sampling 1000 pairs of agents (N).

Only in 2011 (published the year after) a mathematical formulation was presented and proven by Lanchier (Lanchier, 2012). First, to rigorously present the result, we introduce the notion of compatibility. The Deffuant process on $\mathbb{Z}$ with random initial opinions identically and independently distributed uniformly on $[0, 1]$ achieve consensus if all pairs of neighbors are compatible, i.e.:

$$\lim_{t\to\infty} \mathbb{P}(|\eta_t(x) - \eta_t(x+1)| \leq \theta) = 1 \text{ for all } x \in \mathbb{Z}. \tag{2.2}$$

In other words, as time goes to infinity, the probability that agents will maintain an open

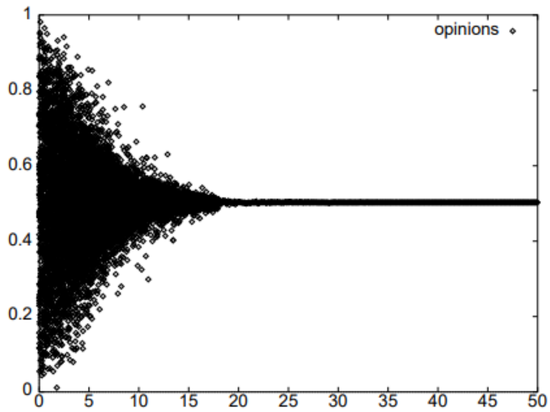

Figure 2.1: *Time chart of opinions on the Deffuant model with parameters* $\mu = 0.5, \theta = 0.5, N = 2000$. ***Source***: *Deffuant et al., 2000.*

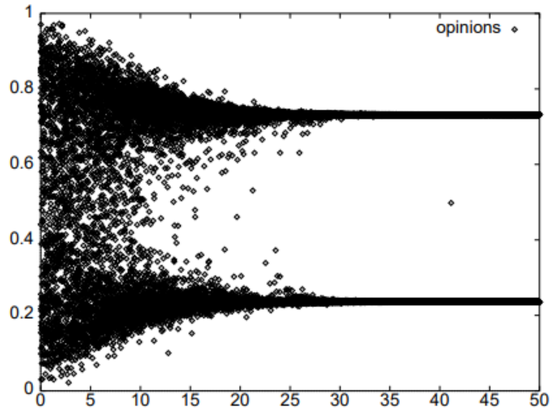

Figure 2.2: *Time chart of opinions on the Deffuant model with parameters* $\mu = 0.5, \theta = 0.2, N = 1000$. ***Source***: *Deffuant et al., 2000.*

line of communication (interactions) tends to 1. Now, let

$$\theta_c := \inf\{\theta \in [0,1] : \text{ consensus (2.2) holds}\},$$

using non-traditional geometric techniques, the following theorem was proved by Lanchier:
Theorem 2.2.1 (Lanchier, 2012). *Consider the Deffuant model on the graph* $(\mathbb{Z}, E)$ *where* $E = \{< v, v+1 >: v \in \mathbb{Z}\}$ *with i.i.d unif*$([0,1])$ *initial configuration and fixed* $\mu \in (0, \frac{1}{2}]$. *Then, regardless of the value of* $\mu \in (0, 1/2]$, $\theta_c = 1/2$.

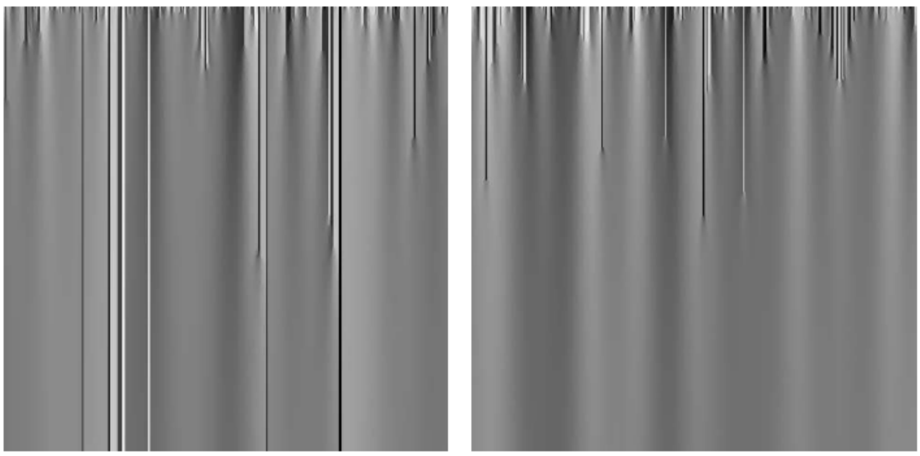

Figure 2.3: *Realizations of the Deffuant model on the one-dimensional torus with 400 vertices. Time goes from 0 to 400 and opinions are represented as functions of brightness, such that* $1 =$ *white and* $0 =$ *black. In both realizations,* $\mu = 0.25$ *and* $\theta = 0.48$ *and* 0.52 *(from left to right).* ***Source***: *Lanchier, 2012.*

In the same year, (Häggström, 2012) proved similar results using more traditional tools. He defined a deterministic process called *Sharing a Drink* and sharpened the difference in (2.2) to any fixed $\varepsilon > 0$, not only $\theta \in (0,1)$. Moreover, for the case $\theta > 1/2$, the almost sure convergence was proved, i.e., for any $x \in \mathbb{Z}$:

$$\mathbb{P}\left(\lim_{t\to\infty} |\eta_t(x) - \eta_t(x+1)| = 0\right) = 1, \tag{2.3}$$

which was called asymptotic consensus.

Later on, in a joint work with Timo Hirscher (Häggström and Hirscher, 2013), the result from Theorem 2.2.1 was extended to others initial distributions. Below, we present the types of convergence that will be explored in this article, along with the generalized theorem.

Definition 2.2.1. *[Häggström and Hirscher, 2013]*

(i) *No consensus. There will be finally blocked edges, i.e., edges $e = \langle u, v\rangle$ such that for all times t large enough:*

$$|\eta_t(u) - \eta_t(v)| > \theta.$$

Hence the vertices fall into different opinion groups.

(ii) *Weak consensus. Every pair of neighbors $\{u, v\}$ will finally concur, i.e.,*

$$\lim_{t\to\infty} |\eta_t(u) - \eta_t(v)| = 0.$$

(iii) *Strong consensus. The value at every vertex converges, as $t \to \infty$, to a common limit ℓ, where*

$$\ell = \begin{cases} \text{the average of the initial opinion values,} & \text{if } G \text{ is finite} \\ \mathbb{E}\eta_0, & \text{if } G \text{ is infinite.} \end{cases}$$

Using the key points of the first article with the Sharing a Drink process, Häggström and Hischer demonstrated the following generalization:

Theorem 2.2.2 (Häggström and Hirscher, 2013)**.** *Consider the Deffuant model on $\mathbb{Z}$ with identically and independent distributions as the initial configuration (not only uniform on $[0, 1]$)*

(a) Suppose the initial opinion of all the agents follows an arbitrary bounded distribution $\mathcal{L}(\eta_0)$ with expected value $\mathbb{E}\eta_0$ and $[a, b]$ being the smallest closed interval containing its support. If $\mathbb{E}\eta_0$ does not lie in the support, there exists some maximal, open interval $I \subset [a, b]$ such that $\mathbb{E}\eta_0$ lies in I and $\mathbb{P}(\eta_0 \in I) = 0$. In this case let h denote the length of I, otherwise set $h = 0$.

Then the critical value for θ, where a phase transition from a.s. no consensus to a.s. strong consensus takes place, becomes $\theta_c = \max\{\mathbb{E}\eta_0 - a, b - \mathbb{E}\eta_0, h\}$. The limit value in the supercritical regime is $\mathbb{E}\eta_0$.

(b) Suppose the initial opinions' distribution is unbounded but its expected value exists, either in the strong sense, i.e. $\mathbb{E}\eta_0 \in \mathbb{R}$ or the weak sense, i.e. $\mathbb{E}\eta_0 \in \{-\infty, +\infty\}$. Then the Deffuant model with an arbitrary fixed parameter $\theta \in (0, \infty)$ will a.s. behave subcritically, meaning that no consensus will be approached in the long run.

2.2.1 Sharing a Drink on $\mathbb{Z}$

The *Sharing a Drink* on $\mathbb{Z}$ was introduced by Olle Häggström (Häggström, 2012) as an auxiliary process to prove the main result concerning the critical value of the Deffuant model. In this process, individuals (represented as $x \in \mathbb{Z}$) act as glasses of water, and interactions occur deterministically, leading to a sharing of the drink. The initial profile consists of the glass at 0 full ($\xi_0(0) = 1$) and all the others are empty. Then iteratively the dynamic goes as follows: the glass with more water pours some into the adjacent glass, up to a maximum of half of its contents.

Mathematically, define $\{\xi_0(x)\}_{x\in\mathbb{Z}}$ by setting:

$$\xi_0(x) = \begin{cases} 1 \text{ if } x = 0, \\ 0 \text{ otherwise.} \end{cases} \tag{2.4}$$

Given the parameters $x_1, x_2, \ldots \in \mathbb{Z}$ and $\mu_1, \mu_2, \ldots \in (0, 1/2]$, for any $i = 1, 2, \ldots$ the profiles $\{\xi_i(x)\}_{x\in\mathbb{Z}}$ are obtained by:

$$\xi_i(x) = \begin{cases} \xi_{i-1}(x) + \mu_i(\xi_{i-1}(x+1) - \xi_{i-1}(x)) \text{ if } x = x_i, \\ \xi_{i-1}(x) + \mu_i(\xi_{i-1}(x-1) - \xi_{i-1}(x)) \text{ if } x = x_i + 1, \\ \xi_i(x) \text{ if } x \notin \{x_i, x_i + 1\} \end{cases} \tag{2.5}$$

The similarities with the Deffuant model are evident, however there are many differences. They lie in the type of dynamics since no stochastic component exists. Moreover, interactions take place without a limiting parameter, which allows any two adjacent glasses to exchange. Finally, the initial configuration is very different. Figure 2.4 illustrates the model.

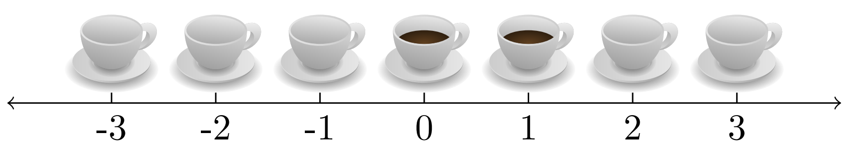

Figure 2.4: *Illustration of the Sharing a Drink (SAD) model, with the first interaction between agents at positions* 0 *and* 1*, and a convergence parameter* $\mu = 1/2$ *at time* $t = 1$*. In that case* $\xi_1(0) = \xi_1(1) = 1/2$ *and* $\xi_1(x) = 0$ *for all* $x \notin \{0, 1\}$*.*

From the definition, it is clear that many profiles can be reached. Fixing $\mu = 1/2$, choosing only non-negative $x \in \mathbb{Z}$ and interactions to happen with the next neighbor, i.e $x_i = x_{i-1} + 1$, guarantees that $\{\xi_t(x)\}_{x\in\mathbb{Z}}$ will be monotonous. For all fixed $t \in \mathbb{N}$:

$$\xi_t(0) \geq \xi_t(1) \geq \xi_t(2) \geq \cdots.$$

This shows one of many types of configuration achievable by the SAD structure. This and many other properties will be explored in the next section that will serve as auxiliary tools to prove the main result about the critical value for θ on the Deffuant model.

Properties

The first result is a direct consequence of the definition: preservability. The pairwise averaging does not change the total quantity of water. At any epoch, the sum will always be one.

Lemma 2.2.3. *For all $i \in \mathbb{N}$:*

$$\sum_{j \in \mathbb{N}} \xi_i(j) = 1. \tag{2.6}$$

Proof. By induction. For $i = 0$, the result follows immediately from the initial configuration (2.4). Suppose now that for a fixed $k \in \mathbb{N}$:

$$\sum_{j \in \mathbb{N}} \xi_k(j) = 1.$$

For $i = k+1$, x_k and $x_k + 1$ exchange liquids. Fix $\mu_{k+1} \in (0, 1/2]$. First, see that $\xi_k(x_k) + \xi_k(x_k + 1) = \xi_{k+1}(x_k) + \xi_{k+1}(x_k + 1)$, since by the SAD procedure (2.5):

$$\begin{aligned} \xi_{k+1}(x_k) + \xi_{k+1}(x_k + 1) &= \xi_k(x_k) + \mu_{k+1}(\xi_k(x_k + 1) - \xi_k(x_k))) \\ &\quad + \xi_k(x_k + 1) + \mu_{k+1}(\xi_k(x_k) - \xi_k(x_k + 1)) \\ &= \xi_k(x_k) + \xi_k(x_k + 1). \end{aligned}$$

Moreover, $\xi_{k+1}(x_j) = \xi_k(x_j)$ for all $x_j \notin \{x_k, x_{k+1}\}$. Hence,

$$\begin{aligned} \sum_{j \in \mathbb{N}} \xi_{k+1}(j) &= \sum_{\substack{j \in \mathbb{N} \\ j \neq \{x_k, x_k+1\}}} \xi_k(j) + \xi_{k+1}(x_k) + \xi_{k+1}(x_k + 1) \\ &= \sum_{\substack{j \in \mathbb{N} \\ j \neq \{x_k, x_k+1\}}} \xi_k(j) + \xi_k(x_k) + \xi_k(x_k + 1) \\ &= \sum_{j \in \mathbb{N}} \xi_k(j) = 1. \end{aligned}$$

which completes the proof. □

The next result involves an important profile achieved when interactions are restricted to non-negative integers:

Lemma 2.2.4 (Häggström, 2012)**.** *If $\{\xi_i(x)\}_{x \in \mathbb{Z}}$ is obtained via SAD procedure and $x_j \neq -1$ for $j = 1, \dots, i$, then:*

$$\xi_i(0) \geq \xi_i(1) \geq \xi_i(2) \geq \cdots. \tag{2.7}$$

Given the assumptions, the model is decreasing on the non-negative integers. Additionally, it has (only) one mode at 0, meaning that $\xi_j(0)$ takes the higher value in the configuration. This last result can be generalized to any configuration of the SAD model:

Lemma 2.2.5 (Häggström, 2012)**.** *Any $\{\xi_i(x)\}_{x \in \mathbb{Z}}$ obtained via the SAD procedure is unimodal, meaning that there exists $y \in \mathbb{Z}$ such that*

$$\cdots \leq \xi_i(y-2) \leq \xi_i(y-1) \leq \xi_i(y) \tag{2.8}$$

and

$$\xi_i(y) \geq \xi_i(y+1) \geq \xi_i(y+2) \geq \cdots . \tag{2.9}$$

It is evident that the mode can move within the integers depending on the set of choices for interactions. For example, taking $x_1 = -1$, $x_2 = -2$, and $x_3 = 1$ with $\mu_i = 1/2$ for all $i \in \{1, 2, \dots\}$ results in a configuration $\{\xi_3(x)\}_{x\in\mathbb{Z}}$ with $\xi_3(-2) = \xi_3(-1) = \xi_3(0) = \xi_3(1) = 1/4$.

The next lemma shows that the value of the mode of any SAD profile decreases over time:

Lemma 2.2.6. *Let ξ be a SAD procedure and $M(\xi_i)$ its supremum over all x at time i; $x \in \mathbb{Z}$, $i \in \mathbb{N}$. Then,*

$$M(\xi_i) \geq M(\xi_j), \ \forall j > i, \ j \in \mathbb{N}. \tag{2.10}$$

In words, the height of the mode decreases over time.

Proof. By induction. For $i = 0$, by the model definition, $M(\xi_0) = 1$. The next time, either one of the following options happens: 0 exchanges liquids or not.

In the first option, it exchanges with 1 or -1. In each case, $\xi_1(0)$, $\xi_1(\pm 1)$ will be less than $\xi_0(0)$, for $\mu_1 \in (0, \frac{1}{2}]$. Therefore, $M(\xi_1) = \max\{1-\mu_1, \mu_1\} \leq 1 = M(\xi_0)$.

In the second option, the inequality remains, for $M(\xi_0) = M(\xi_1) = 1$.

Assume now that (2.10) holds for $i = k-1$ and fix $y \in \mathbb{Z}$ as the mode. We have:

$$\xi_{k-1}(y) \geq \xi_{k-1}(y+1) \geq \xi_{k-1}(y+2) \geq \cdots \tag{2.11}$$

and

$$\cdots \leq \xi_{k-1}(y-2) \leq \xi_{k-1}(y-1) \leq \xi_{k-1}(y). \tag{2.12}$$

For $i = k$, again, either y exchanges liquids or not. The second case ($y < x_k$ or $y > x_k + 1$) is trivial, and with a similar calculation as in Lemma 2.2.4, we conclude that $M(\xi_{k-1}) = M(\xi_k) = \xi_k(y)$. In the first one, fix μ_k and let $y = x_k$. Then,

$$(\xi_k(x_k), \xi_k(x_k+1)) = (\xi_{k-1}(x_k)(1-\mu_k), \mu_k\xi_{k-1}(x_k))$$

as x_k did not exchange liquids before time k. Hence, since

$$M(\xi_{k-1}) = \xi_{k-1}(x_k) \geq \max\{\xi_{k-1}(x_k)(1-\mu_k), \mu_k\xi_{k-1}(x_k)\} = M(\xi_k),$$

(2.10) is valid. A similar calculation is done when $y = x_k + 1$. □

The proof of the lemma shows a crucial behavior of the model. Since the water can only move to one of the adjacent neighbors, the mode can only move to the right or left by one step per time. We have a similar situation with the Deffuant model, which will be explored in section 2.2.2, with the concept of ε-flatness.

Before presenting the final result about the mode, let us introduce the concept of dominance between two SAD profiles — or any two sequences that sum to one:

Definition 2.2.2. *We say that $\xi' = \{\xi'(x)\}_{x\in\mathbb{Z}}$ dominates $\xi = \{\xi(x)\}_{x\in\mathbb{Z}}$ if $\sum_{x=k}^{\infty} \xi(x) \leq \sum_{x=k}^{\infty} \xi'(x)$ for all $k \in \mathbb{Z}$. When this occurs, we will write $\xi \preceq \xi'$.*

Note that if we deal with stochastic sequences, the definition resembles the notion of stochastic dominance.

