Physics of Collective Behaviour

UCD School of Physics

This thesis is submitted to University College Dublin in partial fulfilment of the requirements for the degree of MSc in Applied Mathematics and Theoretical Physics

Davis Sandefur 12250836

Supervised by Dr. Vladimir Lobaskin

22 August 2022

Abstract

Opinion dynamics is a growing and increasingly important field as our world seems to be heading towards more and more polarised extremes. It has been growing mostly using the tools of statistical physics, and has seen a variety of models from discrete time bounded-confidence once, as introduced by Deffuant, to continuous-time models that utilise differential equations. In this paper, we will first look at the Deffuant model and then update it in two ways. First, we will introduce a noise factor, which allows an agent to change it's opinion within a bounded range after the interaction. Second, we look at how the Deffuant model with noise behaves under a mean-field value approach on two dimensional square lattice with periodic boundary conditions. We generate phase diagrams for these and look at various statistical parameters of their final stages. Finally, we apply these models to real-world situations as taken from sentiment scores of data obtained via Twitter and try to parameterise political opinions within Ukraine; we then see where this falls and discuss what can be done to possibly help push past the polarised society that Ukraine, and the world, is currently experiencing.

Declaration

I declare that all material in this project is my own work, except where there is clear acknowledgement and appropriate reference to the work of others.

Contents

List of Figures

List of Tables

1 Introduction

Opinions are a major part of our daily lives. Everyone has opinions on pretty much every topic, even if they claim not to, people with different opinions are constantly interacting, debating and changing their opinions across all topics. These changes in opinions can have drastic – sometimes disastrous – effects across all levels of society, so attempts have been made to model them and predict what factors can cause polarisation or consensus in the opinions of a population. To this end, the tools of statistical physics, in particular the physics of collective behaviour, has proven to be a tremendous boon.

In the past half a century, scientists from a variety of fields outside physics – ranging from biology to sociology to finance – have found that, on a macroscopic scale, collections of agents behave similarly to how collections of particles behave in various statistical physics applications, from liquid flow to the phase change representations of various physical quantities. Starting in the 1990s, it was discovered that the flocking of birds also behaves similarly, and shortly after the first models using statistical physics for opinion dynamics were built. These have since been extended to a variety of fields, such as financial opinion modelling [1], determining how algorithmic filtering can bias our opinions [2] and how opinions change during significantly polarising political events [3].

Most opinion dynamics models work on a bounded confidence model, whereby agents only update their opinions after an interaction if they are within a bounded interval (tolerance) from the other agent's opinion. This has been encapsulated particularly by the Deffuant model, which, since first introduced, has been adapted in a variety of ways. In his original work, Deffuant discovered that, at certain tolerance intervals, polarisation arose spontaneously among interacting agents; furthermore, as the tolerance was decreased, the number of polarised peaks increased. It was also discovered that there is no interaction among these groups – they are completely polarised and will never interact and converge back to the centre of the spectrum again.

In this work, we adapt the model in several distinct ways and use this to determine how the overall system behaves. We hope to then use this to answer such pressing societal questions as what are the factors, in terms of the bounded tolerance interval and "free will" noise, that cause consensus, polarisation or a random distribution of opinions. We will discuss this in the language of phase changes, showing that certain combinations lead to predictable phases in our system. We hope to then extrapolate this to some of the world's most pressing political problems and parameterise society, so that we can hopefully understand what is driving the ever-increasing polarisation we are seeing. We expect to see three different states appear in our model: a polarised state, where there are multiple opinion clusters that do not interact; a consensus state, where there is one major opinion cluster that all others are drawn towards; and an 'anarchy' state, where there is a seemingly random distribution of the opinions without a trend towards consensus or polarisation. We expect that, as noise increases, the likelihood of obtaining the anarchy state also increases regardless of the model.

Particular focus will be paid to Ukrainian language opinion modelling, as found via sentiment analysis on political-related Tweets, in the months leading up to and following the Russian annexation of Crimea, as well as longer term afterwards, leading into the CoViD-19 pandemic. This parameterisation could then be used to help parameterise other societies/topics and see what causes shifts in topics like global warming or the Black Lives Matters movement, and how we can possibly overcome the ever-increasing polarisation of society along contentious topics.

2 Previous Studies

The type of modelling that will be used is discrete-time agent based modelling, which involves creating agents and modelling their interactions with each other to see how the overall state of the system evolves. This type of modelling has its origins in the 1940s, where it was first formalised by Von Neumann. Von Nuemann's posthumous *Theory of self-reproducing automata*, describes a lattice of cells that can have 29 possible states and where each cell is connected to the cells immediately orthogonal to it, and in which he showed biological processes can be modelled using such an automata [4]. Interest and implementations in cellular automata continued to increase over the next few years, with a famous instance being Conway's Game of Life.

Models of opinion dynamics utilising the theory of cellular automata soon followed. Most of these featured binary opinions, where the agents updated their opinions after interactions. It was shown that when the interactions are allowed to happen across the entire population, a consensus opinion forms, with all agents adapting that consensus opinion; this behaviour has been seen in real life and is known as "herd" behaviour among economists [5] [6] [7]. In the cases where agents are limited in who they interact with, clusters of agents with opposite opinions form. If some agents' influence is higher than others, this clustering is enhanced [8] [9] [10].

Continuous opinion dynamics, where the opinions are spread over a closed interval as opposed to two binary values, was first modelled in 2001 [11]. It had already been predicted in the late 1980s that when all agents were allowed to interact in a continuous opinion system that they would converge towards the average opinion of the initial system [12]. By introducing a tolerance parameter, whereby two agents only interacted if their current opinions were within a certain, bounded, distance of each other, Deffuant et al. showed that we do see a convergence of opinions, though complete consensus only occurs if the tolerance parameter is above a certain threshold; indeed, as the tolerance parameter decreases, the number of polarised peaks increases [11].

Deffuant et al., in the same paper, also go on to test how dynamics work in a lattice with continuous dynamics. They put their agents on a two-dimensional lattice and allowed them to interact only with the neighbours directly orthogonal to them, instead of across the whole population. Once again, they found that above a certain tolerance threshold the opinions went towards a converge with large enough time. However, as the tolerance threshold was lowered, things got more interesting. For middle-low range values of the tolerance parameter, most agents still converged completely, with a few isolated extremists remaining; as the number got lower, convergence was reached among the clusters that appeared, with more extreme views tending to cluster together while average views also clustered together [11].

Others have since adapted the Deffuant model to simulate alternate scenarios, such as adding extremist opinions to the initial distribution [13] [14]. It was found that, if tolerance was allowed to adapt along with opinion after the initial interaction, the dynamics of extremism can change compared to what was discovered in Deffuant 2001. For example, if the initial tolerance was quite high except for the extremists, where it was quite low, the extreme values would prevail overall, as the rest of the opinions shifted towards the extremes, while their tolerance for differing opinions also decreased. This would lead to either a bimodal split with two non-interacting groups, or consensus at a single extreme opinion. However, if the initial tolerance of all agents was quite low, the influence of the extremists is limited to those already close to them, so a central consensus emerges [13]. Similarly, models have been created where there are 'stubborn agents', who influence others but are not influenced themselves (asymmetrical tolerance conditions) [15].

The Deffuant model has also been adapted in two other significant ways. The first is to allow repulsive interactions. In these simulations it was found, perhaps unintuitively, that even in a closed-minded society (i.e. one with low tolerance among agents), a central consensus can be adopted by everyone if there are repulsive interactions [16]. Lastly, the concept of noise was added [17]. Here the agents were allowed to update opinions as normal according to the Deffuant model, then there was a randomly determined chance that the agent's opinions could jump anywhere on the bounded spectrum, as an act of 'free will'. It was discovered that the final state depends on two different parameters: the tolerance and the probability of an agent changing its opinion. The two states are an ordered state consisting of opinion clusters, with some variation in them, or a disordered one that has roughly a uniform spread [17].

Various other discrete-time models have been created for opinion dynamics, such as the Hegelsmann-Krause model. In this model, an agent is updated based on a set of compatible neighbouring agents. That is, the opinion is updated based on the average of the opinions of the neighbours who are within a certain tolerance of the original opinion; note that not all neighbours feature in this calculation [18]. It was found that the same two possible states occurred in this model (polarisation and consensus), though the dynamics required to reach them might be quite different. Variants with noise have also been implemented, where the opinion can jump in the entire opinion space or just within a small, bounded space centered at the updated opinion. It was discovered these models behaved similarly to the Deffuant model under noise, though there were different dependencies on parameters leading to cluster formation [19].

The last type of model that has been implemented are those which rely on continuous time steps. These have been done by modelling the dynamics with Ordinary Differential Equations [20] [21], Stochastic Differential Equations and Partial Differential Equations [22] [23]. These have tended to mostly update the Hegelsmann-Krause model to account for continuous time steps as opposed to the discrete time steps initially used in the model. Studies have been done on the effects of boundary conditions on these models, and it shows that they are extremely susceptible to initial and boundary conditions, as well as noise. In fact, they can often be considered non-deterministic under the application of noise [24].

All types of these models have analogous in the physical sciences as well, and the tools from one model can be used to develop new models from the other, thus the applicability of statistical physics to opinion dynamics. Some examples of similar models in other branches of science include slime dynamics [25], as well as Vicsek's model of phase transitions in statistical physics [26].

As expected, people have tried to use bounded confidence opinion dynamics to model and predict real-world events. One such example, highly relevant to the current paper, is that undertaken by Romenskyy et al., which used a bounded confidence model and parameterised it with sentiment analysis from Twitter to view the polarisation state of Ukraine around the events starting in 2014 [3]. By doing this, they were able to determine several of the contributing factors to driving polarisation in Ukraine, notably the level of emotional intensity. High levels of emotional intensity can lead to a sudden onset of polarisation with certain levels of tolerance. Indeed, they discovered that this can be modelled as a phase change where a sudden level of emotional intensity in the opinions of a society can spontaneously lead to a fracturing of that society, increasing the polarisation and activity of the society [3].

3 Models and Statistics Used

3.1 Models

Since our models are mostly based on that of Deffuant, especially in terms of the equations governing the updating opinions, we shall start with explaining the Deffuant Bounded Confidence Model, as presented in [11]. In this model, the opinions of the two agents are x and x'; d is the 'tolerance factor', i.e. how close the opinions must be for them to update; μ is the convergence parameter, which determines how close the two elements converge after updating, and ranges between 0 and 0.5 in Deffuant's simulations. Then, if $\mid x - x' \mid \leq d$, we update the opinions according to the following equations

$$x = x + \mu \cdot (x' - x) \quad (1)$$

$$x' = x' + \mu \cdot (x - x') \quad (2)$$

We can see in this model that, if the two opinions are initially within the tolerance range of each other, they update towards the mean opinion by a factor of μ. For Deffuant, this base equation was used in several different models. The first one as linear, one-dimensional model where each agent randomly interacts pairwise with any other agent. The second was a lattice like model, where the first agent was randomly chosen, then it could only interact with the agents orthogonally adjacent to it. A simplified example of this can be seen in Figure 1 on a 3 by 3 lattice. The yellow star represents the selected agent, while the cyan ones represent the connected agents that it is able to interact with.

