University of Luxembourg

Experimental mathematics 1

Voter models & influencers in social networks

Authors:
Oliver Philip Jack
Francesca Feyereisen
Noah Thielen

Supervisor:
Dr. Simon Campese

June 4, 2021

Abstract

The following paper is a detailed summary of our research about voter models and the impact of influencers in social networks. Based on our findings, we will establish some conjectures. Our main goal is to broaden our and our readers' understanding of the evolutionary behaviour of voter models.
We would like to thank Dr. Campese for all of his help and support throughout our project.

Contents

1 Introduction

The voter model (abbr. VM) can be defined as an ideal system or network of elements, such as people, particles, et al., interacting with one another, combined with a set of well-defined rules, which determine how the system evolves over time. Each of these elements has one or multiple attributes, e.g., an opinion, a colour, a political view, that can be influenced by its surroundings. The main role of the model is to help understand the development of these attributes over time, on an individual level, as well as on a global one.

Considering the classic voter model, these elements, also known as "voters", can be represented as nodes or vertices of a connected static graph. Each node has one of two possible opinions or states. The connection of two nodes, which can be illustrated by edges, indicates a certain interaction between both, allowing them to influence each other. A simple update rule of the opinion evolution can be defined as following:

1. randomly select a voter from the graph
2. randomly select a neighbour, who shares an edge, of different opinion
3. change this voter's opinion (if there is no such neighbour, then nothing happens, however this still counts as a step)

Mathematically, this evolution rule can be expressed as follows:
Let (x_1, y_1) be the coordinates of a chosen square and let (x_2, y_2) be the coordinates of one of its neighbours of different opinion. If $\sigma_{(x_1,y_1)}, \sigma_{(x_2,y_2)} \in \{0, 1\}$ represent their respective opinions, then the update rule will have the subsequent effect:

$$\sigma_{(x_2,y_2)} = 1 - \sigma_{(x_1,y_1)}$$

An interesting thing to analyse is, how this voter model will evolve over time, if we repeatedly apply these same rules. Will a consensus, where the proportions of the different opinions no longer vary more than a given percentage, eventually be reached, and if so, how long will it take? Does this depend on the initial set-up or proportions of the graph? What happens if we change the selection rules or add a third opinion to the mix?

On the other hand, another important model to consider is one, where initially all the voters of the graph share the same "neutral" opinion. Next, we can place so called "influencers", who have their own non-neutral opinion, in a specific position of the graph. By applying similar rules as above, we can then ask ourselves: is there a position that favours the success of an influencer's opinion? What happens if we change the rules, by instead of selecting randomly, choosing the next voter based on a set of probabilities?

These are all questions we asked ourselves at the beginning of our research phase and ones for which we have come up with experimental answers and results, some of which were quite predictable, others rather surprising.

Over the last century the study of voter models has become an increasingly relevant topic, part of the mathematical theory of probability and statistical physics. Nowadays, it has become essential for big companies, political parties, et al., to fully comprehend the impact that influential advertising and campaigning can have on our society.

To help understand the functionality and evolution of such a voter model, we will first give several examples of how these graphs can be represented visually, as well as some modern-day use cases. Next, we will focus on explaining all the details of the basic case of the voter model, followed by an analysis of what happens, if the number of opinions/colours in the grid and the grid's dimension are changed. Furthermore, we will also take a closer look at the so-called "clustering" phenomenon, a special behaviour of the voter model, which reoccurs regularly. Moreover, we will examine a different set-up of the classical voter model, namely one, where so-called "influencers" play a vital role in the propagation of the opinions/colours. Finally, we will give a short summary of some already well-known mathematical results and theorems, before presenting our final conclusion. Last but not least, we will exhibit the Python codes we used for our simulations in the appendix, followed by our references.

2 Different graphical representations of networks

Although there are many ways to visually represent these network graphs, we have chosen three particular examples, to explain more precisely.

2.1 Network diagrams

In graph theory, a network diagram is considered to be a network, where nodes (i.e., voters) are represented by dots and their interconnections or edges by lines running from one node to the other.

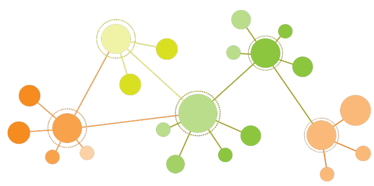


Figure 1: Example of a network diagram

Furthermore, the complexity of the network diagram can be altered, by adding supplementary features to both the nodes, as well as the edges. For example, the nodes can be given certain values and the edges can be given specific weights and directions.

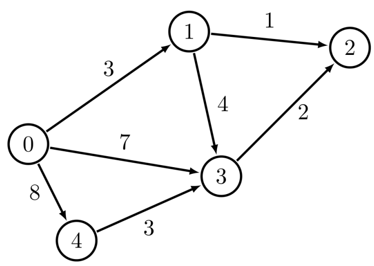


Figure 2: Example of a weighted & directed network diagram

In general, it is an ideal way to map the layout of a network, as its visual representation makes it easier for the outside observer to understand how elements are connected.

2.2 Matrices

Another straightforward way of representing a network can be done by using matrices. If we consider a $n \times m$ matrix ($n, m \in \mathbb{N}$), then we can regard the positions of the coefficients of the matrix as the positions of our voters, while the values of the coefficients represent that voter's opinion. Here the influenceable neighbours of a voter can be seen as the surrounding coefficients, which add up to no more than 8 neighbours. One of the simplest and most common cases is when the entire matrix is filled with the values 0 and 1, both standing for different opinions. Here we show an easy example of how the evolution of a 3×3 matrix, applied with the update rule we defined above, could look like (simulated using one of our programs):

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

2.3 Grids

Finally, a third way to represent networks graphically is by using grids of size $n \times m$ ($n, m \in \mathbb{N}$). Each square of the grid can be seen as a voter and is therefore filled with a specific colour, which stands for their opinion. Similar to the case of matrices, each voter can have a maximum of 8 surrounding neighbours, whose interactions can lead to a change of opinion.

Moreover, we chose this type of graph for most of our visual simulations, as the switch of opinions stands out colourfully and as it is quite simple to code a program to be able to return such a grid, applied with divergent variations of the evolution rule mentioned above. Hence, there will be many examples of this sort of graph throughout our paper.

3 Real-world use cases of network analysis

Let us mention some of the most popular real-world use cases of network analysis.

3.1 Electrical network analysis

In an electrical sense, a network is considered to be a collection of connected elements, which essentially form an electrical circuit. In this case, network analysis is mainly used to understand how and with what properties the electric current flows through the circuit, as well as finding out the exact voltages. This type of circuit analysis is especially important when it comes to setting up and evaluating the efficiency of electric circuits or networks.

3.2 Biological network analysis

Throughout the last century, scientists have been able to gather important scientific data, which has led to an increase of biological network analysis. Biological systems can often be represented by networks, built up of complex connections between two separate entities. The main goal of this research is to find repetitive patterns, as well as abnormalities, to better our understanding of these interconnections. As these interactions happen on all sorts of levels, from microscopic to universal, there are a huge amount of different network models, such as "ecological, neurological, metabolic or molecular interaction networks" ([10], para. 1). This can be seen in the image below:

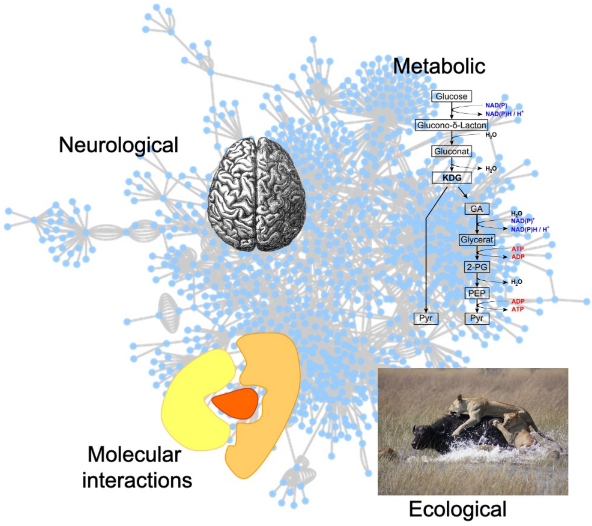


Figure 3: Examples of biological network models

3.3 Social network analysis

Finally, social network analysis focuses on the structure of relationships between social groups. In a world where all of humanity is as connected as never before, it is not surprising that the implementation of social networks has gained of importance in almost every branch (e.g., politics, sociology, geography, economics). By collecting as much data from citizens as possible, governments and companies can become aware of social connections between people. Thanks to social network analysis, many of these social structures, like social media networks, friendship and family networks, and business networks, can be represented visually by graphs as seen below, often referred to as "sociograms".

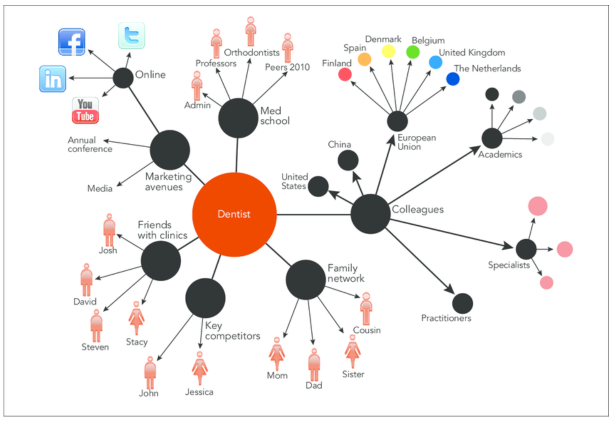


Figure 4: Example of a sociogram

Next, we will take a closer look at different variations of the evolution of the voter model, while using the grid mentioned above for our visual representations. These alterations will for example consist of changing:

- the initial set-up of the grid
- the number of used opinions/colours
- the size & dimension of the grid

To improve the readability of our document, we have chosen to order our experimental results into logical groups.
For all of our simulations, we used the programming language Python, including the plugin "Pygame", to run through our simulations and gather data for our detailed analysis.

4 Basic case

Let us start off with the basic case, to which we will refer to regularly throughout our paper. Consider a square grid of size $n \times n$ ($n \in \mathbb{N}$), filled with squares of two different colours, which can be seen as our voters attributed with a specific opinion. We then get a voter model, by repetitively applying the update rule, mentioned above, to our grid.

Naturally, there were a few simple questions we asked ourselves at the beginning, regarding the evolution of the voter model:

- Given a certain initial set-up, meaning given a specific proportion of voters of "good" opinion or squares of the colour red at the start, will our voter model eventually reach a state of equilibrium/consensus?
- In the case where an equilibrium is achieved, is it always a unanimous one?
- Does the outcome depend on the initial proportion of a certain opinion/colour or on the size of the grid?

To fully understand our experimental results, we first need to lay the groundworks, by giving some necessary definitions. It is important to notice that we came up with these definitions ourselves and that these are by no means the "unique correct" definitions of these terms.

4.1 Definitions

Definition 4.1.1 (Neighbours)**.** A neighbour of a square is considered to be any other square that is directly connected to that square, meaning that it either shares an edge or a corner with that square. In the 2-dimensional case, each square can have up to 8 neighbours in total.

Definition 4.1.2 (Update rule)**.** The update rule, as already mentioned above, is the main rule, which determines the evolution of the voter model. It may be changed from case to case; however, it always has to be clearly defined.

Definition 4.1.3 (Step)**.** A step can be defined as one execution of the update rule.

Definition 4.1.4 (nD)**.** We use nD as the abbreviation for n-dimensional, for $n \in \mathbb{N}$.

Definition 4.1.5 (Dominant opinion/colour)**.** An opinion/colour is defined as dominant, if it holds a larger proportion of the grid than another opinion.

Definition 4.1.6 (Eliminated opinion/colour)**.** An opinion/colour has been eliminated, if there is no longer a voter/square, who has this opinion/colour.

Definition 4.1.7 (Unanimous grid)**.** A grid is unanimous, if all the voters/squares share the same opinion/colour.