Lemma 2.2.7 (Häggström, 2012). *Suppose that $\xi_i = \{\xi_i(x)\}_{x\in\mathbb{Z}}$ and $\xi_i' = \{\xi_i'(x)\}_{x\in\mathbb{Z}}$ are exposed to the same SAD move, i.e. ξ_{i+1} and ξ_{i+1}' are obtained by picking the same pair of vertices $(x, x+1)$ to exchange liquids, with the same $\mu \in \left(0, \frac{1}{2}\right]$. If $\xi_i \preceq \xi_i'$ then $\xi_{i+1} \preceq \xi_{i+1}'$.*

Now, let us define the concept of energy, which will be useful in its proof.
Definition 2.2.3. *The energy W of a SAD profile $\{\xi_i(y)\}_{y\in\mathbb{Z}}$ is given by:*

$$W(i) = W(\{\xi_i(y)\}_{y\in\mathbb{Z}}) = \sum_{y\in\mathbb{Z}} (\xi_i(y))^2. \tag{2.13}$$

Finally, let M_x be the supremum of $\xi_i(x)$ over all i and all possible SAD procedures. Then,
Theorem 2.2.8 (Häggström, 2012). *For any $x \in \mathbb{Z}$, we have $M_x = \frac{1}{|x|+1}$.*

For the complete proof, see Häggström, 2012.

Connection between SAD and Deffuant models

Despite all the differences between the SAD and Deffuant models, the former's usefulness is fundamental. Indeed, established results allow us to express the opinion of site 0 in the Deffuant model at a specific time $t > 0$ as a weighted average of a mimicked SAD profile.

Fix $t > 0$. For each edge in the Deffuant model, as described in 2.1, there is a probability $\exp(-t)$ that its Poisson clock has not rung before time t. Consequently, there are infinitely many edges to the left or right of 0 almost surely whose clocks have not rung by time t. To demonstrate the connection between the models, let us first define:

$$Z_+(t) = \min\{x \geq 0 \,:\, \langle x, x+1\rangle \text{ has not rung by time } t\} \tag{2.14}$$

Similarly,

$$Z_-(t) = \max\{x \leq 0 \,:\, \langle x-1, x\rangle \text{ has not rung by time } t\} \tag{2.15}$$

In words, the edges $\langle Z_- - 1, Z_-\rangle$ and $\langle Z_+, Z_+ + 1\rangle$ determine the smallest region around 0 bounded by edges that did not interact up to time t. This is illustrated in Figure 2.5.

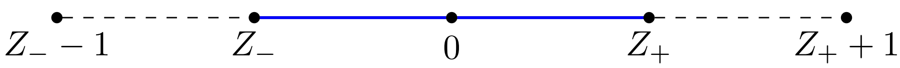

Figure 2.5: *Illustration of a region bounded by non-interacting edges. The thick blue lines represent active interactions, while the dashed lines indicate areas with no interaction up to a fixed time t.*

To simplify the notation, let $Z_+(t) := Z_+$ and $Z_-(t) := Z_-$ for a fixed $t > 0$. As defined, Z_+ and Z_- are almost surely finite for a fixed $t > 0$. Naturally:
Lemma 2.2.9. *Fix $t > 0$. With probability 1, $Z_+(t)$ and $Z_-(t)$ are finite.*

Proof. First, note that for a fixed $t > 0$, $Z_+(t) < \infty$ means that at least one Poisson clock rings on $\mathbb{Z}_+$ up to time t. Now, let T_x be the time of the clock on the edge $\langle x, x+1\rangle$

$(T_x \sim \text{Exp}(1))$. We have:

$$\begin{aligned}\mathbb{P}(Z_+(t) < +\infty) &= 1 - \mathbb{P}(Z_+ = +\infty)\\ &= 1 - \mathbb{P}(T_x \leq t \text{ for all } x \in \mathbb{Z}_+)\\ &= 1 - \prod_{x=0}^{\infty} \mathbb{P}(T_x \leq t)\\ &= 1 - \prod_{x=0}^{\infty}(1 - e^{-t})\\ &= 1 - 0\\ &= 1.\end{aligned}$$

In the last equality, we used the fact that $\lim_{x\to\infty}(1-e^{-t})^x = 0$ for any fixed $t > 0$. Similarly, it can be proven for $Z_-(t)$. □

The main idea is that these values reduce what happens in the Deffuant model to a finite system. Let N be the number of Poisson clocks that has rung in $\{Z_-, Z_- + 1, \dots, Z_+ - 1, Z_+\}$ and write $\tau_N, \tau_{N_1}, \dots, \tau_1$ as the time of these clocks, where τ_i is the time of the Poisson clock attributed to the edge $\langle x_i, x_i{+}1\rangle$. Moreover, consider them in reverse chronological order, i.e:

$$\tau_N \leq \tau_{N_1} \leq \cdots \leq \tau_1$$

Finally, let $\tau_{N+1} = 0$. Then, given the initial state of the Deffuant model $\eta_0(Z_-), \eta_0(Z_- + 1), \dots, \eta_0(Z_+ - 1), \eta_0(Z_+)$ nothing else is necessary to find $\eta_s(Z_-), \eta_s(Z_- - 1), \dots, \eta_s(Z_+ - 1), \eta_s(Z_+)$ for $s \in [0, t]$.

The result will follow from a mimicking SAD process. Given the Poisson rings and edges, define the SAD procedure choosing for $i = 1, \dots, N$ vertices x_i and $x_i + 1$ to exchange liquids (replicating what happened at the stochastic model) and $\mu_i = \mu$. Finally, for each $i = 1, \dots, N$ write $\{\xi_i(y)\}_{y\in\mathbb{Z}}$ for the resulting SAD configuration.

Lemma 2.2.10 (Häggström, 2012). *Consider the Deffuant model on $\mathbb{Z}$ with parameters $\mu \in (0, 1/2]$ and $\theta \in (0, 1)$. For $i = 0, 1, \dots, N$, $\eta_t(0)$ can be decomposed as*

$$\eta_t(0) = \sum_{y\in\mathbb{Z}} \xi_i(y)\eta_{\tau_{i+1}}(y). \tag{2.16}$$

In particular, $\eta_t(0) = \sum_{y\in\mathbb{Z}} \xi_N(y)\eta_0(y)$.

Notice that the definitions of Z_- and Z_+ ensure that the region between them is continuous, meaning that the interaction region has no gaps up to time t. This continuity allows us to replicate the dynamics of the Deffuant model using the SAD procedure. In addition, by analyzing the Poisson clocks in reverse chronological order, we can efficiently trace how the interactions within this interval shaped the current state of the system at time t, thus making the decomposition in (2.16) possible.

Proof of Lemma 2.2.10. By induction. For $i = 0$, by the initial state of the SAD procedure

(2.4) and the fact that τ_1 was the time of the last interaction:

$$\eta_t(0) = \sum_{y\in\mathbb{Z}} \xi_0(y)\eta_{\tau_{i+1}}(y) = 1\cdot\eta_{\tau_1}(0).$$

Suppose, now, that (2.16) is valid for $i = k-1$, i.e:

$$\eta_t(0) = \sum_{y\in\mathbb{Z}} \xi_{k-1}(y)\eta_{\tau_k}(y). \tag{2.17}$$

For $i = k$, we have:

$$\begin{aligned}
\eta_t(0) &= \sum_{y\in\mathbb{Z}} \xi_{k-1}(y)\eta_{\tau_k}(y)\\
&= \xi_{k-1}(x_k)\eta_{\tau_k}(x_k) + \xi_{k-1}(x_k+1)\eta_{\tau_k}(x_k+1) + \sum_{\substack{y\in\mathbb{Z}\\ y\neq\{x_k,x_k+1\}}} \xi_{k-1}(y)\eta_{\tau_k}(y)\\
&= \xi_{k-1}(x_k)\eta_{\tau_k}(x_k) + \xi_{k-1}(x_k+1)\eta_{\tau_k}(x_k+1) + \sum_{\substack{y\in\mathbb{Z}\\ y\neq\{x_k,x_k+1\}}} \xi_{k}(y)\eta_{\tau_{k+1}}(y)
\end{aligned}$$

To conclude the proof, we need to show that

$$\xi_{k-1}(x_k)\eta_{\tau_k}(x_k) + \xi_{k-1}(x_k+1)\eta_{\tau_k}(x_k+1) = \xi_k(x_k)\eta_{\tau_{k+1}}(x_k) + \xi_k(x_k+1)\eta_{\tau_{k+1}}(x_k+1). \tag{2.18}$$

First, recall both procedures of the SAD and Deffuant model respectively:

$$\xi_k(x) = \begin{cases} \xi_{k-1}(x) + \mu(\xi_{k-1}(x+1) - \xi_{k-1}(x)) \text{ if } x = x_k,\\ \xi_{k-1}(x) + \mu(\xi_{k-1}(x-1) - \xi_{k-1}(x)) \text{ if } x = x_k+1,\\ \xi_k(x)\ x \notin \{x_k, x_k+1\} \end{cases}$$

and

$$\eta_{\tau_k}(x) = \begin{cases} \eta_{\tau_{k+1}}(x) + \mu(\eta_{\tau_{k+1}}(x+1) - \eta_{\tau_{k+1}}(x)) \text{ if } x = x_k,\\ \eta_{\tau_{k+1}}(x) + \mu(\eta_{\tau_{k+1}}(x-1) - \eta_{\tau_{k+1}}(x)) \text{ if } x = x_k+1,\\ \eta_{\tau_{k+1}}(x)\ x \notin \{x_k, x_k+1\} \end{cases}$$

Plotting these into (2.18)

$$\begin{aligned}
\xi_{k-1}(x_k)\eta_{\tau_k}(x_k) + \xi_{k-1}(x_k+1)\eta_{\tau_k}(x_k+1) &= \xi_{k-1}(x_k)[\eta_{\tau_{k+1}}(x_k) + \mu(\eta_{\tau_{k+1}}(x_k+1) - \eta_{\tau_{k+1}}(x_k))]\\
&+ \xi_{k-1}(x_k+1)[\eta_{\tau_{k+1}}(x_k+1) + \mu(\eta_{\tau_{k+1}}(x_k) - \eta_{\tau_{k+1}}(x_k+1))]\\
&= \eta_{\tau_{k+1}}(x_k+1)[\xi_{k-1}(x_k+1) + \mu(\xi_{k-1}(x_k) - \xi_{k-1}(x_k+1))]\\
&+ \eta_{\tau_{k+1}}(x_k)[\xi_{k-1}(x_k) + \mu(\xi_{k-1}(x_k+1) - \xi_{k-1}(x_k))]\\
&= \eta_{\tau_{k+1}}(x_k+1)\xi_k(x_k+1) + \eta_{\tau_{k+1}}(x_k)\xi_k(x_k).
\end{aligned}$$

Therefore,

$$\eta_t(0) = \sum_{y\in\mathbb{Z}} \xi_k(y)\eta_{\tau_{k+1}}(y).$$

In particular, for $k = N \,:\, \eta_t(0) = \sum_{y\in\mathbb{Z}} \xi_N(y)\eta_0(y)$. □

Now that the result has been established, an important question arises: can it be extended to sites other than 0? The presented lemma 2.2.10 will not be valid for any $x \in \mathbb{Z}$. However, for any state $x \neq 0$ that belongs to the range from Z_- to Z_+, the equality holds true due to the spatial translation invariance of the Deffuant model.

Corollary 2.2.11. *Let $x \in \{Z_-, Z_- + 1, \dots, Z_+ - 1, Z_+\}$ or be such that $\eta_t(x) = \eta_0(x)$. For $i = 1, \dots, N$:*

$$\eta_t(x) = \sum_{y\in\mathbb{Z}} \xi_t^{(x)}(y)\eta_0(y). \tag{2.19}$$

Proof. By changing the entire opinion profile of the Deffuant model by a fixed integer $x > 0$ such that x becomes the new origin, the outcome remains the same due to translation invariance. Hence, under the circumstances of Lemma 2.2.10, for $x \in \{Z_-, Z_- - 1, \dots, Z_+ - 1, Z_+\}$ or such that $\eta_t(x) = \eta_0(x)$:

$$\eta_t(0) = \sum_{y\in\mathbb{Z}} \xi_i(y)\eta_{\tau_{i+1}}(y) \Rightarrow \eta_t(x) = \sum_{y\in\mathbb{Z}} \xi_i^{(x)}(y)\eta_{\tau_{i+1}}(y)$$

Where $\{\xi_i^{(x)}(y)\}_{y\in\mathbb{Z}}$ is the SAD profile after the translation (x being the new origin). □

2.2.2 Critical value of the Deffuant model on $\mathbb{Z}$

Let us present another important tool that will be used to prove Theorem 2.2.1:

Definition 2.2.4 (Häggström, 2012)**.** *Given $\varepsilon > 0$ and the Deffuant model initial configuration $\{\eta_0(y)\}_{y\in\mathbb{Z}}$, we say that x is an ε-flat point to the right if for all $n \geq 0$ we have*

$$\frac{1}{n+1}\sum_{y=x}^{x+n} \eta_0(y) \in \left[\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\right].$$

Similarly, x is said to be ε-flat point to the left if for all $n \geq 0$ we have

$$\frac{1}{n+1}\sum_{y=x-n}^{x} \eta_0(y) \in \left[\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\right].$$

Finally x is said to be two-sidedly ε-flat if for all $m, n \geq 0$ we have

$$\frac{1}{m+n+1}\sum_{y=x-m}^{x+n} \eta_0(y) \in \left[\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\right].$$

Since the Deffuant profile $\{\eta_0(y)\}_{y\in\mathbb{Z}}$ has the property of translation invariance, the probability of site x being ε-flat does not depend on x. Spatial shifting of x by a fixed $k > 0$ does not remove this property.

The next three Lemmas will show that for any site x, the probability of being ε-flat is strictly positive. These results will provide us with the last tools to finally delve into the critical θ of the Deffuant model.

Lemma 2.2.12 (Häggström, 2012). *For any $\varepsilon>0$ and any $x \in \mathbb{Z}$, we have*

$$\mathbb{P}(x \text{ is } \varepsilon\text{-flat to the right}) > 0. \tag{2.20}$$

Proof. Recall that $\{\eta_0(y)\}_{y\geq x}$ are identically and independently random variables uniformly distributed on $[0,1]$. Hence, by the Strong Law of Large Numbers, their average converges almost surely to $1/2$. This means that for a fixed $\varepsilon > 0$, there exists a natural number N:

$$\mathbb{P}\left(\frac{1}{n+1}\sum_{y=x}^{x+n}\eta_0(y) \in \left[\frac{1}{2}-\frac{\varepsilon}{3}, \frac{1}{2}+\frac{\varepsilon}{3}\right] \text{ for all } n \geq N\right) > 0. \tag{2.21}$$

Fix N. To conclude the demonstration, let us use an important technique known as local modification in percolation theory. Define the following coupling. Let $\{\eta_0(y)\}_{y\geq x}, \{\eta_0'(y)\}_{y\geq x}$ be i.i.d. uniform $[0,1]$ random variables coupled such that for each $y \in \{x, x+1, \dots\}$ the pairs $(\eta_0(y), \eta_0'(y))$ have the correct marginal (uniform $[0,1]$) and:

- $\eta_0(y), \eta_0'(y)$ be independent if $y \in \{x, x+1, \dots, x+N\}$
- $\eta_0(y), \eta_0'(y)$ be equal if $y \in \{x+N+1, x+N+2, \dots\}$

Now define the following events:

$$B := \left\{\frac{1}{n+1}\sum_{y=x}^{x+n}\eta_0'(y) \in \left[\frac{1}{2}-\frac{\varepsilon}{3}, \frac{1}{2}+\frac{\varepsilon}{3}\right] \text{ for all } n \geq N+1\right\}$$

and

$$C := \left\{\eta_0(y) \in \left[\frac{1}{2}-\frac{\varepsilon}{3}, \frac{1}{2}+\frac{\varepsilon}{3}\right] \text{ for all } y \in \{x, x+1, \dots, x+N\}\right\}$$

By (2.21), $\mathbb{P}(B) > 0$. In addition, by construction, $P(C) = (\frac{2\varepsilon}{3})^{N+1}$. Hence, since B and C are independent: $\mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C) > 0$.
Finally, define A as the desired result, i.e:

$$A := \left\{\frac{1}{n+1}\sum_{y=x}^{x+n}\eta_0(y) \in \left[\frac{1}{2}-\frac{\varepsilon}{3}, \frac{1}{2}+\frac{\varepsilon}{3}\right] \text{ for all } n \geq 0\right\}$$

Since $B \cap C \subset A$, we have $\mathbb{P}(A) \geq \mathbb{P}(B \cap C) > 0$, which concludes the proof. □

By symmetry, the probability of any $x \in \mathbb{Z}$ being ε-flat to the right is also strictly positive (and identical), i.e.:
Lemma 2.2.13 (Häggström, 2012). *For any $\varepsilon>0$ and any $x \in \mathbb{Z}$, we have*

$$\mathbb{P}(x \text{ is } \varepsilon\text{-flat to the left}) > 0. \tag{2.22}$$

Finally, now that we have both results, it is easy to show that the two-sided ε-flatness holds.

Lemma 2.2.14 (Häggström, 2012). *For any ε>0 and any $x \in \mathbb{Z}$, we have*

$$\mathbb{P}(x \text{ is two sidedly } \varepsilon\text{-flat}) > 0. \tag{2.23}$$

Proof. Let $\varepsilon > 0$ fixed and consider the following events regarding ε-flatness:

$$\begin{aligned} E_1 &= \{\text{site } x-1 \text{ is } \varepsilon\text{-flat to the left}\} \\ E_2 &= \left\{\eta_0(x) \in \left[\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon\right]\right\} \\ E_3 &= \{\text{site } x+1 \text{ is } \varepsilon\text{-flat to the right}\} \end{aligned}$$

From Lemmas 2.2.12 and 2.2.13, both E_1 and E_3 have positive probability, and, by construction, $\mathbb{P}(E_2) > 0$. Moreover, E_1, E_2 and E_3 are independent, since there are no overlapping regarding $\{\eta_0(x)\}_{x\in\mathbb{Z}}$. Hence $\mathbb{P}(E_1 \cap E_2 \cap E_3) = \mathbb{P}(E_1)\mathbb{P}(E_2)\mathbb{P}(E_3) > 0$.