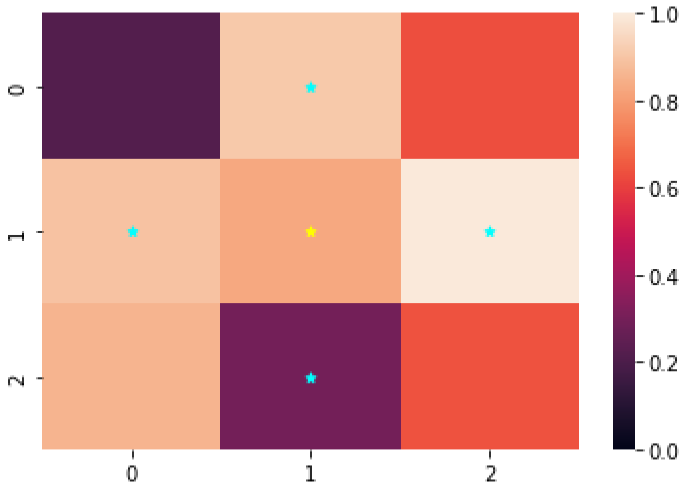

Figure 1: Deffuant Model of Network Interactions

The first update that we have made to the Deffuant model was to allow the process of noise. To do so, we added a ζ-factor to the initial Deffuant equations. As in the Deffuant model, these updates only trigger when the initial opinions x and x' are withing a set tolerance, d from each other, i.e. when $\mid x - x' \mid \leq d$. However, unlike the Deffuant model, the noise factor ζ then can activate. This activates on a 10% chance, and allows the opinion of the first agent to move anywhere in a bounded range of possible noise values, $\zeta \in [-N, N]$, where $N \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$ for a given simulation. The other agent then moves in the opposite direction with the same noise value. This means that the overall opinion of the simulation was constant throughout the entire simulation when noise was introduced, with the rare exception that the change in noise pushed one of the agents past the bounded opinion values, which lay in $[0, 1]$. The updated Deffuant equations for our simulations are seen in equations (3) and (4), below.

$$x = x + \mu \cdot (x' - x) + \zeta \tag{3}$$

$$x' = x' + \mu \cdot (x - x') - \zeta \tag{4}$$

We have used these equations for all our interactions, and, when $N = 0$, $\zeta = 0$ and thus we recover Deffuant's initial model. Our one-dimensional pairwise model worked exactly like Deffuant's, where we randomly picked two agents, updated their opinions and added noise if and as necessary.

We have, however, updated the two-dimensional Deffuant model. Instead of limiting possible interactions to the agents orthogonally adjacent on a square lattice, we had it where each agent's opinion was updated based on the mean value on the 24 nearest neighbours. Unlike the Hegselmann-Krause model, however, our model takes the average of *all* the opinions before testing to see if it updates, rather than just the average of the opinions that would trigger an update. Our lattice was periodic, meaning that it can be viewed as a torus instead of as a simple square lattice where the corners have fewer neighbours. Likewise, all values updated each time stamp instead of randomly picking one and updating it. They did not updated based on the new values, but on the values of their neighbours at the start of the timestamp. When noise was added, the opposite noise was distributed evenly across the 24 neighbours. Figure 2 shows how the neighbours were chosen on a 20 by 20 lattice along one of the periodic boundary conditions; the yellow star is the initial agent, the cyan ones represent the agents it updates it's opinion on.

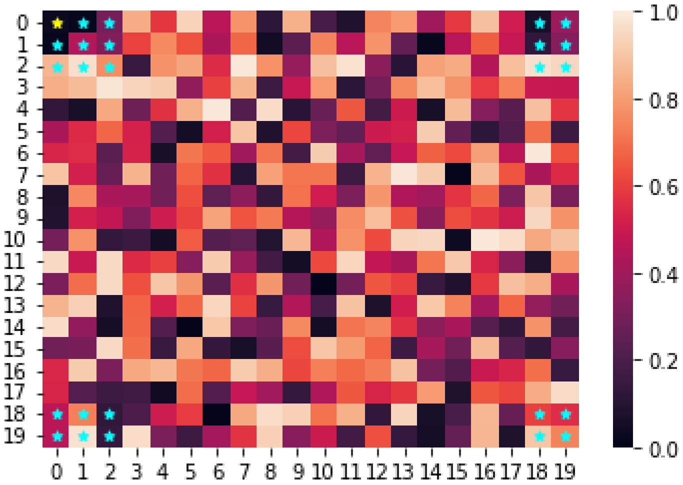

Figure 2: New Model of Network Interactions

3.2 Statistical Parameters

We will now discuss the major statistical parameters we will look at in some of our simulations.

3.2.1 Variance

The first is the variance of the overall system. Often denoted as σ^2, where σ is the standard deviation, or μ_2 (the second standardised moment), the variance is the average of the squared differences of each element from the mean of the total population. Mathematically, it can be defined as in equation (5).

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} \tag{5}$$

where x_i is the value of a given parameter and μ is the population mean. A higher variance value would mean that the opinions of a population are more spread out away from the mean, possibly indicating a higher bimodal distribution, while a lower one means the opinions are closer to the mean, more likely signifying a unimodal distribution.

3.2.2 Skewness

The skewness of a system, often denoted μ_3 (the third standardised moment), is a measure of how asymmetrical the system is; that is, in a unimodal system, it can tell us which direction the majority of the outliers lie in. However, if one tail is long while the other is more intense, the skew could be zero. In a bimodal system, it can tell us if one of the opinion peaks is stronger than the other, or in which direction their tails head. In a perfectly normal system, the skew is zero as both tails of the unimodal distribution are perfectly balanced. Skew is defined mathematically as in equation (6)

$$\mu_3 = \frac{1}{N}[\frac{\sum_{i=1}^{N}(x_i - \mu)^3}{(\mu_2)^{3/2}}] \tag{6}$$

where x_i is the individual element, μ is the population mean, μ_2 is the variance as defined above and N the number of elements in the population.

3.2.3 Kurtosis

The kurtosis of a system, denoted as μ_4 (the fourth standardised moment), is a measure of how extreme the tails of the distribution deviate from its mean. This can help us investigate how strong the outliers of a population are in relation to the mean of the population. The mathematical formulation for kurtosis is given in equation (7) below.

$$\mu_4 = \frac{1}{N} \frac{\sum_{i=1}^{N} (x_i - \mu)^4}{\mu_2^2} \tag{7}$$

where, as before x_i is an agent in the system, μ is the mean, μ_2 is the variance as defined above and N the number of elements in the system.

3.2.4 Bimodality

The last statistical parameter we will look at is the bimodality coefficient. This coefficient is a measure of how bimodal the system is. It ranges from 0 to 1. Any value over 5/9 is extremely likely to be either be bimodal or multimodal. However, the value isn't entirely indicative of whether the system is multimodal, nor does a value less than 5/9 indicate that the system is unimodal; it could, for instance, be random. The mathematical representation of this is given in equation (8).

$$\beta = \frac{\mu_3^2 + 1}{\mu_4} \tag{8}$$

where μ_3 is the skew and μ_4 is the kurtosis, both as defined earlier.

4 Implementation and Methodology

All the code for running the simulations was implemented in Python and can be found in Appendix A.

First, an Agent class was created. This class contains variables that allow for storing of the opinion and tolerance of the agent as well as a method that updates the opinion according to Deffuant's rule with a convergence parameter, passed when the method is called and was used for the basic agent in both the one dimensional pairwise interactions as well as the two dimensional periodic lattice nearest neighbour interactions.

4.1 One Dimensional Model

In the one-dimensional pairwise case, the following algorithm was implemented.

1. Select two random agents

2. Test the opinions to see if they are within a user-provided tolerance condition

3. If they are, update the opinions; if not, move to the next pair

4. If the opinions updated, test to see if the 'free will' factor takes place (10% chance)

5. If free will takes affect, random select the noise, ζ, from [-N,N], where N is user controlled

6. Apply ζ to one agent, $-\zeta$ to the opinion of the other agent, keeping within

boundary conditions as necessary

For the one-dimensional model, 10,000 agents were initiated with a random opinion distribution in [0,1]; a sample can be seen in Figure 3. The algorithm above was then implemented for varying values of tolerance, $d \in \{0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5\}$, as well as noise, $N \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$. All possible combinations were used which enables us to build a phase diagram of noise versus tolerance. The convergence parameter was the default value of 0.5 as used in Deffuant.

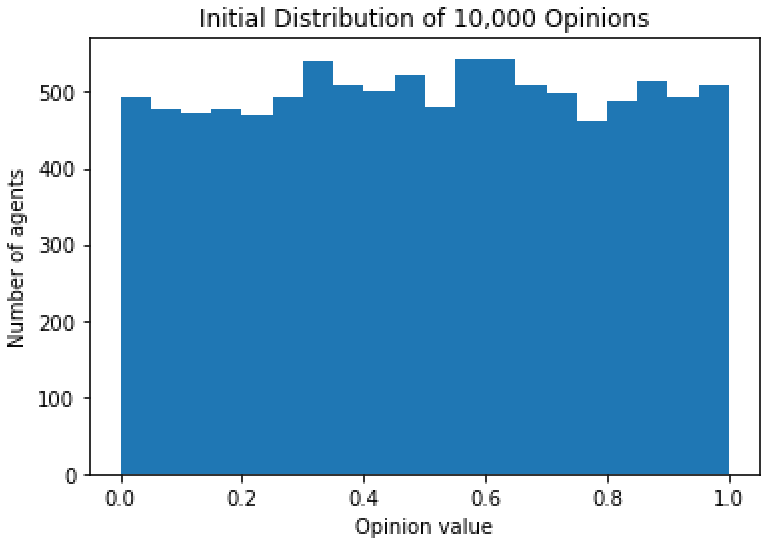

Figure 3: Representative Initial Opinion Distribution of 1D Model

The algorithm above ran 5,000,000 times in each simulation, allowing the system enough time to settle down into a stable state, getting rid of any effects of the initial, random, distribution. For each tolerance-noise combination the simulation was run five times, and the majority result was taken as indicative of the final outcome (polarisation, consensus, random) and used to generate the phase diagram.

4.2 Two Dimensional Model

The algorithm implemented in the two-dimensional case was necessarily a bit more complex. The algorithm can be summarised in the following steps:

1. Loop through all agents, getting the mean value of the opinions of their 24 nearest neighbours, saving these in an array with each opinion in the same position as the agent it will act on

2. Test to see if the 'free will factor' applies (10% chance)

3. If it applies, pick $\zeta \in [-N, N]$ and add it to a 'noise' array at the place of the agent it acts on

4. Add $-\zeta/24$ to the 'noise' array for each of the nearest neighbours of the initial agent

5. Loop over all agents

6. Update opinions based on the opinion array, if within tolerance

7. Add the relevant noise based on the noise array, correcting to make sure all opinions stay within the bounded opinion space

The main thing to notice here is that no agents' opinions were updated until after the mean value of the neighbours had been collected for all of them. This means that each agent updates itself based on the initial state at that time step, and doesn't update itself with regards to what its neighbours are doing at said time-step.

As in the one-dimensional simulations, an initial distribution of 10,000 agents with random opinions in the range of [0,1] was created. A representative example can be seen as a heat map in Figure 4 below; take into account there are periodic boundary conditions and the neighbours of the first element are as shown in Figure 2. The simulation was then run over varying values of tolerance $d \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ and noise, $N \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$. Each pair was tested.

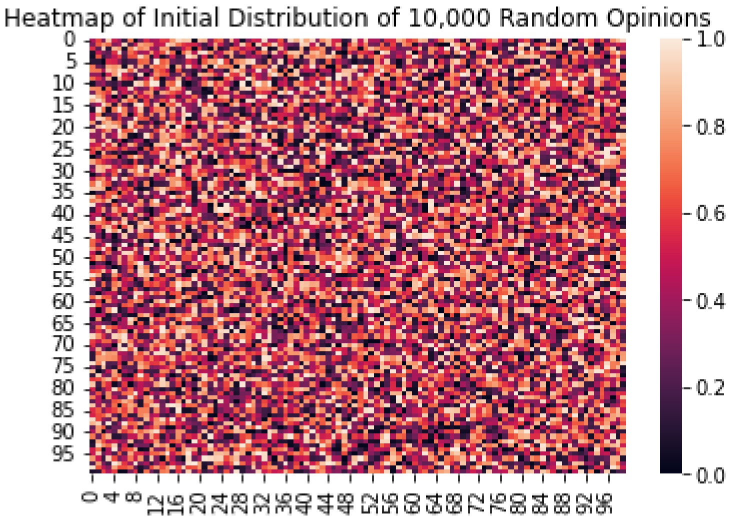

Figure 4: Representative Initial Opinion Distribution of 2D Model

For each tolerance-noise pairing, the simulation was run for 100,000 steps, meaning each agent updated its opinion 100,000 times. However, every 1,000 steps there was a check to see if agents had converged or if their opinions were outside the tolerance range of the mean value of their neighbours' opinions. If so, the simulation stopped as no updating would occur; this was done as the algorithm was quite resource intensive and allowed for simulations to be carried out in less time. Each pairing was repeated a total of three times, and the majority result was taken to be indicative with regards to which result was used to create the phase diagram for this model.