Definition 4.1.8 (Equilibrium/consensus)**.** An equilibrium/consensus can be defined as the state of the grid, when the proportions of the different opinions/colours no longer vary more than a given percentage. For our programs, we decided that an equilibrium is reached if the grid is unanimous or if the proportions of the opinions/colours do not vary more than a specific threshold for at least a certain number of steps. Here, it is important to choose the number of steps (our default value: $10 \times$ the number of squares in the grid), as well as the threshold (our default value: 1%) adequately.

Definition 4.1.9 (Clustering)**.** Clustering can be defined as the grouping of voters/squares who share the same opinion/colour. These tend to form clusters of different shapes and sizes, from where the name originates.

4.2 Visual simulations vs analytical simulations

In order to gather conclusive data, we had to run through two different types of simulations, visual and analytical ones. (The main Python codes of these two programs can be found at the end of our paper.)

On the one hand, the visual simulations allowed us to make some first pre-assessments. By taking screenshots of the grid over regular step intervals, we were able to get a first understanding of how the grid would evolve over time. Moreover, the statistics, concerning the evolution of the visual simulations, were saved in EXCEL files, in order to refer to any data if needed.

On the other hand, the purpose of the analytical simulations was to run through each scenario a certain amount of times (our default value: 100 times) and then to work with the average behaviour of our voter model, to see if this overlapped with our initial assessments. By storing this data in EXCEL files, we were able to look for reoccurring patterns.

(Note: In general, when describing the visual simulations, we use the terms squares and colours, whereas for the analytical ones we use the terms voters and opinions. However, these can be interchanged without loss of meaning.)

4.3 Visual examples

The following images were taken from concrete simulations we ran, for different initial set-ups of the squares and their colours (in these examples, all the grids are of the size 100×100):

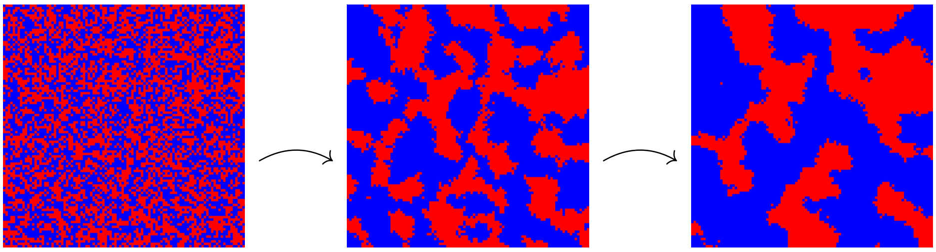


Figure 5: Initial proportions: 50% red & 50% blue

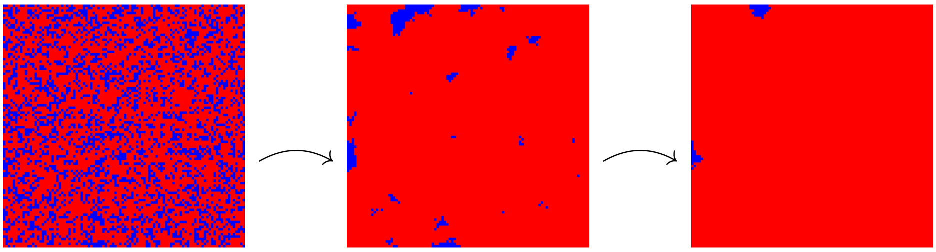


Figure 6: Initial proportions: 65% red & 35% blue

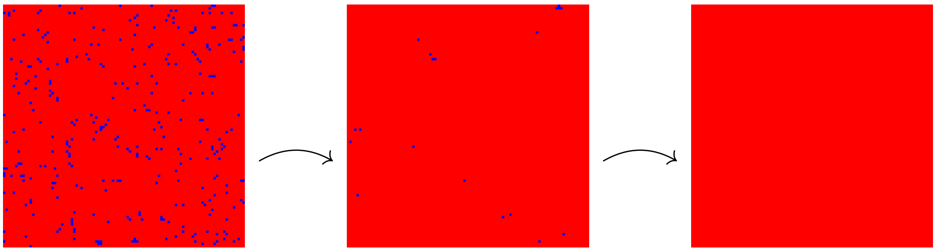


Figure 7: Initial proportions: 95% red & 5% blue

4.4 First observations

After having looked closely at the images with different initial set-ups (initial "good" percentage varied from 50% to 95%) and grid-sizes (varied from 100×100 to 500×500) and comparing the outcomes with the data from our analytical simulations, we were able to come up with our first observations:

- In the case, where the proportions of both opinions were initially equal (so in this case with 2 opinions, 50%/50%), the outcome was inconclusive. Occasionally, one opinion would gather momentum and start to dominate the grid, by gradually growing its proportion. In this case, the opinion having gained momentum early on, would usually go on to fill the grid unanimously. However, the most common case for this initial set-up was that, although both proportions fluctuated randomly, they tended to stay close to their initial proportions of 50%/50%. This can be seen if we check the analytical simulation of this case:

Grid-size	Steps until equilibrium	% at equilibrium
100	36541.94	49.9821
150	47287.88	49.4856
200	64748.87	49.85108
250	74599.23	50.23037
300	55451.93	50.14
350	66120.8	49.89491
400	71403.84	49.965
450	101281.2	50.005
500	67551.43	49.985

Here we can observe that, on average, the percentage of the "good" opinion at our defined equilibrium was around 50%, meaning that the percentage of the "bad" opinion was also around 50%. However, we can see that the number of steps until our consensus was reached, was quite random and unpredictable. (Remember that in this case an equilibrium can also be achieved, if the proportions of the opinions no longer vary more than a given threshold, for a certain number of steps.)

- In all the other cases, where the initial proportions of the opinions were not equal, the opinion which held the bigger initial proportion of the voters tended to be superior to the other opinion, meaning that more often than not this opinion would eventually fill the grid unanimously. Here are some of our analytical results confirming this:

Table 1: Initial proportions: 70% red & 30% blue

Grid-size	Steps until equilibrium	% at equilibrium
100	28198.34	99.414
150	64234.11	99.49462
200	109661.3	99.46635
250	173986.8	99.49211
300	244709.7	99.49
350	339986.6	99.51908
400	431680.3	99.5
450	557307.5	99.51457
500	675512.8	99.5

Table 2: Initial proportions: 90% red & 10% blue

Grid-size	Steps until equilibrium	% at equilibrium
100	11728.58	99.8228
150	15637.54	99.54222
200	33518.24	99.59
250	42591.29	99.5152
300	63317.65	99.515
350	82854.85	99.50776
400	107668.8	99.5
450	136688	99.50469
500	168487.7	99.5

Furthermore, we found out that the probability of this happening grew, when the difference between the two proportions became bigger, e.g., in the case 95%/5%, the chances of a unanimous grid occurring, were higher than in the case 55%/45%.

- A further observation was that, for those cases where the initial proportions of opinions were not equal, the number of steps until an equilibrium occurred depended on both the initial proportions, as well as the size of the grid. On the one hand, the bigger the difference between the two percentages, the fewer amount of steps were needed to reach a unanimous state. On the other hand, the larger the grid-size was, the higher the number of steps needed to reach a unanimous state was. Both these claims can be seen by analysing our analytical results individually, like we already did for the tables above, as well as comparing them with each other:

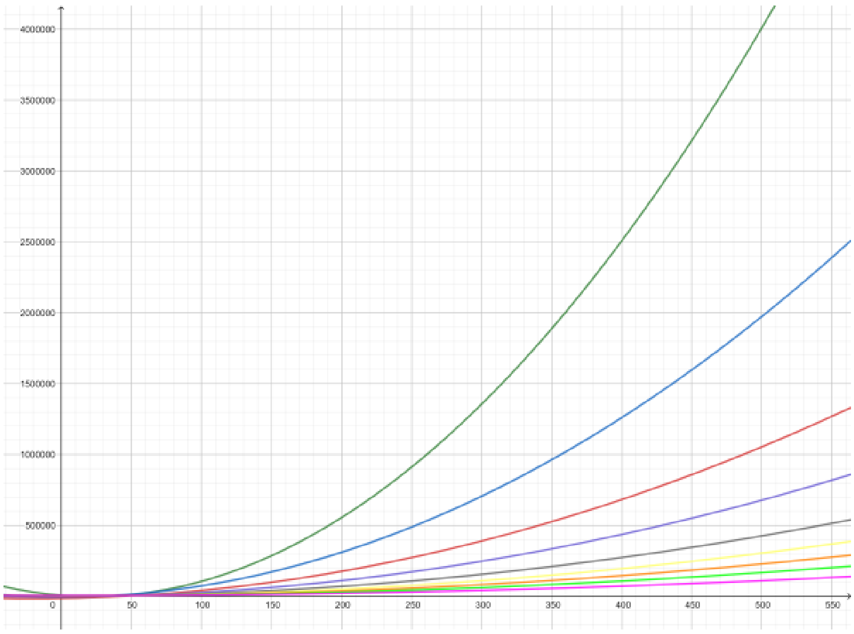


Figure 8: Comparison of the number of steps to reach an equilibrium as a function of the grid-size

In this graph, the dark green function represents the case where the initial proportions of the opinions were 55%/45%. Afterwards, the functions become flatter and more inferior, as the difference between the initial proportions grows, all the way to the pink function, which stands for the case 95%/5%.
Furthermore, by curve fitting the plotted data, we were able to observe that the growth of the number of steps until the equilibrium with regard to the grid-size is quadratic. After having tried to approximate our curve with different types of functions (i.e., linear, exponential), we concluded that a quadratic function gave us a precise enough approximation. Additionally, we can clearly see that the coefficient of the quadratic term of our functions diminishes, as the difference between the two percentages grows. Hence, by curve plotting all of the collected data, we came up with a "general equilibrium function" of two variables x and y, which represent the grid-size and the initial percentage of the "good" opinion respectively (equilibrium threshold: 1%):
For $x \in [100, 500], y \in [0, 1]$:

$$f(x,y) = (1037y^4 - 3477.1y^3 + 4366.7y^2 - 2439.8y + 513.68) \cdot x^2 + (-246322y^4 + 771570y^3 - 895948y^2 + 456249y - 85774) \cdot x$$

Here, it is important to note that this general equilibrium function depends on the number of repetitions, which were executed for each separate scenario. As indicated before, we chose to run each single case 100 times and then work with the averages,

however fewer or more repetitions could lead to variations of this equilibrium. Furthermore, it also depends on the chosen threshold for our equilibrium. For example, we could also look at the general equilibrium function for a threshold of 0.5% and notice that it varies slightly compared to the one above:
For $x \in [100, 500], y \in [0, 1]$:

$$f(x, y) = (3837.7y^4 - 12436y^3 + 15056y^2 - 8078.7y + 1624.3) \cdot x^2 + (-535729y^4 + 2E + 06y^3 - 2E + 06y^2 + 1E + 06y - 200394) \cdot x$$

- Finally, for the visual simulations, we were able to see that in all of our cases, the coloured squares tended to form sorts of "groups" or "colonies", a process called "clustering". Particularly, for the cases where one colour was eliminated in the end, we saw that the last "surviving" squares of the inferior proportion were usually grouped into clusters. The formation of these clusters seemed to be essential for the dominance/survival of the opinions, as they, on the one hand, made the capture of single surrounded squares by the opposition easier, and, on the other hand, offered protection for those squares lying in the core of the clusters. The concept of clustering will be further analysed later on in the paper.

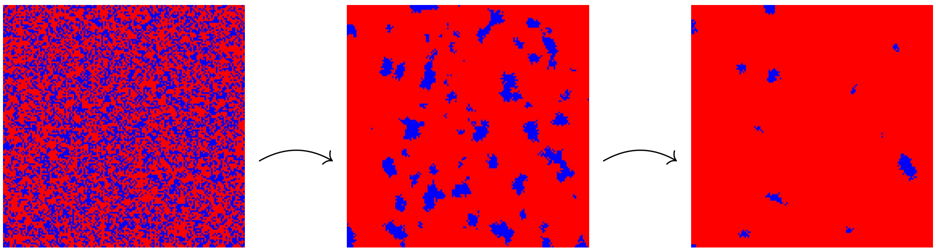


Figure 9: Grid-size: 150×150 / Initial proportions: 60% red & 40% blue

4.5 Conjectures

As a result, we have come up with the following conjectures with regards to our observations:

- If the initial proportions of the two opinions are equal, it takes an indeterminate number of steps to reach an equilibrium, at which, on average, we have the same proportions as at the beginning.