Finally, note that we can achieve the two-sided ε-flatness with a convex combination of $\frac{1}{m}\sum_{y=x-m}^{x-1}\eta_0(y)$, $\eta_0(x)$, $\frac{1}{n}\sum_{y=x+1}^{x+n}\eta_0(y)$ such that:

$$E^+ := \left\{\frac{\frac{1}{m}\sum_{y=x-m}^{x-1}\eta_0(y)+\eta_0(x)+\frac{1}{n}\sum_{y=x+1}^{x+n}\eta_0(y)}{3} \in \left[\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon\right]\right\} \subset \{x \text{ is two sidedly } \varepsilon\text{-flat}\}$$

Consequently, $\mathbb{P}(x \text{ is two sidedly } \varepsilon\text{-flat}) \geq \mathbb{P}(E^+) \geq \mathbb{P}(E_1 \cap E_2 \cap E_3) > 0$. ☐

The case $\theta < 1/2$

Consider, now, the Deffuant model with $\theta \in (0, \frac{1}{2})$. Then, there is a positive probability that there are edges $e = \langle x, x+1\rangle$ that will be forever blocked from interaction.

Proposition 2.2.15 (Häggström, 2012). *For the Deffuant model with parameters $\mu \in (0, 1/2]$ and $\theta \in (0, 1/2)$, let $B_x := \{|\eta_t(x) - \eta_t(x+1)| > \theta, \forall t \geq 0\}$, then for any $x \in \mathbb{Z}$:*

$$\mathbb{P}(B_x) > 0. \tag{2.24}$$

Proof. Fix $\theta \in (0, 1/2)$ and let $\varepsilon > 0$ be such that $\theta = \frac{1}{2} - \varepsilon$. Following the same argument of Lemma 2.2.14, define the events:

$$\begin{aligned} E_x^1 &= \{\text{site } x-1 \text{ is } \varepsilon\text{-flat to the left}\} \\ E_x^2 &= \{\eta_0(x) > 1-\varepsilon\} \\ E_x^3 &= \{\text{site } x+1 \text{ is } \varepsilon\text{-flat to the right}\} \end{aligned}$$

We have $\mathbb{P}(E_x^i) > 0$ for $i = 1, 2, 3$ and $x \in \mathbb{Z}$. Futhermore, $\mathbb{P}(E_x^1 \cap E_x^2 \cap E_x^3) = \mathbb{P}(E_x^1)\mathbb{P}(E_x^2)\mathbb{P}(E_x^3) > 0$. Now define $E_x = E_x^1 \cap E_x^2 \cap E_x^3$. The result will follow by proving that $E_x \subseteq B_x$.

First, assume E_x. Note that $\eta_t(x)$ will not change until interaction occurs in one of the edges $\langle x, x+1\rangle$ or $\langle x-1, x\rangle$. Let T be the first time one of these interactions takes place, meaning that $\eta_t(x) = \eta_0(x)$ for all $0 < t < T$.

By the SAD representation given in Corollary 2.2.11, for $0 < t < T$:

$$\eta_t(x+1) = \sum_{y \in \mathbb{Z}} \xi_t^{(x+1)}(y)\eta_0(y), \tag{2.25}$$

where $\{\xi_t^{(x+1)}(y)\}_{y\in\mathbb{Z}}$ is the SAD profile after the translation with $x+1$ being the new origin. Also, $\xi_t^{(x+1)}(y) = 0$ for all $y \le x$ and $0 < t < T$. Hence, by Lemma 2.2.4:

$$\xi_t^{(x+1)}(x+1) \ge \xi_t^{(x+1)}(x+2) \ge \xi_t^{(x+1)}(x+3) \ge \cdots. \tag{2.26}$$

Since $\{\xi_t^{(x+1)}(y)\}_{y\in\mathbb{Z}}$ are nonzero for only finitely many y, let us define $R := \max\{m : \xi_t^{(x+1)}(x+m) > 0\}$ as the range to the right, where the profile assumes values. In the next steps, we will show that $x+1$ will not be able to interact with x.

Define, now, $\delta_k = \xi_t^{(x+1)}(x+k) - \xi_t^{(x+1)}(x+k+1)$ and $c_k = k\delta_k$ for $k = 1,\dots,R$. Then, $\delta_k \ge 0$ and $c_k \ge 0$ by (2.26) for each k, and $\sum_{n=k}^{R} \delta_k = \xi_t^{(x+1)}(x+k)$. Moreover,

$$\begin{aligned}\sum_{n=1}^{R} c_n = \sum_{n=1}^{R} n\delta_n &= \sum_{k=1}^{R}\sum_{n=k}^{R} \delta_n \\ &= \sum_{k=1}^{R} \xi_t^{(x+1)}(x+k) = 1.\end{aligned}$$

Using that $\xi_t^{(x+1)}(y) = 0$ for $y \le x$, we can rewrite (2.25) as:

$$\begin{aligned}\eta_t(x+1) = \sum_{k=1}^{R} \xi_t^{(x+1)}(x+k)\eta_0(x+k) &= \sum_{k=1}^{R} \eta_0(x+k) \sum_{n=k}^{R} \delta_n \\ = \sum_{n=1}^{R} \delta_n \sum_{k=1}^{n} \eta_0(x+k) &= \sum_{n=1}^{R} c_n \left(\frac{1}{n}\sum_{k=1}^{n} \eta_0(x+k)\right)\end{aligned}$$

where $c_n's$ are non-negative and sum to 1. Hence, $\eta_t(x+1)$ is a convex combination of $\left(\frac{1}{n}\sum_{k=1}^{n} \eta_0(x+k)\right)$. In addition, since for each n, $\left(\frac{1}{n}\sum_{k=1}^{n} \eta_0(x+k)\right)$ is on the event $E_x^3 = \{\text{site } x+1 \text{ is } \varepsilon\text{-flat to the right}\}$, i.e., is in the interval $[1/2-\varepsilon, 1/2+\varepsilon]$, we have $\sum_{n=1}^{R} c_n \left(\frac{1}{n}\sum_{k=1}^{n} \eta_0(x+k)\right) \le \frac{1}{2}+\varepsilon$. Therefore, $\eta_t(x+k)$ will never exceed $\frac{1}{2}+\varepsilon$ before time T, and $|\eta_t(x) - \eta_t(x+1)| > \frac{1}{2} - 2\varepsilon = \theta$ for all $0 < t < T$. Similarly for site $x-1$. In conclusion, $E_x \subseteq B_x$, and $\mathbb{P}(B_x) \ge \mathbb{P}(E_x) > 0$. □

The next lemma will serve as an auxiliary result to finally prove our last theorem, regarding the case $\theta < 1/2$ in the Deffuant model. With results from ergodic theory (see Häggström, 2012) it can be proven that almost surely $\lim_{n\to\infty} \frac{1}{n}\sum_{x=1}^{n} I_{B_x} = \mathbb{P}(B_0)$ and $\lim_{n\to\infty} \frac{1}{n}\sum_{x=1}^{n} I_{B_{-x}} = \mathbb{P}(B_0)$, where I_{B_x} is the indicator function of the event B_x.

Lemma 2.2.16 (Häggström, 2012). *With probability 1, there will be infinitely many sites x to the left of* 0 *such that B_x happens, and infinitely many to the right.*

The Lemma still holds if we substitute B_x with the event E_x. Finally,

Theorem 2.2.17 (Häggström, 2012). *For the Deffuant model with $\theta < 1/2$, we have a.s. that for all $x \in \mathbb{Z}$:*

1. *The limiting value $\eta_\infty(x) = \lim_{t\to\infty} \eta_t(x)$ exists;*

2. *The limiting configuration $\{\eta_\infty(x)\}_{x\in\mathbb{Z}}$ satisfies $\{|\eta_\infty(x) - \eta_\infty(x+1)|\} \in \{0\} \cup [\theta, 1]$ for all $x \in \mathbb{Z}$.*

Proof. Under the conditions of the statement, consider the initial configuration $\{\eta_0(y)\}_{y\in\mathbb{Z}}$. Now recall that for any vertex y_1, E_{y_1} has positive probability to happen. Let $y_1 - 1$ be such event and $y_2 := \min\{y_2 > y_1 : E_{y_2}\}$. In other words, any vertex $x \in \{y_1+1, \dots, y_2-1\}$ can only interact with agents inside the same set. Since this can happen with any $x \in \mathbb{Z}$, it suffices to prove the result for all $x \in \{y_1+1, \dots, y_2-1\}$. Figure 2.6 illustrates the scenario described, where the vertical lines represent opinions and $\theta = \frac{1}{2} - 2\varepsilon$, for $\varepsilon > 0$.

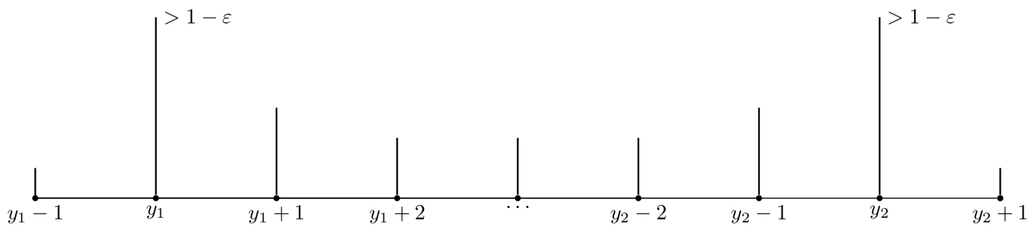

Figure 2.6: *Deffuant model with interaction region limited.*

First, recall the concept of energy presented in 2.2.1. Define

$$W_t = \sum_{y\in\{y_1,y_1+1,\dots,y_2-1\}} (\eta_t(y))^2$$

Note that, W_t is non-negative and each time t that two vertices in $\{y_1+1, \dots, y_2-1\}$ interact, W_t drops its value by $2\mu(1-\mu)|\eta_{t^-}(x) - \eta_{t^-}(x+1)|^2$. Indeed, let $\eta_{t^-}(x) := a$ and $\eta_{t^-}(x+1) = b$, suppose that at time t the edge $\langle x, x+1\rangle$ is selected. Then,

$$\begin{aligned} W_t = \sum_{y\in\{y_1+1,\dots,y_2-1\}} (\eta_t(y))^2 &= \sum_{\substack{y\in\{y_1+1,\dots,y_2-1\}\\ y\neq\{x,x+1\}}} (\eta_t(y))^2 + (\eta_t(x))^2 + (\eta_t(x+1))^2 \\ &= \sum_{\substack{y\in\{y_1+1,\dots,y_2-1\}\\ y\neq\{x,x+1\}}} (\eta_t(y))^2 + (a + \mu(b-a))^2 + (b + \mu(a-b))^2 \\ &= \sum_{\substack{y\in\{y_1+1,\dots,y_2-1\}\\ y\neq\{x,x+1\}}} (\eta_{t^-}(y))^2 + a^2 + b^2 - 2\mu(1-\mu)|a-b|^2 \\ &= W_{t^-} - 2\mu(1-\mu)|a-b|^2. \end{aligned}$$

Hence, W_t is decreasing in time.

Now, define

$$F_t = \max\{I_{\{|\eta_t(x)-\eta_t(x+1)|\leq\theta\}}|\eta_t(x) - \eta_t(x+1)| : x \in \{y_1+1, \dots, y_2-1\}\}. \tag{2.27}$$

To prove the second statement of the Theorem, it suffices to show that $\lim_{t\to\infty} F_t = 0$. Note that, in this scenario:

$$\begin{aligned}&\text{either } |\eta_t(x) - \eta_t(x+1)| > \theta \text{ for all sufficiently large } t,\\ &\text{or } \lim_{t\to\infty} |\eta_t(x) - \eta_t(x+1)| = 0.\end{aligned} \tag{2.28}$$

for the alternative would suggest that there is some edge $\langle x, x+1\rangle$ such that for any $\delta > 0$, $|\eta_t(x) - \eta_t(x+1)|$ would jump repeatedly in the intervals $[0, \delta]$ and $(\theta, 1]$. Due to the Deffuant dynamics, this would not be possible, since for δ small enough, the gap between these intervals would be bigger than the maximum increase per interaction, which is given by $\mu\theta$.

Let us now prove that as time increases F_t goes to zero. Suppose by contradiction that $F_t \geq \delta$ for $\delta > 0$ and a large t. Consider the edge $e_x = \langle x, x+1\rangle$ for $x \in \{y_1 + 1, \ldots, y_2 - 1\}$ and T_{e_x} the time of its Poisson clock. Now, at each time a Poisson clock rings check the value of F_t. If it exceeds δ, suppose without loss of generality, that e_x is such that $|\eta_t(x) - \eta_t(x+1)| \in (\delta, \theta]$. Then, the probability of e_x being the next to exchange is given by:

$$\begin{aligned}\mathbb{P}(T_{e_x} = \min\{T_{e_y} : y \in \{y_1 + 1, \ldots, y_2 - 1\}) &= \frac{1}{|\{y_1, y_1 + 1, \ldots, y_2 - 1\}|}\\ &= \frac{1}{y_2 - y_1},\end{aligned}$$

for $T_{e_x} \sim \exp(1)$ for all $x \in \mathbb{Z}$. Furthermore, in this case, W_t will decrease. Using the conditional Borel-Cantelli Lemma, we can affirm that this will happen infinitely often with probability one, leading to $\lim_{t\to\infty} W_t = -\infty$. This contradicts the fact that $W_t \geq 0$. Therefore, $\lim_{t\to\infty} F_t = 0$.

Finally, to prove the existence of $\eta_\infty(x) = \lim_{t\to\infty} \eta_t(x)$, recall that by (2.28) there are two scenarios for each edge $\langle x, x+1\rangle$ in $\{y_1 + 1, \ldots, y_2 - 1\}$. If $\lim_{t\to\infty} |\eta_t(x) - \eta_t(x+1)| = 0$, then $\sum_{y\in\{y_1+1,\ldots,y_2-1\}} \eta_t(y)$ is preserved over time since interaction only occurs in the interval. Consequently $\eta_t(y_1 + 1), \ldots, \eta_t(y_2 - 1)$ must all converge to their initial average $\frac{1}{y_2+y_1-1} \sum_{y\in\{y_1+1,\ldots,y_2-1\}} \eta_0(y)$.

Conversely, if the parameter θ censors an edge $\langle x, x+1\rangle$, where $|\eta_t(x) - \eta_t(x+1)| > \theta$, then there will be a subinterval $\{y_1', y_1' + 1, \ldots, y_2'\}$ containing the edge with $y_1 + 1 \leq y_1' \leq y_2' \leq y_2 - 1$. In this subinterval, the first scenario happens for each edge. As a result, there will be a time $T > 0$ such that from that point onwards, y_1' and y_2' will not exchange with neighbors outside the subset. This establishes the preservation property again, and agents will converge to their average $\frac{1}{y_2'+y_1'+1} \sum_{y\in\{y_1',\ldots,y_2'\}} \eta_T(y)$, which concludes the proof. □

The case $\theta > 1/2$

To explore the Deffuant model with $\theta > \frac{1}{2}$, let us first define an auxiliary process to track the energy loss at each site $x \in \mathbb{Z}$ over time $t > 0$. First, consider the energy at site x as $W_t(x) := (\eta_t(x))^2$ and define:

Definition 2.2.5 (Cumulative energy loss of site x)**.** *Let $W_t^\dagger(x)$ be the cumulative energy loss of site x, defined as follows:*

- $W_0^\dagger(x) = 0$;

- $W_t^\dagger(x) = W_{t^-}^\dagger(x) + 2\mu(1-\mu)|\eta_{t^-}(x) - \eta_{t^-}(x+1)|^2$ *whenever an interaction occurs at the edge* $\langle x, x+1\rangle$.

Otherwise, it stays constant.

