INDIAN STATISTICAL INSTITUTE, KOLKATA

A REPORT

ON

VOTER MODEL

STOCHASTIC PROCESS GROUP PROJECT

SUBMITTED BY

ANJALI (BS1706)
DIGANTA BHATTACHARYA(BS1718)

UNDER THE GUIDANCE OF

Prof. Soumendu Sundar Mukherjee

(Academic Year: 2020-2021)

DATE OF SUBMISSION: **February 26, 2022**

Contents

1 Introduction

1.1 Why was this introduced?

Researchers have been studying the question of opinions evolution in social networks since the fifties. We mainly focus on a basic question of social dynamics on consensus formation.
We have a population of interacting particles having their own individual opinions. They can change their opinions by interaction with their neighbours. The basic questions in these kind of models are: what will be the final state or opinion, and how is the transient to this state?
The interaction amongst the individuals of the population and the topology of the interactions (who interacts with whom) defines typically what the final state will be: either all agents having the same state (consensus) or coexistence of agents in the different states.
We basically want to study the spread of opinions in the population.

1.2 Who introduced Voter Model ?

Namely, we study the well-known voter model, in which each user holds one of two possible opinions and updates it randomly under the distribution of others beliefs.
Independently introduced by Clifford and Sudbury (1973)(where it was called the invasion process) and Holley and Liggett (1975) in the context of particles interaction, this model has since been used to describe in a simple and intuitive manner social dynamics where people are divided between two parties and form their opinion by observing that of others around them.

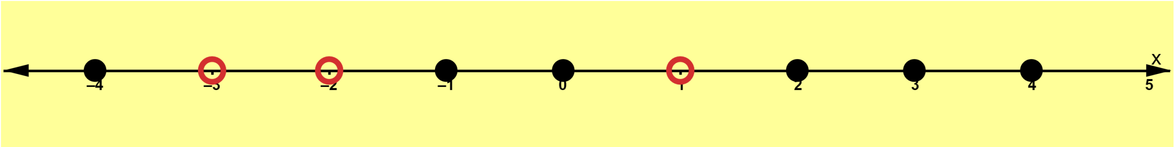
2 Clifford and Sudbury(1973)

2.1 Assumptions

- Two species compete for territory along their mutual boundary. The species are fairly matched and the result of conflict is the invasion by one of the species of territory held by the other.
- We only consider the case where conflict takes place only along the frontier, victory and defeat resulting in the acquisition and loss of territory.
- A further specification is that the chance of a particular position being overrun depends only on the disposition of the enemy in the immediate neighbourhood.
- Note that throughout the paper we refer to members of the opposing species as black and white cells.

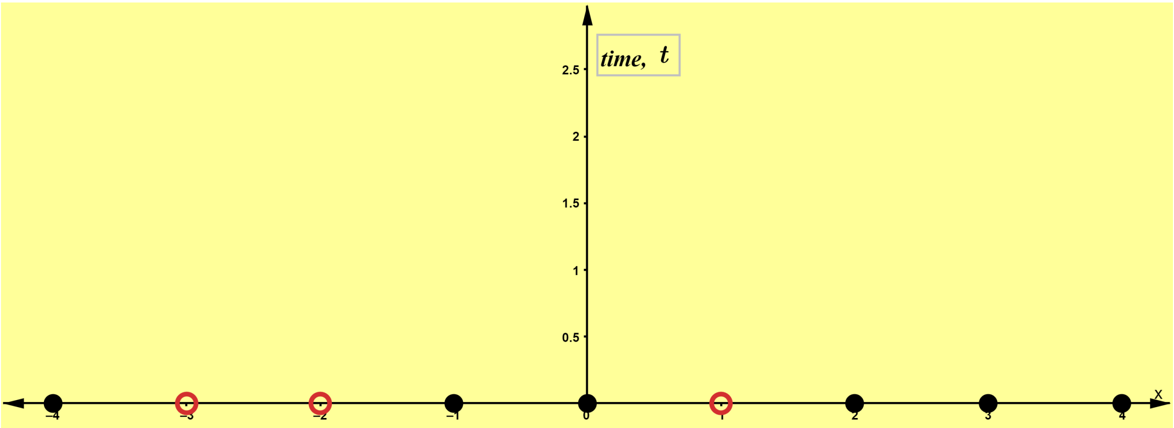
2.2 Swapping Process on $\mathbb{R}^1$

In swapping processes confrontations are resolved by the exchange of territory.
In the one-dimensional case each integral point on a line is occupied either by a black or white cell and the following process occurs in time. In any small interval of time $(t, t+\tau)$ the probability of adjacent black and white cells swapping positions is $\tau + o(\tau)$.