- If the initial proportions of the two opinions are not equal, the dominant opinion tends to eventually fill the grid unanimously, with the probability of this happening increasing, when the difference between the initial proportions grows.

- In those cases, where the initial proportions of the opinions are not equal, the amount of steps to reach a state of equilibrium grows, as the difference between the initial proportions diminishes.

- In those cases, where the initial proportions of the opinions are not equal, the number of steps to reach a state of consensus grows quadratically with regard to the grid-size.

- In all of the cases, the voters tend to group into clusters over time.

5 Variation of the number of opinions/colours

After having come up with our first observations and conjectures for the basic case of the voter model, we continued our research by asking ourselves: would the evolution of the voter model be fundamentally different, if our model had more than two opinions/colours at the start? To answer this question, we took a closer look at the following cases.

5.1 3 opinions/colours

First of all, in total, we tested 4 initial set-ups based on different proportions of the three opinions, each for grid-sizes ranging from 100×100 to 500×500. Let us enumerate these cases:

1. proportions of all three opinions are the same (relative proportions: $(1, 1, 1)$)
2. one dominant opinion & two weaker opinions (relative proportions: $(2, 1, 1)$)
3. two dominant opinions & one weaker opinion (relative proportions: $(2, 2, 1)$)
4. regular decreasing dominance of the opinions (relative proportions: $(3, 2, 1)$)

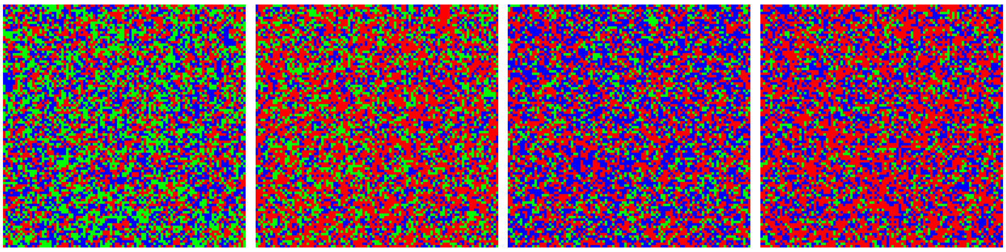


Figure 10: Ordered examples of the 4 cases at the start (grid-size: 100×100)

Let us go through the different cases:

- In case 1., we saw a similar outcome as with the 2-opinions case, where the initial proportions were equal. Although the behaviour of our voter model was quite unpredictable and the number of steps to reach a state of equilibrium fluctuated randomly, on average, the proportions at the equilibrium were equivalent to those at the beginning.

Grid-size	Steps until equilibrium	Opinion 1 % at equilibrium	Opinion 2 % at equilibrium	Opinion 3 % at equilibrium
100	10933.78	33.6965	32.8439	33.4596
150	10982.16	33.22964	33.37013	33.40022
200	12519.51	32.98913	33.4799	33.53098
250	3252.1	33.27797	33.35584	33.36619
300	638.62	33.29257	33.33422	33.37321
350	4760.11	33.32479	33.36329	33.31192
400	2776.12	33.30058	33.36456	33.33486
450	4567.2	33.33637	33.39167	33.27196
500	9812.1	33.32086	33.28961	33.38953

- In case 2., we know that the dominant opinion occupies 50% of the grid at the start, whereas the two weaker opinions each occupy 25% of the grid. At first one might think that the evolution of this case should be similar to the one of two opinions with equal initial proportions, as the dominant opinion and the two weaker opinions combined both cover half of the grid. However, we immediately observed that there was no type of coalition between the weaker opinions, as they not only had to protect themselves from the dominant opinion, but also from each other, which inevitably lead to the equally timed elimination of both in the end. As a result, it is interesting to see that even without occupying over 50% of the grid, an opinion can still fill the grid unanimously over time.

Grid-size	Steps until equilibrium	Opinion 1 % at equilibrium	Opinion 2 % at equilibrium	Opinion 3 % at equilibrium
100	57436.15	97.8837	1.022	1.0943
150	122823.9	98.01907	0.956711	1.024222
200	218720.7	98.07745	0.942725	0.979825
250	347990.6	98.13101	0.920144	0.948848
300	505053.5	98.2019	0.927133	0.870967
350	682399.4	98.1779	0.925829	0.896269
400	894126.2	98.29222	0.8533	0.854481
450	1118539	98.16727	0.918731	0.913995
500	1371900	98.17864	0.917292	0.904064

- In case 3., we have that the dominant opinions both possess 40% of the grid, while the weaker opinion only possesses 20%. Here, we were able to observe that the majority of the time, the weaker opinion was eliminated over time, leaving the two dominant opinions to battle it out between each other. From then onwards, on average, the voter model evolved similarly to the case of two opinions with equal initial proportions. This shows that most of the time, the two dominant opinions were equally as responsible for the elimination of the weaker opinion, otherwise they would not both have a proportion of around 50% at the state of equilibrium.

Grid-size	Steps until equilibrium	Opinion 1 % at equilibrium	Opinion 2 % at equilibrium	Opinion 3 % at equilibrium
100	27464.98	48.6943	47.9969	3.3088
150	62276.37	48.69956	48.24391	3.056533
200	114646.3	48.8424	48.48265	2.67495
250	180635.8	48.66736	48.64357	2.689072
300	256013.2	48.74803	48.57873	2.673233
350	355238	48.97673	48.4497	2.573576
400	448565.9	48.56604	48.77615	2.657806
450	585873.9	48.90412	48.49016	2.605719
500	707765.7	48.64669	48.73488	2.618428

- In case 4., the initial percentages of the three opinions are around 50%, $33.\overline{3}\%$ and $16.\overline{6}\%$. We saw that this case evolved similarly to case 2., in the sense that the most dominant opinion, which occupied half of the grid at the start, could eventually eliminate the other two weaker opinions and fill the grid unanimously. However, there was one main difference, namely the fact that, on average, the opinion with an initial percentage of $33.\overline{3}\%$ was able to survive for a longer period than the opinion with an initial percentage of $16.\overline{6}\%$.

Grid-size	Steps until equilibrium	Opinion 1 % at equilibrium	Opinion 2 % at equilibrium	Opinion 3 % at equilibrium
100	65917.8	94.8901	4.9967	0.1132
150	156351.6	95.55547	4.371556	0.072978
200	270535.9	95.5536	4.3725	0.0739
250	436329.7	95.58258	4.355344	0.06208
300	623886.8	95.7671	4.174867	0.058033
350	841572.8	95.72924	4.203869	0.06689
400	1093568	95.65631	4.27595	0.067744
450	1388146	95.847	4.087901	0.065096
500	1749352	95.7943	4.146108	0.059592

(Note: In all four of these cases, we ran through each single scenario 100 times, while using a consensus threshold of 1%. By running through each individual case more often and by reducing the threshold, the obtained results become even more precise.)

5.2 4 opinions/colours

Similarly, to the previous case, we also experimented with 4 initial set-ups based on different proportions of four opinions (each for grid-sizes ranging from 100×100 to 500×500). Again, we can enumerate these cases:

1. proportions of all four opinions are the same (relative proportions: $(1, 1, 1, 1)$)
2. one dominant opinion & three weaker opinions (relative proportions: $(2, 1, 1, 1)$)
3. two dominant opinions & two weaker opinions (relative proportions: $(2, 2, 1, 1)$)
4. regular decreasing dominance of the opinions (relative proportions: $(4, 3, 2, 1)$)

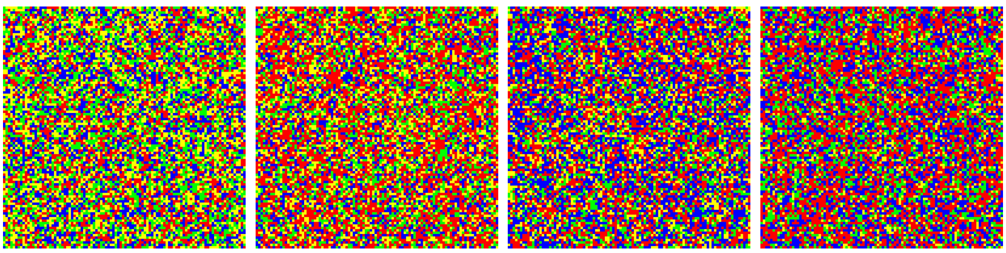


Figure 11: Ordered examples of the 4 cases at the start (grid-size: 100×100)

Let us go through the different cases shortly:

- In case 1., we came up with the exact same results as for the 3-opinions case 1., only that our initial and equilibrium percentages were around 25% each, which makes sense as we are working with four different opinions here.

Grid-size	Steps until equilibrium	Opinion 1 %	Opinion 2 %	Opinion 3 %	Opinion 4 %
100	5307.98	24.7669	25.1842	24.7191	25.3298
150	8330.89	25.09218	24.89427	24.97884	25.03471
200	4850.92	25.08008	24.95948	24.9454	25.01505
250	1935.33	24.98427	25.04059	24.95227	25.02286
300	3247.81	24.97142	24.93052	25.04339	25.05467
350	2783.22	24.95894	24.98477	25.00426	25.05203
400	3456.72	25.01014	25.01117	24.98869	24.99
450	5238.04	24.99998	25.00319	25.00778	24.98905
500	6347.39	24.97425	25.01211	25.01642	24.99723

- In case 2., our conclusions were similar to the 3-opinions case 2.. Although this time, the three weaker opinions combined ($3 \times 20\%$) held a larger proportion of the grid than the dominant opinion alone (40%), on average, the dominant opinion was still able to eliminate the three weaker opinions equally as fast, leading to a unanimous grid eventually.

Grid-size	Steps until equilibrium	Opinion 1 %	Opinion 2 %	Opinion 3 %	Opinion 4 %
100	83735.53	95.5138	1.5241	1.5538	1.4083
150	191257.2	95.97276	1.292978	1.399822	1.334444
200	343151.2	96.3605	1.146375	1.270325	1.2228
250	536773.8	96.25944	1.27848	1.198576	1.263504
300	776336.6	96.36452	1.206533	1.241178	1.187767
350	1058307	96.30841	1.252367	1.242155	1.197069
400	1381098	96.45833	1.192675	1.162269	1.186731
450	1750268	96.56842	1.155151	1.118316	1.158109
500	2169658	96.52827	1.199808	1.141232	1.130688

- In case 3., again, we came up with an almost identical outcome as for the 3-opinions case 3., namely that the two dominant opinions would eliminate both weaker opinions and then battle it out evenly between each other. This then led back to the equally proportioned 2-opinions case, which we have already described in detail.

Grid-size	Steps until equilibrium	Opinion 1 %	Opinion 2 %	Opinion 3 %	Opinion 4 %
100	42673.94	47.8213	45.6059	3.4518	3.121
150	97133.75	47.45249	46.98133	2.797956	2.768222
200	167049.3	46.88138	47.50263	2.7742	2.8418
250	266243.7	46.96472	47.41883	2.85624	2.760208
300	394436.2	47.47612	47.16757	2.6921	2.664211
350	523497.1	47.33715	47.30411	2.649249	2.709494
400	684553.1	47.26323	47.35659	2.682469	2.697706
450	855703.4	47.33466	47.27869	2.675398	2.711259
500	1072385	47.35888	47.33962	2.636588	2.664908

- In case 4., we came up with the same interpretation as for the 3-opinions case 4.. Over time, the weaker opinions tended to be eliminated one by one, while the proportion of the dominant opinion grew. We were also able to see that, on average, the disappearance of the weaker opinions happened in an increasing order of their initial proportion.