4.3 Ukrainian Twitter Data

To obtain the data we needed in order to parameterise Ukrainian society, we relied on the Twitter researcher API. A list of 426 Ukrainian-language query terms was created and can be found in Appendix B. The list was split into sets of 50 query terms, each joined by an OR operator for the Twitter search. The search was then conducted, restricting the tweets to those that Twitter's data had labeled as being in Ukrainian and those that were not retweets. Queries were conducted across all nine groups daily for several time periods from July 2014 onward. Queries were done monthly from July 2014 to February 2015, at which time the Second Minsk Agreement was signed. From February 2015 to December 2016 queries were conducted from March-June and from September-December. From 2017 to 2020, queries were only conducted on three months out of the year, September-November.

Until the Autumn 2017 set of queries, each query was allowed to collect 1,000 responses a day, giving a max total of 9,000 tweets/day. It was ran over every single day of the period. From Autumn 2017, due to tweet limits imposed by Twitter, it was limited to 800 Tweets data. Thus, over the periods where Tweets were collected monthly, roughly 270,000 tweets were collected per period. Over the four month span, over 1 million were collected per period, and over the last four time spans (September - November in 2017-2020), approximately 650,000 were collected. Regardless of the varying numbers, this gave us enough tweets to see statistics and evolution of the system over time. No metadata was saved from the tweets except for the raw text of the tweet. Thus we could not group by users and find the users' overall opinions; only the overall opinion of the tweets themselves.

To do this, a sentiment analysis was run on the tweets we had saved. First, we converted the tweets from UTF-16 to UTF-8 encoding of the Cyrillic alphabet. We then compared this to a dictionary list of keys, each with an associated score on how pro-/anti-Ukraine the word was. The entire score of the tweet was then summed, and added to a list for that month. This was repeated for every tweet in the month. A sample spread of opinions, for July 2014, can be seen in Figure 5 below. The range of the y-axis has been capped at 3,000, otherwise the neutral tweets drown out the majority of the others.

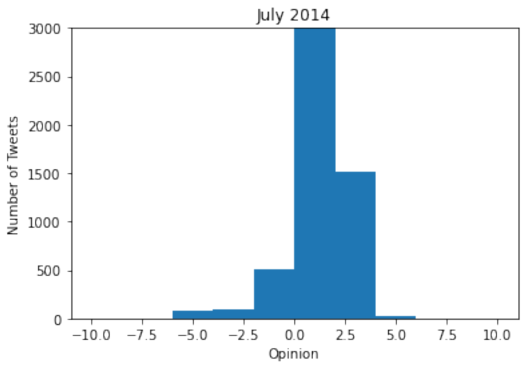

Figure 5: Tweet Opinion Distribution for July 2014

We repeated this for all the time periods collected, for a grand total of 16 periods. The statistical parameters for each of these were calculated. We also took the July 2014 data and ran it through our simulations, seeing if any would give the results of August that we would expect. Finally, we repeated this process without the neutral peaks. We removed them on the assumption that they are possibly news agents, thus they just broadcast and do not interact and change their opinions with the others. We believe this would give us a better approach at how the users who are driving polarisation, those with extreme emotions tend to behave.

After this data was collected, we worked on trying to parameterise the model. Focus was given to the one-dimensional Deffuant model.. To test this and parameterise it, the data without the neutral actors was run through the simulations, turning off any boundary conditions. The tolerance and noise were scaled to match the new distribution and the simulations were run to test them. Each month was simulated over 1,000,000 steps and then compared to the empirical results obtained from the next month.

The two dimensional model was tested on the July 2014 data with a smaller number of steps – only 5,000. This is because each agent updates based on the mean number at each step, so they all get 5,000 updates. This was also partially due to time and computational constraints as it's a much more computationally heavy process. When testing this, the non-neutral tweets were put in a randomised order in the lattice and then excess was purged to take it down to the nearest perfect square to make the square lattice. For instance, the July 2014 data had 3,051 non-neutral tweets; we used 3,025 of these, cutting off the last 26 at random in order to make the square lattice work. The code was updated to include the boundaries for the new lattice.

5 Results

5.1 One Dimensional Model

The most pertinent result is that we discovered the three states we were predicting to find: anarchy (random distribution of opinions), polarisation (two or more peaks that don't interact) and consensus (one peak that draws most to it). This confirms that we will be able to build a phase diagram representation out of the noise-tolerance combinations tested, enabling us to predict what settings will result in a given state and hopefully allowing for parameterisation of the Twitter data from Ukraine. Representative states can be seen below in Figure 6

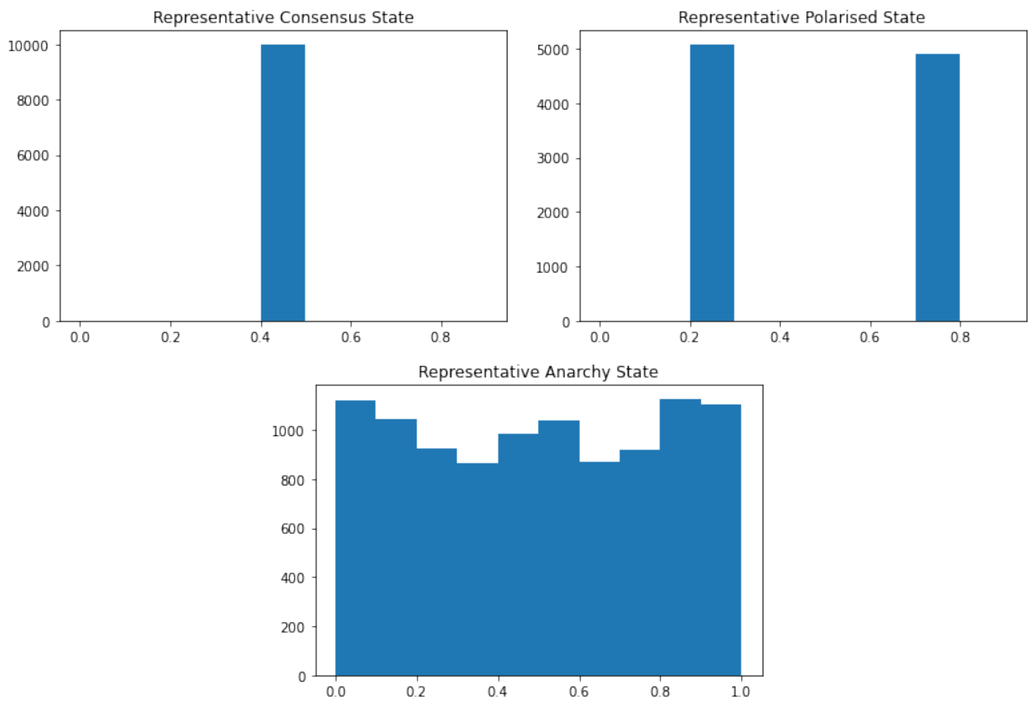

Figure 6: Representative Final States

A more thorough inspection of the results of the one-dimensional model for N=0 at

$d = 0.2$ and $d = 0.5$ gave the results as shown in Figure 7. This accords with the results obtained by the Deffuant model, thus proving that our algorithm, at least in the case with no noise, was implemented correctly.

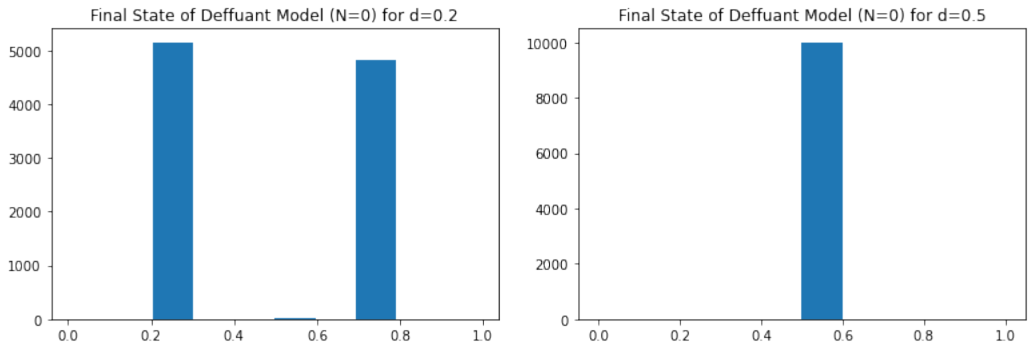

Figure 7: Representative Final Outputs when N=0 for d = 0.2 and d=0.5

As the noise in the simulation increased, several distinct things were seen to happen, regardless of tolerance level. First, the tails of the peaks started to spread out, much as would be expected. This can be seen by comparing the two images in Figure 8. The left one is $d = 0.5$ and $N = 0$, whereas the right is $d = 0.5$ and $N = 0.3$.

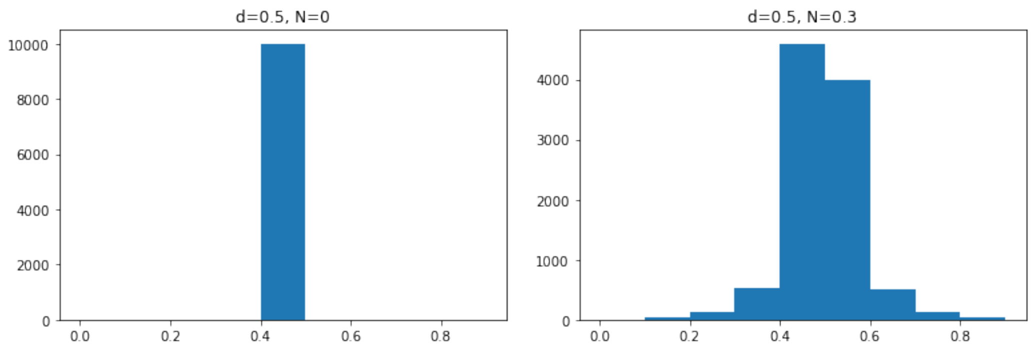

Figure 8: Effect of Noise, d=0.5

At high enough tolerances, there was no fracturing into an anarchy state at the lower noise levels, with it only occurring at higher noise levels. In some instances, for the highest tolerances, the fracturing into an anarchy state never occurred at all, the tolerance condition being high enough to constantly "draw" the moving opinions back in before they could cause it to fracture.

At lower tolerances, however, the fracturing of the polarised state happened fairly quickly. In some instances, there was a phase transition into a state of consensus where one hadn't been achieved before, as illustrated in Figure 9. In others, even lower, the fracturing of the polarised state led straight into an anarchy state, as the tolerance was lower than the possible noise values, meaning any change could easily have pushed two agents away from interacting.