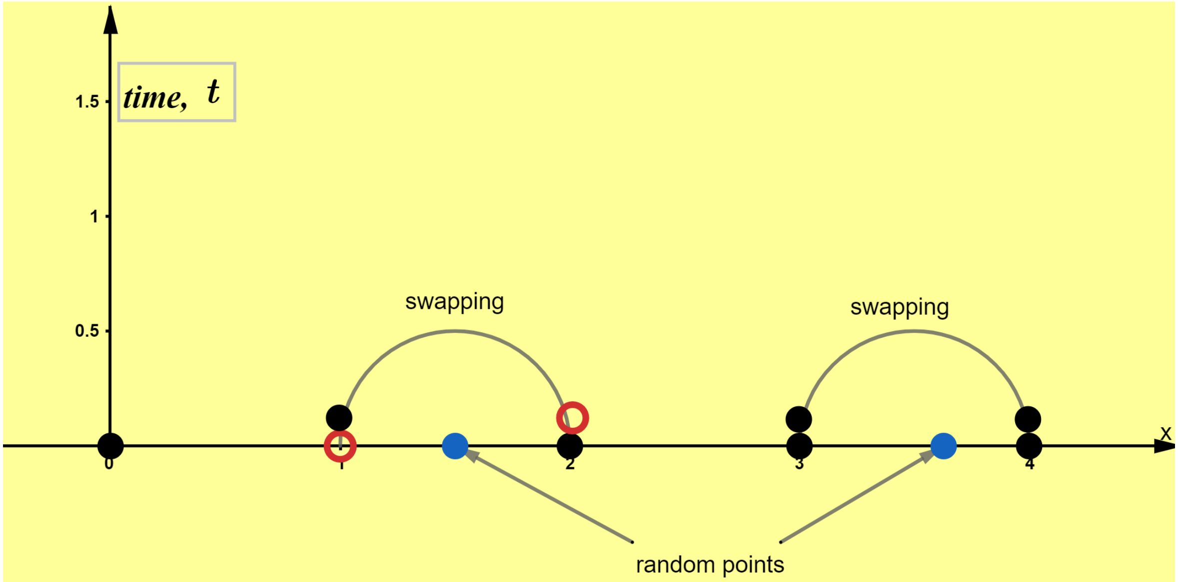


2.2.1 Representation of swapping process

- Consider a plane with the original line of cells along one axis and continuous time along the other, with points thrown down randomly with unit density.



- We now envisage, that the line of black and white cells sweeps forward in time and that swapping of adjacent cells take place at every instant when a random point intercedes between them.
- If a random point passes between two cells of same colour then swapping occurs but the pattern of colours remains unchanged.
- If a random point passes between two cells of different colours then swapping occurs and the pattern of colours changes.



Note

- The probability that adjacent black and white cells swap position in any time interval is the same in both representations.
- With respect to the development of the pattern of colours the processes are stochastically identical.

Since the points in the plane determine a unique path back to some position in the original configuration and each path is a realization of a unit step symmetric random walk in continuous time with rate two. Before summarizing this as a lemma, let us first define some notations that will be useful later on,

- $\mathbb{P}_k(n,t)$: The probability that position n at time t is occupied by the cell originally at position k
- $\mathbf{X}_t$: unit step symmetric random walk in continuous time with rate 2, on integer lattice
- $\mathbb{P}(\mathbf{X}_t = n \mid \mathbf{X}_0 = k)$: Probability that random walk $\mathbf{X}_t$ starting at k, is at n at time t

Lemma 2.1 *In a swapping process on* $\mathbb{Z}$,

$$\mathbb{P}_k(n,t) = \mathbb{P}(\mathbf{X}_t = n \mid \mathbf{X}_0 = k)$$

In words, this lemma basically states that in a swapping process on the lattice of integers, the probability of the position n being occupied at time t by the cell originally at k is the probability that the unit step symmetric random walk in continuous time with rate 2, starting at k, is at n at time t.
We arrive at a theorem, which is actually a consequence of the above Lemma 2.1. The notations used are,

- A = Set of positions initially occupied by black cells
- A_t = Set of positions occupied by black cells at time t
- $\mathbb{P}_A(x_0 \in A_{t_0})$ = Probability that at time t_0 the position x_0 is occupied by a black cell

Theorem 2.2 *In a swapping process,*

$$\mathbb{P}_A(x_0 \in A_{t_0}) = \mathbb{P}(\mathbf{X}_{t_0} \in A \mid \mathbf{X}_0 = x_0)$$

The theorem basically states that at time t_0 the probability that a position x_0 is occupied by a black cell is the probability that a random walk starting from x_0 at time $t = 0$ takes one of the values in the set A at time $t = t_0$.

