Hierarchies in Humans & Opinion dynamics

Bioinspired Systems - 2021 Sept 29

Technical information

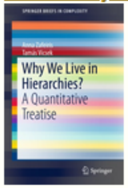
The presentations (pdf versions of the ppt files) and exam topics are available on https://hal.elte.hu/~lanna/

Main Page

Teaching

Collective Motion (Review)

Hierarchy book



Statistical Physics of Biological Systems

Bio-inspired Systems

2021

Information sheet

Number	Lecture slides	Exam-topic slides
1	Collective Motion (Sept 9)	Topic 1 - Collective Motion
2	Hierarchy formation - Part 1(Sept 15)	Topic 2 - Hierarchy formation Part 1
3	Hierarchy formation - Part 2(Sept 22)	

Part 1
Hierarchies in Humans

(Continuation of the previous lecture)

Large-scale human hierarchies: from small goups to ultrasocieties

Problem:

- What enforced the transition from small, genetically related cooperative H-G groups to huge anonymous, hierarchically organized societies, typically organized as states, “ultrasocieties”?
  - small, „traditional” HG societies: kin selection + reciprocal altruism
  - Only 10-12,000 years ago (vs. 200,000 y)
- Neolithic transition
- Dunbar Number

Turchin et al, War, space, and the evolution of Old World complex societies, PNAS, 2013

Existing theories

- Many theories, but non of them completely satisfactory
- Mostly anthropological, historical approaches (qualitative)
- Quantitative approaches are rare (but existent)
  - a field of science in its infancy

  - Mostly agent-based models:
    - Barceló and Castillo (eds) 2016: Simulating Prehistoric and Ancient Worlds (Computational Social Sciences). Springer, Cham, Switzerland
    - Grinin and Korotayev (eds) 2014: History & Mathematics: Trends and Cycles. Uchitel, Volgograd
    - Pumain and Reuillon 2017: Urban Dynamics and Simulation Models (Lecture Notes in Morphogenesis). Springer, Cham, Switzerland
  - AB models combined with game theory
    - Boix 2015, Political Order and Inequality. Cambridge Univ. Press, New Jersey
    - Greif 2006. Institutions and the Path to the Modern Economy: Lessons from Medieval Trade. Cambridge Univ. Press, New York
  - The book by Turchin (2003) Historical Dynamics: Why States Rise and Fall. Princeton Univ. Press, New Jersey - offers one of the deepest analysis

Premise: Costly institutions that enabled large human groups to function without splitting up evolved as a result of :

1. Warfare
2. Multilevel selection

Warfare intensity depends on

- the spread of historically attested military technologies (e.g., chariots and cavalry) and
- geographic factors (e.g., rugged landscape).

Multilevel selection:

- group selection „on the top of" individual selection

Simplified train of thought

- **Small H-G societies:** Throughout most of human history, people lived in small-scale, mostly egalitarian societies.
- **Warfare over resources:** These tribes often engaged in warfare with each other, over various resources.
- **Selfishness vs. Group behavior:** Although selfish behavior can be beneficial for the individuals within a group, when groups intensively compete with each other (for example, during warfare), those groups that have more cooperative and less selfish members have the advantage. Thus, human societies are subject **to multilevel selection**.

The effects of warfare on social evolution:

- Groups become internally more cohesive
- Technological progress, including military and organizational applications
- „God always favors the big battalions" (Napoleon / Turenne) → **Enlargement of group sizes**

The **capacity of the human brain** has its limits,

- it cannot handle social relations in detail among more than around 150 people (Dunbar number).
- → there is a limit to the size of egalitarian, face-to-face human groups.

Simplified train of thought – cont.

Pressure on the group size to grow $\leftrightarrow$ Dunbar no.

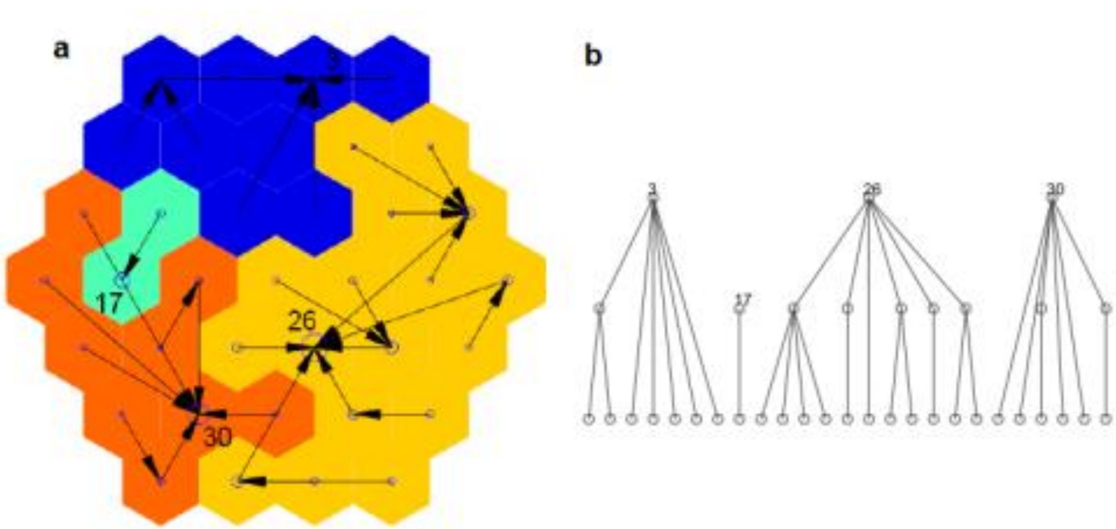
Assumption: the evolutionary response to this dilemma:

1. the ability to demarcate group membership based on cultural traits (language, dialect, clothing, etc.)
2. hierarchical organization, allowing group sizes to grow basically ad infinitum

Each element within a given level of a strictly hierarchical system needs to have, at most, n+1 connections: n :„span of control”; +1: its superior

Turchin-model:

- Nodes stand for a political entity (e.g., villages)
- Numerical experiments with AB model:
  - The modelled area is divided into hexagonal cells (autonomous local communities „villages”)
  - Each of these villages are characterized by:
    - **a base-line resource level**, accounting for the heterogeneous environment, defining the productive/demographic potential of the region (a tunable parameter)
    - **actual resource level**, the base-line resource level minus the costs of the various actions in which the given community participates



A system of 37 communities organizing themselves into four polities.
The numbers in the hexagons mark the chief communities.

a. Spatial view.
b. The hierarchical structure

The Turchin-model in detail:

- Polities are organized in a hierarchical way
- Subordinate communities pay „tribute”to their superiors (a fixed portion of their total resources)
  → the total resource level of a community =

  = base resource level - tribute + the tribute it receives from its subordinates
- Polities may engage in warfare
  - Rebel
  - Conquest
  - Being attacked

Probability of warfare

A polity will attack its weakest neighbor if

i. it estimates that the attack will be successful

ii. it is ready to pay the corresponding costs and

iii. it is not too devastated from previous wars.