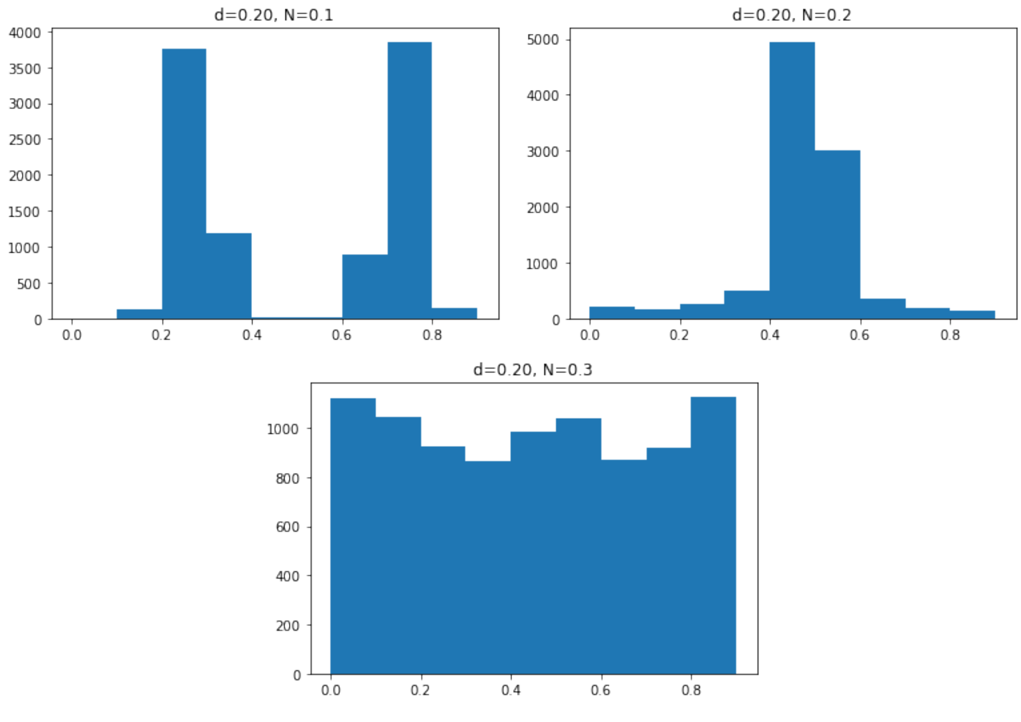

Figure 9: Phase Change from Polarisation to Consensus to Anarchy

Overall, these results allow us to build a phase diagram of the expected states with the Deffuant model given a set tolerance and noise value. If a state starts off polarised, as all with $d < 0.3$ do, it will either cross straight into an anarchy state with increasing noise or will go into a consensus state before heading into an anarchy state. If the system initially would evolve towards a consensus state, two possibilities can happen in the range we looked at. At high noise values, it can tend towards anarchy, though this only happens for one and only at a high noise level mostly, at the noise values we looked at (which could, theoretically, push you halfway across the spectrum of opinions), most systems that would form a consensus with no noise continue to form a consensus with increasing amounts of noise. The complete phase diagram can be seen in Figure 10.

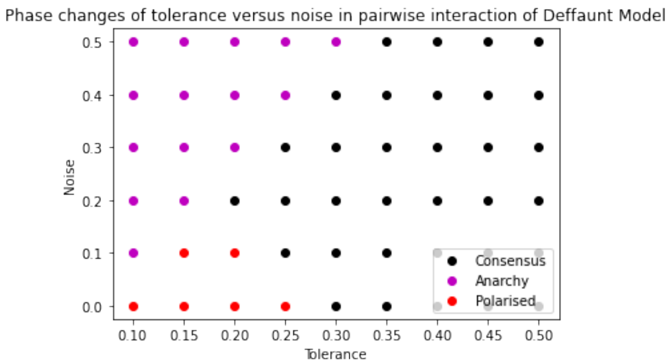

Figure 10: Phase Diagram for Pairwise Deffuant Model With Noise

We also analysed the variance, skewness, kurtosis and bimodality coefficient for each of these simulations. For each set of five simulations, the average of each of the statistical parameters was taken. As expected, as the noise increased the variance of the system generally tended to increase as well, except in the cases where there was a phase change from polarised to consensus before shifting to anarchy. Also, as noise increased the bimodality coefficient tended to stay in the middle, with it leaning towards the extremes (1 in the case of polarisation, 0 in the case of consensus) only when it was really strong at low noise levels, or right after a phase change from polarisation to consensus. The skew tends towards zero in all cases, showing that there isn't really a favoured side for the opinions to shift towards. This also concurs with expectations, as when we added noise in one direction we also added noise in the other, keeping the overall value of the opinions constant. The kurtosis tended to show that outliers lie close to the mean opinion, except in the cases where there was consensus after a phase change, in which some extremists exist far away from the mean values that weren't drawn back in but were too far away to influence others.

5.2 Two Dimensional Model

As in the case of the one dimensional model, all three states were found in the two dimensional model. However, because of the way the model interacts, we don't necessarily see the clustering of extreme opinions in the polarised state that is seen under the Deffuant model. Thus it can only be viewed from looking at the overall histogram of opinions. Representative examples from the three states are shown in Figure 11 and the corresponding histograms of opinions can be seen in Figure 12. It should be noted that the consensus state does not necessarily reach all agents, as there can be extreme outliers based on initial configuration of opinions, but it reaches all but one or two.

In contrast to the one-dimensional pairwise system, none of the tolerances displayed a polarised state when there was no noise applied. Further investigation revealed that this is because of the mean value calculation of the updated opinion, as well as the initial random distribution of opinions. Each agent quickly reached a point where it's

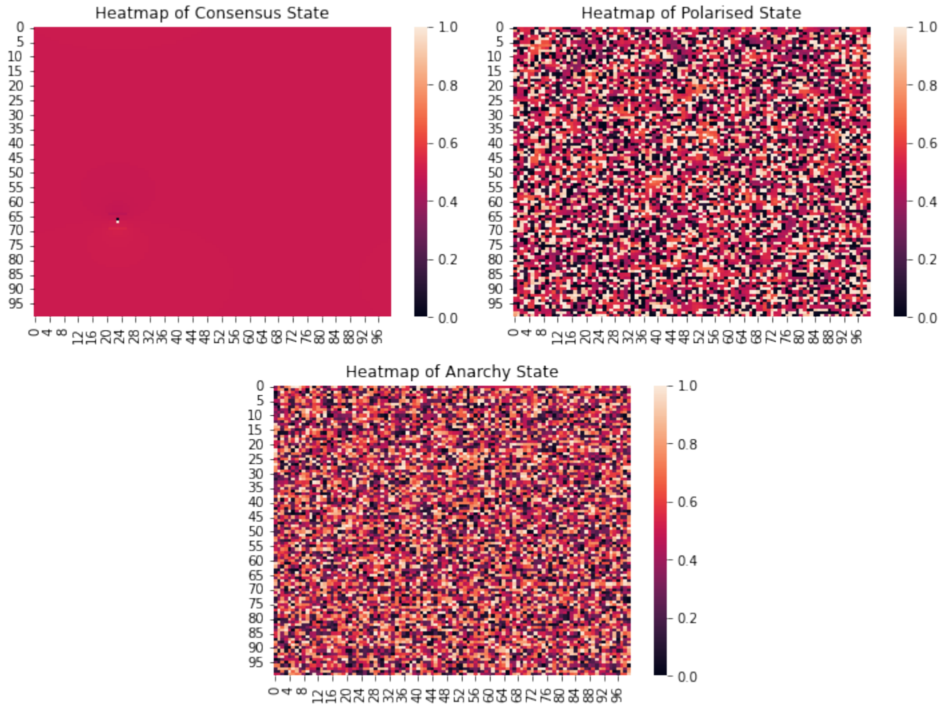

Figure 11: Heatmaps of Representative Final States

opinion was equal to the average of all its neighbours, in which case there would be no change during the update, or it reached a point where its opinion was outside the tolerance range for all its neighbours. This was only prevalent for low tolerances, and as tolerance increase, we saw the shift towards consensus opinion forming as expected.

Interestingly, we observed that increasing noise levels does not tend to push the system towards an anarchic state, instead pushing it first, if in an anarchic state, towards consensus before going back to anarchy and then finally leading towards polarisation with multiple peaks located further than d apart from each other, as shown in Figure 12. If the system would achieve a consensus state when no noise is applied, the application of noise would simply push it straight towards polarisation; the strength of the noise needed for this varies with the tolerance, with higher tolerances requiring

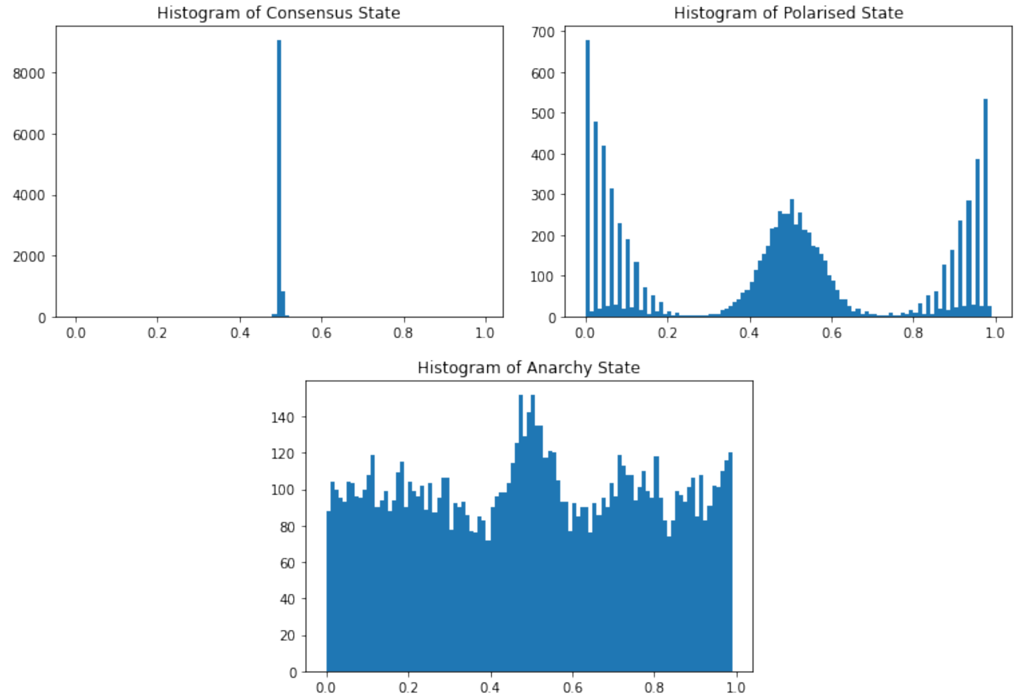

Figure 12: Histograms of Representative Final States

higher noise values to reach the polarisation state. In the pairings we simulated, none of the tolerance values which achieved consensus at first ended up in anarchy; nor did the polarised states shift to anarchy within our testing range. The full phase diagram of the two-dimensional model can be seen in Figure 13.

With regards to the statistical parameters we tested against, the systems as detailed above behave as expected. When the anarchy state was in control, the variance was at its highest, around 0.12. At the lower noise levels, when consensus was achieved the variance was negligible, being on the order of 10^{-5} in some cases. However, when the variance was this low, the kurtosis was also extremely high, as there were one or two outliers on the edge of the parameter space, as far away from the mean as it was possible to get. The skew showed that there was hardly any bias in the system,

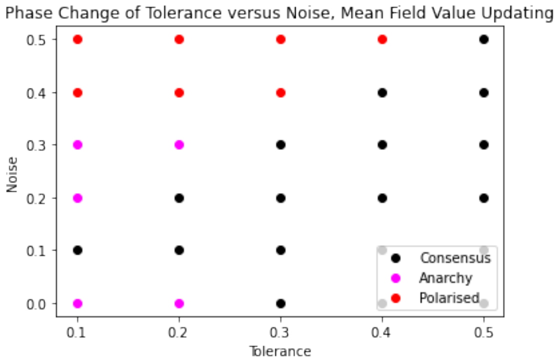

Figure 13: Phase Diagram of 2D Mean Neighbour Value Updating

no matter the result; it was mostly on the order of 10^{-3} or lower, and seemed to equally apply in both the positive and negative directions. This implies that starting conditions likely determined which side of the mean would have more opinions, but it was never a high level of skewness. In consensus, the bimodality parameter was basically zero, in some cases getting as low as 10^{-4}. When there was a polarised state, we saw values above 0.6, which, as mentioned in Section 2.3 indicates multiple peaks in the system. For the anarchy states, it hovered anywhere between 0.4 and 0.6

We found that the effect of locality is extremely important in the two-dimensional system, as well as the initial values of each agent. Unlike Deffuant's model, where they found that we could see spontaneous polarisation and clusters arising [11], we found that this does not generally occur in a mean field value, even at low noise levels. This is because the mean value was too great for the selectivity (tolerance) of the agent, thus it didn't get updated.