Note

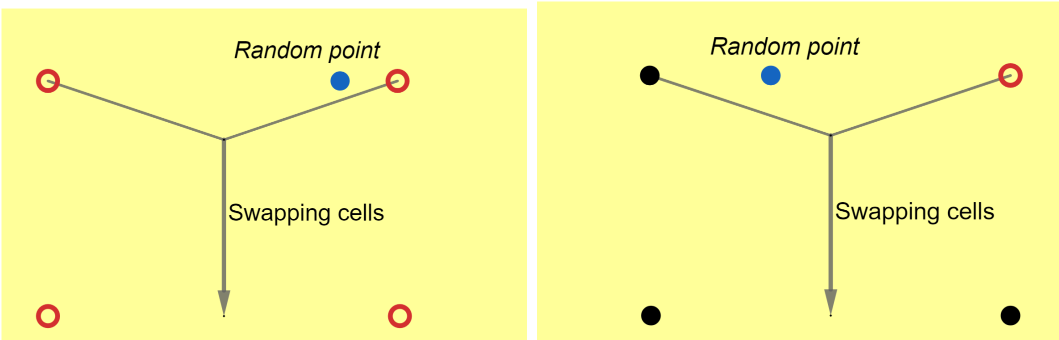
Note that the theorem, with the appropriate random walk, is true for arbitrary graphs. Thus, for an arbitrary set of positions, it will be necessary to decide which pairs of positions will be able to swap cells. This will determine the arcs of a graph and the random points will be thrown down on the product space formed by the arcs of the graph and the time dimension.

2.3 Invasion Process on $\mathbb{R}^1$

- In the invasion process a pair of adjacent black and white cells do not swap, but either cell may generate a new cell of the same colour; the neighbour is then eliminated and replaced by the newly born cell.
- Thus a black-white combination may become either black-black or white-white. The process is fair if both these events occur at a unit Poisson rate.

2.3.1 Representation of Invasion process

- The representation of invasion process is essentially the same as the representation of swapping process.
- The differences are
  - The random points have density two
  - If at some time s, say, there is a random point between two positions, then both positions become of one colour, namely, the colour of the cell at the position nearer the point immediately prior to s.



Our observation from the representation of Invasion Process:

- We observe that with regard to the development of patterns of colours we have two equivalent processes.
- The random walk taken to determine the probability of a particular position being occupied by a black cell is much the same, except in that a jump along an arc of the lattice is made only if a random point is found more than half way along that arc.
- Because the points have been thrown down with density two the random walk has the same prob- abilities as that in the swapping case.

This gives us the following theorem

Theorem 2.3 *Starting from the same initial configuration, the probabilities that a certain position will be occupied by a black cell in a swapping process and a fair invasion process are equal.*

3 Voter Model on Complete Graphs

3.1 Basic Voter Model - Overview

Let us have N agents, each of which can be in one of q states (opinions) $\sigma \in S$.

1. Each agent sits on a node of a fully connected network, and they can interact along the edges with their nearest neighbours.
2. The opinion of the agent i is denoted σ_i.
3. The state of the system is described by the set of opinions of all the agents, $\Sigma = [\sigma_1, \sigma_2, \cdots, \sigma_N]$.
4. We choose an agent i `(speaker)` at random out of the N agents.
5. Then, choose j `(listener)` randomly among neighbours and set

$$\sigma_j(t+1) = \sigma_i(t).$$

6. The process can be summarized as:

$$AB \to AA \quad , \quad BA \to BB.$$

3.2 Voter Model (Pictorial Representation)

Figure 1: Initial Configuration

Figure 2: First Step

If we had taken the brown and the blue vertices then, the colors would have been swapped, but if we had taken the two blue vertices there would be no change in color.

3.3 Opinion Sets

Definition 3.1 $V_i(t) := \{j : j th\ agent\ has\ opinion\ i\ at\ time\ t\}$.

The number of non empty $V_i(t)$'s for $i = 1, \ldots, N$ can only decrease with time. The `ultimate behaviour` of the system is one of the absorbing states which are for any fixed opinion i,

$$\sigma_j = i \quad \forall 1 \leq j \leq N$$

The following is a random partition of the set $\{1, 2, \ldots, N\}$ for each t,

$$\{V_i(t) : 1 \leq i \leq N\}$$

An important variable of interest is the `consensus time` which is time taken to reach one of the absorbing states i.e,

$$T_N^{\text{voter}} := min\{t : V_i(t) = \{1, 2, \ldots, N\} \quad \text{for some} \quad 1 \leq j \leq N\}$$