Grid-size	Steps until equilibrium	Opinion 1 %	Opinion 2 %	Opinion 3 %	Opinion 4 %
100	93075.54	86.6964	12.1777	1.0443	0.0816
150	226996.8	87.80644	11.17044	0.981556	0.041556
200	410076.2	88.32623	10.8546	0.764075	0.0551
250	649234.8	87.88166	11.24325	0.834384	0.040704
300	930594.7	87.97989	11.1393	0.852233	0.028578
350	1291916	88.4623	10.72762	0.780547	0.029527
400	1679137	88.2917	10.85999	0.819031	0.029275
450	2177885	88.4781	10.70764	0.784277	0.02998
500	2671358	88.55875	10.65568	0.758992	0.026576

5.3 Conjectures

As a result, we have come up with the following conjectures for a grid with n opinions initially ($n \in \mathbb{N}_{\geq 2}$), with regards to our observations:

- If the initial proportions of the opinions are equal, then, on average, an equilibrium state will occur after an unpredictable number of steps, with the different opinions all still holding equal proportions of the grid.
- If there is one single dominant opinion in the voter model, without necessarily covering more than 50% of the grid, then, on average, it will be able to obtain a unanimous grid eventually, after eliminating the weaker opinions.
- If there are multiple equally dominant opinions in the voter model, then, on average, all the weaker opinions (if there are any) will be eliminated over time, leaving the dominant opinions to battle it out evenly among each other. In this case, we tend to have the exact same scenario as explained in our first conjecture.
- If the different opinions have decreasing dominance in the voter model, meaning if the initial percentages of the opinions can be ordered in a decreasing way (e.g., 40%, 30%, 20% & 10%), then, on average, the most dominant opinion will eventually become unanimous, while the opinions with the lowest initial proportions tend to be the first to be eliminated.

6 Variation of the grid's dimension

Following the conjectures based on the variation of the number of opinions/colours in the voter model, we asked ourselves a further interesting question: would the evolution of the voter model be fundamentally different, if we considered our grid to have a different dimension (i.e., 1, 3 or 4 dimension(s))? Therefore, we took a closer look at the following cases. As it is extremely hard, or even impossible in some cases, to represent the grid, if its dimension is greater than 2, we decided to primarily use our analytical program to collect our data.

6.1 1 dimension

Let us start off by considering a 1-dimensional grid, which can also be represented by a simple line segment of squares. In a mathematical sense, this means that we are working with a coordinate system, where every square only has one coordinate. As it is still possible to simulate this case visually, we decided to do so for grid-sizes ranging from 100 to 500 and for initial opinion proportions ranging from 50%/50% to 95%/5%. Hence, we came up with the following observations:

- In most cases, we were able to see the formation of 1D clusters, represented as continuous line segments, where all squares share the same opinion. These line segments were of arbitrary lengths, with the dominant opinion usually having the longest clusters.

- In the majority of the cases, after letting the simulation run for over 10 $\times$ the total number of squares, the initial proportions remained almost preserved, with the only change coming from the distribution of the squares.

Figure 12: Grid-size: 100 / Initial proportions: 50% red & 50% blue

Figure 13: Grid-size: 100 / Initial proportions: 90% red & 10% blue

Following some first visual observations, we ran our analytical program, working with grids of different set-ups of 2 opinions and of grid-sizes ranging from 10˙000 to 100˙000, while using a consensus threshold of 1% and repeating each single case 50 times. Let us go through the obtained results:

- The cases, where the opinions held close to equal proportions initially, behaved in a similar way to the same cases of the 2D voter model. In general, the percentages of the opinions remained almost identical at the equilibrium, however, the number of steps it took to reach this state was quite random, as we can see in the table below.

Table 3: Initial proportions: 50%/50%

Grid-size	Steps until equilibrium	% of dominant opinion at equilibrium
10000	3261.55	49.9932
20000	2052.49	49.97995
30000	1969.21	49.97
40000	1821.5	49.98
50000	696.35	50.02
60000	1204.98	50.01
70000	25.21	49.99
80000	120.21	50.03
90000	36.23	49.99
100000	17.88	50.02

- The major differences could be seen for the cases, where the opinions did not possess the same proportions at the beginning and where the dominant opinion initially occupied over 55% of the grid. First of all, we were able to observe that, on average, a state of equilibrium was reached, while both opinions had close to the same proportions as at the start.

Table 4: Initial proportions: 80%/20%

Grid-size	Steps until equilibrium	% of dominant opinion at equilibrium
10000	11492.23	83.9185
20000	22209.97	83.82
30000	32498.69	83.76
40000	44956.04	83.81
50000	57451.95	83.81
60000	67824.28	83.79
70000	79319.14	83.78
80000	91189.22	83.82
90000	106506.7	83.81
100000	117183.9	83.84

Furthermore, the number of steps it took to reach an equilibrium as a function of the grid-size grew linearly in general, meaning it could be approximated optimally by a linear function.

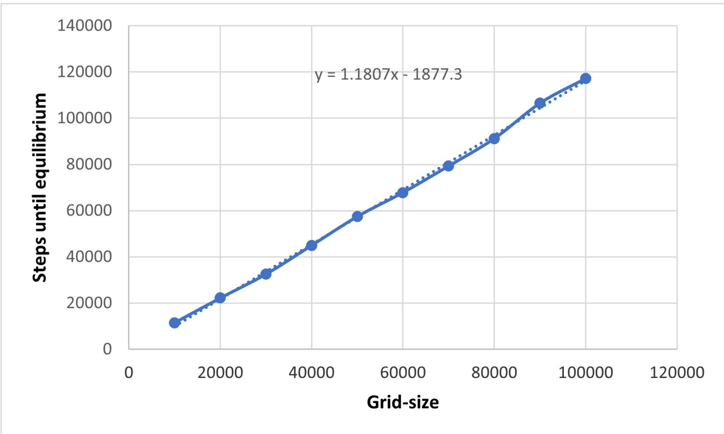
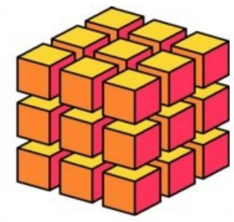


Figure 14: Initial proportions: 80%/10%

6.2 3 dimensions

Next, let us consider a 3-dimensional grid, which can also be seen as a basic cube, filled with smaller cubes.



Mathematically, this means that each single voter simply has an additional third coordinate. As the number of voters in our grid is no longer the side-length of the grid squared, but cubed, we were forced to work with grids of smaller side-length. Nonetheless, the actual number of cubes in our grid roughly stays the same as for the number of squares in the 2D cases.

Overall, we worked with grids of different initial set-ups of 2 opinions and of sizes ranging from $50 \times 50 \times 50$ to $90 \times 90 \times 90$, while using a consensus threshold of 1% and repeating each single case 50 times. As a result, we found out the following properties:

- By looking at the gathered data and graphs, we can conclude that for the cases where the opinions are initially equally proportioned, as well as for those, where they are not, we get almost identical results as for the 2D model in general.

Table 5: Initial proportions: 75%/25%

2D grid-size	2D % at equilibrium	3D grid-size	3D % at equilibrium
250	99	50	100
300	99	60	100
350	99	70	100
400	99	80	100
450	99	90	100
500	99	100	99.9

  Here we can see that the average percentages at the equilibrium of the dominant opinion for the 3D voter model are close to those of the 2D model.

- The only main difference one can observe, is that generally, it takes the 3D voter model a larger amount of steps to reach a state of equilibrium, if we compare the 2D and 3D models using grids with a similar number of voters. We can only assume that this is because the 3D grid has more edges (12) and faces (6), which essentially means that a larger proportion of voters, which lie on these edges and faces, have a weakened probability of finding a neighbour of different opinion to influence, as there are fewer neighbours to choose from. Let us show with an example that the proportion of these deprived voters is in fact larger than the one for the 2D grid: Consider a 2D grid of size 300×300 and a 3D grid of size $45 \times 45 \times 45$. We can see that the total number of voters in both grids are close to each other ($300^2 \approx 45^3$). Hence, we can calculate the proportion of squares lying on the respective boundaries (edges for the 2D model, edges & faces for the 3D model) in both grids:

  2D: $\frac{\text{\# of voters lying on edges}}{\text{total \# of voters}} = \frac{300^2 - 298^2}{300^2} \approx 0.0133 \approx 1.33\%$
  3D: $\frac{\text{\# of voters lying on edges \& faces}}{\text{total \# of voters}} = \frac{45^3 - 43^3}{45^3} \approx 0.1275 \approx 12.75\%$

- Finally, by curve fitting the plotted results, we were able to conclude that a quadratic function no longer approximates the average number of steps to reach a consensus as a function of the grid-size precisely enough, as the average relative error is above 3%. However, cubic functions and polynomials of higher degree approximate this function with an average relative error of less than 1%.

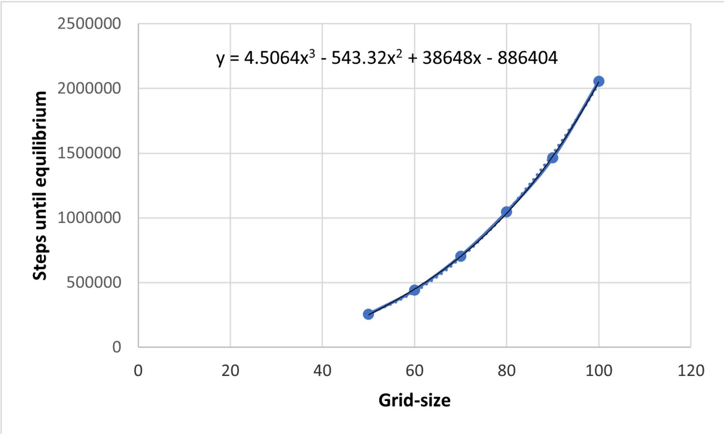


Figure 15: Initial proportions: 55%/45%

6.3 4 dimensions

Finally, we can consider a 4-dimensional grid, which one can also imagine to be a tesseract, filled with smaller tesseracts. Adding a forth coordinate to each voter, we worked with grids of different initial set-ups of 2 opinions and of sizes ranging from $5 \times 5 \times 5 \times 5$ to $25 \times 25 \times 25 \times 25$, while using a consensus threshold of 1% and repeating each single case 50 times. Let us enumerate our assessments:

- Comparing this case to the 2D and 3D case, we hardly saw any differences. On average, the general evolution of the 4D voter model was almost identical to the other cases, with the average proportions of the dominant opinion at the consensus being similar again.

Table 6: Initial proportions: 75%/25%

2D grid-size	2D % at equilibrium	4D grid-size	4D % at equilibrium
300	99	5	99.68
350	99	10	100
400	99	15	99.03407
450	99	20	100
500	99	25	99.00442

- The only dissimilarity came, yet again, from the larger number of steps to reach a state of consensus. Here we can make a similar assumption as for the 3D voter model, specifically that this average larger number of steps is related to the increased number of edges, faces and cubic cells of our 4D grid, which has the exact same consequence as mentioned in the 3D case.

Hence, we can see that as the dimensions increase, so to do the amount of "boundaries" of the grids, which have an immediate knock-on effect, when it comes to the number of steps it takes to achieve a state of equilibrium.

- Last but not least, by curve fitting the plotted data, we were able to observe that a cubic function performs poorly, when it comes to approximating the average number of steps to reach an equilibrium as a function of the grid-size. However, a quartic function (forth degree polynomial) and polynomials of higher degree perform ideally, having an average relative error below 1%.