We also briefly ran simulations on locality outside of the mean field value. That is, we set a square where the agent can interact with any of its nearest 24 neighbors, but it only interacts with one in the manner of the pairwise Deffuant interaction. Here we saw clustering as observed by Deffuant. The main difference is that it took longer for it to settle into an equilibrium state as opposed to when they were able to interact with any neighbour in their region. We also see instant clustering, as expected. The results of a small 20x20 simulation of this with tolerance 0.2 can be seen in Figure 14. Increasing the number of interacting neighbours, without switching to a mean field value interaction model, merely increases the speed of convergence.

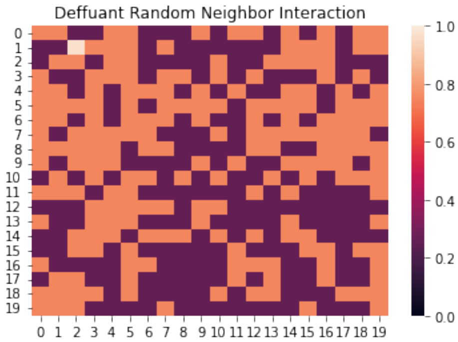

Figure 14: Two Dimensional Clustering Under Locality

5.3 Ukrainian Data

The histograms for each time period, as well as a table summarising the number of tweets for each opinion value can be found in Appendix C.

First, we look at the Ukrainian data as obtained based on the sentiment analysis from Twitter. Figure 15 below shows the data, both with and without the neutral peak, for the month of July 2014 using a kernel density plot. As we can easily see, the vast majority of the tweets tended to be neutral. Whether this was from people having neutral opinions or from newscasters we cannot say without explicit analysis of the tweets themselves. However, when we take out the neutral elements, we can easily see there are multiple polarised peaks, leaning slightly positive and slightly negative. There is also one strongly negative peak, albeit it is smaller than the other peaks. Because of this, and because we wanted to see the dynamics of polarisation and extremism, we worked with the neutral-excluded peaks, on the assumption the neutral peak would overrule all dynamics of the system and draw everything towards itself. Similar phenomenon are known to occur in magnets, where there's not a strong enough field to cause alignment of spins over the presence of the noise and neutral ones. Initial test simulations with the neutral-included data confirmed our opinions.

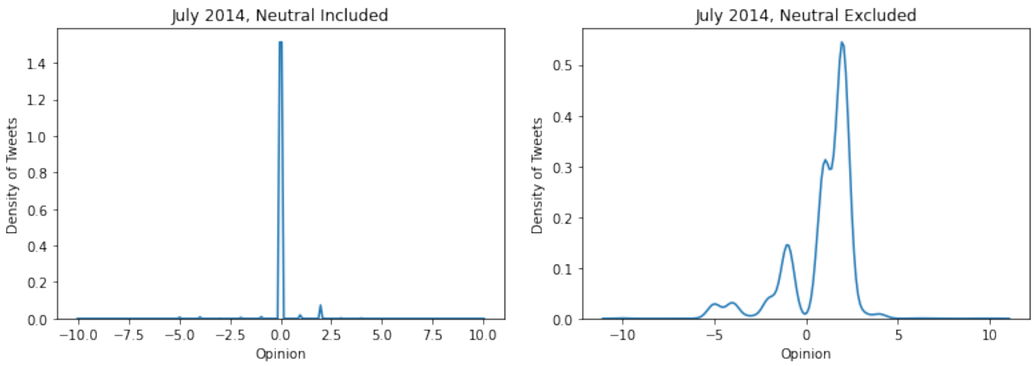

Figure 15: Density Plots of Tweet Opinion for July 2014

Next we examined how the opinion of Twitter changed over time, as we can see in Figure 16. This is the data, neutral-excluded, from Autumn 2017 and thus covers the months of September, October and November from that time period. We can see immediately that there was a huge jump in the concentration of negative opinions, and that the split is stronger between some of the positive opinions. Likewise, the strongest negative opinions experienced the most significant jump between July 2014 and Autumn 2017.

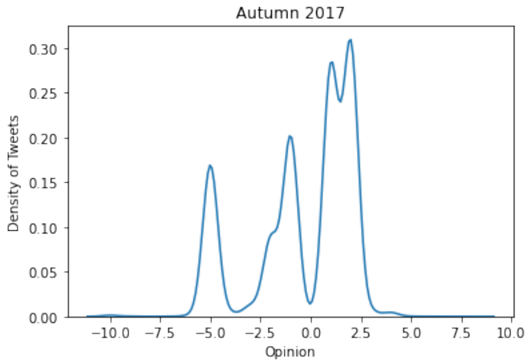

Figure 16: Density Plot of Tweet Opinion for July 2017

To classify and describe how this change happened, we looked at the results of the statistical parameters, as well as the mean, for each month. Their trajectory over time can be seen on the plots in Figure 17. We can see that, until the Spring (March-June) of 2015, overall the tendency was towards a positive view towards Ukraine and then a shift happened that made the view negative. It hovered around neutral overall for a few periods, before becoming negative again in the Autumn of 2017. However, it's worth noting that the opinion never became as negative as it was positive. What we're possibly seeing reflected here are attitudes towards the Second Minsk Agreement (Minsk II), which was signed on the 12 February, 2015. Twitter generally had pro-Ukraine tweets when it happened, but it shifted towards negativity in the following months, rising back to neutrality as time wore on. It is worth noting that, in the period following Minsk II that Poroshenko, the then-president of Ukraine, experienced a significant drop in approval, and it is possible that we are seeing the reflection of this in the Twitter data as users express their dissatisfaction with him.

We can see that, during the same time period, the variance was reasonably low, but experienced a sharp jump in the Spring of 2015 after Minsk II. This is due to the mean shifting to negative results, increasing the distance from the mean of the extremes, which results from the increasing numbers of negative opinions. Interestingly enough, we see the kurtosis drop in said time period, which makes sense as our negative outliers are now closer to the mean and so aren't considered as extreme.

The skew stays negative at all times, meaning that there is a higher prevalence of negative opinions away from the mean. This makes sense when the mean was positive, but even when the mean becomes negative, the skew stays negative. Thus, even when the mean is slightly negative, there are more strong negative opinions than there are positive ones to counteract that.

Overall, the bimodality shows a fairly polarised state starting from the beginning, with jumps happening in August 2014 and Spring 2015. It's interesting to note that polarity goes down to a range that could be comparable to our 'anarchy' states as described above in February 2015, suggesting a coming together of opinions, though not a complete convergence, directly before and after Minsk II, and then jumps to a polarised state, from which it slowly recovers to a more 'anarchy'-like state with various opinions. This suggests that things were polarised for a bit, culminating in Autumn 2016, then trended back towards the median overall, without clear polarisation with no interaction.

Next we turned to see how well our models were able to predict the changes that were seen in Ukraine, and if we could parameterise the society as a whole. We started with the data from July 2014 and ran a shortened version (26 random opinions were dropped) of it in the 2D lattice-like environment, with no noise interactions. This produced the density graph as seen in Figure 18. The actual density graph for

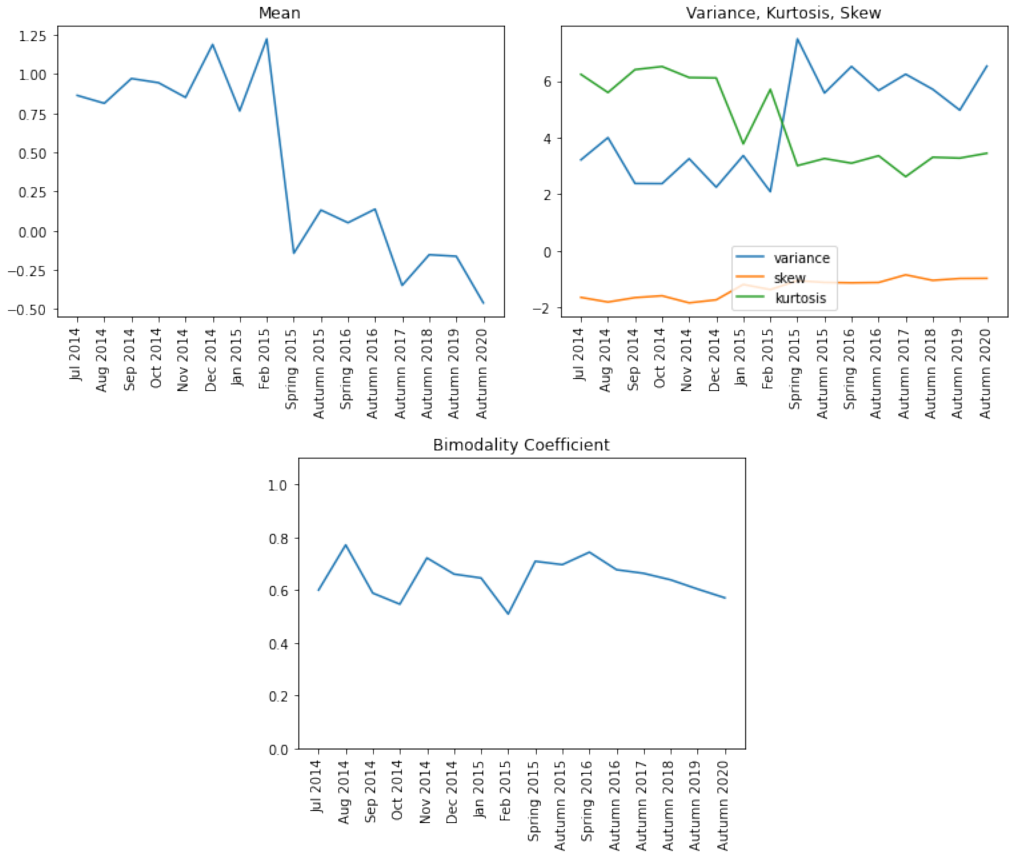

Figure 17: Statistical Parameters on Twitter Data

August 2014 is also shown in Figure 17. We can instantly see this that prediction was inaccurate, tending to have a consensus opinion at around 0, with one peak around 2 and another around -1, whereas the actual data had several peaks, with the highest being at around 2, and several being negative. Other tests of noise/tolerance combinations with the 2D lattice model yielded similar results, nothing as close to approximating the real-world data as the one-dimensional model. This held true regardless of the start month, leading us to conclude that, if agents interacted in a network, it was likely different than the average of their nearest neighbours.