3.4 Coalescing MC model

For each agent a set of tokens $C_i(t)$ is assigned varying with time. Initialization is done by

$$C_i(0) = \{i\} \quad \forall 1 \leq j \leq N$$

For i^{th} and j^{th} agent interacting with direction $i \to j$ at time t,

$$C_i(t+1) = \phi, \qquad C_j(t+1) = C_i(t) \cup C_j(t)$$

Now, the following is also random partition of the set $\{1, 2, \ldots, N\}$ for each t,

$$\{C_i(t) : 1 \leq i \leq N\}$$

An important variable of interest is `coalescence time` which is the time taken to reach for some i,

$$T^C := min\{t : C_i(t) = \{1, 2, \ldots, N\} \quad \text{for some} \quad 1 \leq j \leq N\}$$

3.5 The duality relationship

Result 3.2 *For fixed* t, $\{V_i(t) : 1 \leq i \leq N\} \stackrel{d}{=} \{C_i(t) : 1 \leq i \leq N\}$

Also, the following holds. The above means that the sets are same but ordering might change.

Result 3.3 *Coalescence time and consensus time are same in distribution.*

$$T^C = T^V$$

For fixed i, $|V_i(t)|$ changes by ± 1, whereas $|C_i(t)|$ may jump to 0 from a higher value. Hence they are different processes.

3.6 Representation of the above model

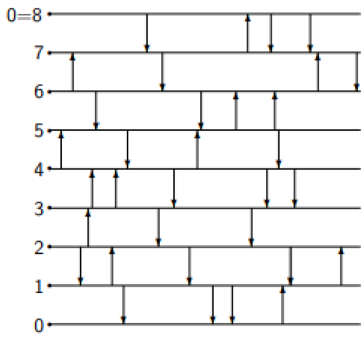


Figure 3: Moving left to right in time

We move from left to right in time. The voter model and the coalescing model can be visualized as follows: When an arrow touches a horizontal line, it propagates the speaker's opinion to the line the arrow touches. If we trace back to the starting points from right to left we get the coalescing set and if we trace to the end from left to right we get the voter set.

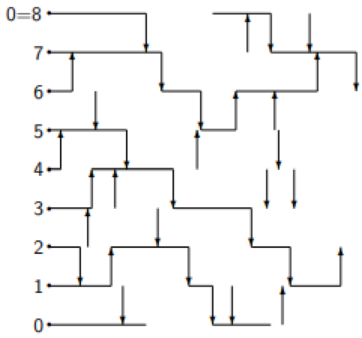


Figure 4: Left to right and right to Left

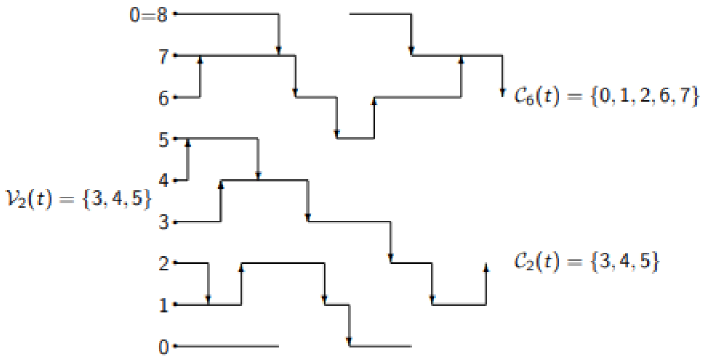


Figure 5: Duality

3.7 Easy Approach for asymptotics

A slight continuous time variation of the voter model with 2 opinions and N agents is given by the continuous time birth and death chain on $\{1, 2, \ldots, N\}$ with

$$X(t) = \text{number of agents with first opinion at time } t.$$

$$\text{Rates:} \quad \lambda_k = \mu_k = \frac{k(N-k)}{2(N-1)}$$

Clearly, the absorbing states are 0 and N. The rates are proportional to the number of pairs $k(N-k)$ since, the minimum of exponentials will follow exponential with parameter as the sum of individual parameters. We want to study,

$$T_{0,N}^{\text{hit}} := \min\{t : X(t) = 0 \text{ or } N\}$$

By general birth-and-death formulas,

$$\mathbb{E}_k T_{0,N}^{\text{hit}} = \frac{2(N-1)}{N} \left(k\left(h_{N-1} - h_{k+1}\right) + (N-k)\left(h_{N-1} - h_{N-k+1}\right)\right)$$

where $h_m := \sum_{i=1}^{m} 1/i$. This is maximized by $k = \lfloor N/2 \rfloor$, and

$$\max \mathbb{E}_k T_{0,N}^{\text{hit}} \sim (2 \log 2) N.$$

3.8 Framing the inequality

For the true voter model with N different opinions, assign two dummy opinions to k and $N-k$ agents respectively. The process is continued with the changes in opinions carried out in both the original and dummy opinion. The number of agents having first dummy opinion follow the 2 opinion model. The original opinion being same for any two agents always implies same dummy opinion but not vice versa. This implies the dummy opinions become all same if all the original opinions become same.