Note that $W_t^\dagger(x)$ increases by the exact amount that $W_t(x) + W_t(x+1)$ decreases whenever edge $\langle x, x+1\rangle$ is selected (see the proof of Theorem 2.2.17 for details). We have:
Lemma 2.2.18 (Häggström, 2012)**.** *For some (hence any)* $x \in \mathbb{Z}$ *we have*

$$\mathbb{E}[W_t(x)] + \mathbb{E}[W_t^\dagger(x)] = \frac{1}{3}. \tag{2.29}$$

Proof. At $t = 0$, $\mathbb{E}[W_0(x)] = \int_0^1 u^2 du = \frac{1}{3}$, since $\eta_0(x) \sim U[0,1]$. Also, by definition, $\mathbb{E}[W_0^\dagger(x)] = 0$. For $t > 0$, let $W_t^{\text{tot}}(x) := W_t(x) + W_t^\dagger(x)$. Due to the ergodic nature of the Poisson process and the translation invariance of the model, the auxiliary process $\{W_t^{\text{tot}}(x)\}_{x\in\mathbb{Z}}$ inherits ergodic properties. This implies that almost surely the spatial average converges to its ensemble average. Mathematically,

$$\mathbb{P}\left(\lim_{\substack{y_1\to-\infty\\ y_2\to\infty}} \frac{1}{y_2 - y_1 + 1}\sum_{x=y_1}^{y_2} W_t^{\text{tot}}(x) = \mathbb{E}[W_t^{\text{tot}}(0)]\right) = 1. \tag{2.30}$$

To finish the proof, it suffices to show that $\mathbb{E}[W_t^{\text{tot}}(0)]$ is constant over time. To this end, we will use an argument similar to what was done in Section 2.2.1, with Z_- and Z_+. Note that, for any $t > 0$ there are $y_1 < 0$ and $y_2 > 0$ such that the edges $\langle y_1 - 1, y_1\rangle$, $\langle y_2, y_2+1\rangle$ have not rung. Now, consider the set $\{y_1, \dots, y_2\}$. By induction over the Poisson rings in it, the sum of the total energy remains constant for any s ranging from 0 up to t. In particular,

$$\sum_{x\in\{y_1,\dots,y_2\}} W_s^{\text{tot}}(x) = \sum_{x\in\{y_1,\dots,y_2\}} W_0^{\text{tot}}(x)$$

Indeed, for $t = 0$ no interaction occurred. Now, suppose that the energy sum remains constant until some time s. Then, interaction takes place at edge $\langle x, x+1\rangle$ inside the interval, we have:

$$W_s(x) = W_{s^-}(x) - \mu(1-\mu)|\eta_{t^-}(x) - \eta_{t^-}(x+1)|^2$$

and

$$W_s^\dagger(x) = W_{s^-}^\dagger(x) + \mu(1-\mu)|\eta_{t^-}(x) - \eta_{t^-}(x+1)|^2$$

Similarly for $x+1$, which makes $W_{s^-}^{\text{tot}}(x) + W_{s^-}^{\text{tot}}(x+1) = W_s^{\text{tot}}(x) + W_s^{\text{tot}}(x+1)$. Thus,

$$\sum_{x\in\{y_1,\dots,y_2\}} W_t^{\text{tot}}(x) = \sum_{x\in\{y_1,\dots,y_2\}} W_0^{\text{tot}}(x).$$

Since y_1 and y_2 could be taken arbitrarily far from 0, we have for the limit (2.30):

$$\lim_{\substack{y_1\to-\infty\\ y_2\to\infty}} \frac{1}{y_2 - y_1 + 1}\sum_{x=y_1}^{y_2} W_t^{\text{tot}}(x) = \lim_{\substack{y_1\to-\infty\\ y_2\to\infty}} \frac{1}{y_2 - y_1 + 1}\sum_{x=y_1}^{y_2} W_0^{\text{tot}}(x)$$

Hence,

$$\begin{aligned}\mathbb{E}[W_t(x)] + \mathbb{E}[W_t^{\dagger}(x)] &= \mathbb{E}[W_t^{\text{tot}}(x)] \\ &= \lim_{\substack{y_1\to-\infty \\ y_2\to\infty}} \frac{1}{y_2 - y_1 + 1} \sum_{x=y_1}^{y_2} W_0^{\text{tot}}(x) \\ &= \mathbb{E}[W_0^{\text{tot}}(x)] = \frac{1}{3}.\end{aligned}$$

□

Proposition 2.2.19 (Häggström, 2012)**.** *For the Deffuant model with arbitrary threshold parameter $\theta \in (0,1)$, we have almost surely that for each $x \in \mathbb{Z}$:*

$$\begin{aligned}&\textit{either } |\eta_t(x) - \eta_t(x+1)| > \theta \textit{ for all sufficiently large } t, \\ &\textit{or } \lim_{t\to\infty} |\eta_t(x) - \eta_t(x+1)| = 0.\end{aligned} \tag{2.31}$$

Proof. The demonstration of this proposition is very similar to what was done in the proof of Theorem 2.2.17 involving F_t. Fix $\delta > 0$ and $x \in \mathbb{Z}$. First, let $e_x = \langle x, x+1 \rangle$ and T_{e_x} be the time of its ring. The probability that the next ring among e_{x-1}, e_x and e_{x+1} will be at e_x is positive, for $\mathbb{P}(T_{e_x} = \min\{T_{e_{x-1}}, T_{e_x}, T_{e_{x+1}}\}) = 1/3$.

Now, if $\delta < |\eta_t(x) - \eta_t(x+1) \leq \theta$ for a large t, then each time interaction occurs at e_x the cumulative energy loss of site x increases by $2\mu(1-\mu)|\eta_{t^-}(x) - \eta_{t^-}(x+1)|^2$, by definition. Hence, by conditional Borel-Cantelli Lemma, $\lim_{t\to\infty} W_t^{\dagger}(x) = \infty$, which must have probability 0, since as a consequence of Lemma 2.2.18 and the Markov inequality:

$$\mathbb{P}(W_t^{\dagger}(x) \geq n) \leq \frac{\mathbb{E}[W_t^{\dagger}(x)]}{n} \leq \frac{1}{3n}$$

goes to zero when $n \to \infty$. Therefore, we have almost surely that

$$|\eta_t(x) - \eta_t(x+1)| \in [0,\delta] \cup (\theta, 1] \text{ for all sufficiently large } t. \tag{2.32}$$

Finally, recall that for small enough $\delta > 0$, $|\eta_t(x) - \eta_t(x+1)|$ can not cross the gap of the intervals above, which suffices to end the proof. □

The next Lemma is essential for proving the main result of the model. It states that, despite future Poisson rings, if a site x starts with an initial opinion that is two-sided ε-flat, it will remain so for all $t > 0$.

Lemma 2.2.20 (Häggström, 2012)**.** *Suppose, given $\varepsilon > 0$, that site $x \in \mathbb{Z}$ is two-sidedly ε-flat for the initial configuration $\{\eta_0(y)\}_{y\in\mathbb{Z}}$. Then, regardless of all future Poisson rings, we have*

$$\eta_t(x) \in \left[\frac{1}{2} - 6\varepsilon, \frac{1}{2} + 6\varepsilon\right] \textit{ for all } t \geq 0. \tag{2.33}$$

The proof is centered on the relationship between the Deffuant model and the SAD profile, similar to what was done in 2.2.15. Auxiliary functions are first defined to rewrite the opinion of site x at time t as a weighted sum of initial opinions. Through algebraic manipulation and the properties of the SAD profile, it is demonstrated that this weighted sum is bounded, which leads to the desired result. For the complete proof, see Häggström, 2012.

With this result established, we can sharpen Proposition 2.2.19 for the case $\theta > 1/2$:

Proposition 2.2.21 (Häggström, 2012)**.** *For the Deffuant model with $\theta > \frac{1}{2}$, we have almost surely for all $x \in \mathbb{Z}$ that*

$$\lim_{t\to\infty} |\eta_t(x) - \eta_t(x-1)| = 0. \tag{2.34}$$

Proof. Let $\varepsilon > 0$ be such that $\theta = \frac{1}{2} + 6\varepsilon$. We will show that almost surely for any site x, $|\eta_t(x) - \eta_t(x-1)| = 0$ for all sufficiently large t.

Suppose by contradiction that for any $x \in \mathbb{Z}$

$$\mathbb{P}(|\eta_t(x) - \eta_t(x-1)| > \theta \text{ for all sufficiently large } t) > 0 \tag{2.35}$$

Then, by applying the Borel-Cantelli Lemma, (2.35) will happen for infinitely many $x \in \mathbb{Z}$ with probability 1. Now, following the steps on the proof of Theorem 2.2.17, we can conclude that $\eta_\infty(x) = \lim_{t\to\infty} \eta_t(x)$ exists and satisfies $\{|\eta_\infty(x) - \eta_\infty(x+1)|\} \in \{0\} \cup [\theta, 1]$ for all $x \in \mathbb{Z}$.

From Lemma 2.2.14, with positive probability there exists $z \in \mathbb{Z}$ such that $\eta_0(z)$ is two-sided $\varepsilon-$flat. Combined with Lemma 2.2.20, we know that, regardless the future Poisson rings, $\eta_t(z)$ will be stuck in the interval $\left[\frac{1}{2} - 6\varepsilon, \frac{1}{2} + 6\varepsilon\right]$ for all t. Therefore, it will also be the case of $\eta_\infty(z)$.

Finally, by Proposition 2.2.19 there are two scenarios for $z+1$:

$$\begin{aligned} &\text{either } |\eta_\infty(z) - \eta_\infty(z+1)| > \theta \text{ for all sufficiently large } t, \\ &\text{or } \lim_{t\to\infty} |\eta_t(z) - \eta_t(z+1)| = 0. \end{aligned} \tag{2.36}$$

From the first option there are two scenarios: either $\eta_\infty(z+1)$ exceeds or is less than $\eta_\infty(z)$ by at least θ. These can not be possible, since by the choice of θ, $\eta_\infty(z+1)$ would be greater than 1 or negative, respectively. Therefore, $\eta_\infty(z) = \eta_\infty(z+1)$, and similarly $eta_\infty(z-1) = \eta_\infty(z)$. By iteration, we extend this to all $y \in \mathbb{Z}$, contradicting 2.35 as desired. □

Theorem 2.2.22 (Häggström, 2012)**.** *For the Deffuant model with $\theta > \frac{1}{2}$, we have a.s. for all $x \in \mathbb{Z}$ that*

$$\lim_{t\to\infty} \eta_t(x) = \frac{1}{2}. \tag{2.37}$$

Proof. Fix $x \in \mathbb{Z}$ and $\varepsilon > 0$. Combining the Lemmas 2.2.14 and 2.2.20 we have almost surely a site $z \in \mathbb{Z}$ such that its initial configuration is two-sided $\varepsilon-$flat and $\eta_t(z) \in [1/2 - 6\varepsilon, 1/2 + 6\varepsilon]$ for all t. Now, applying Proposition 2.2.21 in all finitely many edges between z and x, we can extend the bounds to other sites and affirm that $\eta_t(x) \in [1/2 - 8\varepsilon, 1/2 + 8\varepsilon]$ for all sufficiently large t. Finally, since the choice of ε was arbitrary, we have $\lim_{t\to\infty} \eta_t(x) = \frac{1}{2}$. □

2.2.3 Extension to other initial distributions

Based on the work of LANCHIER, 2012 and HÄGGSTRÖM, 2012, a joint article, from 2014, written by Olle Häggström and Timo Hischer presented a generalization of consensus formation in the Deffuant model. They extended Theorem 2.2.1 to cover other cases of independent and identically (i.i.d) distributed initial opinions, including those with gaps in their support or infinity expected value.

The proof followed a similar structure to what was done in the previous section. Using key tools such as the SAD profile and the concept of ε-flatness, the next Theorem arose:
Theorem 2.2.23 (HÄGGSTRÖM and HIRSCHER, 2013)**.** *Consider the Deffuant model on* $\mathbb{Z}$ *with identically and independent distributions as the initial configuration (not only uniform on* $[0,1]$*)*

(a) Suppose the initial opinion of all the agents follows an arbitrary bounded distribution $\mathcal{L}(\eta_0)$ *with expected value* $\mathbb{E}\eta_0$ *and* $[a,b]$ *being the smallest closed interval containing its support. If* $\mathbb{E}\eta_0$ *does not lie in the support, there exists some maximal, open interval* $I \subset [a,b]$ *such that* $\mathbb{E}\eta_0$ *lies in* I *and* $\mathbb{P}(\eta_0 \in I) = 0$*. In this case let* h *denote the length of* I*, otherwise set* $h = 0$*.*

Then the critical value for θ*, where a phase transition from a.s. no consensus to a.s. strong consensus takes place, becomes* $\theta_c = \max\{\mathbb{E}\eta_0 - a, b - \mathbb{E}\eta_0, h\}$*. The limit value in the supercritical regime is* $\mathbb{E}\eta_0$*.*

(b) Suppose the initial opinions' distribution is unbounded but its expected value exists, either in the strong sense, i.e. $\mathbb{E}\eta_0 \in \mathbb{R}$ *or the weak sense, i.e.* $\mathbb{E}\eta_0 \in \{-\infty, +\infty\}$*. Then the Deffuant model with an arbitrary fixed parameter* $\theta \in (0,\infty)$ *will a.s. behave subcritically, meaning that no consensus will be approached in the long run.*

With this theorem established, it is possible to explore the circumstances under which consensus is achieved in populations with other initial opinions, which is closer to real-world dynamics. For example, using a $Beta(0.1, 0.3)$ distribution, consensus is achieved for a confidence parameter θ greater than $3/4$. Figure 2.7 illustrates two different scenarios with this distribution.

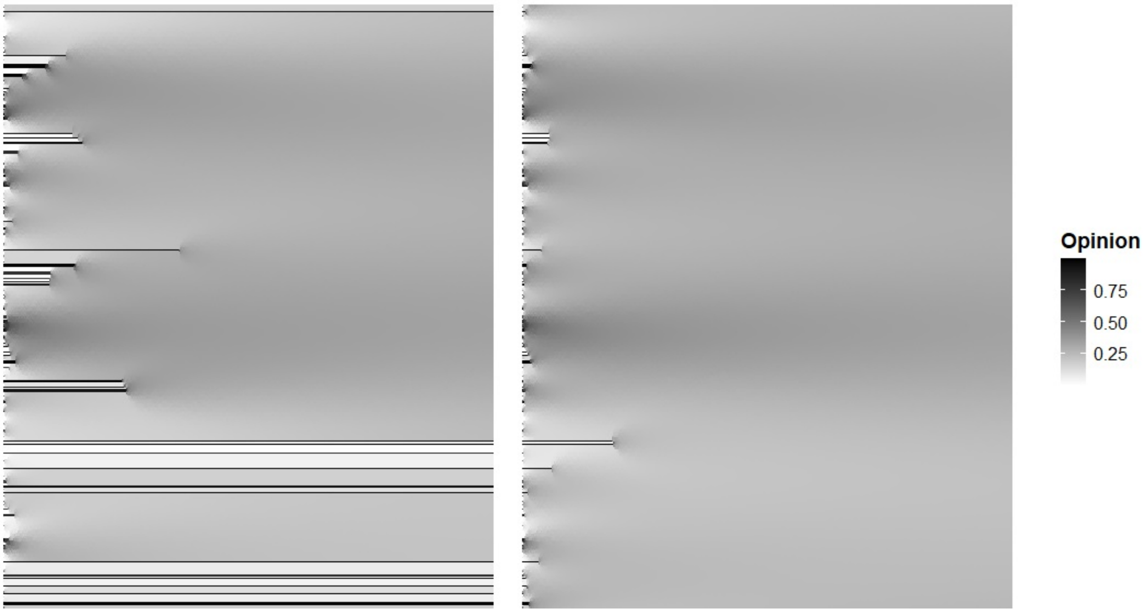

Figure 2.7: *Realizations of the Deffuant model in the one-dimensional torus with* 400 *agents and initial opinions* $\eta_0 \sim Beta(0.1, 0.3)$. *Parameters* $\theta = 0.7$ *and* 0.8 *(from left to right) and* $\mu = 0.5$. *Results after* 500 *thousands iterations.*

The image represents the evolution of opinions over time in the Deffuant model on a one-dimensional torus. The vertical axis corresponds to the agents arranged in a circular topology, while the horizontal axis represents time (iterations). Each pixel's shade of gray indicates the opinion of an agent at a specific time, with darker shades representing higher opinion values and lighter shades representing lower opinion values. As the simulation progresses (moving to the right), the interactions between agents lead to the formation of distinct opinion clusters (image on the left), visible as horizontal bands. Conversely, consensus formation is represented by smooth transitions and a uniform pattern of shading (image on the right).

It is still an open question whether strong consensus can be achieved at criticality $\theta = \theta_c$. However, for special cases such as bounded initial distributions with a large gap around their mean, an important proposition holds:

Proposition 2.2.24 (Häggström and Hirscher, 2013)**.** *Let the initial opinions be again i.i.d. with* $[a, b]$ *being the smallest closed interval containing the support of the marginal distribution, and the latter feature a gap* (α, β) *of width* $\beta - \alpha > \max\{\mathbb{E}\eta_0 - a, b - \mathbb{E}\eta_0\}$ *around its expected value* $\mathbb{E}\eta_0 \in [a, b]$.

At criticality, that is for $\theta = \theta_c = \max\{\mathbb{E}\eta_0 - a, b - \mathbb{E}\eta_0, \beta - \alpha\}$, *we get the following: If both* α *and* β *are atoms of the distribution* $L(\eta_0)$, *i.e.* $\mathbb{P}(\eta_0 = \alpha) > 0$ *and* $\mathbb{P}(\eta_0 = \beta) > 0$, *the system approaches a.s. strong consensus. However, it will a.s. lead to no consensus if either* $\mathbb{P}(\eta_0 = \alpha) = 0$ *or* $\mathbb{P}(\eta_0 = \beta) = 0$.

Let us analyze a scenario where it is being evaluated how much one agrees or disagrees on a topic (without a neutral opinion). For example, consider a population with independent initial distribution of opinions uniformly distributed on $\{-0.5, -0.3, 0.3, 0.5\}$ (from left to right: strongly disagrees, disagrees, agrees, strongly agrees). We have $\mathbb{E}\eta_0 = 0$

and $[-0.5, 0.5]$ is the smallest closed interval containing its support. Moreover, the gap $(-0.3, 0.3)$ around the mean has length 0.6 and $\mathbb{P}(\eta_0 = 0.3)$ and $\mathbb{P}(\eta_0 = -0.3)$ are both positive. Therefore, by Proposition 2.2.24, $\theta = \theta_c = 0.6$ and consensus will be established at the neutral opinion $\eta_\infty = 0$.

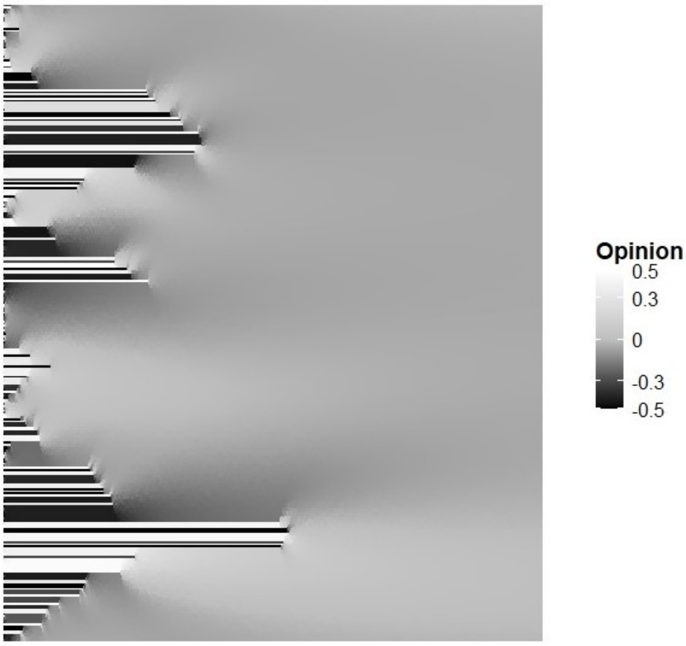

Figure 2.8: *Realizations of the Deffuant model in the one-dimensional torus with* 300 *agents and initial opinions* $\eta_0 \sim Unif(\{-0.5, -0.3, 0.3, 0.5\})$. *Parameters* $\theta = 0.6$ *and* $\mu = 0.5$. *Results after* 500 *thousands iterations.*

2.3 Extension to higher-dimensional opinion spaces

So far, all results concerning the Deffuant model assumed the dynamics over the integers. Nicolas Lanchier and Hsin-Lun Li presented one great extension of the model in the article titled "Probability of consensus in the multivariate Deffuant model on finite connected graphs" from 2020. In the joint work, they extended the social network to any finite connected graph, and the opinion space to any bounded convex subset of a normed vector space.