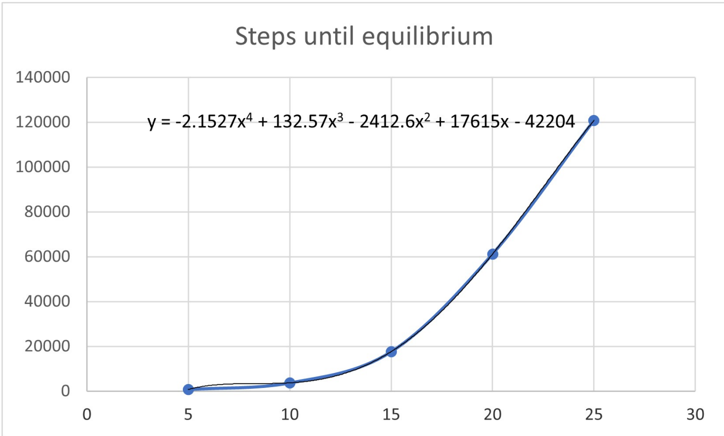


Figure 16: Initial proportions: 80%/20%

6.4 Conjectures

Following the detailed analysis of our observations, we would like to make some conjectures, concerning the variation of the grid's dimension:

- In the case, where our grid is 1-dimensional, the initial and final proportions of the two opinions commonly lie close to each other.
- If the grid's dimension is greater than 1, the variation of its dimension generally does not change the overall evolution of the voter model, all behaving in a similar way to the basic 2D case.
- The one main difference is that, on average, the number of steps it takes to reach a state of consensus becomes larger as the grid grows in dimensions. This is presumably due to the fact, that the number of edges, faces and further hindering obstacles increases with the rising dimensions, which deprives a growingly larger proportion of voters lying on these boundaries.
- If the dimension of the grid is equal to n ($n \in \mathbb{N}$), then in general, after curve fitting the plotted data, we can see that the average number of steps it takes to reach a state of equilibrium as a function of the grid-size can be optimally approximated by a polynomial function of degree n or higher.

7 Clustering analysis

As already briefly mentioned in the basic case, for every single scenario we simulated visually, no matter the initial set-up of the grid, we were always able to observe the formation of so-called clusters. To show that this process, known as "clustering", is a common behaviour of the voter model, we decided to further analyse the 2D model, to understand as much as possible about these clusters.

7.1 First observations of formations of clusters

In order to gather some first observations concerning the formation of clusters, we decided to keep track of the number and sizes of these groups for all types of single cases of the 2-opinions 2D voter model, with the help of a program that simulated each individual scenario analytically.

After having run through each individual case several times, with the initial proportions of the two opinions ranging from 50%/50% to 95%/5% and with grid-sizes ranging from 50×50 to 85×85, we were able to observe some frequent patterns.

Generally, we found that at the very beginning of each simulation, there tended to be a few bigger clusters and many smaller clusters or single squares. The sizes of these, as well as their frequency largely depended on the grid-size and initial proportions of the opinions. However, apart from the sizes and the frequency of these clusters, on average, their overall behaviour was quite similar over time. The majority of the smaller groups tended to attach themselves to the larger ones of same opinion or were influenced by larger clusters of different opinion, allowing these to grow significantly in size.

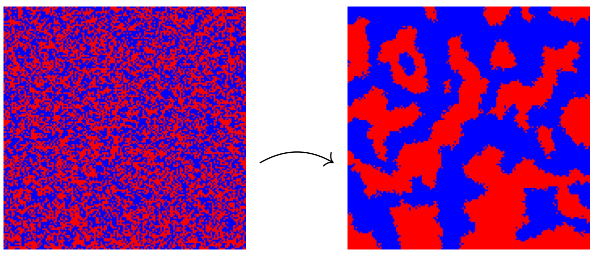


Figure 17: Grid-size: 150×150 / Initial proportions: 50% red & 50% blue

Here we can see that after a certain amount of steps, our grid began to almost look like a map of 2 colours. Meanwhile, the existence of smaller or even single isolated clusters became less common over time, a topic which we will further analyse in our next paragraph.

Let us now take a closer look at a few examples of the evolution of clusters for some specific cases for the grid-size 75×75:

Table 7: Example of clusters at the start for initial proportions 50%/50%, 70%/30% & 90%/10%

Cluster size	Frequency	Cluster size	Frequency	Cluster size	Frequency
1247	1	1797	1	2247	1
1218	1	35	1	6	2
6	1	34	1	5	1
4	1	27	1	4	6
3	1	25	1	3	9
3	2	22	2	2	30
2	2	19	2	1	125
2	2	18	1		
1	6	17	1		
1	2	13	4		
		12	2		
		11	1		
		10	5		
		9	3		
		8	3		
		7	5		
		6	9		
		5	6		
		4	4		
		3	16		
		2	21		
		1	52		

At first it might be surprising that for almost equally proportioned cases, the grid is initially filled with a couple of clusters of such a large size. For example, if we look back at Figure 17, overall, we seem to see very few clusters in the grid at the start, while almost all the squares seem to be fully isolated, having nearly no neighbours of the same opinion. However, it is important to keep our definition of a "neighbour" in mind, as we stated earlier that two squares are also considered as neighbours if they share a corner.

Figure 18: All squares surrounding the centred square are its neighbours

Table 8: Example of clusters after 5˙625 steps for initial proportions 50%/50%, 70%/30% & 90%/10%

Cluster size	Frequency	Cluster size	Frequency	Cluster size	Frequency
2895	1	5223	1	5621	1
2312	1	48	1	2	1
190	1	26	2	1	2
88	1	22	1		
53	1	20	1		
11	1	19	1		
8	1	14	2		
7	1	12	1		
6	1	10	1		
5	2	9	2		
3	2	8	2		
2	4	7	4		
2	3	6	1		
1	18	5	6		
1	7	4	5		
		3	6		
		2	10		
		1	35		

(Note: The reason why we work with smaller grid-sizes here, is because the code we used for our program counted the sizes of the clusters in a recursive way. Therefore, we were forced to work with smaller grid-sizes, to avoid a "`RecursionError`".)

7.2 Average cluster occurrence frequency

After having collected some first assessments based on multiple single cases, we wanted to have a better understanding of the average frequency with which a cluster of size n would appear after k steps ($n, k \in \mathbb{N}$) in general.

Therefore, we decided to simulate each single case of the 2-opinions 2D voter model 50 times, counting the number of clusters of different sizes after regular step intervals and then working with the average results. Here our initial proportions ranged from 50%/50% to 95%/5% and our grid-sizes from 50×50 to 85×85 once more.

Comparing the obtained outcome with our previous assessments, we can see that they overlap. First of all, we were able to observe that, on average, at the beginning of our simulations, there is a large number of smaller clusters and only very few bigger clusters. In most cases, if we plot the average occurrence frequency of a cluster with regard to its size after 0 steps, we get a function whose graph resembles the graph of the inverse exponential function.

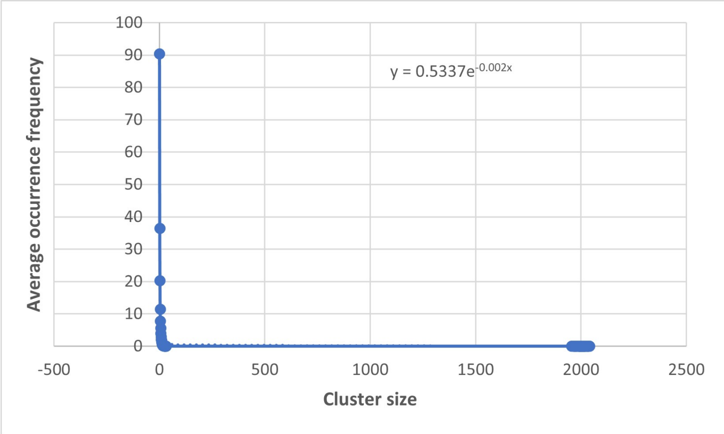


Figure 19: Example of clusters at the start for grid-size 50×50 & initial proportions 80%/20%

As our voter model continued to evolve, we were able to observe that, on average, the clusters of smaller size became less and less frequent, while larger and middle-sized clusters tended to appear more often. In terms of the graph of our data, we realised that this curve started to flatten down over time in general. After curve fitting the plotted data, we still found the inverse exponential function to represent the optimal approximation for our data, but this time with a slightly bigger negative exponent of the exponential.

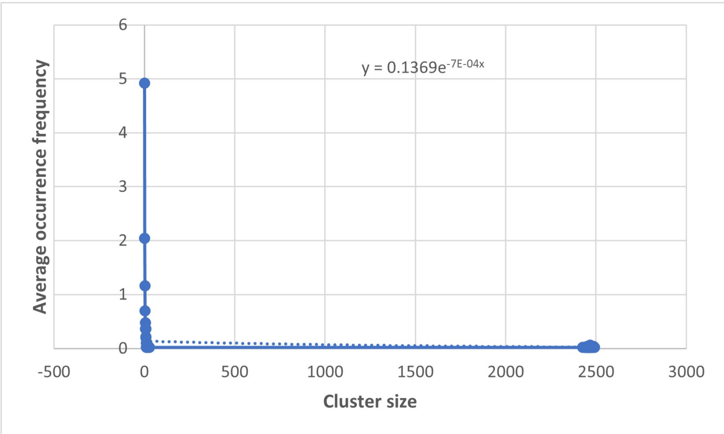


Figure 20: Example of clusters after 2'500 steps for grid-size 50×50 & initial proportions 80%/20%

Once the voter model for our different scenarios approached a state of equilibrium, we found that the clusters of larger size were as common, if not even more common, than those of smaller size. This is the exact clustering phenomenon we already discussed before and which we have now been able to support even further.

In terms of the graphical representation of our data, as we approach a state of consensus, we can see that the function we used beforehand is no longer relevant. By reflecting our previous graph with regard to the y-axis, we get a more appropriate graph to approximate our plotted data, namely the standard exponential function.

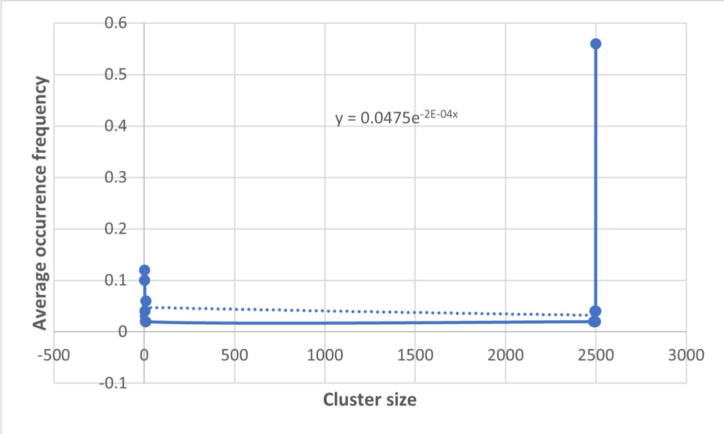
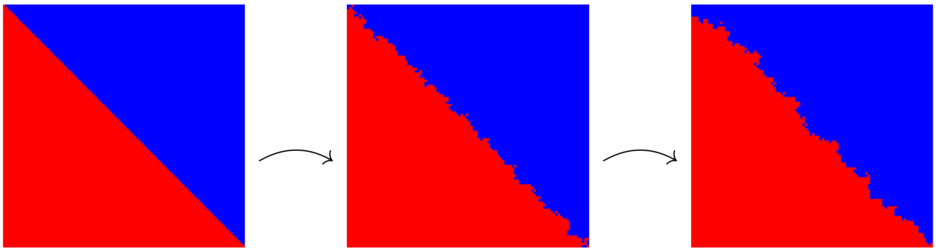


Figure 21: Example of clusters after 7'500 steps for grid-size 50×50 & initial proportions 80%/20%