4 Calculations and Mean Field Theory

4.1 Occupation Numbers

Definition 4.1 *Occupation numbers are defined as*

$$N_j = \sum_{i=1}^{N} \mathbf{1}(\sigma_i = j) = |V_j|$$

or equivalently the densities for each opinion $j \in \{1, 2, \cdots, N\}$.

$$n_j = N_j/N,$$

In fully-connected network the state of the system is fully described by the occupation numbers. The dynamics of these occupation numbers fully describes the evolution of the system. As the total number of nodes is conserved, there are $q-1$ independent dynamical variables. Let us start with only two opinions, $q = 2$. The variable σ can assume only two values, denoted $\sigma = \pm 1$ for convenience. The state is described by one dynamical variable only, which will be taken as a *magnetisation*,

$$m = \frac{N_+ - N_-}{N}$$

In one step of the dynamics, three events can happen. The magnetisation may remain constant or it can change by $\pm\frac{2}{n}$. The probabilities of these three events can be easily calculated $\longrightarrow$

$$\mathbb{P}\left\{m \to m + \frac{2}{N}\right\} = \frac{(N_+) \times (N_-)}{N \times (N-1)} = \frac{1}{4}\left(1 - m^2\right)\left(\frac{N}{N-1}\right)$$

$$\mathbb{P}\left\{m \to m - \frac{2}{N}\right\} = \frac{(N_-) \times (N_+)}{N \times (N-1)} = \frac{1}{4}\left(1 - m^2\right)\left(\frac{N}{N-1}\right)$$

$$\mathbb{P}\{m \to m\} = \frac{(N^2{}_+) + (N^2{}_-)}{N \times (N-1)} = \left(\frac{1}{2}\left(1 + m^2\right) - \frac{1}{N}\right)\left(\frac{N}{N-1}\right)$$

4.2 Simulation Studies

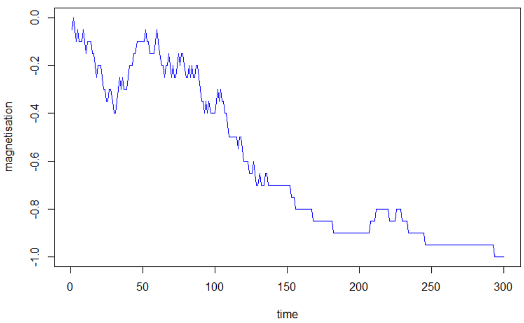


Figure 6: Model with $N = 40, k = 20$

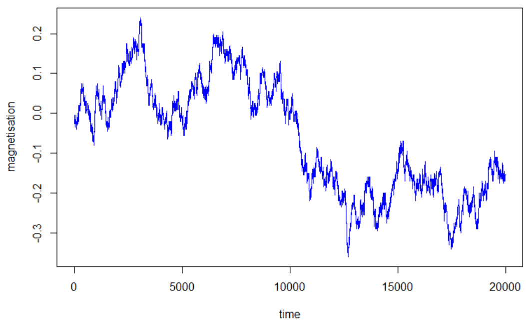


Figure 7: Model with $N = 4000, k = 2000$

4.3 Partial Differential Equation

Our objective is writing the master equation for the probability density of the random variable $m(t)$, which we denote P_m.
$P_m(m, \tau)$ is the probability of the random variable $m(t)$ having magnetization m after time τ.
The probability density $P_m(m, \tau)$ for $m(t)$ evolves according to the partial differential equation where time is re scaled as $N \to \infty$ by τN^2

$$\frac{\partial}{\partial \tau} P_m(m, \tau) = \frac{\partial^2}{\partial m^2} \left[\left(1 - m^2\right) P_m(m, \tau) \right] \qquad \text{where } \tau = \frac{t}{N^2} \tag{1}$$

This equation describes in principle fully the evolution of the Voter Model on a complete graph.

Say that, we have $q \gg 1$. Let us define the distribution of occupation numbers as follows

Definition 4.2 *The distribution of occupation numbers say, $D(n)$:*

$$D(n) = \frac{N}{q} \sum_{\sigma=1}^{q} \delta\left(n - n_\sigma\right) \quad ; \delta(x) = 1 \ \textit{if} \quad x = 0 \tag{2}$$

where $\delta(x) = 1$ for $x = 0$ and zero elsewhere.