It goes as follows: let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a finite connected graph and $\Delta \subset \mathbb{R}^d$ be a bounded convex subset endowed with a norm $||\cdot||$. The multivariate Deffuant model is a continuous-time Markov chain whose configuration at time t is given by the function ξ_t such that:

$$\gamma_t(x) = \begin{cases} \gamma_{t^-}(x) + \mu(\gamma_{t^-}(y) - \gamma_{t^-}(x)) \text{ if } ||\gamma_{t^-}(y) - \gamma_{t^-}(x)|| \leq \tau, \\ \gamma_{t^-}(x) \text{ otherwise.} \end{cases}$$

and

$$\gamma_t(y) = \begin{cases} \gamma_{t^-}(y) + \mu(\gamma_{t^-}(x) - \gamma_{t^-}(y)) \text{ if } ||\gamma_{t^-}(x) - \gamma_{t^-}(y)|| \le \tau, \\ \gamma_{t^-}(y) \text{ otherwise.} \end{cases}$$

Their main result gives a lower bound for the probability of consensus with a wide range of applications. Before presenting it, let $\boldsymbol{d} = \sup_{a,b\in\Delta} ||a - b||$ be the diameter of the opinion space; and $\boldsymbol{c} \in \Delta$ be the center of the convex set Δ, i.e., $\sup_{a\in\Delta} ||a - \boldsymbol{c}|| = \boldsymbol{d}/2$. Furthermore, let the initial opinions γ_0 be independent and identically distributed, and let X be a random variable with distribution:

$$\mathbb{P}(X \in B) = \mathbb{P}(\gamma_0 \in B) \text{ for all } x \in \mathcal{V} \text{ and all Borel sets } \mathcal{B} \subset \Delta\,.$$

We have:

Theorem 2.3.1 (Lanchier and H. Li, 2020). *For all $\tau > \boldsymbol{d}/2$,*

$$P(C) \ge 1 - \frac{E\|X - c\|}{\tau - \boldsymbol{d}/2} \quad \textit{where} \quad C = \left\{ \lim_{t\to\infty} \sup_{x,y\in V} \|\gamma_t(x) - \gamma_t(y)\| = 0 \right\}. \tag{2.38}$$

Note that the result does not depend on the size or the topology of the network space, but rather the initial distribution, confidence threshold, and the opinion space (the norm). This is crucial as it can be used in many different scenarios, making it closer to real-world situations.

The proof was done by constructing auxiliary processes such as bounded martingales to keep track of the cumulative disagreement between opinions and other properties derived from the norm. See Lanchier and H. Li, 2020 for a complete demonstration.

To better visualize (2.39), let us present a numerical example. Let $\Delta = [0, 1]^2$ be the unit square and $X \sim Unif(\Delta)$. Consider the Euclidian norm on $\mathbb{R}^2$, then we have $d = \sqrt{2}$ and $c = (1/2, 1/2)$. Finally, it remains to compute $\mathbb{E}||X - c||$.

Without any loss, consider Δ_0 the unit square centered at the origin $(0, 0)$. Of course, the expected value of $||X - c||$ will be the same of $||X_0 - c||$ where $X_0 \sim Unif(\Delta_0)$. Hence,

$$\begin{aligned} \mathbb{E}||X_0 - c|| = \mathbb{E}||X_0|| &= \int_{\mathbb{R}^2} ||\boldsymbol{x}|| f_{X_0}(\boldsymbol{x})\, d\boldsymbol{x} \\ &= \int_{\Delta_0} ||\boldsymbol{x}|| \cdot 1\, d\boldsymbol{x} \\ &= \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \sqrt{x^2 + y^2}\, dx\, dy \\ &= \frac{1}{6}(\sqrt{2} + \ln(1 + \sqrt{2})). \end{aligned}$$

where f_{X_0} is the probability density function of X_0. Therefore,

$$P(C) \ge 1 - \frac{\frac{1}{6}(\sqrt{2} + \ln(1 + \sqrt{2}))}{\tau - \sqrt{2}/2}.$$

Figure 2.9 shows this lower bound for different values of τ. Due to the constraints of the probability measure, the plot is only exhibited for $\tau \ge \frac{1}{6}(\sqrt{2} + \ln(1 + \sqrt{2})) + \frac{\sqrt{2}}{2}$. This scenario will be more explored in chapter 4, with slight changes in the initial distribution.

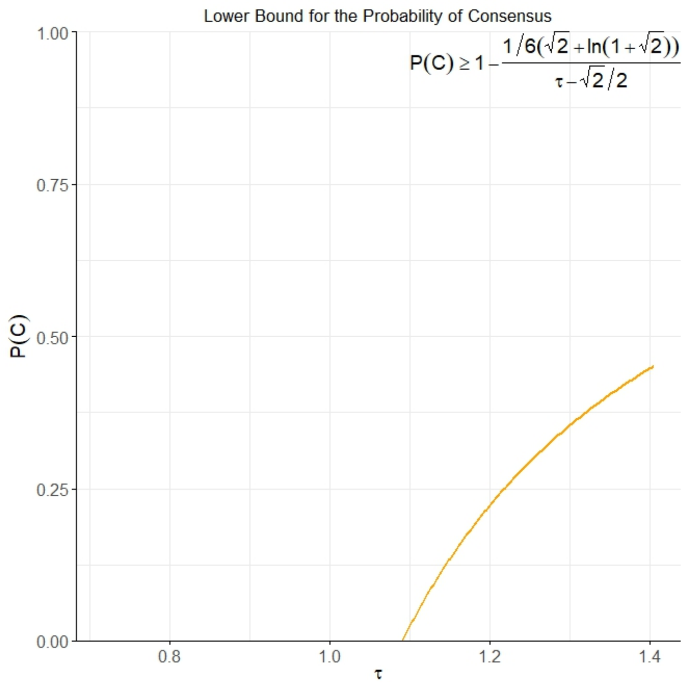

Figure 2.9: *Lower bound for probability of consensus of the multivariate Deffuant model on the unit square.*

This result can be generalized to any other square. We have:

Corollary 2.3.2. *Let X be uniformly distributed on the square with sides of length $L > 0$. Then, for all $\tau > \boldsymbol{d}/2$,*

$$P(C) \geq 1 - \frac{\frac{L}{6}(\sqrt{2} + \ln(1+\sqrt{2}))}{\tau - \boldsymbol{d}/2} \quad \textit{where} \quad C = \left\{ \lim_{t\to\infty} \sup_{x,y\in V} \|\gamma_t(x) - \gamma_t(y)\| = 0 \right\}. \tag{2.39}$$

Proof. For the demonstration, it suffices to compute $\mathbb{E}||X - c||$ for a general square with sides of length L. Under the conditions of the statement, assume without loss of generality that the square Δ is centered at the origin $(0,0)$, we have:

$$\begin{aligned}
\mathbb{E}||X - c|| = \mathbb{E}||X|| &= \int_{\mathbb{R}^2} ||\boldsymbol{x}|| f_X(\boldsymbol{x})\, d\boldsymbol{x} \\
&= \int_{\Delta} ||\boldsymbol{x}|| \cdot \frac{1}{L^2}\, d\boldsymbol{x} \\
&= \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \sqrt{x^2 + y^2} \frac{1}{L^2}\, dx\, dy \qquad (2.40) \\
&= \frac{L}{6}(\sqrt{2} + \ln(1+\sqrt{2})). \qquad (2.41)
\end{aligned}$$

To obtain the last expression, consider the region $R = \{(x, y) \mid 0 \leq y \leq x \leq L/2\}$, which represents an eighth of the total area of the square. We can re-write 2.40 as:

$$
\begin{aligned}
8\int_0^{L/2}\int_0^{x}\sqrt{x^2+y^2}\frac{1}{L^2}\,dy\,dx &= \frac{8}{L^2}\int_0^{\frac{\pi}{4}}\int_0^{\frac{L/2}{\cos\theta}} r^2\,dr\,d\theta \\
&= \frac{L}{3}\int_0^{\frac{\pi}{4}}\frac{1}{\cos^3\theta}\,d\theta \\
&= \frac{L}{3}\left(\frac{\sqrt{2}}{2}+\frac{\ln(1+\sqrt{2})}{2}\right)
\end{aligned}
$$

Where the second expression is obtained by the change of variables $x = r\cos\theta$ and $y = r\sin\theta$. □

2.4 Universality of the Critical Value

Before the formal demonstration of the critical value of the Deffuant model was published, an interesting finding was presented by Santo Fortunato in the article titled "Universality of the Threshold for Complete Consensus for the Opinion Dynamics of Deffuant et al." from 2004. In this work, Fortunato studied the Deffuant model with initial opinions independently and uniformly distributed on the interval $[0, 1]$, while keeping the parameter $\mu = 1/2$. He investigated the conditions that lead to achieving consensus in opinion for different social network structures. The numerical evidence demonstrated that, regardless of the network structure, the critical threshold for achieving consensus, denoted as ε_c, is always $1/2$.

Different from the traditional edge selection in the model, the simulation took place by selecting one agent at a time. After the first agent was chosen, a neighboring agent was randomly selected for potential interaction. The population size, denoted as N, varied across simulations, ranging from $2,500$ to $100,000$ agents. The opinions were adjusted based on the compatibility defined by the confidence bound parameter ε. To verify if consensus was achieved, the program would stop if, after an iteration, no agent changed their opinion by more than 10^{-9}. Once the system reached its final configuration, the program checked if the agents belonged to the same cluster. Finally, the fraction of samples with agents in a single opinion cluster, denoted as P_c, was evaluated as a function of ε.

The results were presented for four graph structures:

- Complete graph;
- Square lattice;
- Random graph *à la* Erdös-Rényi;
- Scale-free graph *à la* Barabási-Albert.

Figure 2.10 shows the result for a society with $10,000$ agents. The change in behavior is pronounced at $\varepsilon \approx 0.46$, where the value of P_c starts increasing until it reaches 1 at $\varepsilon = 0.5$.

A similar scenario happens with the square lattice with periodic boundary conditions, for different values of N in Figure 2.11. Both results suggest that in the limit $N \to \infty$, P_c

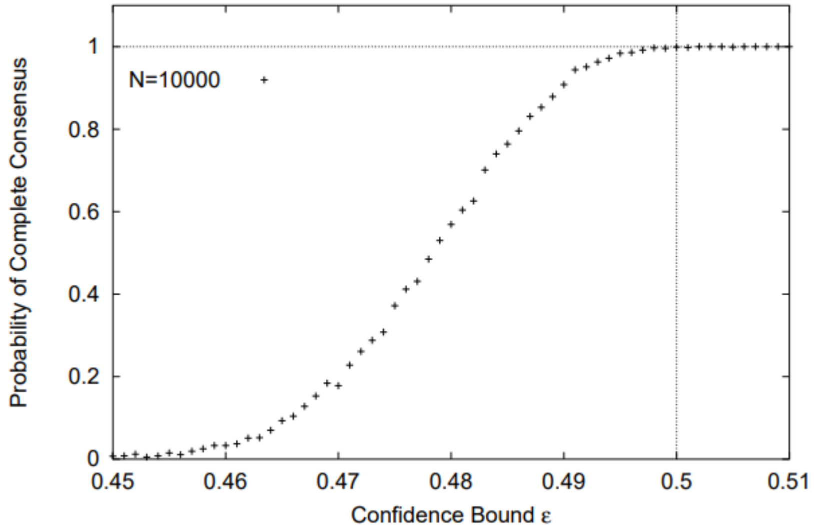

Figure 2.10: *Fraction of samples with a single cluster of opinions in the final configuration in the complete graph.* ***Source:*** *FORTUNATO, 2004.*

converges to a step function, in which $P_c = 0$ for $\varepsilon < 0.5$ and $P_c = 1$ for $\varepsilon > 0.5$.

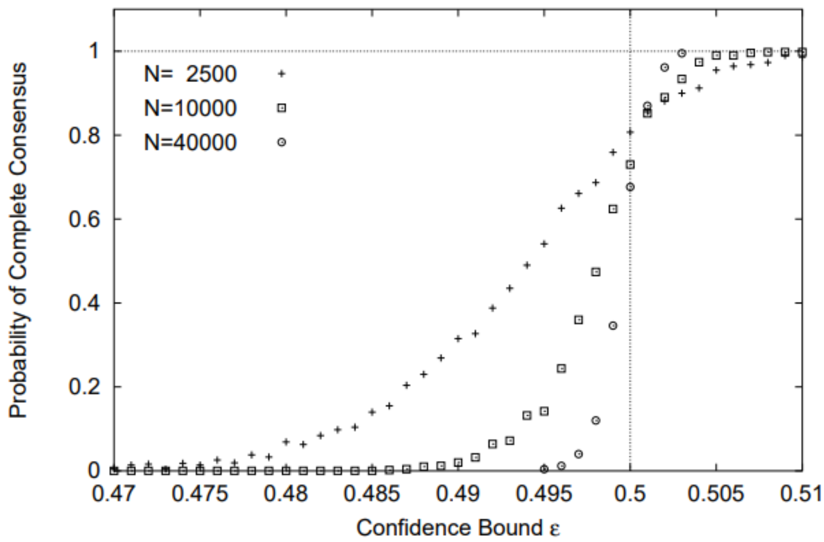

Figure 2.11: *Fraction of samples with a single cluster of opinions in the final configuration in the square lattice for different values of N.* ***Source:*** *FORTUNATO, 2004.*

To approximate the exploration to more realistic situations, Fortunato also examined the model for random networks. He considered two examples of random graphs: Erdös-Rényi and Barabási-Albert.

The Erdős-Rényi random graph (ERDÖS and RÉNYI, 1959), represented in Figure 2.12, is constructed in the following manner: given a set of vertices, each pair has a probability $p \in (0, 1)$ of being connected. The total number of edges m is given by $p\binom{N}{2}$. Furthermore, the expected degree of the graph, which is the expected number of connections for each agent, is given by $k = p(N-1)$. When $N \to \infty$, this result can be approximated to $k \approx pN$.

For the realizations, the graphs had the same average degree $k = pN = 400$ for different numbers of agents N, varying from $40,000$ to $100,000$. Figure 2.12 show the result for this framework.

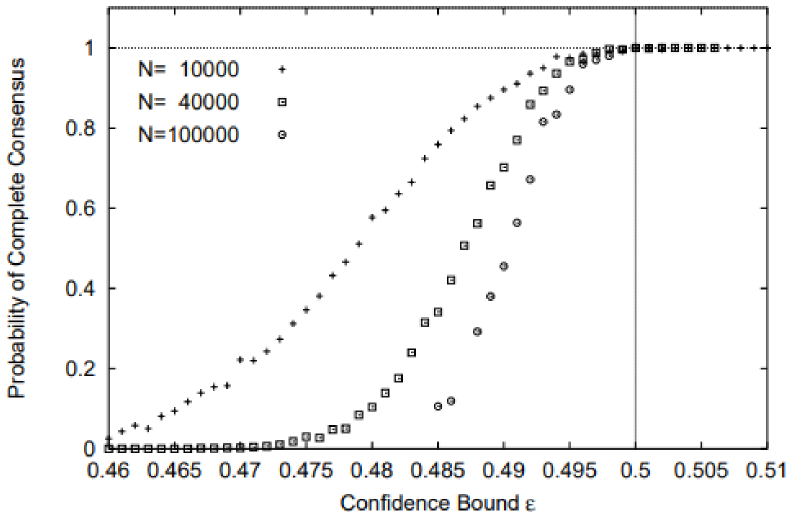

Figure 2.12: *Fraction of samples with a single cluster of opinions in the final configuration in the Erdös-Rényi random graph for different values of N.* ***Source:*** *FORTUNATO, 2004.*

Finally, the scale-free graph *à la* Barabási-Albert (ALBERT and BARABÁSI, 2002), represented in Figure 2.13. To construct this graph, the algorithm goes as follows: fix $m > 0$, the number of edges originating from a node, and start the procedure with $m_0 \geq m$. In this case, the procedure started with m connected nodes and added one node at each step until the total of N was achieved. Each node added is connected to m sampled preexisting nodes, so the probability of being connected to a node is proportional to the number of its neighbors. In all graphs generated, the degree distribution was a power law of the form $P(k) \sim k^{-3}$, independently of m, chosen to be $m = 3$. Figure 2.13 shows the results. Again, the same pattern was observed. For a broader exploration of this model see STAUFFER and MEYER-ORTMANNS, 2003.

The simulations reveal a consistent pattern in the Deffuant model: regardless of the social network, the critical value of the confidence threshold is always $1/2$. The author suggested that this would also hold true for theoretical proofs based on the model's dynamics.

By examining the opinion change, the extremists' opinions would fall within the ranges $[0, \varepsilon]$ and $[1 - \varepsilon, 1]$. After iterations, because of the symmetry of the model, the agents at the extreme ends would have opinions around $\varepsilon_l \approx \varepsilon/2$ and $\varepsilon_r \approx 1 - \varepsilon/2$. For $\varepsilon < 1/2$, $\varepsilon_r - \varepsilon_l$ would be greater than ε. Conversely, when $\varepsilon > 1/2$, the opposite holds true, leading to a formation of a single cluster. This argument is independent of the underlying social topology.