Quantitatively, the probability of an attack is:

$$A_{ij} = P_{ij} \cdot e^{-\beta c_{ij}} \cdot \frac{F_i}{F_{i,0}}$$

$P_{i,j}$: the probability of success (an attack by community *i* on community *j*)

$$P_{ij} = \frac{F_i^a}{F_i^a + F_j^a}$$

F_i : the power of polity *i*

$F_{i,0}$: the maximum possible power of polity *i*

a : is the „success probability exponent”

$c_{i,j}$: cost of warfare

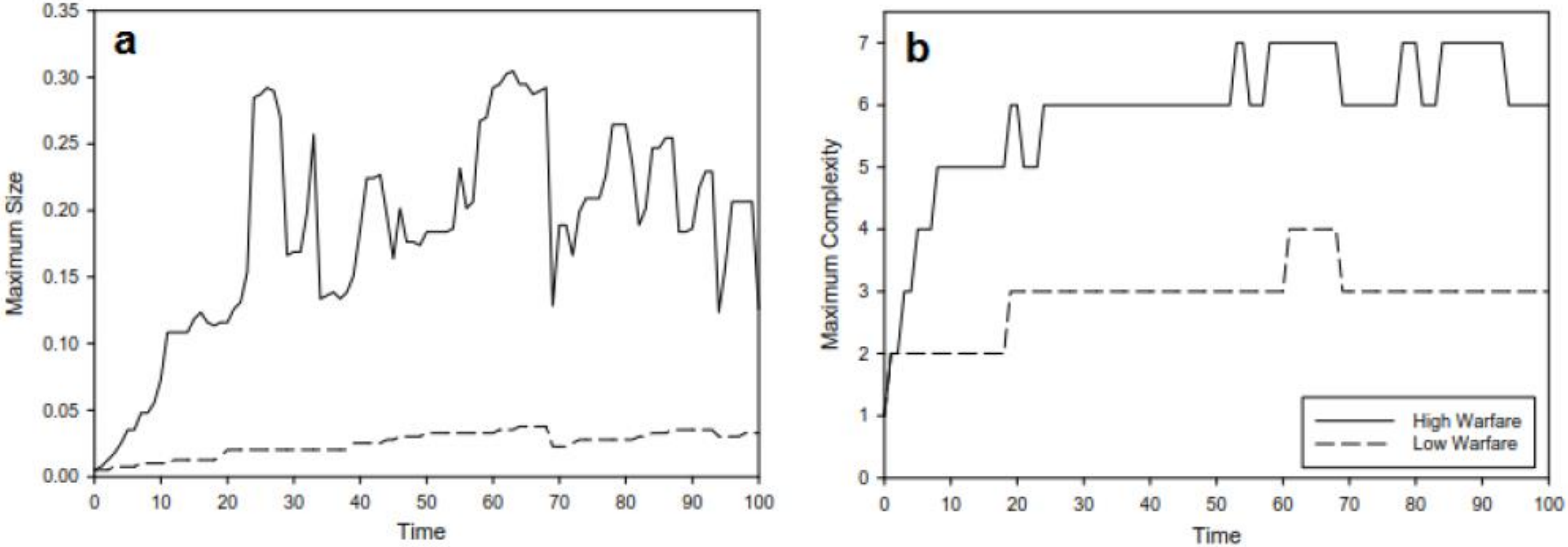
β : parameter

The Turchin-model in detail:

- Each time step is considered to be a year.
- Each year, the chief community decides whether to launch an attack on its weakest neighbor.
- If it decides to go to war:
  - it first attempts to conquer the bordering communities, followed by a series of „battles", until it either suffers a defeat or the chief community of the victim polity falls.
  - Annexing the conquered communities may require restructuring the hierarchical organization of the winner polity (the number of max. subordinates is a parameter varying between 4 and 10)
  - the direct subordinates of the aggressor chief community might decide to **secede** if they estimate that the attack will be unsuccessful.

    → spatial separation from the master state, together with all the subordinate communities of the rebelling village

Results



(a) The size and
(b) the hierarchical complexity of the polities under low and high pressure of war.

Intense warfare results in larger and more complex polities.

Provides a fission-fusion cycle reminiscent of the dynamics characterising early states of humans.

The model with realistic historical data

- A **more detailed version**
- **Afroeurasian** landmass divided into **a grid of 100 × 100-km squares**
- **Grid cells are characterized** by existence of agriculture, biome (e.g., desert), and elevation
- **At the beginning of the simulation**, each agricultural square is inhabited by an **independent polity**
- Cells adjacent to the steppe are **"seeded" with military technology** (MilTech) traits, which gradually diffuse out to the rest of the landmass
- Each cell is inhabited by a community that has a "**cultural genome**," a vector taking values of 1 or 0, depending on whether an ultrasocial trait is present.
  - such traits are costly: the probability of losing it is big, thus, in the absence of other evolutionary forces, they are present in the landscape at a very low frequency. The force that favors their spread is warfare
- Agricultural cells can conquer other such squares, building multicell polities. **The probability of winning** depends on relative powers, determined by the **polity size** (number of cells) and the average **number of ultrasocial traits**.
- The losing cell may copy the cultural genome of the victor.

causal chain: spread of military technologies → intensification of warfare → evolution of ultrasocial traits → rise of large-scale societies