It would be much more difficult to write the full dynamic equation for $D(n)$. Therefore, we use the approximation which replaces the distribution $D(n)$ by its configuration average $P_n(n) = \langle D(n) \rangle$.
In the limit $N \to \infty$ and $q \to \infty$ and substituting the variable $x = 2n - 1$ we arrive at the equation

$$\frac{\partial}{\partial \tau} P_n(x, \tau) = \frac{\partial^2}{\partial x^2} \left[\left(1 - x^2\right) P_n(x, \tau) \right] \tag{3}$$

Now, the equation

$$\frac{\partial}{\partial\tau}P(x,\tau) = \frac{\partial^2}{\partial x^2}\left[\left(1-x^2\right)P(x,\tau)\right] \tag{4}$$

describes both the cases, when $q = 2$ and when $q \gg 1$, only the interpretation of the variable x differ: in the former case it corresponds to the magnetisation, while in the latter case it is shifted percentage of votes. By solving the above equation we treat simultaneously both cases.
Now we look for the solution using the expansion in eigen vectors. We can write the equation in the Linear Operator Form

$$\frac{\partial}{\partial\tau}P(x,\tau) = \mathcal{L}P(x,\tau) \tag{5}$$

where the linear operator $\mathcal{L}$ acts as

$$(\mathcal{L}f)(x) = \frac{\partial^2}{\partial x^2}\left[\left(1-x^2\right)f(x)\right] \tag{6}$$

We need to find the set of eigen vectors of $\mathcal{L}$. Denoting $\Phi_c(x)$ the eigen vector corresponding to the eigenvalue $-c$, we have the following equation

$$\left(1-x^2\right)\Phi_c''(x) - 4x\Phi_c'(x) + (c-2)\Phi_c(x) = 0 \tag{7}$$

The full solution can be then expanded a

$$P(x,\tau) = \sum_c A_c e^{-c\tau}\Phi_c(x) \tag{8}$$

where A_c comes from the initial conditions.

We look for the solution of Equation (7) in the space of distributions (i.e. linear functionals on sufficiently differentiable functions) with support restricted to the interval [-1,1].

- **Solution when c =0**
  We want to find the eigenvectors corresponding to eigenvalue $c = 0$ i.e. the stationary solutions of the Equation (4). They are composed of δ functions only. The corresponding eigen subspace is two-dimensional and the base vectors can be chosen as
  $$\Phi_{01} = \delta(x-1) \quad , \quad \Phi_{02} = \delta(x+1)$$

- **Solution when c $\neq$ 0**
  Now for $c \neq 0$ we first decompose the solution in ordinary function of x plus a pair of δ -functions, namely
  $$\Phi_c = \phi_{c+}\delta(x-1) + \phi_{c-}\delta(x+1) + +\phi_c(x)\theta(x-1)\theta(x+1) \tag{9}$$
  where ϕ_{c+} and ϕ_{c-} are real numbers and $\phi_c(x)$ is a real doubly differentiable function. Then, the Equation (2) translates into equation for $\phi_c(x)$ given as
  $$\left(1-x^2\right)\phi_c''(x) - 4x\phi_e'(x) + (c-2)\phi_c(x) = 0 \tag{10}$$

accompanied by conditions:

$$\lim_{x\to+1}\phi_c(x) = -\frac{c}{2}\phi_{c+}$$

$$\lim_{x\to-1}\phi_c(x) = -\frac{c}{2}\phi_{c-}$$

The general solution of the equation exhibits behaviour $\phi_c(x) \sim (1\mp x)^\alpha$ at $x\to\pm1$, where either $\alpha=0$ or $\alpha=-1$.

We want to find the coefficients in the solution of equation(7) given by equation(8) *i.e;*

$$P(x,\tau) = \sum_c A_c e^{-c\tau}\Phi_c(x)$$

with initial conditions $P(x,0)=P_0(x)$. This is calculated as

$$\text{Initial Configuration, } A_c = \frac{\int P_0(x)\psi_c(x)\mathrm{d}x}{\int \phi_c(x)\psi_c(x)\mathrm{d}x} \tag{11}$$

Solution (8) help us to deduce an important feature for the distribution of waiting times needed to reach the stationary state.

Let us denote $P_{\mathrm{st}}(\tau)$ is the probability density for ending at time τ in the stationary frozen configuration with all agents in the same state.
The probability that the stationary configuration was not reached before time τ :

$$P_{st}^{>}(\tau) \equiv \int_\tau^\infty P_{\mathrm{st}}(\tau')\,\mathrm{d}\tau' = 1 - \lim_{\varepsilon\to0^+}\left(\int_{-1-\varepsilon}^{-1+\varepsilon} + \int_{1-\varepsilon}^{1+\varepsilon}\right) P(x,\tau)\mathrm{d}x \tag{12}$$

More explicitly, we find

$$P_{st}^{>}(\tau) = \sum_{c>0} 2A_c \frac{\phi_c(-1)+\phi_c(1)}{c} e^{-c\tau} \tag{13}$$

The distribution of waiting times will have an exponential tail

$$P_{st}^{>}(\tau) \sim e^{-2\tau}, \tau\to\infty \tag{14}$$

For initial condition $P_0(x) = \delta(x-x_0)$ we can also compute the pre factor in the leading term for large τ.