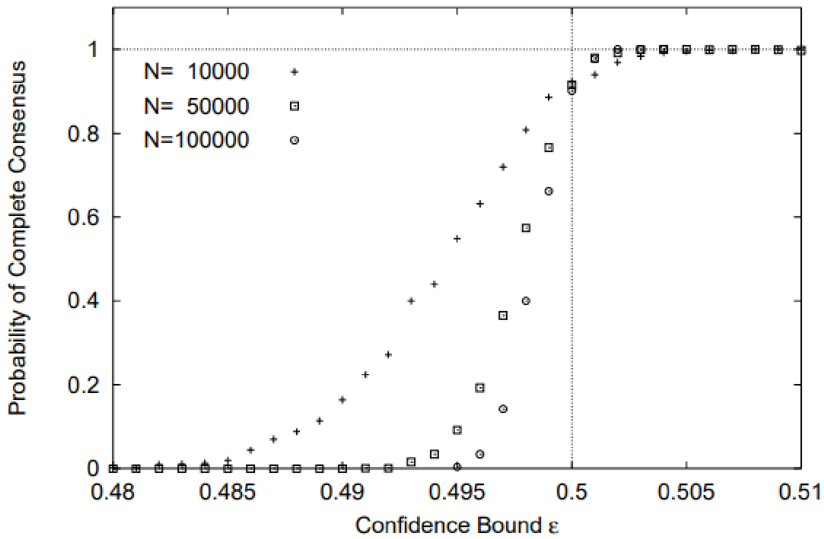

Figure 2.13: *Fraction of samples with a single cluster of opinions in the final configuration in the Barabási-Albert graph for different values of N.* ***Source:*** *Fortunato, 2004.*

Chapter 3

The Role of Moderates and Extremists on Consensus Formation

The Deffuant model describes consensus formation where pairs of agents interact only if their opinions differ by no more than a given threshold, leading to an approximation of their beliefs. Traditionally, this model is defined on a graph $G = (\mathbb{Z}, E)$ with $E = \{(x, x+1) : x \in \mathbb{Z}\}$ and initial opinions uniformly distributed on the unit interval. In this chapter, we change the model to work with different initial distributions, to assess the role of moderates — agents whose opinions lie within the opinion space — in addition to extremists — positioned at the boundaries of the opinion space — on consensus formation. The analysis is conducted on one and two-dimensional social networks, as well as the complete graph, using both unidimensional and bidimensional opinion spaces. Our results suggest that consensus rapidly occurs once a critical proportion of moderates is reached.

3.1 Critical value of moderates in one-dimensional opinion space on the complete graph

The article "Opinion dynamics and consensus formation in the Deffuant model with extremists and moderates in the population," written by L. Marconi and F. Cecconi in 2020, explores an extension of the classical Deffuant model by assuming the presence of moderates and extremists within the population. Moderates are those with opinions in the open interval $(0, 1)$, while extremists have opinions at the extreme values of 0 or 1. Through computational methods, the authors reported that moderates are crucial to the emergence of consensus: when a critical number is reached, opinions progressively converge.

To understand the importance of moderates, let us analyze an encounter between agents u and v. If, for any given time $t > 0$, $\eta_t(u) = 0$ and $\eta_t(v) = 1$, then regardless of the values of $\theta \in (0, 1)$ or $\mu \in (0, 1/2]$, their interaction will not result in a compromise because it will be blocked by the confidence threshold. Therefore, it is clear that having intermediate opinions is essential for reaching a consensus.

To delve into the role of the moderates, the authors considered four different scenarios, regarding the values of the parameters μ and θ, described in Figure 3.1.

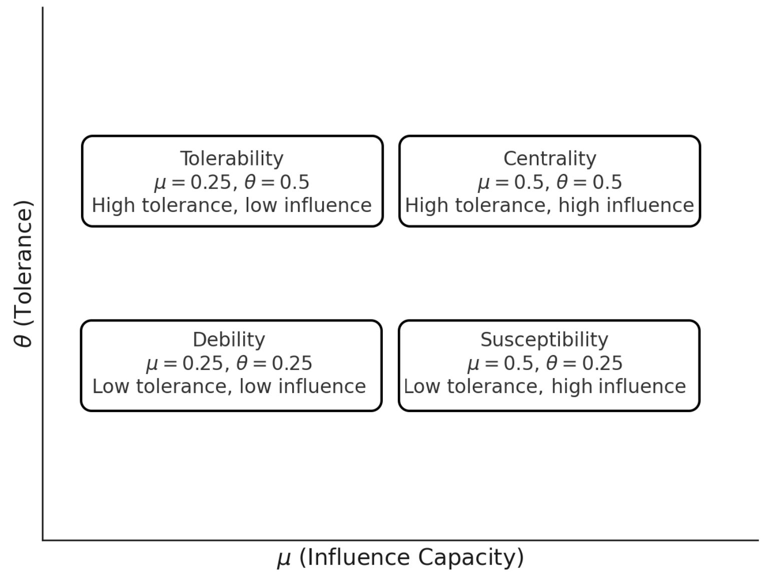

Figure 3.1: *Configurations of the Deffuant model on the cartesian plane.*

The realizations were performed through NetLogo and the algorithm started with $N = 500$ agents equally distributed between extremists over the full graph. Then, varied the number of moderates and recorded results. For each configuration, it performed 10.000 iterations and 50 runs to store the results. To extend the experiment to other social networks, we reproduced the experimental setup in Python, summarized in Table 3.1.

Minimum number of moderates	Maximum number of moderates	Steps	Parameters
50	350	1	$\theta = 0.5$ $\mu = 0.5$
50	350	1	$\theta = 0.25$ $\mu = 0.25$
50	350	1	$\theta = 0.5$ $\mu = 0.25$
50	350	1	$\theta = 0.25$ $\mu = 0.5$

Table 3.1: *Number of moderates and runs for each configuration of the Deffuant model on the complete graph.*

Furthermore, the primary information collected was:

- Standard deviation of opinions;
- Mean of the opinions;
- Number of agents with opinions 0 and 1 over the population;
- Proportion of agents with opinions lower than 0.25 and greater than 0.75 over the population.

- Maximum distance between agent's opinions.

The following plots show the results, which are consistent with those of the original study (Marconi and Cecconi, 2020) — except for the last plot, which was not originally considered.

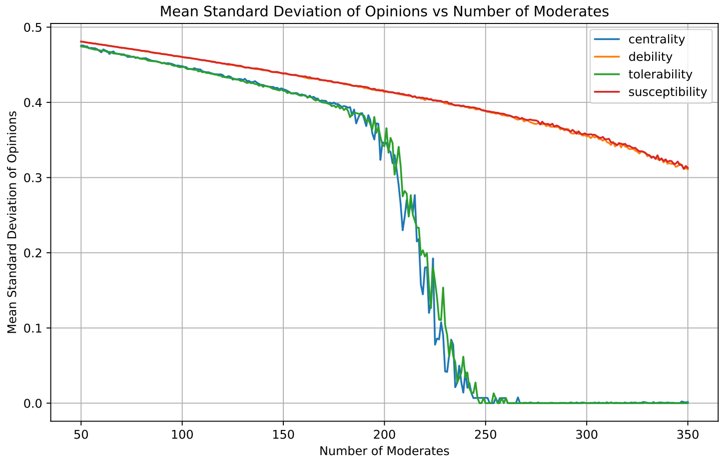

Figure 3.2: *Mean standard deviation of opinions for different scenarios of the Deffuant model with extremists on the complete graph.*

Figure 3.2 shows the standard deviation for different values of μ and θ. The scenarios with $\theta = 0.25$, namely *susceptibility* and *debility*, showed no critical value for the number of moderates. In contrast, when $\theta = 0.5$, *centrality* and *tolerability*, a rapid decrease of the mean of the standard deviation was observed when the number of moderates reached 200. Also, it reached zero when the population was composed of half moderates. Therefore, in these specific configurations for μ and θ, the critical value of moderates, which progressively reduce the distance of opinions from the mean, is 200.

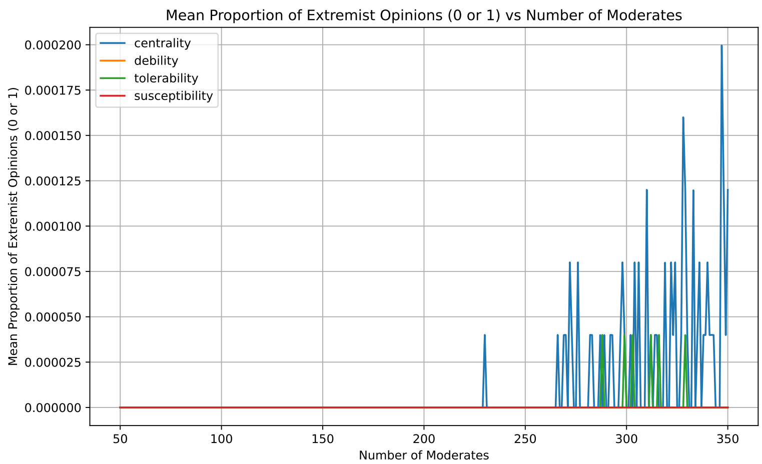

Figure 3.3: *Mean of the proportion of agents with opinion* 0 *and* 1 *for different scenarios in the Deffuant model with extremists on the complete graph.*

Figure 3.3 shows that, besides the high tolerance ($\theta = 0.5$), the presence of extremists is almost none in the population, with small sporadic peaks around 300 moderates.

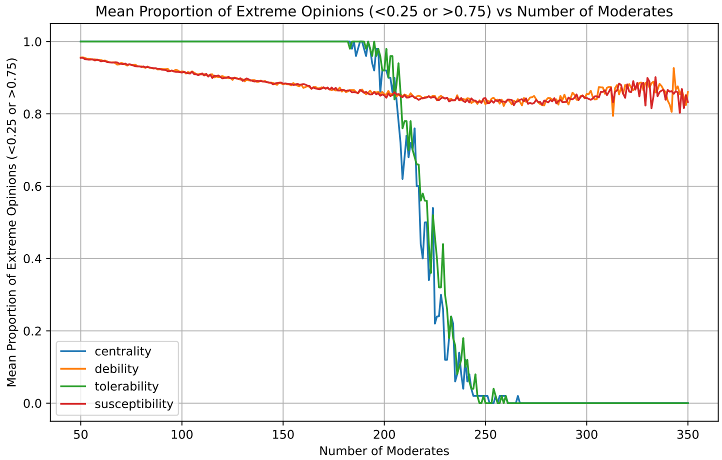

Figure 3.4: *Mean of the proportion of agents with opinion lower than* 0.25 *and higher than* 0.75 *for different scenarios in the Deffuant model with extremists on the complete graph.*

Finally, the proportion of agents with opinions lower than 0.25 and higher than 0.75 follows along with what was seen with the standard deviation. In configurations with a higher tolerance, there is a rapid decrease when the number of moderates is 200. Also, the proportion reaches zero when the population is equally divided between moderates and extremists. However, when $\theta = 0.25$, no such behavior was observed.

Another important data obtained was the maximum distance between agents. This helps us determine when the model reached compatibility (see section 2.2), which is when every agent can interact with any other agent. Figure 3.4 suggests that compatibility is reached around the mean with half of the population as moderates. Below, figure 3.5 confirms the scenario.

The results indicate that, in populations with a high tolerance to interact (θ fixed by at least 0.5), 40% of the population as moderates (200 out of 500) can significantly impact the speed of achieving agreement. Furthermore, the data collected showed that when the number of moderates reaches 50%, compatibility is reached, meaning that there are no barriers for interaction to occur, except those depending on the social network. In other words, no interaction is censored by θ, which is a great achievement for any democratic population. In opposition, no such behavior was observed when the tolerance parameter was 0.25.

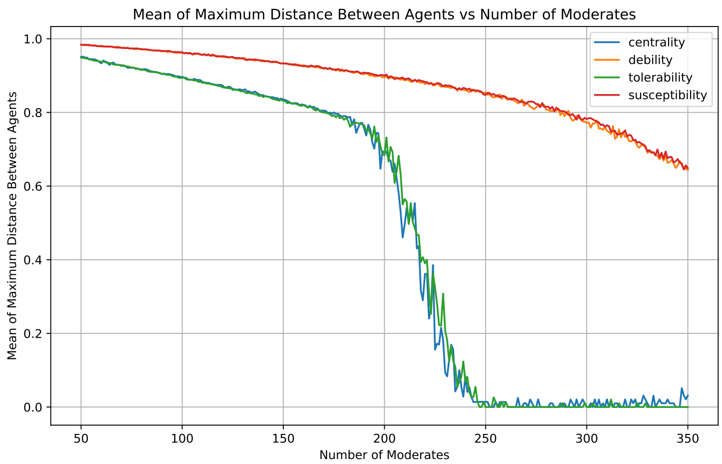

Figure 3.5: *Mean of the maximum distance between agent's opinions for different scenarios of the Deffuant model with extremists on the complete graph.*

3.2 Critical value of moderates in one-dimensional opinion space on the one and two-dimensional lattices

Based on the work of Marconi and Cecconi, 2020 we extend the Deffuant model with extremists to both the integer lattice and the square lattice, incorporating periodic boundary conditions. In this model, N agents are considered, with their opinions distributed between extremists (agents with opinions 0 or 1) and moderates (agents with opinions within the interval $(0, 1)$).

With this extension, we aim to determine the critical quantity of moderates necessary for consensus formation in different social networks. By analyzing the behavior of the model on $\mathbb{Z}$ and $\mathbb{Z}^2$, we can assess the robustness of the results discussed in chapter 2. In other words, how much can we change the initial distribution to maintain the convergence of the model? Is there a critical quantity of moderates that leads the population towards agreement?

The experiments were conducted in Python. Four configurations, introduced in the previous section, in figure 3.1, were considered: *debility, tolerability, susceptibility, centrality*. For each configuration, 10.000 iterations and 50 runs were performed to store the results. The data collected were: standard deviation of opinions, mean of the opinions, number of agents with opinions 0 and 1 over the population, proportion of agents with opinions lower than 0.25 and greater than 0.75 over the population, and the maximum distance between agent's opinions.

3.2.1 Integer lattice

For the integer lattice, the total number of agents was fixed at 500 and we varied the number of moderates through the realizations. The experiments conducted are summarized in Table 3.2.

Minimum number of moderates	Maximum number of moderates	Steps	Parameters
0	500	1	$\theta = 0.5$ $\mu = 0.5$
0	500	1	$\theta = 0.25$ $\mu = 0.25$
0	500	1	$\theta = 0.5$ $\mu = 0.25$
0	500	1	$\theta = 0.25$ $\mu = 0.5$

Table 3.2: *Number of moderates and runs for each configuration of the Deffuant model on the one-dimensional lattice.*

In addition, Algorithm 1 shows the initial setup to run the model.

Algorithm 1 Initialization and Neighbor Computation for a One-Dimensional Integer Lattice

```
Input: line size L, number of moderates M
Output: Initialized opinions, neighbor relationships
procedure InitializeOpinions(L, M)
    N ← L                                         ▷ Total number of agents
    E ← N − M                                     ▷ Number of extremists
    E_0 ← ⌊E/2⌋                                   ▷ Number of extremists with opinion 0
    E_1 ← ⌈E/2⌉                                   ▷ Number of extremists with opinion 1
    Initialize E_0 agents with opinion 0
    Initialize E_1 agents with opinion 1
    Initialize M agents with opinions uniformly distributed in (0, 1)
    Combine all initialized opinions into a single list
    Randomly shuffle the opinions
    return opinions
end procedure
procedure PrecomputeNeighbors(L)
    N ← L                                         ▷ Total number of agents
    Initialize an array neighbors of size N × 2
    for each index i ← 0 to N − 1 do
        if i = 0 then
            neighbors[i] ← [1, 1]                 ▷ Only one neighbor at the beginning of the line
        else if i = N − 1 then
            neighbors[i] ← [N − 2, N − 2]         ▷ Only one neighbor at the end of the line
        else
            neighbors[i] ← [i − 1, i + 1]         ▷ Neighbors in both directions
        end if
    end for
    return neighbors
end procedure
```

Since the number of moderates varies by steps of one, the remaining agents are almost equally divided among the group of extremists, using the floor division.

The dynamics occurred over $\mathbb{Z}$. This social network is more restrict since one agent can be forever stuck, if its only two neighbors are distant more than θ, regarding the opinion value. Therefore, it is expected a more linear behavior towards consensus, when the number of moderates increases.

Results

The mean of the Standard deviation presented in Figure 3.6 shows a slow decrease over the number of moderates. For the configurations with $\theta = 0.5$, it approximates zero as the number of moderates approximates the maximum value 500. Additionally, the configurations with $\theta = 0.25$ show similar behavior, however, it stays around 0.3, maintaining a wider range of opinions in the population.

Moreover, the higher rate change for *centrality* and *tolerability* occurs at $300 - 400$ of moderates, indicating that at 300 moderates, the population is set to achieve consensus faster.

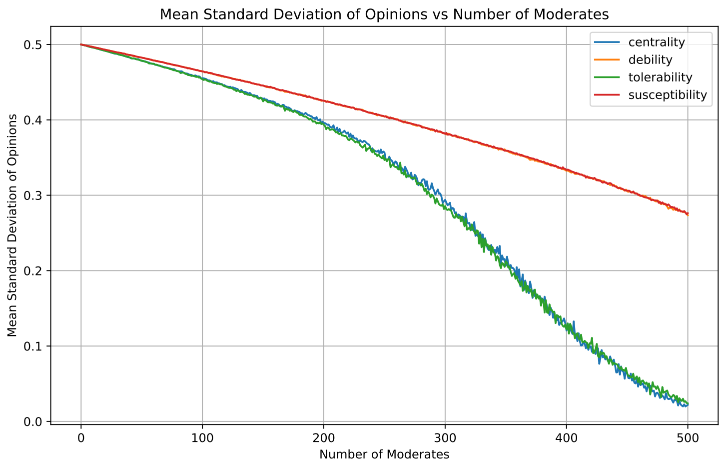

Figure 3.6: *Mean of the standard deviation of opinions in the Deffuant model on the one-dimensional lattice with* 500 *agents.*

The mean proportion of extremist opinions over the population (Figure 3.7) shows a slow decrease for all configurations. Unlike the full graph, extremists remain present in the population, only approximating zero when the number of moderates is closer to its maximum.

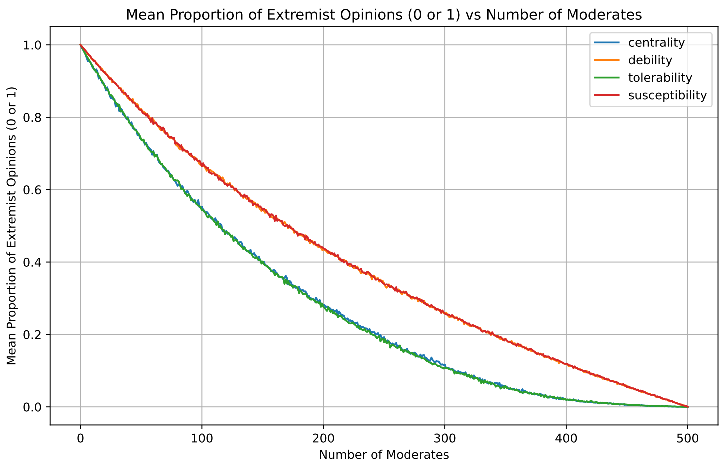

Figure 3.7: *Mean proportion of agents with opinions 0 or 1 over the population in the Deffuant model on the one-dimensional lattice with* 500 *agents.*

Similarly, the mean proportion of extreme opinions over the population (Figure 3.8) does not show a rapid decrease. Although for the configurations *centrality* and *tolerability* it approximates 0 faster, compatibility is only reached when the population is closer to the moderates-only state.