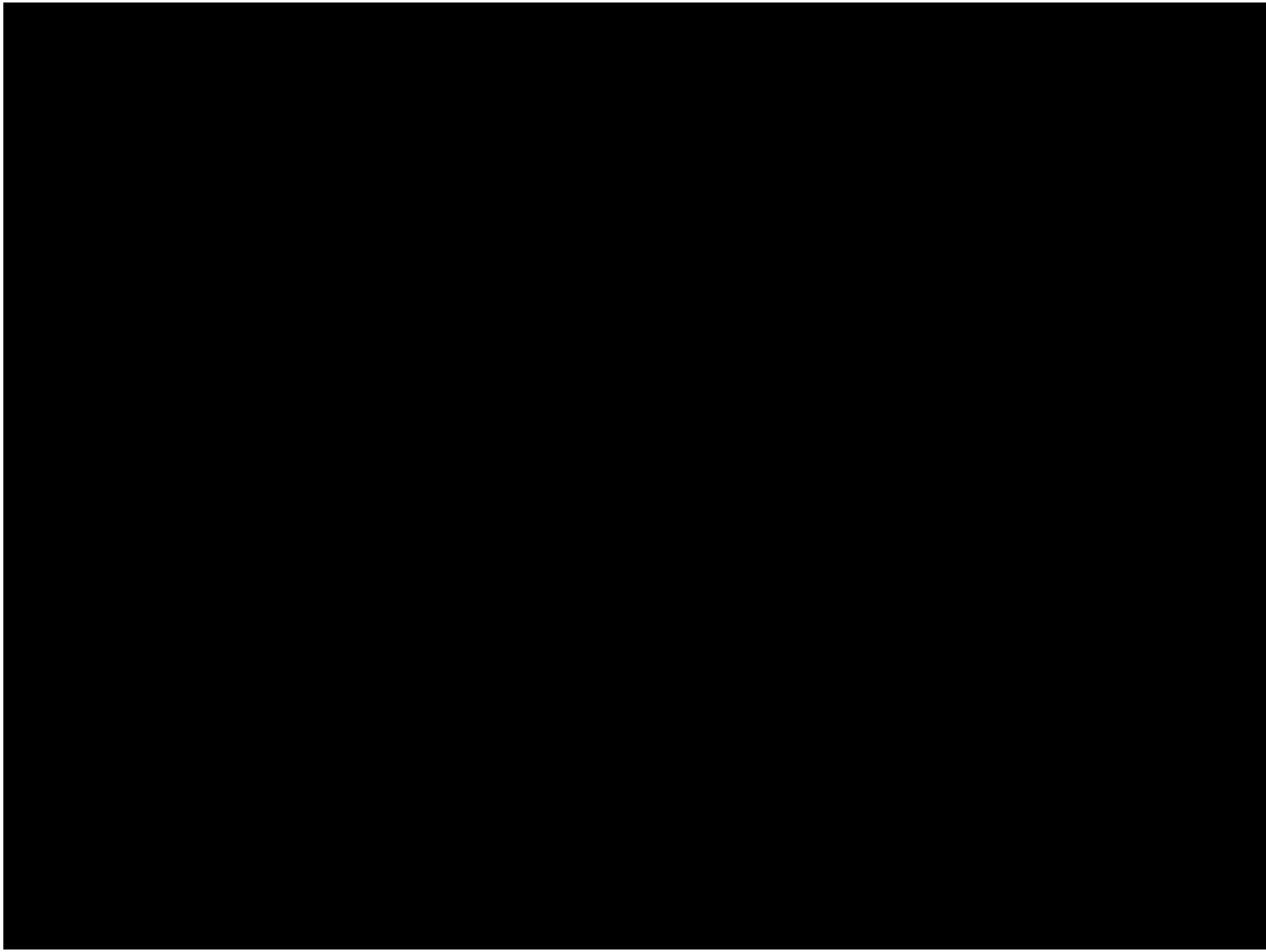
Data

- polities that controlled territories greater than ~100,000 km^2 between 1,500 BCE and 1,500 CE
- on the Afroeurasian landmass
- by taking 100-year time windows, *imperial density maps* indicating the frequency and distribution of large-scale societies
- 7,941 empirical points

- predicting where and when the largest-scale complex societies arise

- hotspots appear in Mesopotamia, Egypt, and North China

- near the steppe frontier, where MilTech diffuse first, tipping the selection in favor of ultrasocial traits



Part 2 - Opinion dynamics

Opinion dynamics

- The scientific field aiming to understand the way „opinions” spread in human communities.
  - The community is usually described by means of networks
    - Nodes are individuals
    - Links are the ties (connections)
      - Direction
      - Strength
  - „opinion” (or „state” of the node) is usually described by a scalar (binary or continuous)

- Close relation to fields studying other spreading phenomena
  - Infection spreading

- Relevant general questions include:
  - What are the fundamental interaction mechanisms („local rules”) that allow for the emergence of
    - Consensus / polarization / fragmentation
    - a shared culture
    - a common language, etc. …
  - What favors the homogenization process? What hinders it?

- “Unfortunately”: Opinion formation is a complex process affected by the interplay of different elements, including the
  - Individual predisposition / family background
  - Background knowledge
  - External information (e.g. public media)
  - Etc.

Typical models

- Consider a finite number of connected agents
- each possessing opinions, described by variables,
- Assume certain *local rules* by which opinions change
  - Change of opinions result from interactions, either with peers or other sources.

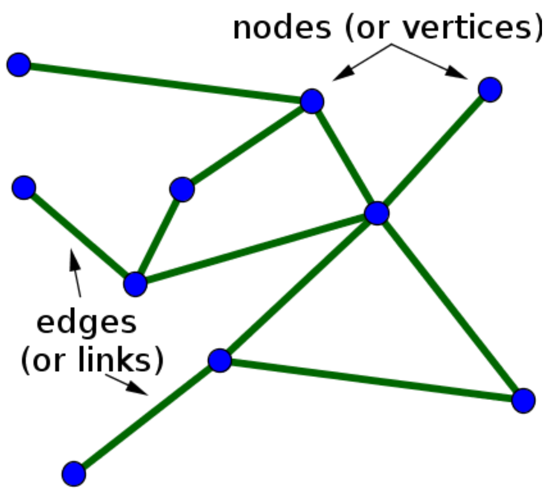
- Opinions:
  - Variables:
    - one dimensional/multidimensional vector
    - discrete (the components can assume a finite number of states)
    - or continuous (values in the domain of real numbers)

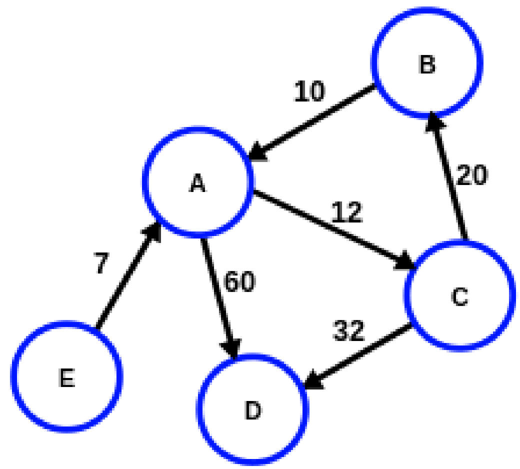
- Connections:
  - Topology of the interaction NW (what is realistic?)
  - “Heritage” from physics: lattices or all-to-all (MF);
    (hardly realistic in social context)