$$P_{\mathrm{st}}^{>}(\tau) \simeq \frac{6}{4}\left(1-x_0^2\right)e^{-2\tau}, \tau\to\infty \tag{15}$$

4.3.1 Stationary State Probability

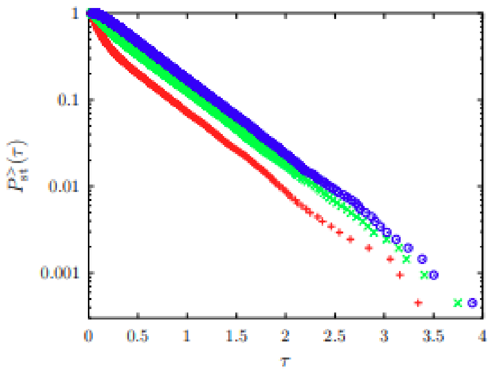


Figure 8: Probability of reaching the stationary state in time larger than τ , q = 2, N = 2000. The values of initial fraction p of opinions +1 are 0.1 (+) 0.2 ($\times$) and 0.7 (.)

5 Results and Simulations

Master Equation for the probability density $P(m, t')$ for complete graphs :

$$\frac{\partial P(m, t')}{\partial t'} = \frac{\partial^2}{\partial \mu^2} \left[\left(1 - m^2\right) P(m, t') \right]$$

The natural scaling of time with the number of sites is taken as$t' = \frac{t}{N^2}$. By standard methods for the Fokker-Planck equation, for large $t' = \frac{t}{N^2}$, the fraction of edges connecting nodes with opposite values of the variable (active edges) is

$$n_A(t') = \frac{(1 - m_0^2)}{2} e^{-2t'}$$

where m_0 is the initial magnetization.

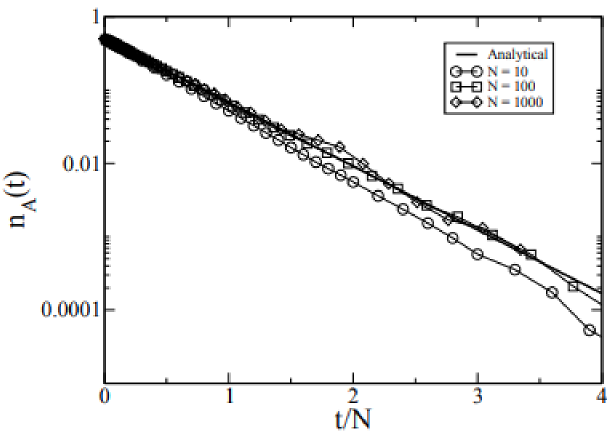


Figure 9: Fraction of active edges $n_A(t)$ for voter dynamics on a complete graph

5.1 Model for Simulation Code:

Each node (voter) takes one of the finite discrete states (say, black , white, red and blue). The voter model considers only one opinion transfer event at a time between a pair of connected nodes that are randomly chosen from the network. First, a (listener node) is randomly chosen from the network, and then a "speaker" node is randomly chosen from the listener's neighbors.

We have not taken a complete graph while running the visual simulation, so it took longer iterations to reach consensus. We have also tweaked with the probability that when two vertices are chosen with some probability they get changed, otherwise with some probability they do not change.

From the visual simulation we can see the absorbing states are reached eventually. We can check the following from the simulation.

1. Random fluctuations bring eventually all surviving runs to the fully ordered absorbing state; however, as long as the runs survive they do not order on average.