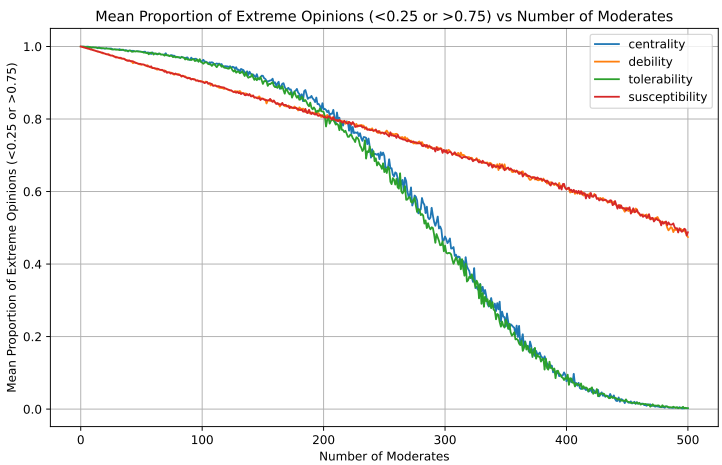

Figure 3.8: *Mean proportion of agents with opinions greater than* 0.75 *or less than* 0.25 *over the population in the Deffuant model on the one-dimensional lattice with* 500 *agents.*

Finally, the mean of maximum distance between agents over the population (Figure 3.9). It confirms what was seen in the last three plots, regarding the presence of extremists in the population. Only at 400 moderates, the quantity of extremists starts decreasing for the configurations with $\theta = 0.5$. Conversely, for $\theta = 0.25$, the distance remained at one (the maximum) over the population.

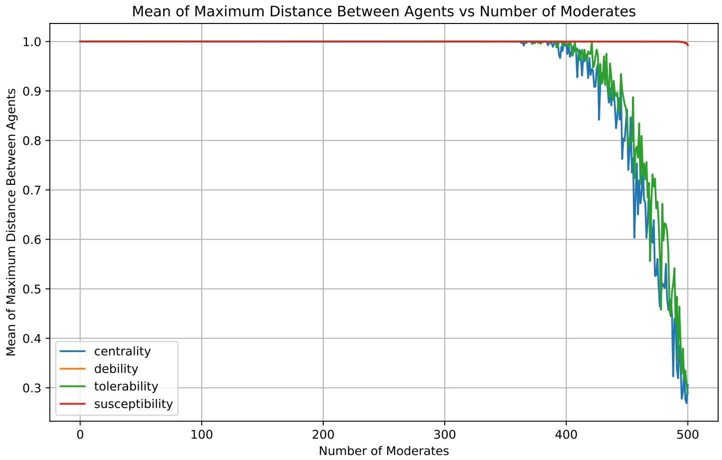

Figure 3.9: *Mean of the maximum distance between agent's opinions in the Deffuant model on the one-dimensional lattice with* 500 *agents.*

Therefore, the data suggests 300 as the critical value of moderates that progressively set the population towards agreement in the Deffuant model on the one-dimensional lattice. However, even with the acceleration at this proportion of moderates, the population slowly approximates consensus. This is a consequence of the great restriction of the social network. Indeed, each agent has only two neighbors, except for the first and last agent on the line, which makes it difficult to approach consensus. For instance, one can be easily trapped, if surrounded by neighbors with very different opinions. Additionally, compatibility was not reached, but only with 480 moderates.

3.2.2 Square lattice (torus)

For the two-dimensional torus, the total number of agents was fixed at 900 and we varied the number of moderates through the realizations. The experiments conducted are summarized in Table 3.3.

Minimum number of moderates	Maximum number of moderates	Steps	Parameters
0	600	1	$\theta = 0.5$ $\mu = 0.5$
0	600	1	$\theta = 0.25$ $\mu = 0.25$
0	600	1	$\theta = 0.5$ $\mu = 0.25$
0	600	1	$\theta = 0.25$ $\mu = 0.5$

Table 3.3: *Number of moderates and runs for each configuration of the Deffuant model on the two-dimensional lattice.*

The initial setup are illustrated in Algorithm 2. The experimental setup is similar to the previous one, with a major change in neighbor computation, since each agent has four neighbors on the two-dimensional lattice.

Algorithm 2 Initialization and Neighbor Computation for a Two-Dimensional Lattice

Input: grid size G, number of moderates M
Output: Initialized opinions, neighbor relationships
procedure InitializeOpinions(G, M)
    $N \leftarrow G^2$ ▷ Total number of agents
    $E \leftarrow N - M$ ▷ Number of extremists
    $E_0 \leftarrow \lfloor \frac{E}{2} \rfloor$ ▷ Number of extremists with opinion 0
    $E_1 \leftarrow \lfloor \frac{E}{2} \rfloor$ ▷ Number of extremists with opinion 1
    Initialize E_0 agents with opinion 0
    Initialize E_1 agents with opinion 1
    Initialize M agents with opinions uniformly distributed in $(0, 1)$
    Combine all initialized opinions into a single list
    Randomly shuffle the opinions
    return opinions
end procedure
procedure PrecomputeNeighbors(G)
    $N \leftarrow G^2$ ▷ Total number of agents
    Initialize an array $neighbors$ of size $N \times 4$ ▷ 4 neighbors per agent
    for each index $i \leftarrow 0$ to $N - 1$ **do**
        $row \leftarrow \lfloor \frac{i}{G} \rfloor$ ▷ Row of the agent on the grid
        $col \leftarrow i \mod G$ ▷ Column of the agent on the grid
        $down\_neighbor \leftarrow ((row + 1) \mod G) \times G + col$ ▷ Down neighbor (wraps around)
        $up\_neighbor \leftarrow ((row - 1 + G) \mod G) \times G + col$ ▷ Up neighbor (wraps around)
        $right\_neighbor \leftarrow row \times G + (col + 1) \mod G$ ▷ Right neighbor (wraps around)
        $left\_neighbor \leftarrow row \times G + (col - 1 + G) \mod G$ ▷ Left neighbor (wraps around)
        $neighbors[i] \leftarrow [down\_neighbor, up\_neighbor, right\_neighbor, left\_neighbor]$
    end for
    return neighbors
end procedure

In comparison with the one-dimensional lattice, this setup is closer to the complete graph, since each agent has four neighbors. Therefore, it is expected a behavior similar to the original study, regarding consensus formation.

Results

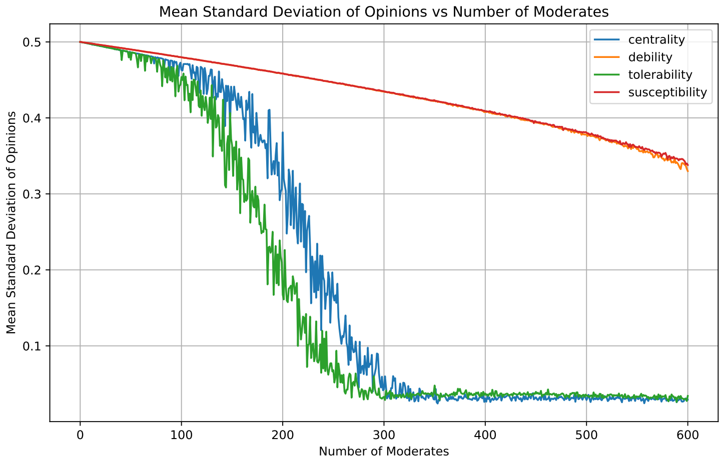

Figure 3.10: *Mean of the standard deviation of opinions in the Deffuant model on the two-dimensional lattice with* 900 *agents.*

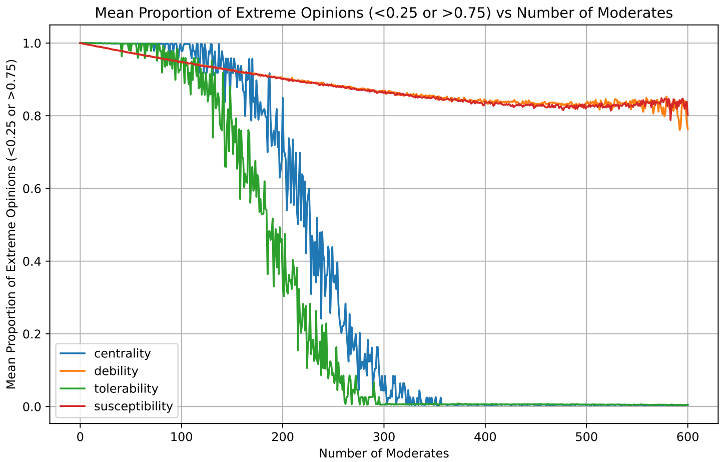

Figure 3.11: *Mean proportion of agents with opinions greater than* 0.75 *or less than* 0.25 *over the population in the Deffuant model on the two-dimensional lattice with* 900 *agents.*

Figures 3.10 and 3.11 suggest a critical value of moderates around 100, which represents 1/9 of the population. Indeed, both the mean of the standard deviation and the mean proportion of agents with opinions greater than 0.75 or less than 0.25 start decreasing at 100 moderates, for the configurations with $\theta = 0.5$. Furthermore, when the number of moderates reaches 300, 1/3 of the population, compatibility is achieved: every agent can interact with any other agent. The same behavior was not observed for populations with lower tolerance.

Interestingly, the value of moderates to achieve criticality was lower, in percentage, than the original study (Marconi and Cecconi, 2020). Although this criterion was not considered in the study, from the results it is clear that on the complete graph, compatibility was reached at 250 moderates, which represented 50% of the population. On the two-dimensional torus, approximately 33% of moderates (300 out of 900) were necessary.

In addition, it is important to highlight that the mean standard deviation of the configuration *tolerability* ($\mu = 0.25$) reached zero faster than *centrality*, which has $\mu = 0.5$. This outcome may be from the inherent randomness of the simulations, but it also suggests that, in this type of social network, smaller values of μ cause smoother updates which could lead to smaller variations in the standard deviation. Moreover, the higher rate change for *centrality* occurs at $200 - 300$, whereas for *tolerability* it occurs at $100 - 200$ moderates.

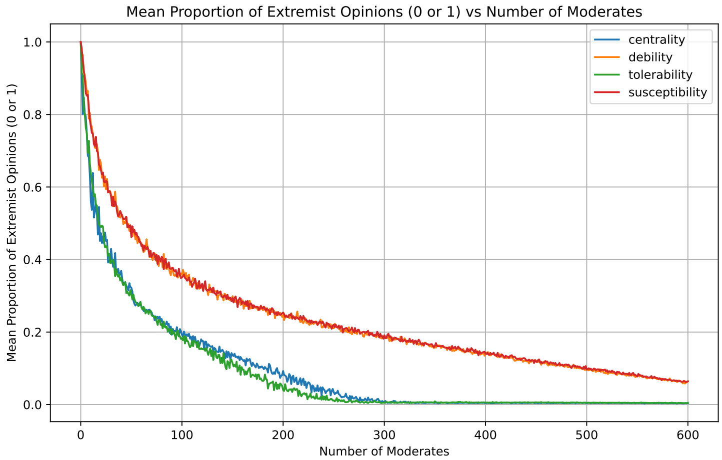

Figure 3.12: *Mean proportion of agents with opinions 0 or 1 over the population in the Deffuant model on the two-dimensional lattice with* 900 *agents.*

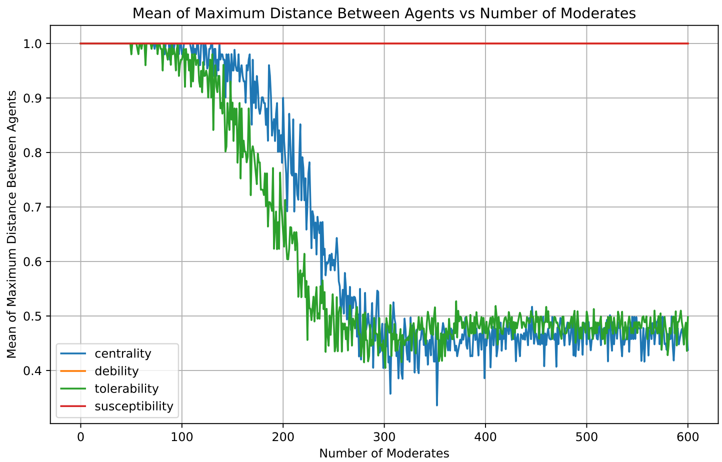

Figure 3.13: *Mean of the maximum distance between agent's opinions in the Deffuant model on the two-dimensional lattice with* 900 *agents.*

Similarly, Figures 3.12 and 3.13 show the decrease of extremists, as the number of moderates reaches 300. Additionally, the mean of the maximum distance between agents does not reach zero, which confirms the presence of a wider range of opinions - but not greater than θ for the configurations *centrality* and *tolerability*. For the configurations with $\theta = 0.25$, the presence of extremists decreased to almost 10% of the population, when the model started with 600 moderates.

The data collected indicates two important values for moderates regarding consensus formation in the population. For configurations with higher tolerance, the critical value of moderates is 100 for *tolerability* and 200 for *centrality*, indicating that even a small number of moderates can have a significant impact on the population. The second important value is 300: when the population consists of 33% moderates, compatibility is achieved. Conversely, for configurations with $\theta = 0.25$, no critical value was observed.

3.3 Critical value of moderates in two-dimensional opinion space on the complete graph

In this section, we extend the traditional Deffuant model to work on the complete graph with bidimensional opinions. Again, we aim to asses the importance of moderates and extremists in consensus achieving.

The dynamics will follow what was described in section 2.3, in equation (3.3), with discrete time (iterations):

$$\gamma_t(u) = \begin{cases} \gamma_{t-1}(u) + \mu(\gamma_{t-1}(v) - \gamma_{t-1}(u)) \text{ if } \|\gamma_{t-1}(v) - \gamma_{t-1}(u)\| \leq \tau, \\ \gamma_{t-1}(u) \text{ otherwise.} \end{cases}$$

and

$$\gamma_t(v) = \begin{cases} \gamma_{t-1}(v) + \mu(\gamma_{t-1}(u) - \gamma_{t-1}(v)) \text{ if } \|\gamma_{t-1}(u) - \gamma_{t-1}(v)\| \leq \tau, \\ \gamma_{t-1}(v) \text{ otherwise.} \end{cases}$$

where $\|\cdot\|$ represents the Euclidean norm on $\mathbb{R}^2$ and $\gamma_t(u)$ the opinion of agent u at time t.

We classify individuals holding opinions at the corners of the unit square $(0,0), (0,1), (1,0), (1,1)$ as extremists. Conversely, moderates will hold opinions within the square. The areas around the four corners of the square, which represent extreme opinions, are illustrated in Figure 3.14 below.

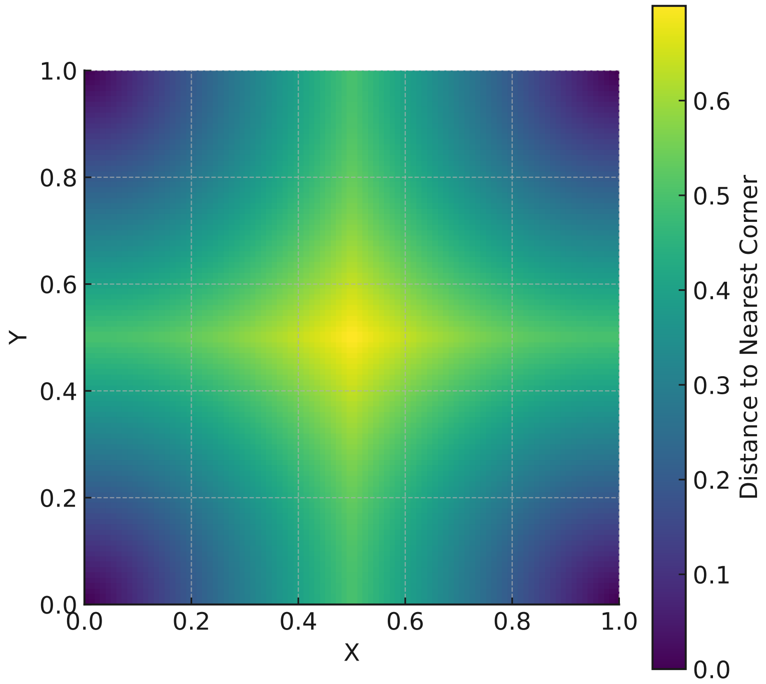

Figure 3.14: *Color-coded grid representing the distribution of opinions in the unit square for the multivariate Deffuant model. The intensity of the color indicates the proximity to the extreme regions located at the corners.*

For the complete graph, three scenarios were considered, based on the maximum distance of the opinion space d, which on the unit square is the length of its diagonal $\sqrt{2}$. Based on the work of Marconi and Cecconi (Marconi and Cecconi, 2020), we fixed μ at 0.5 and varied τ as follows:

- *Debility:* Low tolerance configuration. Here agents interact only with neighbors that are distant not more than the fourth of the maximum distance. Namely, $\tau = d/4$.
- *Centrality:* Medium tolerance. In this scenario, $\tau = d/2$: interactions occur up until half the maximum distance.
- *Flexibility:* High tolerance. For this configuration, $\tau = 3d/4$, which is approximately 1. Therefore, even with an extremist-only population, interaction can still occur from vertex to vertex.

Table 3.4 summarizes the experiments conducted.

In addition, Algorithm 3 shows the initialization. Here, since the opinions are bidimensional, four groups of extremists were considered. And, as the number of moderates varied, the remaining agents were almost equally divided using the floor function.