- Drawbacks of the models:
  - many simplifications;
  - many of the omitted parameters (most probably) have a fundamental effect in the final dynamics
  - Hard to say when are the results „good” (polarization)

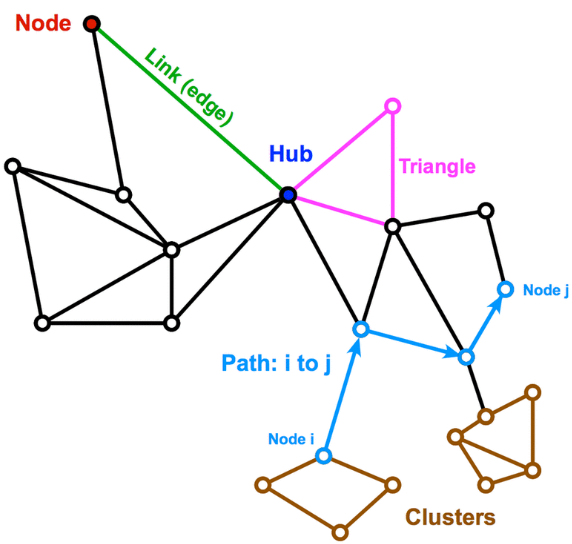
- (It is said to have) Success in:
  - Agreement
  - Cluster formation
  - Transition between order (consensus) and disorder (fragmentation)

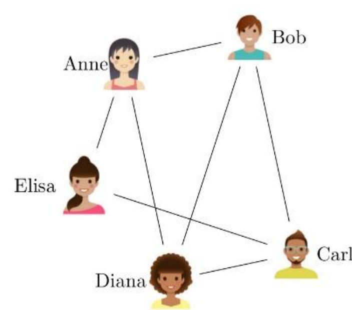
Basic concepts of networks

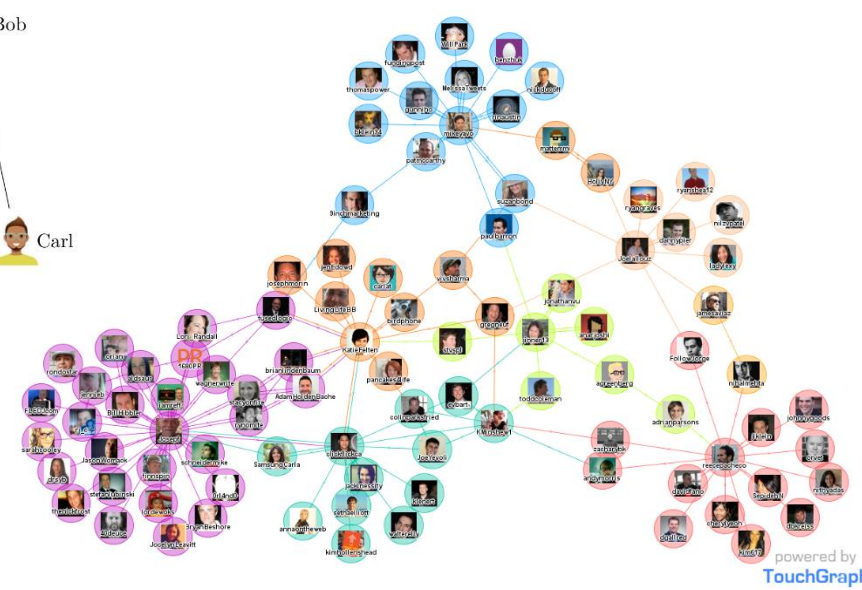




With some network analysis







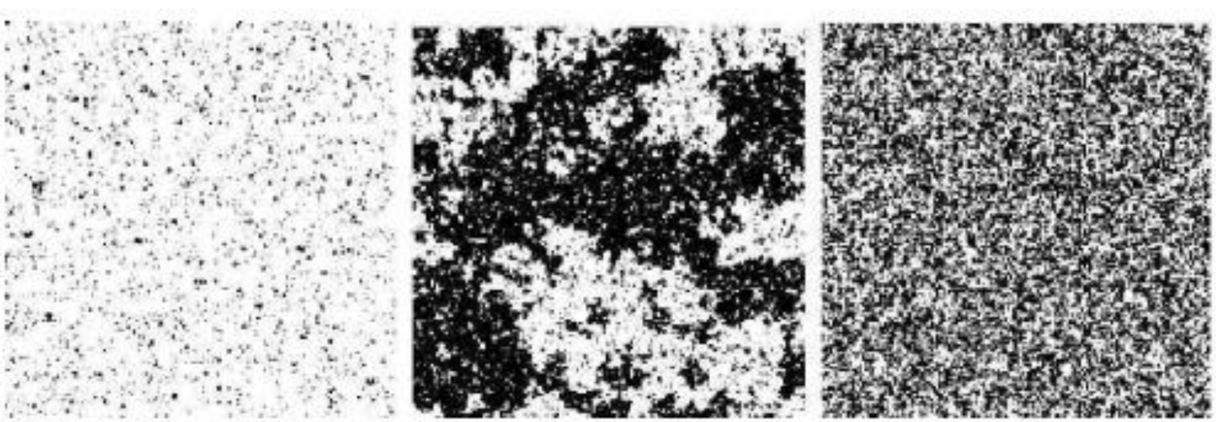
Binary opinions

- Discrete, one dimensional
- 0/1; yes/no; etc
- Interpretation in op. dyn: political questions
  infection models: infected / not
  market behavior: selling/buying

Very first opinion dynamic model by physicist:
1971, Weidlich

Ising model metaphor

- Consider a collection of N spins (agents): s_i
- They can assume two values: +/- 1
- Each spin is energetically pushed to be aligned with its nearest neighbors.
- The total energy is:
  (the sum runs on the pairs of nearest-neighbors) $$H = -\frac{1}{2} \sum_{<i,j>} s_i s_j$$
- Elementary move:
- a single spin flip is accepted with probability $\exp(-\Delta E / k_B T)$
  - ΔE: change in the energy
  - T: temperature (In ferromagnetic systems thermal noise injects fluctuations – tends to destroy order)
  - Critical temperature T_c : above: the system is macroscopically disordered
    under: long-range order is established



Snapshots of equilibrium configurations of the Ising model (from left to right) below, at and above T_c.