2. The decay of $n_A(t)$ is just a consequence of the decay of $P^{>}_{st}(t)$.

5.2 Simulation Results:

Figure 10: Iteration=128

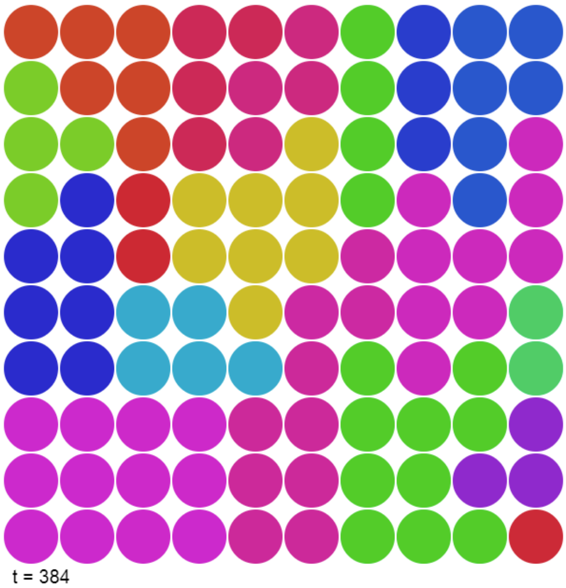


Figure 11: Iteration=384

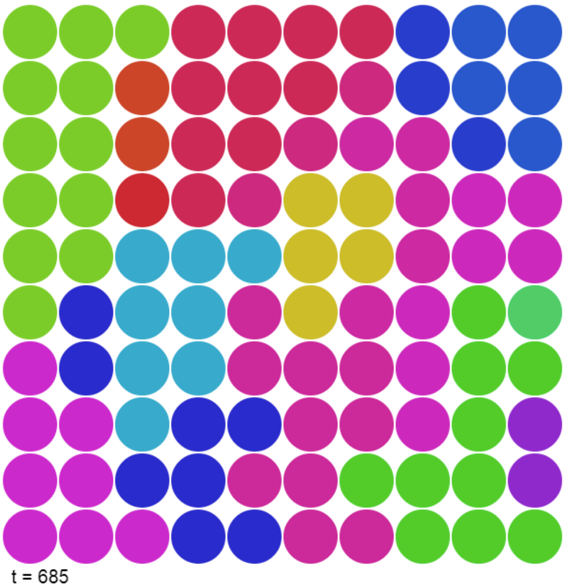


Figure 12: Iteration=685

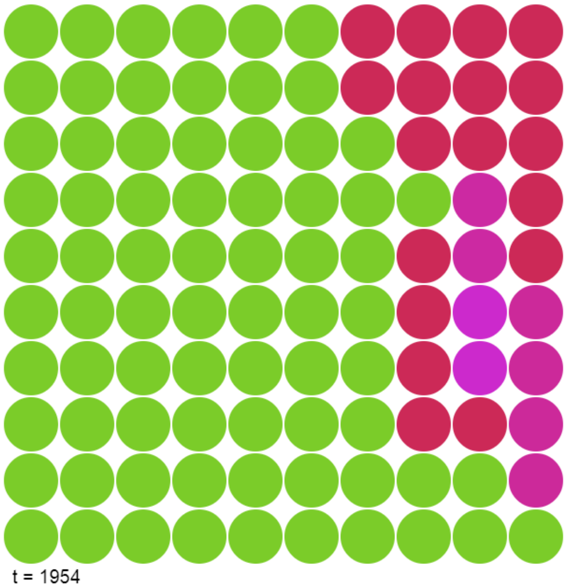


Figure 13: Iteration=1954

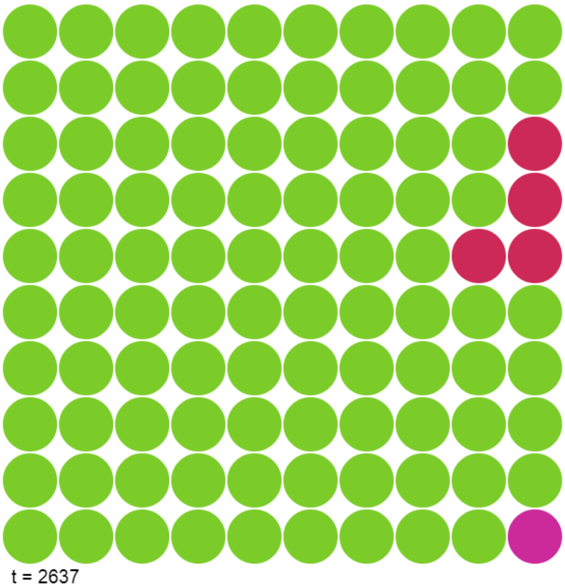


Figure 14: Iteration=2637

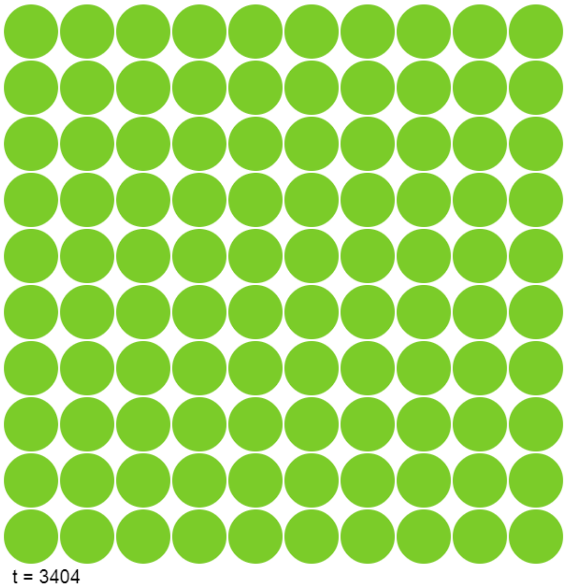


Figure 15: Iteration=3404