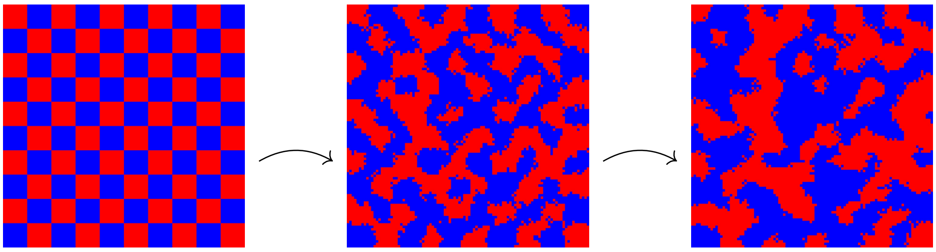
7.3 Special cluster scenarios

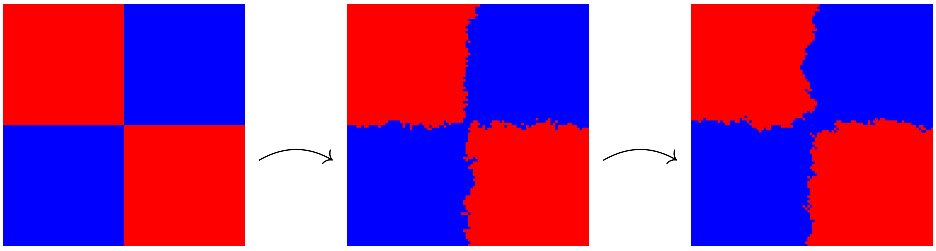
Finally, we were interested in, if instead of filling the grid randomly, we would predetermine the initial configuration of the clusters of different sizes and shapes, and then see how the voter model would develop over time. At first one might think that this would have little to no impact on our results, however, we did come up with some interesting observations.
Using the basic 2-opinions 2D voter model with sizes ranging from 100×100 to 500×500, we simulated the following cases visually:

- First of all, we started off by considering the case, where all the squares in the left half of our grid have one colour (in our case red) and all the squares in the other half have another colour (in our case blue). Similarly, we looked at the case, for which a diagonal splits our grid into two halves of different colours (red & blue). For both cases, we came up with similar results, namely that the proportions of the two colours hardly vary at all, while the border line is pretty much preserved throughout the entire simulation. One could almost compare it to a battlefield, where two equally matched armies are fighting each other on the front line, with neither one of them gaining an advantage over the other.
  Here are some examples for both cases (grid-size: 100×100), of how our grid evolved over time:

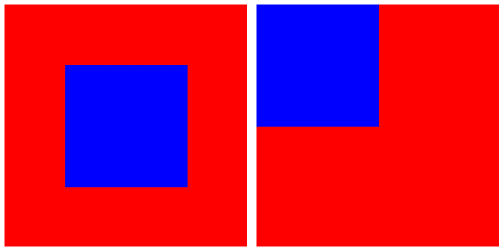


- Next, we considered the case, where our grid is filled with square clusters of sizes 10×10, as well as 50×50. In this sense, our grid looks very similar to a chessboard, only with squares of the colours red and blue. As the simulations evolved, we noticed that, although the squares lost their shape and transformed into randomly configured clusters, their cores tended to survive for a certain number of steps. Furthermore, by looking at the analytical report of these simulations, we could also see that the initial proportions, which are 50%/50%, were preserved most of the time.
  Here are some examples for both cases (grid-size: 100×100), of how our grid evolved over time:

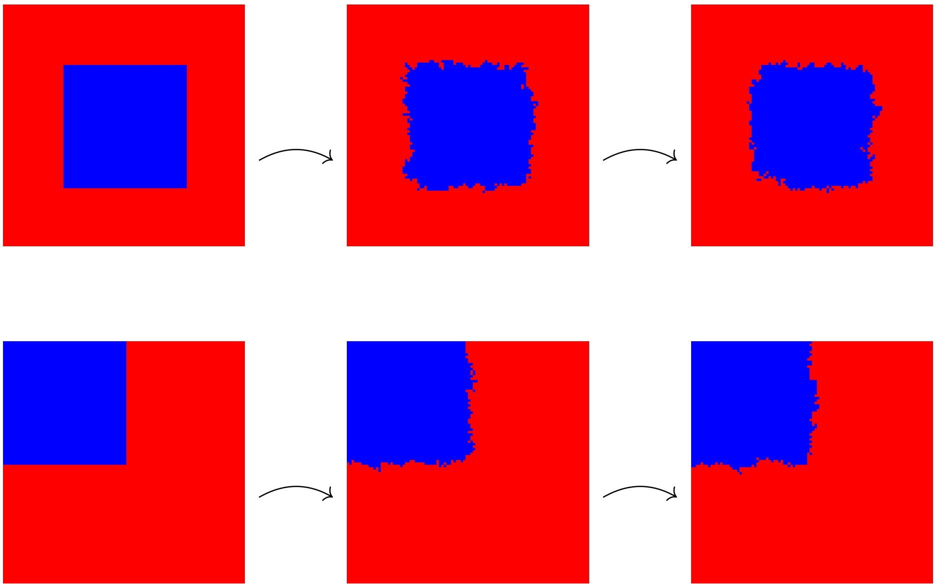




- Last but not least, we wanted to test a couple of disproportional cases, meaning cases where one colour is clearly more dominant. Therefore, we chose the following two cases, whose initial set-ups look like this:



In both cases, we saw that, although the weaker colour had an obvious proportional disadvantage, it was able to maintain its initial shape for a long period of time. By looking at the analytical report, we can see that the proportion of the weaker colour only shrunk minimally over time, which means that the dominant colour struggled to break down the border line rapidly.
Here are some examples for both cases (grid-size: 100×100), of how our grid evolved over time:



7.4 Conjectures

Following the observations explained in the previous paragraphs, we came up with our very own conjectures, concerning the clustering process of the 2D voter model:

- If we consider an initial random set-up of our grid, then after a certain amount of time, clusters usually will begin to form. In addition, smaller clusters tend to attach to larger clusters of both opinions.
- Over time, the occurrence frequency of smaller clusters decreases in general, while larger clusters become more frequent. This is simply known as the "clustering" phenomenon.
- At the beginning of a randomly configurated grid, the average occurrence frequency of different clusters as a function of their size generally behaves like the inverse exponential function, whereas as the grid approaches a state of equilibrium, it tends to behave like the standard exponential function.

- Once a cluster with a solid core has formed, it generally takes a large number of steps to break down the outer layer of the cluster and dissolve it. The larger the cluster is, the harder this becomes. Moreover, larger clusters are favoured to growth, as opposed to smaller clusters.

8 Influencers in social networks

After having experimented and analysed multiple variations of the basic voter model, we wanted to study one final type of voter model, slightly different from all the others seen so far. Let us change the dynamics of our model a bit.

8.1 Basic set-up

Consider a model, where all the voters share the same opinion at the beginning. One could also simply see this as a "blank" grid. Next, we are allowed to choose k voters ($k \in \mathbb{N}_{\geq 2}$), who all have their own unique opinion and are positioned in the grid in a predetermined way. These voters can be referred to as "influencers", as initially they are the only voters able to influence the other neutral voters. Afterwards, we apply the same update rule as for the basic model, however with one slight difference. Instead of choosing any voter at random, we need to choose a non-neutral voter, meaning someone who shares the same opinion as an initial influencer, based on a given probability. This voter can then influence one of his neighbours of different opinion (neutral or non-neutral). If they do not have a neighbour to influence, then the last influenced voter of the same opinion, who has a neighbour they can influence, is selected to influence one of their neighbours. Finally, we can apply this rule until there are no neutral voters left. Therefore, we can also give a new definition of what an "equilibrium" means.

Definition 8.1.1 (Equilibrium/consensus)**.** An equilibrium/consensus has occurred, if the grid no longer has any neutral voters.

A natural question we asked ourselves was the following: is there a certain position of the grid, which favours the dominance of an influencer?

Interestingly, we found out that the initial placement of the different influencers is important, however, it is not the only ruling factor. The final outcome also heavily depends on the selection process, which determines how the next voter, who can influence one of his neighbours, is chosen.

8.2 Different selection processes

Although there infinitely many ways to select the next voter of our grid, we have decided to concentrate on two specific manners.

On the one hand, we can choose the next voter corresponding to the proportions of the different opinions in the grid. By this we mean that the relative probability of a voter being selected is linked to the proportion of the grid that is filled by the opinion of that same voter. Let us explain with a simple example:
If we consider a grid with two influencers at the beginning, then at the very first step, each influencer has a 50% chance of being selected. Then after a certain amount of

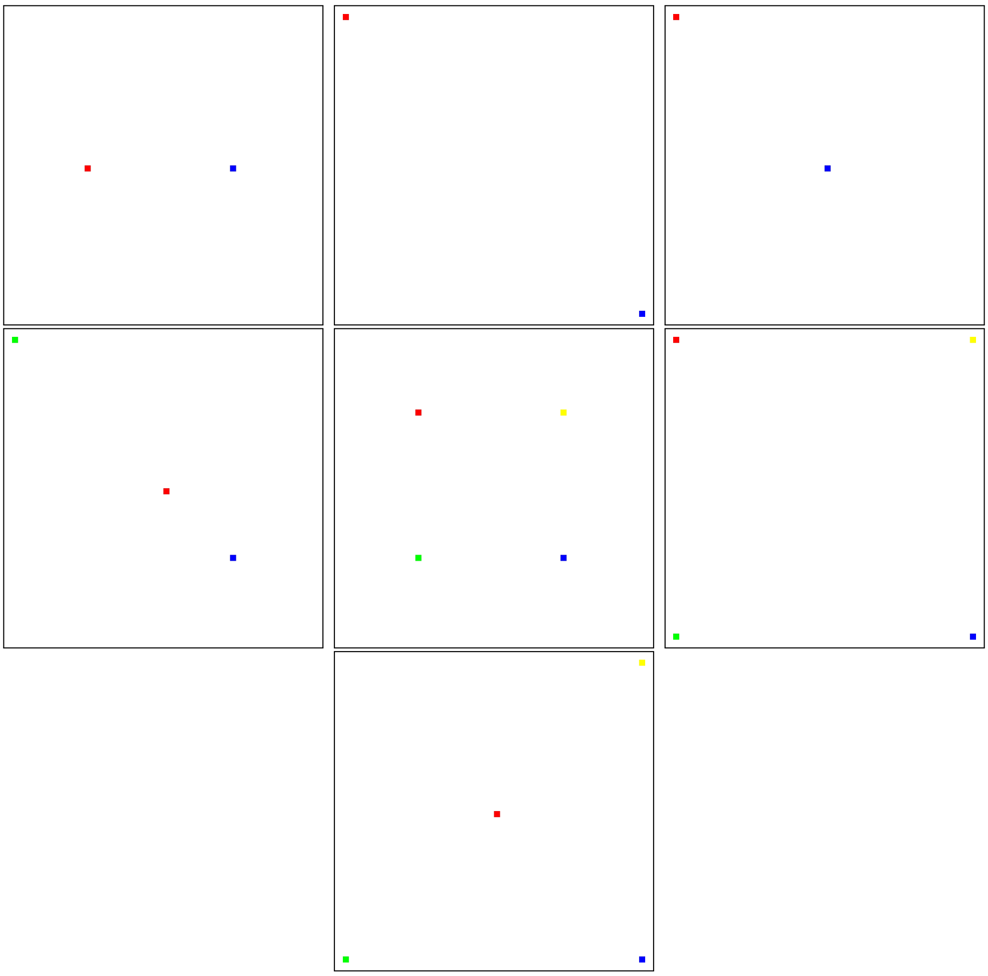
steps, if for example one opinion has 25 voters and the other opinion has 75 voters, the first opinion is chosen with a probability of 25% and the second with a probability of 75%.

On the other hand, we can select the next voter corresponding to the number of neutral neighbours. In that sense, the relative probability of a voter being selected is based on the total amount of neutral neighbours the voters of that opinion have. Let us consider another simple example:
Consider a grid with two influencers at the start. If after a certain number of steps, the cluster of one opinion has 25 adjacent neutral neighbours and the other one has 75 adjacent neutral neighbours, the first opinion is chosen with a probability of 25% and the second with a probability of 75%.

In the next paragraphs, the first selection process will be referred to as "square area", whereas the second one will be referred to as "blank neighbours".