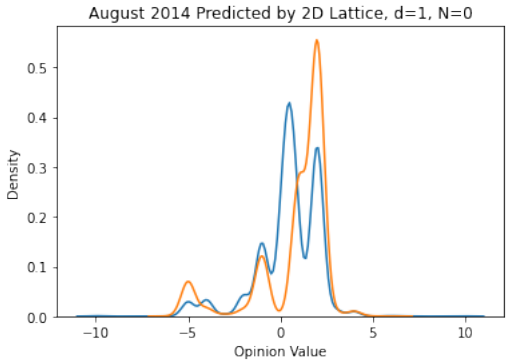

Figure 18: 2D Lattice Extrapolated August 2014 versus Twitter Data August 2014

We get better results when we try to parameterise it with the modified Deffuant one-dimensional model, with pairwise interactions. Running the initial data from July 2015 over 1,000,000 pairwise steps produces the graph in Figure 19. The actual results from August 2014 are included in the figure as well. The first thing to see is that the major peaks that were observed were expected by the model, though the most negative one was slightly more negative in real life than it was in the simulation. The simulated values were also more intense than the actual observed ones across all intensities but the most negative of the three peaks. Likewise, a smaller peak centered around an opinion of 1 was smoothed out, merged into the overall opinion of 1.5 or 2 that dominates the positive values in the expected opinions.

There are several causes for why our simulated opinions overestimate the value of the that positive peak. For one, all opinions within an interaction range tend to be drawn towards their mean. Thus the initial spread of opinions at 1 and 2 in the July 2104 data will be drawn in to the mean, somewhere between 1 and 2, in the simulated data. This is exactly what we see in Figure 18, and can account for why the peak at 1 and the peak at 2 don't coincide exactly with the observed values for August

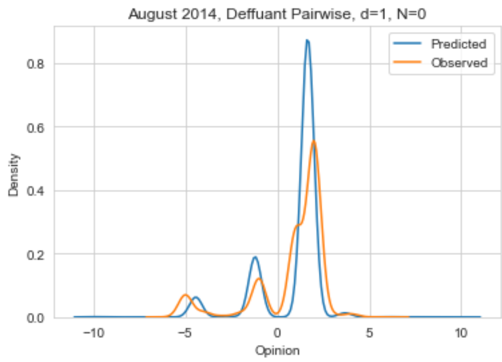

Figure 19: 1D Pairwise Extrapolated August 2014 versus Twitter Data August 2014

2014. In terms of the statistical parameters, the variation in the simulated data is 2.97 while the variation of the observed data was 3.99, so reasonably close even if different. This is likely due to the fact that the most negative peak in the observed data is more negative than ours, and the most positive peak, even though it's smaller, is slightly more positive. The skew was the same for both, -1.8, while the kurtosis was 6.7 for our simulated data but only 5.58 for the observed data. This means that the elements of the data our model extrapolated deviated more from the mean than did those of the actual observed data. Finally, we analysed the bimodality coefficient for both: our simulations gave us 0.65 while the real value for the bimodality was 0.77. We can see from here that both are simulated to be bi- or multi-modal, and that's exactly what we observe with looking at our graphs, with the observed data giving a slightly stronger coefficient because of the greater distance between the most negative and the positive peak.

The same test was run for all our other time periods – starting from the observed data from the period before, we ran 1,000,000 steps and compared the results. The Root Mean Squared Error (RMSE) was calculated for each time period. The overall average of all the RMSEs calculated was 0.89, though that was biased with a much higher RMSE (around 3) in the period of Spring 2015. This coincides with the sudden jump in the mean, and likely something shifted after Minsk II so that the general interaction which was simulated would not hold true. The RMSE over time can be seen in Figure 20.

Apart from the jump in RMSE in Spring 2015, our simulations were quite accurate to the observed data we had. The main issue was that often two close peaks in the observed data were smoothed together to one peak in the simulated results. This happened most often with peaks at opinion values 1 and 2 in the actual data. There are several ways this could perhaps be evaluated and fixed. The main one would be to change the tolerance and noise parameters. However, none of the combinations that we tried were able to replicate the close dual peaks that are seen in the actual data. Even lowering tolerance to 0.5 and adding 0.5 in noise started leading immediately towards convergence, representing a phase change similar to that seen with a tolerance of 0.2 in the Deffuant model under the application of noise.

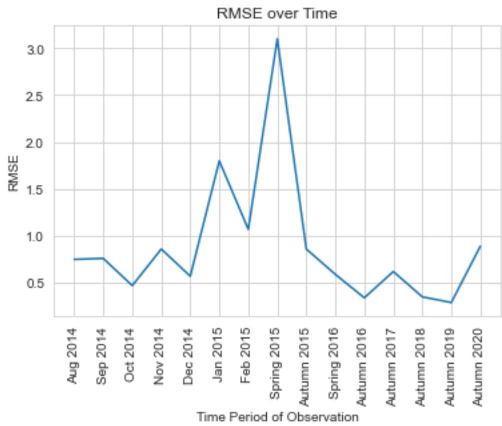

Figure 20: RMSE

Having tested the combination of several tolerance levels and noise patterns on our data, the one that provided the closest match to the actual data was the one discussed above, with a tolerance of 1, meaning each agent interacts only with their nearest neighbours. Because of this we do get issues where two neighbouring peaks in the data get merged together in the simulated results. Attempts were made to correct this by changing tolerance versus noise, but none of those work; as soon as noise was added, the systems started to evolve towards convergence, leading us to conclude there is no noise in the Ukrainian system. Overall, we are thus able to roughly parameterise Ukrainian society as having a tolerance value of 1, which in Deffuant's scale of [0,1] would be roughly a tolerance value of 0.1, and a noise value of 0. We see that in the phase change diagram as simulated by the Deffuant simulations we get polarisation in these states, which is also exactly what we observe in the Ukrainian data.

6 Conclusions and Recommendations for Further Study

In this work, we were able to classify two new updates to the Deffuant Model of opinion dynamics, using the tools of statistical physics. We were able to update his model to include noise, where an agent can move anywhere on a bounded subset of the full opinion interval after having updated its opinion, and we were able to classify how mean field interactions behave in a torus-like lattice under the various tolerance and noise conditions. We were able to succinctly map phase changes across both of these, seeing what levels of tolerance and noise lead to polarisation, consensus or 'anarchy'-like states. We were able to collect valuable statistical data about the end results of these states and can use that to compare to how a real-world system would behave. We were able to see that as noise increased in a pairwise interaction system, the system tended towards anarchy, with some systems that are polarised under no noise first passing through a consensus state, while the two-dimensional lattice systems went from consensus to anarchy or from consensus to polarisation in the noise levels we looked at. This means if people are constantly changing their opinions in these systems, it's quite likely many fractured groups will appear that aren't inclined to interact with each other.

We also downloaded Ukrainian-language tweets which related to politics and managed to use sentiment analysis to adapt the data to a continuous scale of opinions. We ran this data with our simulations to parameterise Ukrainian Twitter on our models. It was discovered that the two-dimensional lattice model didn't work well at all for this system, but the one dimensional one with a tolerance of 1 (corresponding to roughly

0.1 in the Deffuant model) and a noise of 0 was reasonably accurate, having a root square mean error value of less than 1 on average.

This is hopeful because it shows that, if we could increase the noise level in Ukraine, such as how much people change their opinions after interacting with another user, there is the possibility of being able to push back against the polarisation that the sentiment data shows exists in the society. We were also able to see that this could lower the bimodality coefficient, which for Ukraine hung constantly around a bimodal state, though it sometimes crossed into the upper edges of the 'anarchy'-like territory of our simulations. This could hopefully bring Ukrainian language Twitter back towards a more neutral stance, where they interact with each other and, not entirely forming a consensus, provides less extremism overall.

When comparing our study to those previously done, we can see that we see the same states that they observed. For instance, Romenskyy et al [3], showed that in their models and simulations they observed the same three types of states, and along the same parameters we did. When there was high noise, there was a globally disordered behaviour (anarchy state), while high tolerance allows for only states with 'global consensus'. Low tolerance states produced the bipolar states [3]. All of this matches the results described by Romenskyy, which, as they said, isn't too dissimilar from the appearance of phase separation in dissimilar liquids. Thus we can see that our new model with noise matches the XY-model they used.

However, where are model would fail is in predicting spatial onset and clustering of opinion dynamics. This is because of the inherent issues that arose in the two-dimensional model when the real-world sentiment scores were randomly assigned on the lattice. Because there was no initial clustering, they tended to head towards the centre over the time scales we looked at. This is because it's equally likely for there to be positive and negative ones surrounding it, meaning the mean opinion is likely closer towards 0 than towards either the positive or negative side. While the one-dimensional model gave reasonably accurate results, and predicted the increase and onset of opinion ranges, because it was a pairwise interaction it didn't necessarily lead to spatial clustering of opinions. This was similar to the drawback we saw in initial two-dimensional simulations, where the original output for the mean value Deffuant model determined the evolution of opinions, even in the polarised case, making it difficult to determine where opinion domains would form spatially.

There are a few major drawbacks with our study, however. The first is that we only used Ukrainian language data. Because we only queried Tweets in Ukrainian, this naturally excluded all Russian-language tweets in Ukraine. This could easily have induced a bias into our system, as it's likely that tweets in Ukrainian are more likely to come from those who identify as pro-Ukrainian whereas those in Russian are more likely to come from those who identify as pro-Russian, even if they are Ukrainian. We thus have likely left out a substantial part of Ukraine's population, and quite likely a huge source of polarisation within the country. Before any firm recommendations can be made, the Russian-language tweets from Ukraine would need to be analysed and a more complete picture of the society would need to be drawn.

We were also limited by the number of tweets we were able to access. Only being granted ten million tweets a month, we had to settle for only a few hundred thousand per time period. If we wanted to get full sentiment data of Twitter, we would likely need to spend longer collecting tweets to get as many per time period as we possible can. This would give us more opinions to draw from, potentially allowing us to see the close peaks that get smoothed out in our current model via interactions. We also only collected tweet data; if we had saved user metadata, we could have analysed the overall opinions of actual people based on their tweets, getting an average tweet opinion score per-person as opposed to per-tweet. This could likely help us fine-tune the model as if one person tweets something hugely positive then something less positive, their overall opinion is likely in the middle, instead of having two opinions for it. This is something that further study and more time and access to Twitter data could give us.

Other possible avenues for further study could include looking at adding certain political events which greatly tend to influence opinion. Some residue of this could be seen in the failure of our model to predict what would happen in March-June 2015 after the Second Minsk Agreement, the area where we saw the widest jump in the mean of the sentiment data and the highest RSME from our model. Furthermore, we could look into the work that is currently taking place in non-closed statistical system to see what would be applicable to opinion dynamics, which inherently takes place in an open system. Overall, there are multiple avenues for possible future research, and hopefully some will be done to help us better understand what is fostering polarisation in society, as well as fight against it.

7 Acknowledgements

I would like to acknowledge and thank Dr. Vladimir Lobaskin. His thoughtful ideas and ways to look at the Deffuant model, as well as knowledge of statistical physics and the events surrounding the Ukraine-Russian conflicts of the past decade have been seminal in making this work come to light, as has his work with Romemskyy. His feedback has also been very much appreciated, and has led to me learning a lot about opinion dynamics and how we can use it to model real world events as well as how to possibly fight against the polarisation we see in our world.

References

[1] Q. Zha, G. Kou, H. Zhang, H. Liang, X. Chen, C.-C. Li, and Y. Dong, "Opinion dynamics in finance and business: a literature review and research opportunities," *Financial Innovation*, vol. 6, no. 1, pp. 1–22, 2020.

[2] N. Perra and L. E. Rocha, "Modelling opinion dynamics in the age of algorithmic personalisation," *Scientific reports*, vol. 9, no. 1, pp. 1–11, 2019.

[3] M. Romenskyy, V. Spaiser, T. Ihle, and V. Lobaskin, "Polarized ukraine 2014: Opinion and territorial split demonstrated with the bounded confidence xy model, parametrized by twitter data," *Royal Society open science*, vol. 5, no. 8, p. 171935, 2018.

[4] J. v. Neumann, "Theory of self-reproducing automata," *Edited by Arthur W. Burks*, 1966.