Relation to opinion dynamics models

- Each agent has one opinion represented as a spin:
  a choice between two options
- Spin couplings: peer interactions (social conformity)
- Magnetic field: external information / propaganda
- Simple, but attractive model

Potts model (1951)

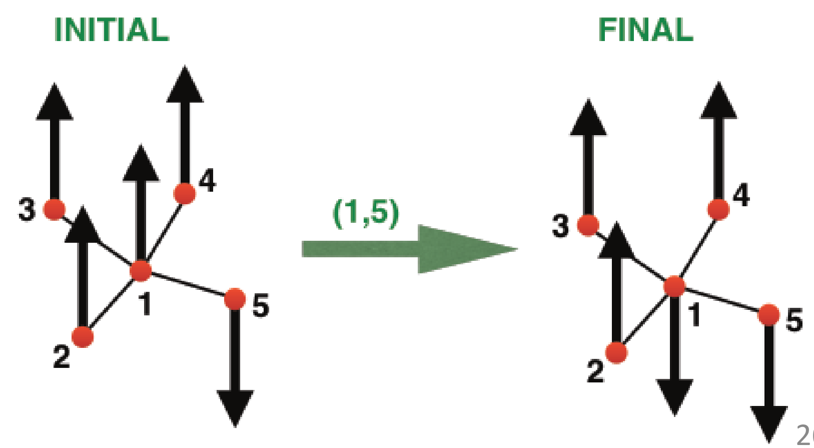
- a generalization of the Ising model
- Each spin can assume one out of q values
- equal nearest neighbor values are energetically favored.
- The Ising model corresponds to the special case $q=2$

Voter model

- Originally introduced to analyze competition among species, early 1970s
- Rather crude description of any real process
- Popular: it is one of the very few non-equilibrium stochastic processes that can be solved exactly in any dimension
- its name stems from its application to electoral competitions
- The model:
  - each agent in a population of *N* holds one of two discrete opinions: $s = +/- 1$
  - agents are connected by an underlying graph (topology)
  - At each time step:

a random agent *i* is selected (1) along with one of its neighbors *j* (5) and the agent takes the opinion of the neighbor: $s_i=s_j$

(alignment *not* to the majority, but to a random neighbor)

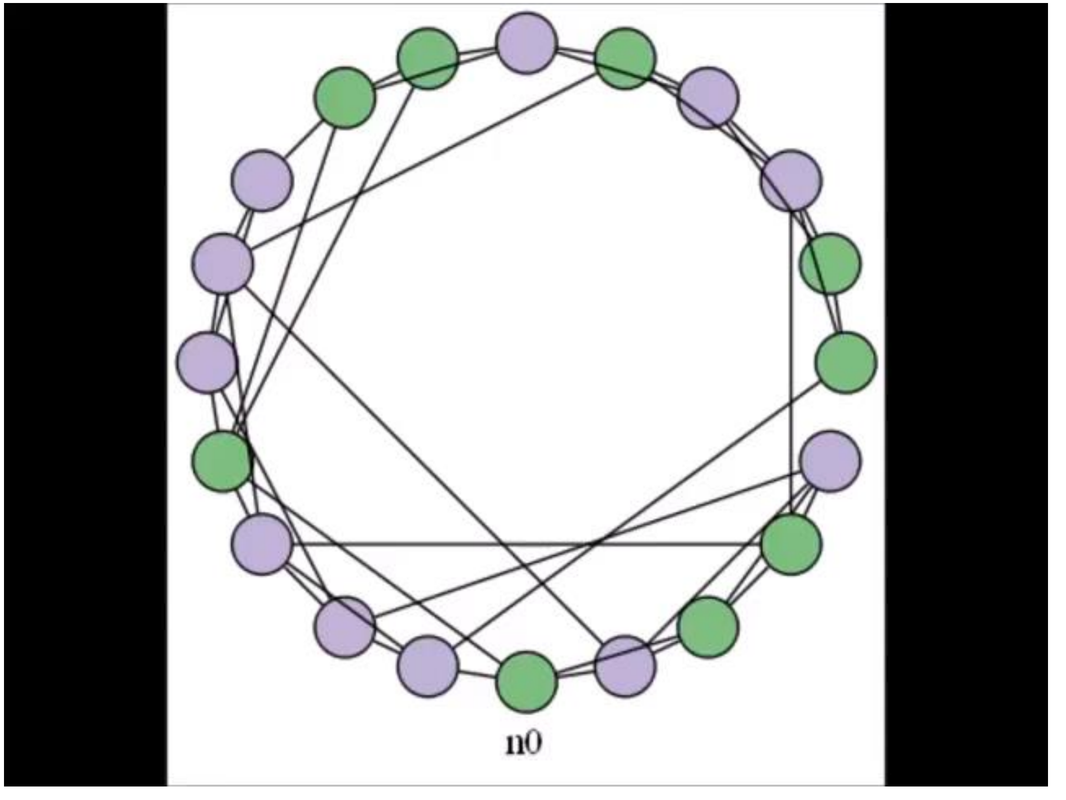


Behavior of the Voter model

- Has been extensively studied
- If people are modeled as vertices in a *d*-dimensional hyper-cubic lattice.
  - For finite system: for any dimension *d* of the lattice, the voter dynamics always leads to one of the two possible consensus states: each agent with the same opinion s = 1 or s = −1.
  - The probability of reaching one or the other state depends on the initial state of the population.
  - Time needed for reaching the consensus state:
    - d = 1: $T_N \sim N^2$
    - d = 2: $T_N \sim N \ln N$
    - d > 2: $T_N \sim N$
  - For infinite systems: consensus is reached only if d ≤ 2

Extensions of the voter model

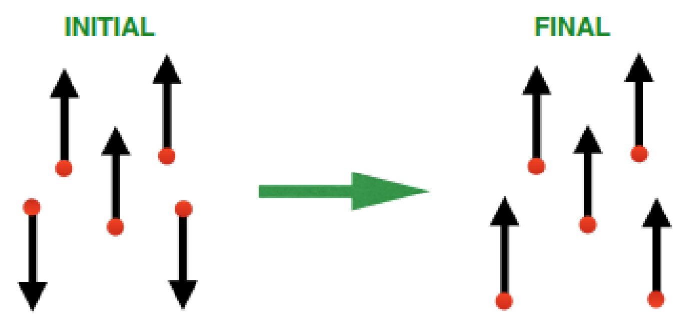
- Introduction of “zealots”: individuals who do not change their opinion
- Constrained voter model:
  - agents can be leftist, rightist, centralist;
  - Extremists do not talk to each other (discrete analogue of the bounded confidence model)
- Communication is based on various NW