Minimum number of moderates	Maximum number of moderates	Steps	Parameters
0	500	1	$\theta = d/4$ $\mu = 0.5$
0	500	1	$\theta = d/2$ $\mu = 0.5$
0	500	1	$\theta = 3d/4$ $\mu = 0.5$

Table 3.4: *Number of moderates and runs for each configuration of the multivariate Deffuant model on the complete graph.*

Algorithm 3 Initialization of Opinions and Social Network for Complete Graph

```
Input: Total number of agents N, number of moderates M
Output: Initialized opinions, complete graph
procedure InitializeOpinions(N, M)
    E00 ← ⌊N/4 − M/4⌋                          ▷ Number of extremists at (0, 0)
    E01 ← ⌊N/4 − M/4⌋                          ▷ Number of extremists at (0, 1)
    E10 ← ⌊N/4 − M/4⌋                          ▷ Number of extremists at (1, 0)
    E11 ← ⌊N/4 − M/4⌋                          ▷ Number of extremists at (1, 1)
    Initialize E00 agents with opinion (0, 0)
    Initialize E01 agents with opinion (0, 1)
    Initialize E10 agents with opinion (1, 0)
    Initialize E11 agents with opinion (1, 1)
    Initialize M agents with opinions uniformly distributed in (0, 1)^2
    Combine all initialized opinions into a single list opinions
    Randomly shuffle the opinions
    return opinions
end procedure
procedure CreateCompleteGraph(N)
    Initialize a N × N adjacency matrix A      ▷ Complete graph represented by adjacency matrix
    for each pair of agents i, j where i ≠ j do
        Set A[i][j] ← 1                        ▷ Each agent is connected to every other agent
    end for
    return A                                   ▷ Return adjacency matrix of complete graph
end procedure
```

Before collect data, it performed 25.000 iterations and 50 runs for each variation of moderates. The data collected follows what was done in the previous section, with the necessary adaptations caused by the space and metric changes:

- Standard deviation of opinions;
- Mean of the opinions;
- Number of agents with opinions $(0,0), (0,1), (1,0)$ and $(1,1)$ over the population;
- Proportion of agents with opinions outside the central circle with radius $r = \frac{1}{2\sqrt{2}}$;
- Maximum distance between agent's opinions.

The circle centered at $(0.5, 0.5)$ with radius $r = \frac{1}{2\sqrt{2}}$ is an attempt to find equivalent data as the proportion of agents with opinion less than 0.25 and higher than 0.75 for the Deffuant model with unidimensional opinion. Indeed, the circle has a diameter of $d/2 = \tau$, which allows us to check when the model achieved compatibility for the configuration *centrality* — similar to the proportion outside the central interval, which has a length $1/2 = \theta$.

Results

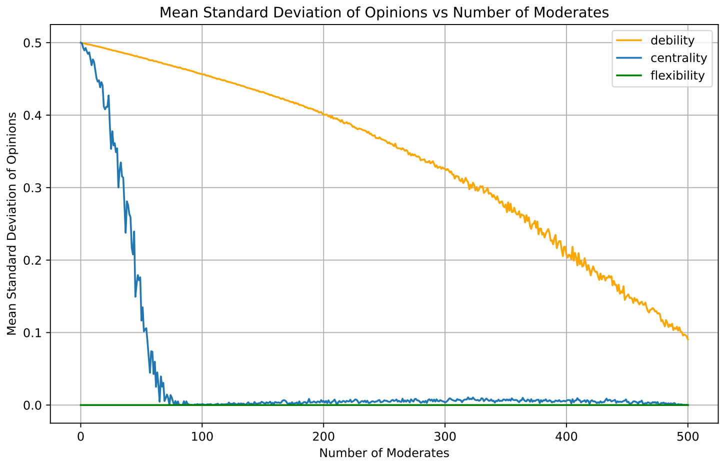

Figure 3.15: *Mean of the standard deviation of opinions in the multivariate Deffuant model on the complete graph with* 500 *agents.*

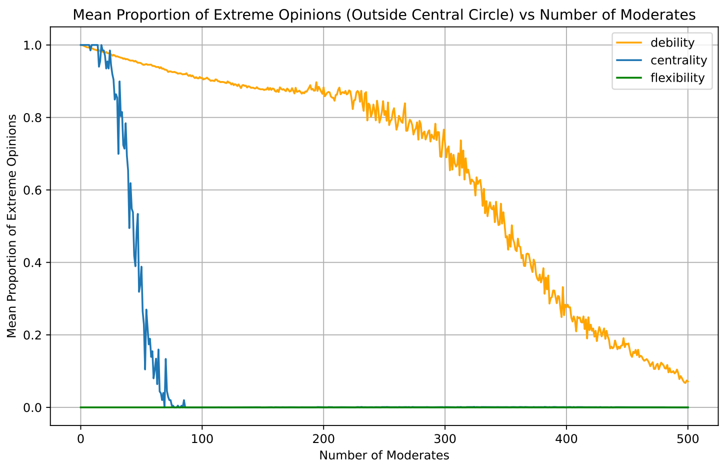

Figure 3.16: *Mean proportion of agents with opinions outside the central circle with* $r = \frac{1}{2\sqrt{2}}$ *in the multivariate Deffuant model on the complete graph with* 500 *agents.*

Figures 3.15 and 3.16, respectively the mean of the standard deviation of opinions and the mean proportion of agents with opinion outside the central circle with $r = \frac{1}{2\sqrt{2}}$, show similar behavior, especially for the configuration *centrality*. Both measures start decreasing as moderates are inserted in the population, reaching zero when the number of moderates approaches 100. These data indicate that compatibility is reached around 50 moderates and consensus is rapidly achieved when the population has medium tolerance.

For the configuration with $\tau = d/4$, the decay of both statistics are slower. The mean standard deviation approaches zero as the number of moderates approaches the maximum number of agents 500. The proportion of agents outside the central circle follows the same, with some fluctuations. Conversely, when the population has a high tolerance, convergence is reached at any quantity of moderates.

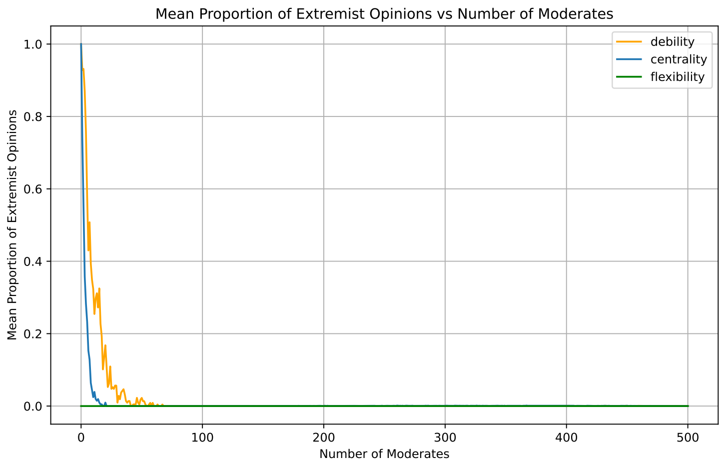

Figure 3.17: *Mean proportion of agents with opinions* $(0, 0), (0, 1), (1, 0), (1, 1)$ *over the population in the multivariate Deffuant model on the complete graph with* 500 *agents.*

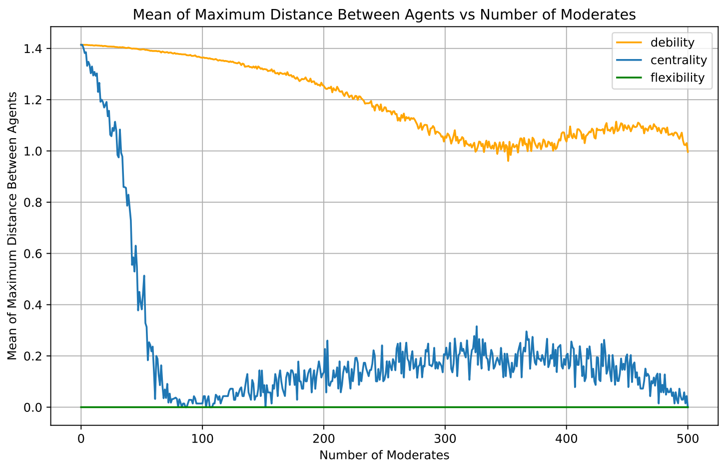

Figure 3.18: *Mean of the maximum distance between agent's opinions in the multivariate Deffuant model on the complete graph with* 500 *agents.*

In addition, Figures 3.17 and 3.18, the mean proportion of agents with opinions $(0, 0), (0, 1), (1, 0), (1, 1)$ and the mean of the maximum distance between agents respectively, suggest that extremists rapidly vanish from the population in all three configurations. However, there is still a wider range of opinions when $\tau = d/4$, which censors interactions between agents. Again, the configuration *flexibility* remained unaltered: consensus is achieved for any number of moderates.

Moreover, at 100 moderates, consensus was achieved for the configuration *centrality*, and a small fluctuation was observed until the number of moderates reached its maximum. Around 50 moderates, the population reached compatibility, confirming what the data from Figure 3.16 suggested.

Therefore, the results suggest that in the multivariate Deffuant model consensus is achieved with 100 moderates, which represents 20% of the population analyzed (500 agents). The state of compatibility reached around 50 moderates is rapidly surpassed in the configuration with $\tau = d/2$. This indicates that in a model with the exchange of two opinions, the velocity to consensus is higher than in the unidimensional model. Conversely, in populations with lower and higher tolerance, $\tau = d/4$ and $3d/4$, no critical value was observed: the former reached consensus at any state; the second maintained a wider range of opinions.

Chapter 4

Conclusion

This thesis aims to understand the role of moderates in the Deffuant model of opinion dynamics over different social networks, such as $\mathbb{Z}$, $\mathbb{Z}^2$, and the complete graph. Our exploration took place for models with uni and bidimensional opinions.

The first chapter focused on introducing the area. We briefly presented Interacting Particle Systems (IPS) and two famous models: the Ising model for magnetism, and the voter model for opinion dynamics.

The second chapter explored the theoretical aspects of the Deffuant model in different scenarios. We proved the pivotal convergence results for the traditional model on the one-dimensional lattice, based, mainly, on the findings of Lanchier and Häggström. (Lanchier, 2012; Häggström, 2012) Through the construction of an auxiliary deterministic process, we demonstrate the consensus formation when the tolerance parameter (θ) is greater than 1/2. Furthermore, we presented a brief review of recent results, such as the lower-bound probability of consensus for the bidimensional model, (Lanchier and H. Li, 2020) and other extensions such as the convergence for other initial distributions. (Häggström and Hirscher, 2013) Simulation results were also presented, for instance, the (possible) universality of the threshold in one-dimensional opinion spaces.(Fortunato, 2004)

The third chapter introduces an extension of the model, considering moderates - agents with opinion within the opinion space - and extremists initially. Through an extension of the work of L. Marconi and F. Cecconi, (Marconi and Cecconi, 2020) we tested how strong the results from chapter two were. As a result, critical values of moderates, that progressively set the population towards consensus, were found. Table 4.1 summarizes the exploration of this chapter, showing results for different setups of the Deffuant model.

Social Network	Opinion space	Number of moderates at criticality | compatibility	Total number of agents	Parameter
Complete graph	Unidimensional	200 | 250	500	$\theta = 1/2$ $\mu = 0.5$
$\mathbb{Z}$	Unidimensional	300 | 480	500	$\theta = 1/2$ $\mu = 0.5$
$\mathbb{Z}^2$ (torus)	Unidimensional	200 | 300	900	$\theta = 1/2$ $\mu = 0.5$
Complete graph	Bidimensional	1 | 50	500	$\tau = d/2$ $\mu = 0.5$

Table 4.1: *Critical number of moderates for different configurations of the Deffuant model.*

Future research could explore the role of moderates in higher-dimensional lattices or under different social network structures, assuming more opinions, as well as incorporating real-world data to test the robustness of the model in practical scenarios.

In summary, this thesis offers new perspectives on opinion dynamics models, particularly the role of moderates on the Deffuant model on different social networks. The findings have implications for understanding real-world social dynamics, especially for polarized societies where centrist views play a pivotal role in consensus formation.

References

[Albert and Barabási 2002] Réka Albert and Albert-László Barabási. "Statistical mechanics of complex networks". *Rev. Mod. Phys.* 74 (1 Jan. 2002), pp. 47–97. doi: 10.1103/RevModPhys.74.47. url: https://link.aps.org/doi/10.1103/RevModPhys.74.47 (cit. on p. 35).

[Bernardo 2016] Ivan Costa Bernardo. "Sistemas de partículas interagentes aplicados a dinâmicas sociais: modelos de confiança limitada". Dissertação de Mestrado. MA thesis. São Paulo, Brazil: Universidade de São Paulo, Instituto de Matemática e Estatística, Feb. 2016 (cit. on p. 10).

[Castellano *et al.* 2009] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. "Statistical physics of social dynamics". *Reviews of Modern Physics* 81.2 (2009), pp. 591–646 (cit. on p. 5).

[Clifford and Sudbury 1973] Peter Clifford and Aidan Sudbury. "A model for spatial conflict". *Biometrika* 60.3 (1973), pp. 581–588. issn: 00063444, 14643510. url: http://www.jstor.org/stable/2335008 (visited on 09/05/2024) (cit. on p. 5).

[Deffuant *et al.* 2000] Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. "Mixing beliefs among interacting agents". *Advances in Complex Systems* 03.01n04 (2000), pp. 87–98. doi: 10.1142/S0219525900000078. eprint: https://doi.org/10.1142/S0219525900000078. url: https://doi.org/10.1142/S0219525900000078 (cit. on pp. 9, 11).

[Dobrušin 1971] R. L. Dobrušin. "Markov processes with a large number of locally interacting components: existence of a limit process and its ergodicity". *Problemy Peredači Informacii* 7.2 (1971), pp. 70–87 (cit. on p. 3).

[Erdös and Rényi 1959] P Erdös and A Rényi. "On random graphs i". *Publicationes Mathematicae Debrecen* 6 (1959), pp. 290–297 (cit. on p. 34).

[Fortunato 2004] Santo Fortunato. "Universality of the threshold for complete consensus for the opinion dynamics of deffuant et al." *International Journal of Modern Physics C* 15.09 (Nov. 2004), pp. 1301–1307. issn: 1793-6586. doi: 10.1142/s0129183104006728. url: http://dx.doi.org/10.1142/S0129183104006728 (cit. on pp. 34–36, 55).

[Häggström 2012] Olle Häggström. "A pairwise averaging procedure with application to consensus formation in the deffuant model". *Acta Applicandae Mathematicae* 119 (June 2012). DOI: 10.1007/s10440-011-9668-9 (cit. on pp. 11, 13, 14, 16, 17, 19–22, 25–28, 55).

[Häggström and Hirscher 2013] Olle Häggström and Timo Hirscher. "Further results on consensus formation in the deffuant model". *Electronic Journal of Probability* 19 (Nov. 2013). DOI: 10.1214/EJP.v19-3116 (cit. on pp. 12, 28, 29, 55).

[Harris 1974] Theodore E. Harris. "Contact interactions on a lattice". *Annals of Probability* 2.6 (1974), pp. 969–988 (cit. on p. 6).

[Holley and Liggett 1975] Richard A. Holley and Thomas M. Liggett. "Ergodic theorems for weakly interacting infinite systems and the voter model". en. *Ann. Probab.* 3.6 (1975), pp. 643–663. URL: http://dml.mathdoc.fr/item/1176996306 (cit. on pp. 5, 7).

[Lanchier 2012] Nicolas Lanchier. "The critical value of the deffuant model equals one half". *ALEA. Latin American Journal of Probability and Mathematical Statistics [electronic only]* 9 (Jan. 2012) (cit. on pp. 10, 11, 28, 55).

[Lanchier 2017] Nicolas Lanchier. *Stochastic Modeling*. 1st ed. Universitext. Includes 57 b/w illustrations, 6 color illustrations. Springer Cham, 2017, pp. XIII, 303. ISBN: 978-3-319-50037-9. DOI: 10.1007/978-3-319-50038-6. URL: https://doi.org/10.1007/978-3-319-50038-6 (cit. on pp. 6, 7).

[Lanchier and H. Li 2020] Nicolas Lanchier and Hsin Lun Li. "Probability of consensus in the multivariate deffuant model on finite connected graphs". English (US). *Electronic Communications in Probability* 25 (2020), pp. 1–12. ISSN: 1083-589X. DOI: 10.1214/20-ECP359 (cit. on pp. 31, 55).

[Lenz 1920] Wilhelm Lenz. "Beiträge zum verständnis der magnetischen eigenschaften in festen körpern". *Physikalische Zeitschrift* 21 (1920), pp. 613–615 (cit. on p. 4).

[H.-L. Li 2021] Hsin-Lun Li. *Mixed Deffuant Dynamics*. 2021. arXiv: 2112.02956 [math.PR] (cit. on p. 10).

[Liggett 1985] Thomas M. Liggett. *Interacting Particle Systems*. Vol. 276. Grundlehren der mathematischen Wissenschaften. New York: Springer-Verlag, 1985. ISBN: 978-0-387-96069-4. DOI: 10.1007/978-1-4613-8542-4 (cit. on pp. 4, 10).

[Marconi and Cecconi 2020] Luca Marconi and Federico Cecconi. *Opinion dynamics and consensus formation in a Deffuant model with extremists and moderates*. 2020. arXiv: 2010.01534 [physics.soc-ph] (cit. on pp. 39, 41, 47, 50, 55).

[Spitzer 1970] Frank Spitzer. "Interaction of Markov processes". *Advances in Mathematics* 5 (1970), pp. 246–290 (cit. on p. 3).

[Stauffer and Meyer-Ortmanns 2003] D. Stauffer and Hildegard Meyer-Ortmanns. "Simulation of consensus model of deffuant et al on a barabasi-albert network". *International Journal of Modern Physics C* 15 (Aug. 2003), pp. 241–246. doi: 10.1142/S0129183104005644 (cit. on p. 35).

[Vazquez *et al.* 2003] F Vazquez, P L Krapivsky, and S Redner. "Constrained opinion dynamics: freezing and slow evolution". *Journal of Physics A: Mathematical and General* 36.3 (Jan. 2003), pp. L61–L68. issn: 0305-4470. doi: 10.1088/0305-4470/36/3/103. url: http://dx.doi.org/10.1088/0305-4470/36/3/103 (cit. on p. 7).

Index