[5] H. Föllmer, "Random economies with many interacting agents," *Journal of mathematical economics*, vol. 1, no. 1, pp. 51–62, 1974.

[6] W. B. Arthur *et al.*, *Increasing returns and path dependence in the economy.* University of michigan Press, 1994.

[7] A. Orléan, "Bayesian interactions and collective dynamics of opinion: Herd behavior and mimetic contagion," *Journal of Economic Behavior & Organization*, vol. 28, no. 2, pp. 257–274, 1995.

[8] B. Latané and A. Nowak, "Self-organizing social systems: Necessary and suffi-

cient conditions for the emergence of clustering, consolidation, and continuing diversity," *Progress in communication sciences*, pp. 43–74, 1997.

[9] S. Galam, "Rational group decision making: A random field ising model at t= 0," *Physica A: Statistical Mechanics and its Applications*, vol. 238, no. 1-4, pp. 66–80, 1997.

[10] G. Weisbuch and G. Boudjema, "Dynamical aspects in the adoption of agri-environmental measures," *Advances in Complex Systems*, vol. 2, no. 01, pp. 11–36, 1999.

[11] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents," *Advances in Complex Systems*, vol. 3, no. 01n04, pp. 87–98, 2000.

[12] J.-F. Laslier, *Diffusion d'informations et évaluations séquentielles.* 1989.

[13] G. Deffuant, F. Amblard, G. Weisbuch, and T. Faure, "How can extremism prevail? a study based on the relative agreement interaction model," *Journal of artificial societies and social simulation*, vol. 5, no. 4, 2002.

[14] G. Weisbuch, G. Deffuant, and F. Amblard, "Persuasion dynamics," *Physica A: Statistical Mechanics and its Applications*, vol. 353, pp. 555–575, 2005.

[15] E. Yildiz, D. Acemoglu, A. E. Ozdaglar, A. Saberi, and A. Scaglione, "Discrete opinion dynamics with stubborn agents," *Available at SSRN 1744113*, 2011.

[16] T. V. Martins, M. Pineda, and R. Toral, "Mass media and repulsive interac-

tions in continuous-opinion dynamics," *EPL (Europhysics Letters)*, vol. 91, no. 4, p. 48003, 2010.

[17] M. Pineda, R. Toral, and E. Hernandez-Garcia, "Noisy continuous-opinion dynamics," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2009, no. 08, p. P08001, 2009.

[18] R. Hegselmann, U. Krause, *et al.*, "Opinion dynamics and bounded confidence models, analysis, and simulation," *Journal of artificial societies and social simulation*, vol. 5, no. 3, 2002.

[19] M. Pineda, R. Toral, and E. Hernández-García, "The noisy hegselmann-krause model for opinion dynamics," *The European Physical Journal B*, vol. 86, no. 12, pp. 1–10, 2013.

[20] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis, "Continuous-time average-preserving opinion dynamics with opinion-dependent communications," *SIAM Journal on Control and Optimization*, vol. 48, no. 8, pp. 5214–5240, 2010.

[21] Y. Yang, D. V. Dimarogonas, and X. Hu, "Opinion consensus of modified hegselmann–krause models," *Automatica*, vol. 50, no. 2, pp. 622–627, 2014.

[22] M. A. S. Kolarijani, A. V. Proskurnikov, and P. M. Esfahani, "Macroscopic noisy bounded confidence models with distributed radical opinions," *IEEE Transactions on Automatic Control*, vol. 66, no. 3, pp. 1174–1189, 2020.

[23] C. Wang, Q. Li, B. Chazelle, *et al.*, "Noisy hegselmann-krause systems: Phase transition and the 2r-conjecture," *Journal of Statistical Physics*, vol. 166, no. 5,

pp. 1209–1225, 2017.

[24] B. D. Goddard, B. Gooding, H. Short, and G. Pavliotis, "Noisy bounded confidence models for opinion dynamics: the effect of boundary conditions on phase transitions," *IMA Journal of Applied Mathematics*, vol. 87, no. 1, pp. 80–110, 2022.

[25] E. F. Keller and L. A. Segel, "Initiation of slime mold aggregation viewed as an instability," *Journal of theoretical biology*, vol. 26, no. 3, pp. 399–415, 1970.

[26] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Physical review letters*, vol. 75, no. 6, p. 1226, 1995.

Appendix A: Model Code

The following is the code used to generate the final states for our two types of models.

[ ]:
```
import numpy as np
import csv
import random
```

[ ]:
```
for tol in tolL:
    for noise in noiseL:
        for i in range(3):
            s = np.random.rand(10000)
            agentsRandomT = [Agent(i,tol,0) for i in s]
            opinionRandomT = [i.opinion for i in agentsRandomT]
            for n in range(100000):
                updates = np.zeros(10000)
                noiseList = np.zeros(10000)

                for i in range(len(agentsRandomT)):
                    avg = 0
                    coord = convert2d(i)
                    neighbors = neighborList(coord)
                    for j in neighbors:
                        oneD = convert1d(j)
                        avg += agentsRandomT[oneD].opinion
                    avg /= 24
                    updates[i] = avg
                    noiseList = addNoise(noise,noiseList,i,neighbors)
                if n%10000 == 0:
                    ops = [agent.opinion for agent in agentsRandomT]
                    difference = np.abs(ops-updates)
                    difference = difference[difference >= 1*10**(-2)]
                    if np.all(difference >tol):
                        break

                for i in range(len(agentsRandomT)):
                    agent = agentsRandomT[i]
                    op = updates[i]
                    if np.abs(agent.opinion-op) < agent.tolerance:
                        agent.update_opinion(op,0.5)
                        agent.opinion = agent.opinion + noiseList[i]
                        if agent.opinion > 1:
                            agent.opinion = 1
                        if agent.opinion < 0:
                            agent.opinion = 0
                    else:
                        pass
            opinionRandomT2 = [i.opinion for i in agentsRandomT]
```

```
            with open(f'tol{int(tol*10)}-noise{int(noise*10)}-{i+1}.
↪csv','w',newline='') as csvfile:
                csvwriter = csv.writer(csvfile, delimiter=',',quotechar='|',␣
↪quoting=csv.QUOTE_MINIMAL)
                csvwriter.writerow(opinionRandomT)
                csvwriter.writerow(opinionRandomT2)
```

1 One Dimensional Pairwise-Deffuant

[ ]:
```
def update(agent1,agent2,alpha,noise):
    op1 = agent1.opinion
    op2 = agent2.opinion
    if np.abs(op1-op2) <= agent1.tolerance:
        # Depends on both having the same tolerance
        agent1.update_opinion(op2,alpha)
        agent2.update_opinion(op1,alpha)
    prob = np.random.rand(1)
    noise = random.uniform(0,noise)
    if prob<= 0.1:
        direction = np.random.randint(0,2)
        if direction:
            agent1.opinion = agent1.opinion + noise
            agent2.opinion = agent2.opinion - noise
            if agent1.opinion>1:
                agent1.opinion = 1
            if agent2.opinion<0:
                agent2.opinion = 0
        else:
            agent1.opinion = agent1.opinion - noise
            agent2.opinion = agent2.opinion + noise
            if agent1.opinion<0:
                agent1.opinion = 0
            if agent2.opinion>1:
                agent2.opinion = 1
```

[ ]:
```
noise = [0,0.1,0.2,0.3,0.4,0.5]
tol = [0.1,0.2,0.3,0.4,0.5]
```

[ ]:
```
for i in noise:
    for j in tolerance:
        for l in range(5):
            s = np.random.rand(10000)
            agentsRandom = [Agent(k,j) for k in s]
            opinionRandom = [k.opinion for k in agentsRandom]
            for n in range(5000000):
                cord1 = np.random.randint(0,len(agentsRandom))
```

```
                cord2 = np.random.randint(0,len(agentsRandom))
                agent1 = agentsRandom[cord1]
                agent2 = agentsRandom[cord2]
                update(agent1,agent2,0.5,i)
                opinionRandom[cord1] = agent1.opinion
                opinionRandom[cord2] = agent2.opinion
            with open(f'tol{int(j*10)}-noise{int(i*10)}-{l+1}.
↪csv','w',newline='') as csvfile:
                csvwriter = csv.writer(csvfile, delimiter=',',quotechar='|',␣
↪quoting=csv.QUOTE_MINIMAL)
                csvwriter.writerow(opinionRandom)
```

2 Two-Dimensional Mean Field Value Lattice

[ ]:

```
def convert2d(x):
    """This function will take a number and give its spot on a 100x100 grid␣
↪(0-99)x(0-99) per Python indexing
    as a tuple"""
    xcord = x//100 # Integer division on the number, thus getting whatever the␣
↪100s column would be, giving row
    ycord = x%100 # Modulo 100, giving the remainder, which is its column
    return (xcord,ycord)

def convert1d(x):
    """Takes a coordinate and gives its spot as a list of 10k elements"""
    return x[0]*100 + x[1]

def findNeighbor(x):
    """Takes a 2D coordinate tuple and finds the neighbor within a set of␣
↪paramenters.
    Here we use +/- 2 on each coordinate. Currently only works on a 100x100␣
↪grid"""

    xnew = (x[0]+ np.random.randint(-2,3))%100 # Half-open intervals for numpy
    ynew = (x[1]+ np.random.randint(-2,3))%100
    return (xnew,ynew)

def neighborList(coord):
    """Takes the 2D coordinate tuple of a neighbor and returns a list of all 24␣
↪neighbors of that agent
    Only works on a 100x100 grid, but can easily be adapted"""
    x = coord[0]
    y = coord[1]
    neighbors = []
    for i in range(-2,3):
```

```
        for j in range(-2,3):
            if i == 0 and j == 0:
                pass
            else:
                xnew = (x + i)%100
                ynew = (y + j)%100
                neighbors.append((xnew,ynew))
    return neighbors

def addNoise(noise,noiseList, coord,neighbors):
        prob = np.random.rand(1)
        noise = noise*np.random.choice([-1,1])
        if prob <= 0.1:
            noiseList[coord] = noise
            for j in neighbors:
                oneD = convert1d(j)
                noiseList[oneD] = -noise/24
        return noiseList
```

[ ]:
```
tolL = [0.1,0.2,0.3,0.4,0.5]
noiseL = [0,0.1,0.2,0.3,0.4,0.5]
```

[ ]:
```
for tol in tolL:
    for noise in noiseL:
        for i in range(3):
            s = np.random.rand(10000)
            agentsRandomT = [Agent(i,tol,0) for i in s]
            opinionRandomT = [i.opinion for i in agentsRandomT]
            for n in range(100000):
                updates = np.zeros(10000)
                noiseList = np.zeros(10000)

                for i in range(len(agentsRandomT)):
                    avg = 0
                    coord = convert2d(i)
                    neighbors = neighborList(coord)
                    for j in neighbors:
                        oneD = convert1d(j)
                        avg += agentsRandomT[oneD].opinion
                    avg /= 24
                    updates[i] = avg
                    noiseList = addNoise(noise,noiseList,i,neighbors)
                if n%10000 == 0:
                    ops = [agent.opinion for agent in agentsRandomT]
                    difference = np.abs(ops-updates)
                    difference = difference[difference >= 1*10**(-2)]
                    if np.all(difference >tol):
```

```
                    break

            for i in range(len(agentsRandomT)):
                agent = agentsRandomT[i]
                op = updates[i]
                if np.abs(agent.opinion-op) < agent.tolerance:
                    agent.update_opinion(op,0.5)
                    agent.opinion = agent.opinion + noiseList[i]
                    if agent.opinion > 1:
                        agent.opinion = 1
                    if agent.opinion < 0:
                        agent.opinion = 0
                else:
                    pass
        opinionRandomT2 = [i.opinion for i in agentsRandomT]
        with open(f'tol{int(tol*10)}-noise{int(noise*10)}-{i+1}.
↪csv','w',newline='') as csvfile:
            csvwriter = csv.writer(csvfile, delimiter=',',quotechar='|',␣
↪quoting=csv.QUOTE_MINIMAL)
            csvwriter.writerow(opinionRandomT)
            csvwriter.writerow(opinionRandomT2)
```