Voter model on a small world network
https://www.youtube.com/watch?v=VmhSTdrsimk

Majority rule model

- Motivation: describing public debates
- (Galam, 2002)
- Definition:
  - Population of *N* agents
  - A fraction p_+ of agents has opinion +1
  - $p_- = 1 - p_+$ has opinion -1
  - Everybody can communicate with everybody else
  - At each interaction:
    - A group of *r* agents are selected at random ("discussion group")
    - Consequence of this interaction: each agents take the majority opinion inside the group
  - *r* is taken from a given distribution at each step
    - If *r* is odd: there is always a clear majority
    - If *r* is even: in case of tie: a bias is introduced in favor of one of the options (Inspired by the principle of "Social inertia" holding that people are reluctant to accept a reform if there is no clear majority in its favor)



Basic features of the MR model

- Original definition:
  - There is a *threshold fraction* p_c such that if $p^+_0 > p_c$, then all agents will have opinion +1 in the long run
  - Time needed for the consensus: $T_N \sim \log N$
    - If the group sizes r are odd: $p_c(r) = 1/2$ (due to the symmetry)
    - If they can be even too: $p_c < 1/2$, that is, the favored opinion will eventually win, even if it was originally in minority

- For fixed odd *r,* group size & mean field approach: analytically solvable for both finite N and for $N \to \infty$

- Many variants and modifications

Social impact theory

- Bibb Latané (psychologist), 1981:
- social impact: any influence on individual feelings, thoughts or behavior that is created from the real, implied or imagined presence or actions of others.

  („Collective” behavior)
- The impact of a social group on a subject depends on:
  - The number of individuals within the group
  - Their convicting power
  - Their distance from the subject (in an abstract space of personal relationships)
- Originally a cellular automata was introduced by Latané (1981) and later refined by Nowak et al (1990).

Social impact theory – the model

- A population of N individuals
- Each individual i is characterized by
  - an opinion $\sigma_i = \pm 1$
  - Persuasiveness p_i : the capability to convince someone **to change** opinion (a real value)
  - Supportiveness: s_i : the capability to convince someone **to keep** its opinion (a real value)
    (these are assumed to be random)
- The distance between agents i and j d_{ij},
- α>2 parameter defining the how fast the impact decreases with the distance

$$I_i = \underbrace{\left[\sum_{j=1}^{N} \frac{p_j}{d_{ij}^{\alpha}} (1 - \sigma_i \sigma_j) \right]}_{\text{Persuasive impact (to change)}} - \underbrace{\left[\sum_{j=1}^{N} \frac{s_j}{d_{ij}^{\alpha}} (1 + \sigma_i \sigma_j) \right]}_{\text{supportive impact (to keep opinion)}}$$

Persuasive impact (to change) supportive impact (to keep opinion)

Opinion dynamics: $$\sigma_i(t+1) = -sgn[\sigma_i(t) I_i(t) + h_i]$$

h_i : personal preference, originating from other sources (e.g. mass media)

a spin flips if the pressure in favor of the opinion change overcomes the pressure to keep the current opinion ($I_i > 0$ for vanishing h_i)

General behavior of the social impact model

- In the absence of individual fields (personal preferences):
  - the dynamics leads to the dominance of one opinion over the other, but not to complete consensus.
  - If the initial magnetization is about zero:
    - large majority of spins in the same opinion with stable domains of spins in the minority opinion state.

- In the presence of individual fields:
  - these minority domains become metastable: they remain stationary for a very long time, then they suddenly shrink to smaller clusters, which again persist for a very long time, before shrinking again, and so on ("staircase dynamics").

- Many modification / extensions:
  - Learning
  - Presence of a strong leader
  - Etc.

Schweitzer and Holyst included: (2000)

- Memory: reflecting past experience
- A finite velocity for the exchange of information between agents
- A physical space, where agents move.

Continuous opinions

- In many cases more realistic
- Requires different framework
  - Concepts like “majority” or “opinion equality” don’t work
  - Has a different ‘history’
- First studies (end of 1970’s and 80’s):
  - Aimed to study the conditions under which a panel of experts would reach a common decision (“consensus”)
  - By applied mathematicians
- Typically:
  - **Initial state**: population of N agents with randomly assigned opinions, represented by real values within some interval.

    discrete op. dyn. ↔ all agents start with different opinions
  - **Possible scenarios**: more complex
    - Opinion clusters emerging in the final stationary state:
      - one cluster: consensus,
      - two clusters: polarization
      - more clusters: fragmentation

Bounded confidence (BC) models

- **In principle:** each agent can interact with every other
- **In practice:** (often) there is a real discussion only if the opinions are sufficiently close:

bounded confidence

- **In the literature:** introducing a real number ε: "*uncertainty*" or "*tolerance*", such that:
- An agent with opinion *x*, only interacts with those whose opinion lies in the interval]x−ε, x+ε[
- („Homophily")

Deffuant model

- population of N agents
- nodes of a graph: agents may discuss with each other if they are connected.
- Initially: each agent *i* is given an opinion x_i randomly chosen from the interval [0, 1].
- Dynamics:
  - random binary encounters, i.e., at each time step, a randomly selected agent discusses with one of its neighbors, also chosen at random.
  - Let *i* and *j* be the pair of interacting agents at time *t*, with opinions $x_i(t)$ and $x_j(t)$
    - if the difference of the opinions $x_i(t)$ and $x_j(t)$ exceeds the threshold ε, nothing happens
    - If $|x_i(t) - x_j(t)| < \varepsilon$, then
    - μ: convergence param. (μ in [0, 1/2])