8.3 Simulations for "square area"

Let us start off by analysing the "square area" case. For this case, we used both visual simulations, to gather some first observations, as well as analytical simulations, to help make conjectures. In total, we chose 7 different initial set-ups, while mainly using grid-sizes ranging from 10×10 to 50×50 and simulating each case 100 times. Let us take a closer look at these 7 cases (for grid-size: 50×50):



As one can see, we selected both symmetric and anti-symmetric cases, while also varying the overall number of initial influencers. The names we attributed to these cases, in the same order as above, are: 2 centres, 2 corners, 2 anti-symmetric, 3 anti-symmetric, 4 centres, 4 corners & 4 anti-symmetric. We came up with the following results:

- First of all, we noticed that in general the number of steps it took, until all the neutral voters had disappeared, grew as the number of initial influencers grew. This can be somewhat surprising, as one might think that, because initially there are already fewer neutral voters in the grid, it should take even fewer steps to reach an equilibrium. We also saw that on average the amount of steps it took to reach a state of consensus seemed not to depend on the initial placement of the influencers.

Table 9: Steps until equilibrium for each case

Grid-size	2 centres	2 corners	2 anti-symmetric	3 anti-symmetric	4 centres	4 corners	4 anti-symmetric
10	136.38	137.22	136.22	152.75	198.21	214.86	197.21
20	527.34	550.15	519.43	697.85	865.57	860.76	831.97
30	1171.81	1228.18	1213.31	1586.61	1733.81	1893.83	1834.57
40	2108.11	2151.33	2223.34	2863.57	3237.05	3556.79	3526.23
50	3293.49	3520.91	3306.8	4254.99	4858.17	5691.32	5449.09

- Furthermore, by looking at the graphs again, we could see that the number of steps it took to reach a consensus as a function of the grid-size can be approximated precisely by a quadratic function.

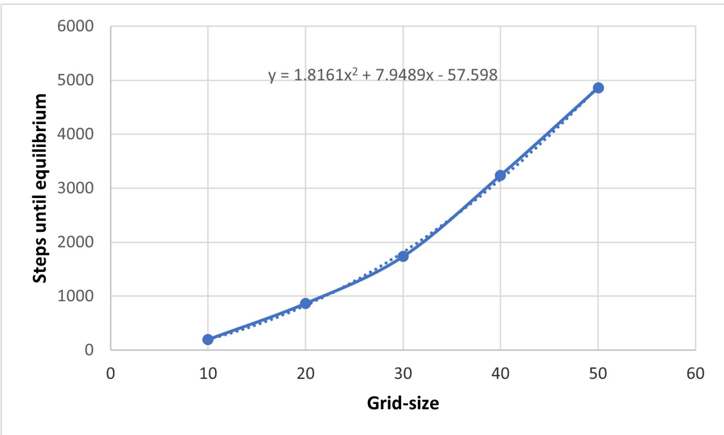


Figure 22: Case: 4 centres

Hence, we can claim that this variant of the voter model behaves similarly to the basic 2D model in that way.

- Finally, we were able to observe that in all 7 cases, there was no repetitive pattern, in terms of which placement of the influencers gave the best chances of becoming the dominant opinion of the grid in the end. We can see by the following charts, that the number of wins of each influencer in the 7 cases was quite random and unpredictable, with an almost even distribution of wins among the influencers.

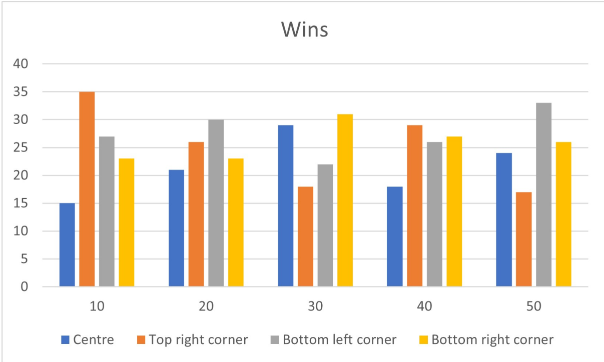


Figure 23: Case: 4 anti-symmetric

This leads to the conclusion that the "square area" selection process does not favour one position of the grid over another, which is quite logical because at the start every single influencer has exactly the same probability of being selected first and as a result gaining momentum.

8.4 Simulations for "blank neighbours"

Next, let us move on to the "blank neighbours" case. Here, we simulated, both visually and analytically, the exact same seven cases as already mentioned above. We ran through these symmetric and anti-symmetric cases, while using grid-sizes ranging from 10×10 to 50×50 and simulating each case 100 times. Let us enumerate our assessments:

- Similarly, to the other case, the number of steps it took to reach an equilibrium generally increased, as the amount of initial influencers increased. However, this number tended not to depend on the placement of these influencers at the beginning.

Table 10: Steps until equilibrium for each case

Grid-size	2 centres	2 corners	2 anti-symmetric	3 anti-symmetric	4 centres	4 corners	4 anti-symmetric
10	188.36	182.41	163.99	254.76	289.25	280.84	280.81
20	868.72	853.37	622.9	1132.04	1450.15	1448.83	1409.46
30	2070.92	1912.72	1469.19	2747.72	3531.32	3460.55	3041.42
40	3585.41	3162	2857.21	5214.03	6496.09	6484.19	5428.58
50	5537.09	4934.23	3858.45	7253.97	10394.96	9057.69	9394.44

- Moreover, we were able to see in these graphs, that likewise to the other selection process, the number of steps it took to reach a consensus as a function of the grid-size can be approximated precisely by a quadratic function.

- The main difference between this selection process can be found, when looking at which influencers had the most success. On the one hand, we observed that for most of the symmetric set-ups, the number of wins was pretty evenly distributed among all the influencers.

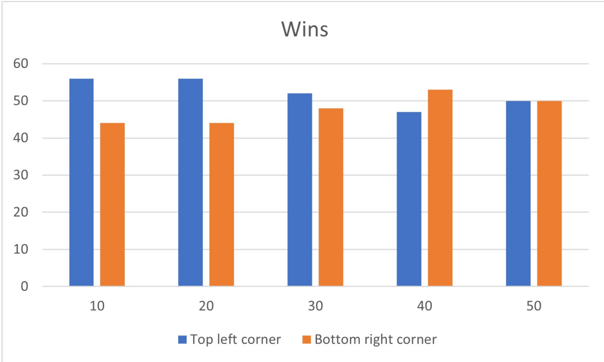


Figure 24: Case: 2 corners

On the other hand, the anti-symmetric cases behaved fundamentally differently.

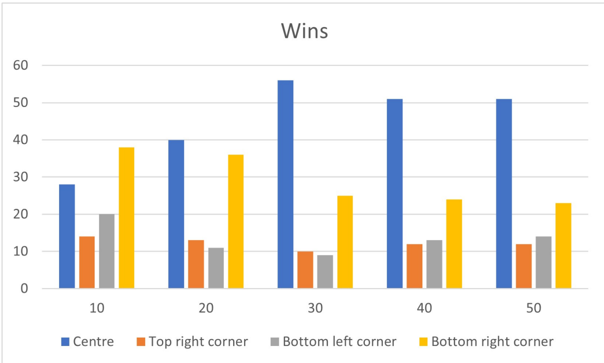


Figure 25: Case: 4 anti-symmetric

By looking at these charts, we can see that for smaller grid-sizes, the win distribution between the influencers still seemed to be quite random. However, as the grid-sizes grew, we noticed a clear pattern evolving, where the anti-symmetric influencers, placed closer to the centre of the grid, tended to claim more wins on average. Another interesting property we found out, was that the closer an influencer was initially placed to the centre, the more likely they were to dominate the grid. This is exactly the situation for the case ”3 anti-symmetric”.

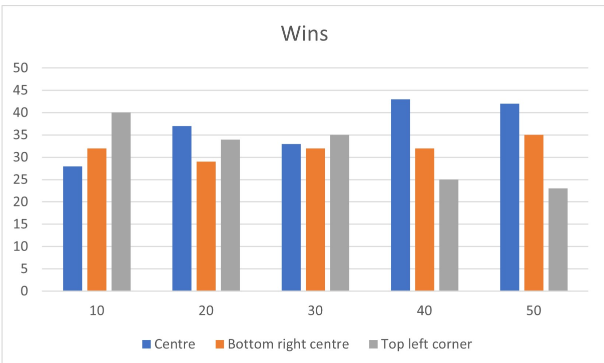


Figure 26: Case: 3 anti-symmetric

This can be explained by the fact that the influencers situated closer to the centre are also automatically further away from the edges. In this sense, the edges can be seen as obstacles, which slow down the growth of an influencer. Due to the lower amount of blank neighbours, the voters lying on the edges are not as efficient when it comes to contributing to the overall count of adjacent neutral neighbours. Therefore, the probability of eventually winning rises, as an influencer is positioned more central.

8.5 Conjectures

Following the detailed observations explained above, we can now state some final conjectures:

- For both selection processes, the number of steps it takes to reach a state of equilibrium on average, grows with the number of initial influencers placed in the grid, however, does generally not depend on how these influencers are placed.
- For both selection processes, a quadratic function can approximate the number of steps it takes to reach a state of consensus as a function of the grid-size precisely.
- For the selection process "square area", the initial positioning of the influencers usually does not have a direct impact on their success, with the probability of each influencer winning being almost equal.
- For the selection process "blank neighbours", the initial placement of the influencers tends to play an important role, when it comes to their win frequency. Here in general, the closer an influencer is to the centre of the grid, the likelier they are to win.

9 Mathematical background

One of the first known introductions of the basic voter model was made by Richard A. Holley and Thomas M. Liggett in 1975, in their research paper "*Ergodic Theorems for Weakly Interacting Infinite Systems and the Voter Model*"[4]. Over the years, it has remained a central research topic in the mathematical theory of probability, leading to many complex variations of the basic voter model.

Let us look at two particular types of voter models, as well as a few interesting theorems related to them:

Consider a continuous voter model in a state space $\boldsymbol{S} = \{0,1\}^{\mathbb{Z}^d}$ with a transition rate function $c(x,\eta)$, where $\mathbb{Z}^d$ is a d-dimensional grid and where $\eta \in \boldsymbol{S}$ is called a configuration. Here $\eta(x)$ stands for the value (e.g, opinion) of a site $x \in \mathbb{Z}^d$ and changes from $\eta(x)$ to $1-\eta(x)$ at a predetermined rate $c(x,\eta)$.
Then an example of the so-called *linear voter model* can be defined, by using the following conversion rate:

$$c(x,\eta) = \sum_{\|y-x\|\leq N} 1_{\{\eta(y)\neq\eta(x)\}}$$

This voter model is referred to as linear because the transition rate at x is a linear function of the number of neighbours y, such that $\eta(y) \neq \eta(x)$. (Note: Here, $\|\cdot\|$ is an arbitrary norm and $N \in \mathbb{N}$.)[6]

Using the same notations as above, we can also give an example of the so-called *threshold voter model*, a non-linear model which has gained a lot of attention in recent years:

$$c(x,\eta) = \begin{cases} 1 & \text{if there is a } y \text{ with } \|x-y\| \leq N \text{ and } \eta(x) \neq \eta(y) \\ 0 & \text{otherwise} \end{cases}$$

In 1985, P. Clifford and A. Sudbury were able to prove one of the most basic theorems about the linear voter model, namely that it becomes unanimous if $d \leq 2$ and coexists if $d \geq 3$ [1]. In other terms, this means that in a voter model of grid-dimension d, multiple values (e.g., opinions) can only survive simultaneously (coexist) over time, if $d \geq 3$. So, if $d = 1$ or $d = 2$, eventually there will only be one single value remaining.
At first, this might seem to be a contradiction with the conjectures we made before, concerning the basic case with initial proportions of 50%/50%. Here, we found out that, on average, a sate of equilibrium would occur, while both opinions retained the same percentage of the grid as at the start, meaning that both opinions did coexist. However, in 1998, D.J. Watts and S. Strogatz were able to prove, that if the voter model is run on a "small-world network", it can temporarily be trapped in a "metastable" state, where different values/opinions can coexist [9]. Nevertheless, they showed that eventually the system would escape this metastable state and allow one value to occupy the entire grid unanimously. Here, the grids we used for our voter model simulations can be regarded as small-world networks, due to the modest grid-sizes we worked with.