[ ]:

Appendix B: Ukrainian Twitter Query Terms

Політика, Україна, РНБО, новини, Кличко, Порошенко, Тимошенко, Турчинов, Янукович, Азаров, Пшонка, Ляшко, Ахмєтов, соцопитування, ОУН, УПА, Донбас, Донецьк, Луганськ, Схід, Захід, Регіони, Регіонів, КПУ, Удар, БЮТ, Батьківщина, країна, Свобода, Путін, армія, бойовики, ЗСУ, РАДА, Крим, МЗС, мир, ОБСЄ, Росія, Харків, АТО, бій, війна, всу, героямслава, Даунбас, Дебальцево, економіка, євромайдан, Єдність, жертви, Заручники, киборги, криза, кіборги, Лугандон, Луганда, майдан, мобілізація, Мінськ, Мінськіугоди, москалі, нацгвардія, небеснасотня, обстріл, ПутінХуйло, переговори, перемога, полонені, ПравийСектор, пропаганда, перемир, санкції, Сепаратизм, славаУкраїні, СБУ, Новоросія, Терористи, Хуйло, укроп, кацап, кацапи, українці, міністр, командування, командуючий, бюджет, верховна, єдність, єдина, Лисенко, сводка, штаб, антитерористична, антитерористичний, операція, контроль, партія, ватник, вата, ватники, москаль, москалям, москалик, москалю, спецоперація, аеропорт, бомбардування, літак, літаки, літаками, бомба, міна, солдат, піхота, полк, взвод, командир, головнокомандуючий, головнокомандувач, верховний, дрг, Маріуполь, Слов'янськ, Краматорськ, Горлівка, Красноармійськ, Артемівськ, Дружківка, Київ, Києва, Дніпропетровськ, Дебальцеве, пропуску, прикордонник, прикордонна, заворушення, протест, протести, Піски, президент, президенту, президентом, євроінтеграція, Бандера, бандерівці, націоналізм, націоналісти, пікет, пікетувальникі, ДОДА, ОДА, адміністрація, політичні, міноборони, закордонний, закордонних, Кернес, Ярош, Семенченко, Симоненко, референдум, анексія, провокація, прем'єр, військова, частина, МВФ, спостерігач, асоціація, євроасоціація, військa, політикум, патріотизм, прем'єрка, Плотницький, Стрєлков, Гіркін, Клімкін, Рогозін, Навальний, Нємцов, Іванов, ЄС, Дещиця, Медведєв, Пургін, Пушилін, Гіві, Толстих, ПАСЕ, ПАРЄ, асамблея, асоціація, Мілонов, Ходаковський, Оланд, Псакі, мінфін, мзс, Коморовський, Эштон, Штайнмаєр, Расмусен, Губарєв, Пономарьов, Боінг, Амвросієвка, Сніжне, Попасна, Попасне, Щастя, Єнакієве, Одеса, Сімферополь, черноморський, Новоазовськ, МДБ, КДБ, волевиявлення, електорат, виборці, виборець, виборча, держпереворот, суверенітет, гривня, інфляція, девальвація, легітимний, нелегітимний, легітимність, нелегітимність, біглий, Бєс, Мюрід, перехоплення, радіоперехоплення, радіорозвідка, розвідка, Тимчук, міліція, міліціант, міліціанти, кіберберкут, окупанти, найманець, найманці, компроміс, спецслужби, спецпризначенці, МІГ, МІ, зброя, Ходорковський, Брідлав, козак, отаман, Керрі, Яценюк, Донецька, Республіка, Захарченко, Дейнего, бойовиків, Інтерфакс, Кучма, Медведчук, вибір, контактної, Моторола, Савченко, нормандському, полон, звільнено, підконтрольна, проукраїнська, активістка, активіст, Лещенко, коаліція, Макрон, зелені чоловічки, кордон, кордону, артилерійських, залпового, Трамп, кібератаки, нацисти, Дуда, Мінським, Авдіївка, НАТО, мирних, батальйон, цілісності, федералізація, Штепа, Самопоміч, Столтенберг, гуманітарка, демонтувати, пам'ятники, Голодомор, Леніну, звірства, тітушки, фашизм, КПРС, Мінські, угоди, бомбардувальник, захоплення, захопили, Містралів, Мінського, протоколу, цивільних, Іловайськом, Іловайськ, Беркут, аеропорт, Інтер, Дебальцевський котел, Меркель, Айдар, Азов, Дебальцевому, припинення, ППО, Парубій, нормандська четвірка, радикальних, націоналіст, моніторингова місія, Небесної Сотні, Революція Гідності, Широкине, агресор, Гройсман, Бородай, Бородая, спостерігачі, зрада, Нормандської четвірки, беркутівець, Саакашвілі, Опоблок, Батьківщина, підрив ЛЕП, СУ-24, збитий, поліція, поліцію, енергоблокада, енергоміст, безвізовий режим, Басурін, Бузини, Бузина, Bellingcat, Поклонської, Поклонськая, Меджліс, Єрофеєв, Александров, Панамські папери, Миротворець, Кремль, Нацполіція, Шеремет, Могеріні, перемир'я, ОРДЛО, заручник, заручників, Антимайдан, загострення, опозиція, Прилєпін, ОРДО, безвіз, Геращенко, Вороненков, стрілянина, Плотницький, Джавелін, Скрипаля, Скрипаль, Кемерові, Зимняя вишня, хакери, Найєм, Бабченка, Бабченко, замінування, Манафорт, Слуга народу, Солідарність, патріарх, Варфоломій, Томос, автокефалія, автокефалію, зеля, московство, павлоштепа, зеленскийпрезидент, зелошары, зессаныч, рашисти, русня, комик, зелохторат, зекоманда, зепрезидент, гетьман, порох, порохобот, держзрада, петяпорошенко, парахоботікі, зрадофіл, свинособаки, свідомий, незалежний

Appendix C: Sentiment Histograms and Summaries

The histograms for each time period can be seen in Figure 21 while the summary for each opinion value can be seen in Table 1.

	≤ -10	≤ -5	-4	-3	-2	-1	0	1	2	3	4	≥ 5	≥ 10
Jul 2014	2	77	84	11	106	400	176820	823	1488	32	24	3	1
Aug 2014	0	270	60	15	49	468	176118	1029	2118	26	39	3	0
Sep 2014	0	68	21	2	118	493	171461	1072	1573	42	30	3	0
Oct 2014	1	51	15	2	78	50	177781	825	1363	32	24	3	0
Nov 2014	0	156	8	6	155	348	168272	950	1563	21	23	1	1
Dec 2014	0	20	64	4	152	343	173587	777	1917	117	28	2	0
Jan 2015	0	83	37	90	309	597	173191	645	1992	64	43	1	0
Feb 2015	0	15	34	8	90	390	159330	818	1622	59	106	1	0
Spr 2015	104	2999	344	120	591	2360	972230	3831	6974	150	131	5	0
Aut 2015	8	1990	81	272	722	3129	958473	3726	6656	140	117	7	0
Spr 2016	30	2889	103	131	762	2373	944139	4644	7495	100	86	20	1
Aut 2016	22	1889	58	289	1002	2581	913515	4004	6198	97	221	4	0
Aut 2017	15	2033	41	135	1007	2383	599544	3281	3591	66	52	4	0
Aut 2018	23	1783	64	271	982	2219	648544	4523	3584	102	66	9	0
Aut 2019	14	1849	106	393	1387	3696	667236	5776	3994	167	74	10	0
Aut 2020	105	2871	47	460	1169	3757	662793	5069	4503	114	67	11	0

Table 1: Table of Opinion Values Per Time Period

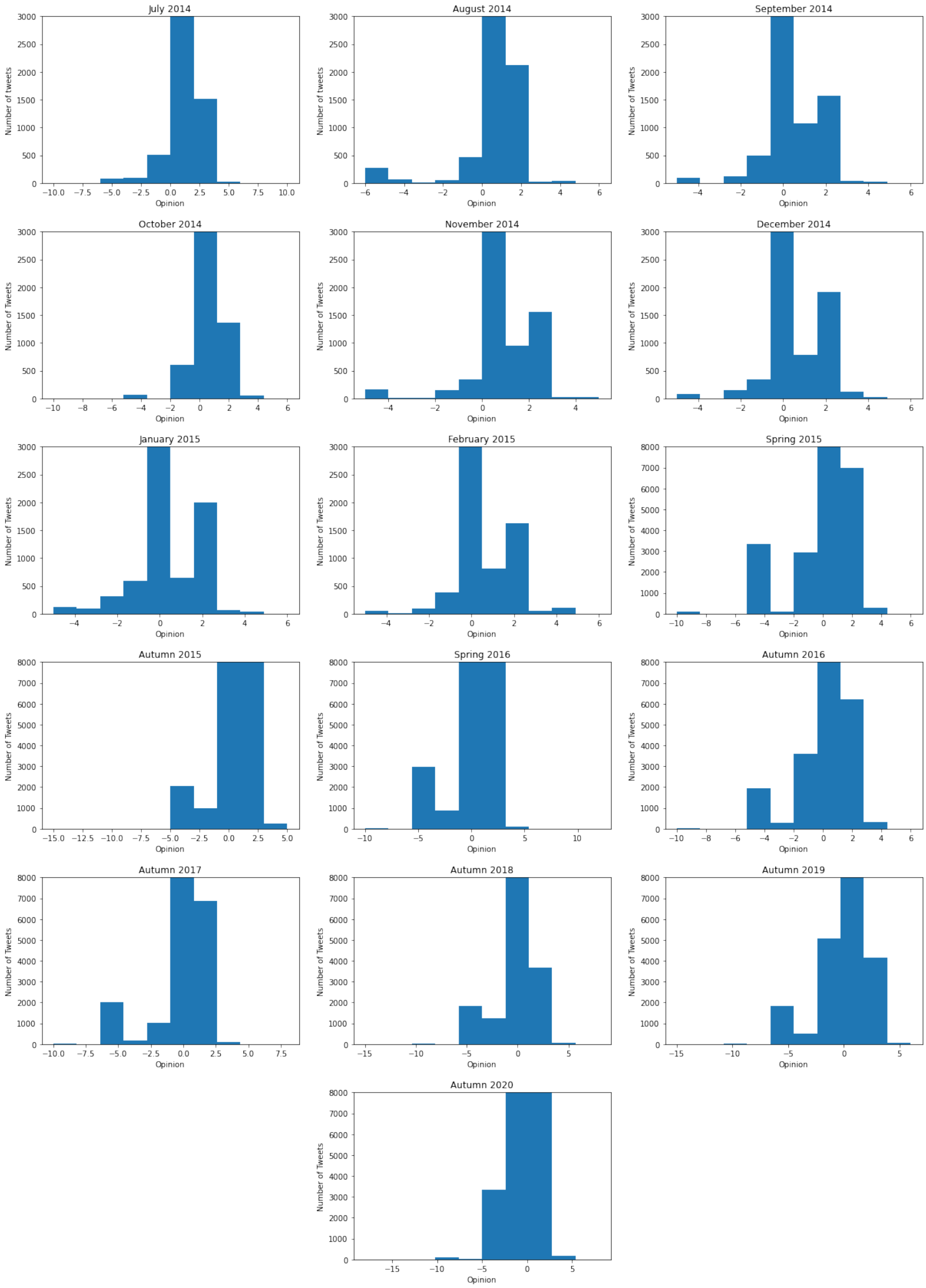

Figure 21: Sentiment Histograms by Time Period