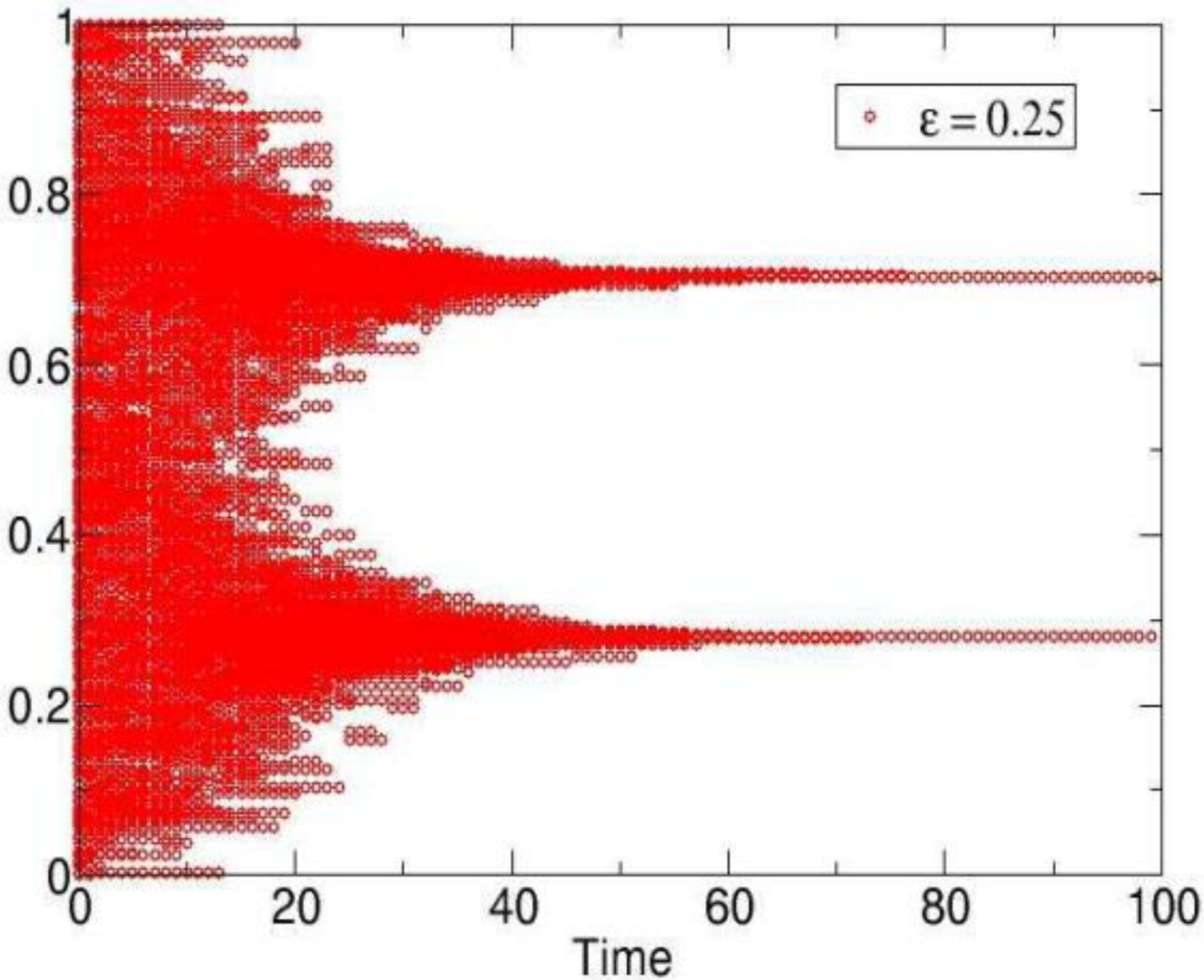
$$x_i(t+1) = x_i(t) + \mu[x_j(t) - x_i(t)]$$

$$x_j(t+1) = x_j(t) + \mu[x_i(t) - x_j(t)]$$

Behavior of the Deffuant model

- For any value of ε and μ, the average opinion of the agents' pair is the same before and after the interaction → the global average opinion (1/2) of the population is invariant
- Patches appear with increasing density of agents
- Once each cluster is sufficiently far from the others (the difference of opinions in distinct clusters exceeds the threshold):
  - only agents *inside* the same cluster interact
  - the dynamics leads to the convergence of the opinions of all agents in the cluster
- In general:
  - the *number and size of the clusters* depend on the threshold ε
    (if ε is small, more clusters emerge)
  - the parameter μ affects the convergence time
  - (when μ is small, the final cluster configuration also depends on μ)

Behavior of the Deffuant model



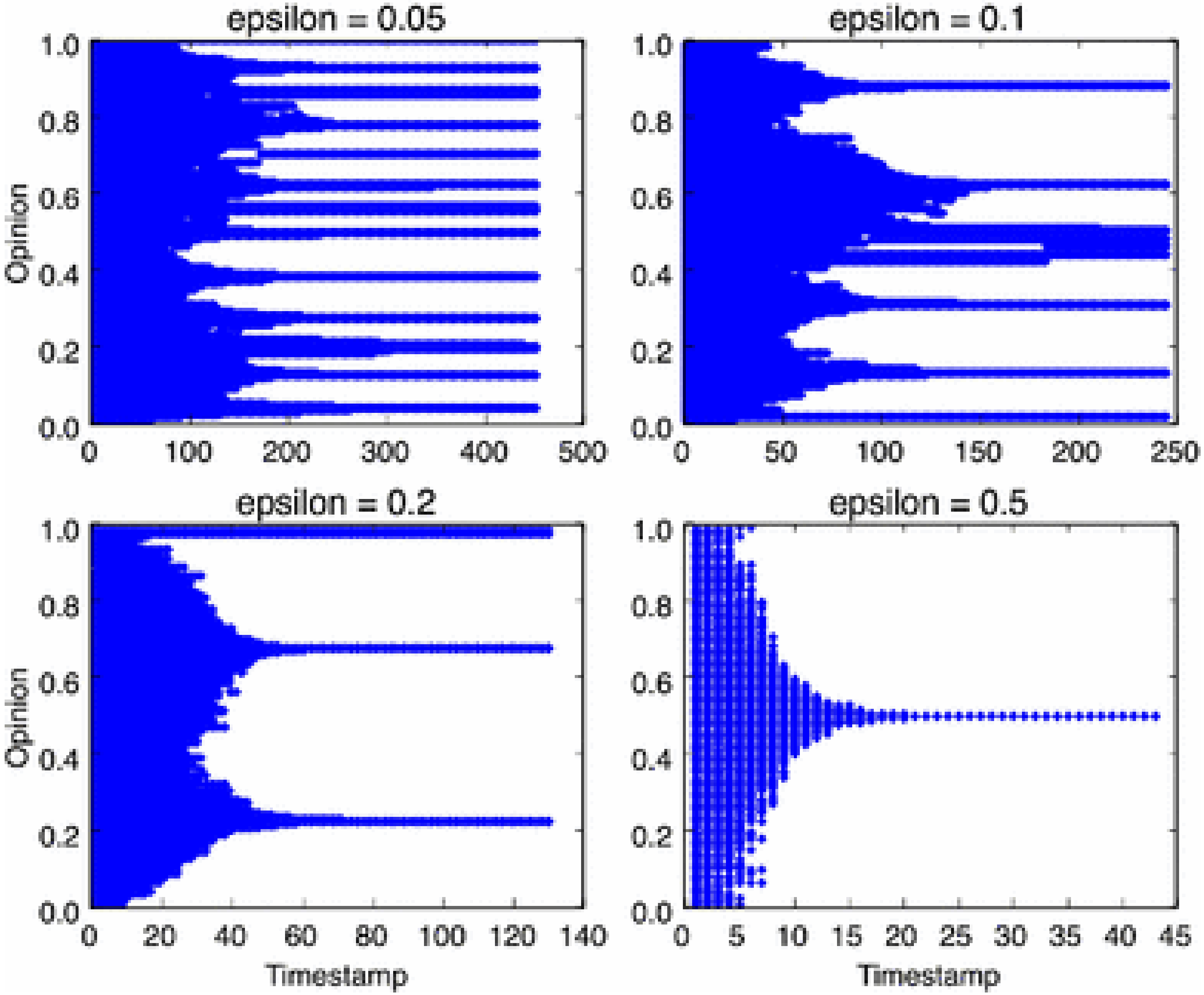
Opinion profile of a population of *N*=500 agents during its time evolution,