On the other hand, in 1991, J.T. Cox and R. Durrett were able to show, that, regarding the threshold voter model, it can coexist in one dimension if $N \geq 4$, in two dimensions if $N \geq 2$ (when $\|\cdot\|$ is the l_∞ norm) or $N \geq 3$ (when $\|\cdot\|$ is the l_1 norm) and in three or more dimensions if $N \geq 1$ [2]. Following this breakthrough, T.M. Liggett was then able to give a computer assisted proof, showing that the threshold voter model coexists in all cases, except for when $N = d = 1$ [6]. These threshold voter models have become an especially popular research topic, as mathematicians have been intrigued to find or modify voter models, such that a state can be reached, where more than one opinion is able to survive long-term.

(Note: As the proofs of these theorems are fairly complicated and require a certain level of mathematical knowledge, we decided to leave them out of our paper. However, they can be found in full detail in the specific papers, listed in the references.)

10 Conclusion

All in all, we can conclude that the simplest versions of the basic voter model lead to astounding results. We found out that the evolution of our model depends on many factors, such as the grid-size, the grid-dimension, the number of different opinions and their initial proportions, as well as the starting configurations of the voters/influencers. Furthermore, we were able to observe repetitive patterns in the behaviour of the voter model in many cases, like the formation of clusters after a certain amount of steps.

The main goal of the course "Experimental mathematics" is to allow young curious mathematicians like us, to gather their first proper experiences of organising and conducting a research paper. Thanks to the creative and imaginative freedom given to the students, we are able to explore unknown topics of mathematics, both with assistance and independently, which makes the end product even more rewarding.

Finally, we would like to encourage anybody, who has some mathematical and programming knowledge, to experiment with these types of voter models themselves and to come up with their very own conjectures.

11 Appendix

The following two Python codes represent the programs used to simulate the visual and analytical versions of the basic voter model. The majority of the other codes that we used were based on these two codes and were obtained by simply changing minor details.

```
import pygame
from pygame.locals import *
from random import choice, choices
from xlwt import Workbook


class Square:
    def __init__(self, x, y, size, opinion):
        self.x = x
        self.y = y
        self.size = size
        self.opinion = opinion

    def change_opinion(self, new_opinion):
        self.opinion = new_opinion

    def draw_square(self, screen):
        pygame.draw.rect(screen, Color(self.opinion), (self.size * self.x, self.size * self.y, self.size, self.size))

    def different_neighbours(self, squares, size):
        outcome = []
        index = size * self.x + self.y
        for a in range(-1, 2):
            for b in range(-1, 2):
                if 0 <= index + a * size + b <= size ** 2 - 1:
                    if self.y == 0 and b == -1 or self.y == size - 1 and b == 1:
                        pass
                    elif squares[index + a * size + b].opinion != self.opinion:
                        outcome.append(squares[index + a * size + b])
        return outcome


class Grid:
    def __init__(self, size, manipulation):
        self.size = size
        self.manipulation = manipulation
        self.squares = []
        self.info = {'red': 0, 'blue': 0}
        for x in range(0, self.size):
            for y in range(0, self.size):
                c = choices(['red', 'blue'], [self.manipulation, 100 - self.manipulation])[0]
                size = 800 // self.size if self.size < 300 else 2
                self.add_square(Square(x, y, size, c))
                self.info[c] += 1

    def add_square(self, square):
        self.squares.append(square)

```

```
    def draw_grid(self, screen):
        screen.fill(Color("white"))
        for s in self.squares:
            s.draw_square(screen)

    def execute(self):
        chosen_one = choice(self.squares)
        neighbours = chosen_one.different_neighbours(self.squares, self.size)
        if neighbours:
            n = choice(neighbours)
            self.info[chosen_one.opinion] += 1
            self.info[n.opinion] -= 1
            n.change_opinion(chosen_one.opinion)

    def unanimous(self):
        return len(self.squares) in self.info.values()


def auto_pygame(grid_size, manipulation, repetitions, step, FPS):
    grid = Grid(grid_size, manipulation)

    clock = pygame.time.Clock()

    pygame.init()
    size = (800 // grid_size) * grid_size if grid_size < 300 else 2 * grid_size
    screen = pygame.display.set_mode((size, size))
    pygame.display.set_caption("Voter models")
    screen.fill(Color("white"))

    grid.draw_grid(screen)
    pygame.display.update()

    rect = pygame.Rect(0, 0, size, size)

    counter = 0
    constant = 100 / (grid_size ** 2)
    red, blue, steps = [grid.info['red'] * constant], [grid.info['blue'] * constant], [0]
    done = False

    while not done:
        for event in pygame.event.get():
            if event.type == QUIT:
                done = True
        grid.execute()
        counter += 1
        if counter % step == 0 or counter > repetitions or grid.unanimous():
            grid.draw_grid(screen)
            pygame.display.update()
            clock.tick(FPS)
            red.append(grid.info['red'] * constant)
            blue.append(grid.info['blue'] * constant)
            steps.append(counter)
            print('', counter, sep='\n')
            sub = screen.subsurface(rect)
```

```
            screenshot = pygame.Surface((size, size))
            screenshot.blit(sub, (0, 0))
            pygame.image.save(screenshot,
                              f'({grid_size}x{grid_size},{manipulation}%-{100 - manipulation}%){counter // step}.jpg')
            if counter > repetitions or grid.unanimous():
                done = True

    pygame.quit()

    wb = Workbook()
    sheet1 = wb.add_sheet('Sheet 1')
    sheet1.write(0, 0, 'Steps')
    sheet1.write(0, 1, 'Red %')
    sheet1.write(0, 2, 'Blue %')

    for i in range(len(red)):
        sheet1.write(i + 1, 0, steps[i])
        sheet1.write(i + 1, 1, red[i])
        sheet1.write(i + 1, 2, blue[i])

    wb.save(f'Pygame stats for ({grid_size}x{grid_size},{manipulation}%-{100 - manipulation}%).xls')


for grid_size in range(100, 501, 50):
    for manipulation in range(50, 96, 5):
        auto_pygame(grid_size, manipulation, 10 * grid_size ** 2, grid_size ** 2 // 10, 100)
```

Listing 1: Code of visual program

```
from random import choice, choices
from xlwt import Workbook


class Square:
    def __init__(self, x, y, opinion):
        self.x = x
        self.y = y
        self.opinion = opinion

    def change_opinion(self, new_opinion):
        self.opinion = new_opinion

    def different_neighbours(self, squares, size):
        outcome = []
        index = size * self.x + self.y
        for a in range(-1, 2):
            for b in range(-1, 2):
                if 0 <= index + a * size + b <= size ** 2 - 1:
                    if self.y == 0 and b == -1 or self.y == size - 1 and b == 1:
                        pass
                    elif squares[index + a * size + b].opinion != self.opinion:
                        outcome.append(squares[index + a * size + b])
        return outcome

```

```

class Grid:
    def __init__(self, size, manipulation):
        self.size = size
        self.manipulation = manipulation
        self.squares = []
        self.info = {'good': 0, 'bad': 0}
        for x in range(0, self.size):
            for y in range(0, self.size):
                o = choices(['good', 'bad'], [self.manipulation, 100 - self.manipulation])[0]
                self.add_square(Square(x, y, o))
                self.info[o] += 1

    def add_square(self, square):
        self.squares.append(square)

    def execute(self):
        chosen_one = choice(self.squares)
        neighbours = chosen_one.different_neighbours(self.squares, self.size)
        if neighbours:
            n = choice(neighbours)
            self.info[chosen_one.opinion] += 1
            self.info[n.opinion] -= 1
            n.change_opinion(chosen_one.opinion)

    def unanimous(self):
        return len(self.squares) in self.info.values()


def auto_equilibrium(manipulation, repetitions, perc, start, end, step):
    gridsizes, values, percentages = [i for i in range(start, end + 1, step)], [], []

    for i in range(start, end + 1, step):
        s1, s2, constant = 0, 0, 100 / (i ** 2)
        for t in range(repetitions):
            grid = Grid(i, manipulation)
            counter, reset, p = 0, 0, manipulation
            while True:
                grid.execute()
                counter += 1
                reset += 1
                if abs(p - grid.info['good'] * constant) < perc:
                    if reset > i ** 2:
                        s1 += counter - reset
                        s2 += p
                        print(t)
                        break
                elif grid.unanimous():
                    s1 += counter
                    s2 += grid.info['good'] * constant
                    print(t)
                    break
                else:
                    p, reset = grid.info['good'] * constant, 0
        values.append(s1 / repetitions)
```

```
        percentages.append(s2 / repetitions)
        print('', i, '', sep='\n')

    print(gridsizes, values, percentages, '', sep='\n')

    wb = Workbook()
    sheet1 = wb.add_sheet('Sheet 1')
    sheet1.write(0, 0, 'Grid-size')
    sheet1.write(0, 1, 'Steps until equilibrium')
    sheet1.write(0, 2, 'Percentage at equilibrium')

    for i in range(len(gridsizes)):
        sheet1.write(i + 1, 0, gridsizes[i])
        sheet1.write(i + 1, 1, values[i])
        sheet1.write(i + 1, 2, percentages[i])

    wb.save(f'Equilibrium function for {manipulation}%-{100 - manipulation}%.xls')


for i in range(50, 96, 5):
    auto_equilibrium(i, 100, 0.5, 100, 501, 50)
```

Listing 2: Code of analytical program

As we are unable to showcase all of the data collected during our research period, we have decided to upload all of our results and Python codes to an online folder, which is linked down below:

https://drive.google.com/drive/folders/15q9PoEtrrSEKAUjt0f2bqN2XnJpvPX54?usp=sharing

References

[1] Clifford, P., Sudbury, A., 1985, *A Sample Path Proof of the Duality for Stochastically Monotone Markov Processes*, Institute of Mathematical Statistics, viewed 15 May 2021

[2] Cox, J.T., Durrett, R, 1991, *Nonlinear Voter Models*, Springer, viewed 15 May 2021

[3] Cox, J.T., Griffeath, D., 1986, *Diffusive Clustering in the Two Dimensional Voter Model*, Institute of Mathematical Statistics, viewed 15 May 2021

[4] Holley, R.A., Liggett, T.M., 1975, *Ergodic Theorems for Weakly Interacting Infinite Systems and the Voter Model*, Institute of Mathematical Statistics, viewed 15 May 2021

[5] Liggett, T.M., 1985, *Interacting Particle Systems*, Springer, viewed 15 May 2021

[6] Liggett, T.M., 1994, *Coexistence in Threshold Voter Models*, Institute of Mathematical Statistics, viewed 15 May 2021

[7] Redner, S., 2019, *Reality-inspired voter models: A mini-review*, Comptes rendus - Physique 20, viewed 13 May 2021

[8] Tang, T., Chorus, C.G., 2019, *Learning Opinions by Observing Actions: Simulation of Opinion Dynamics Using an Action-Opinion Inference Model*, Journal of Artificial Societies and Social Simulation, viewed 14 May 2021

[9] Watts, D.J, Strogatz, H., 1998, *Collective dynamics of 'small-world' networks*, Springer, viewed 15 May 2021

[10] EMBL-EBI, *Network analysis in biology*, viewed 4 May 2021,`https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/network-analysis-in-biology/`

[11] Wikipedia, 2021, *Network theory*, viewed 4 May 2021, `https://en.wikipedia.org/wiki/Network_theory#:~:text=Network%20theory%20is%20the%20study,have%20attributes%20(e.g.%20names)`

[12] Wikipedia, 2021, *Social network analysis*, viewed 4 May 2021, `https://en.wikipedia.org/wiki/Social_network_analysis#:~:text=Social%20network%20analysis%20(SNA)%20is,or%20interactions)%20that%20connect%20them`

[13] Wikipedia, 2021, *Voter model*, viewed 4 May 2021, `https://en.wikipedia.org/wiki/Voter_model#:~:text=A%20voter%20model%20is%20a,in%20the%20product%20topology%20on%20`