ε = 0.25.

The population is fully mixed, i.e., everyone may interact with everybody else.

The dynamics leads to a polarization of the population in two factions.

Behavior of the Deffuant model



Hegselmann-Krause (HK) model

- Hegselmann and Krause, 2002

- Similarities with the Deffuant model:
  - Opinions take real values in an interval, say [0, 1]
  - An agent *i* (with opinion x_i), interacts with neighboring agents whose opinions lie in the range $[x_i-\varepsilon, x_i+\varepsilon]$
- Difference: update rule

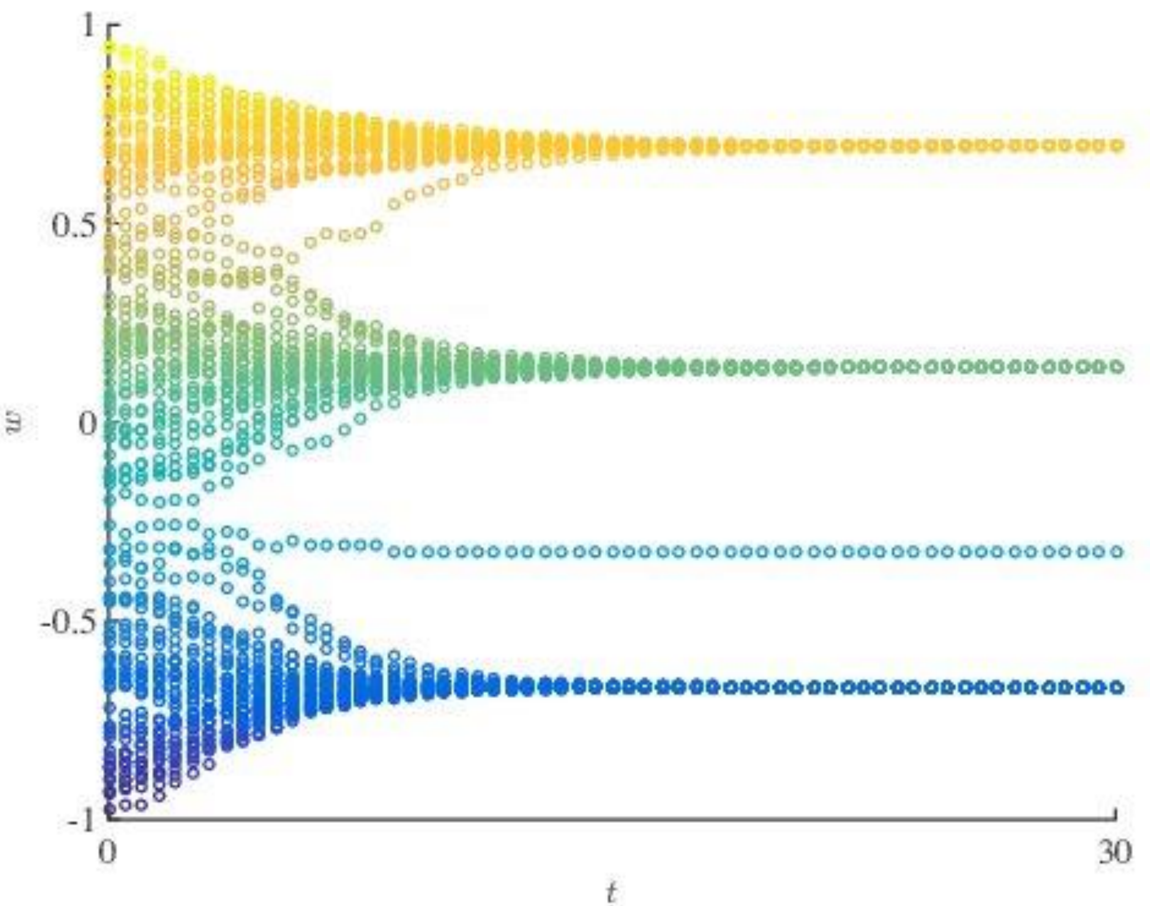
  - An agent *i* does not interact with *one* of its compatible neighbors (like in Deffuant), but with *all* its compatible neighbors at once.
  - intended to describe *formal meetings*

$$x_i(t+1) = \frac{\sum_{j:|x_i(t)-x_j(t)|<\epsilon} a_{ij} x_j(t)}{\sum_{j:|x_i(t)-x_j(t)|<\epsilon} a_{ij}}$$

- a_{ij}: elements of an adjacency matrix describing the communication network.
- Agent *i* takes the average opinion of its compatible neighbors.

Behavior of the Hegselmann-Krause model

- fully determined by the uncertainty ε
- Need lot of computation power (due to the average calculation)



The dynamics develops similarly to the Deffuant model:

- Leads to the same pattern of stationary states, with the number of final opinion clusters decreasing if ε increases.
- for $\varepsilon > \varepsilon_c$ (a threshold) there can be only one cluster