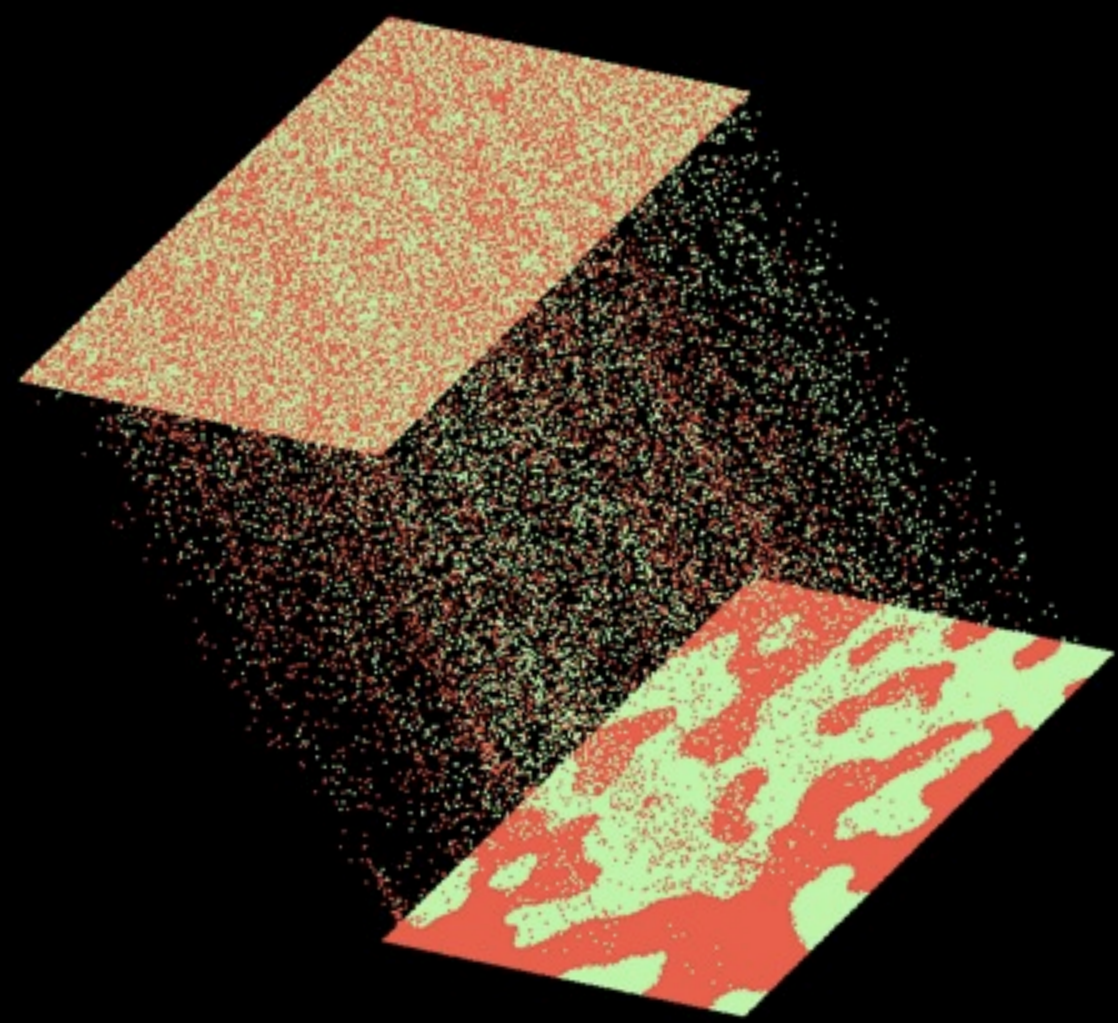
From randomness to segregation

Schelling segregation, Ising models and network cascades



George Barmpalias (Chinese Academy of Sciences)
Richard Elwes (University of Leeds)
Andy Lewis-Pye (London School of Economics)

More than two millennia ago, Greek philosopher

Εμπεδοκλής

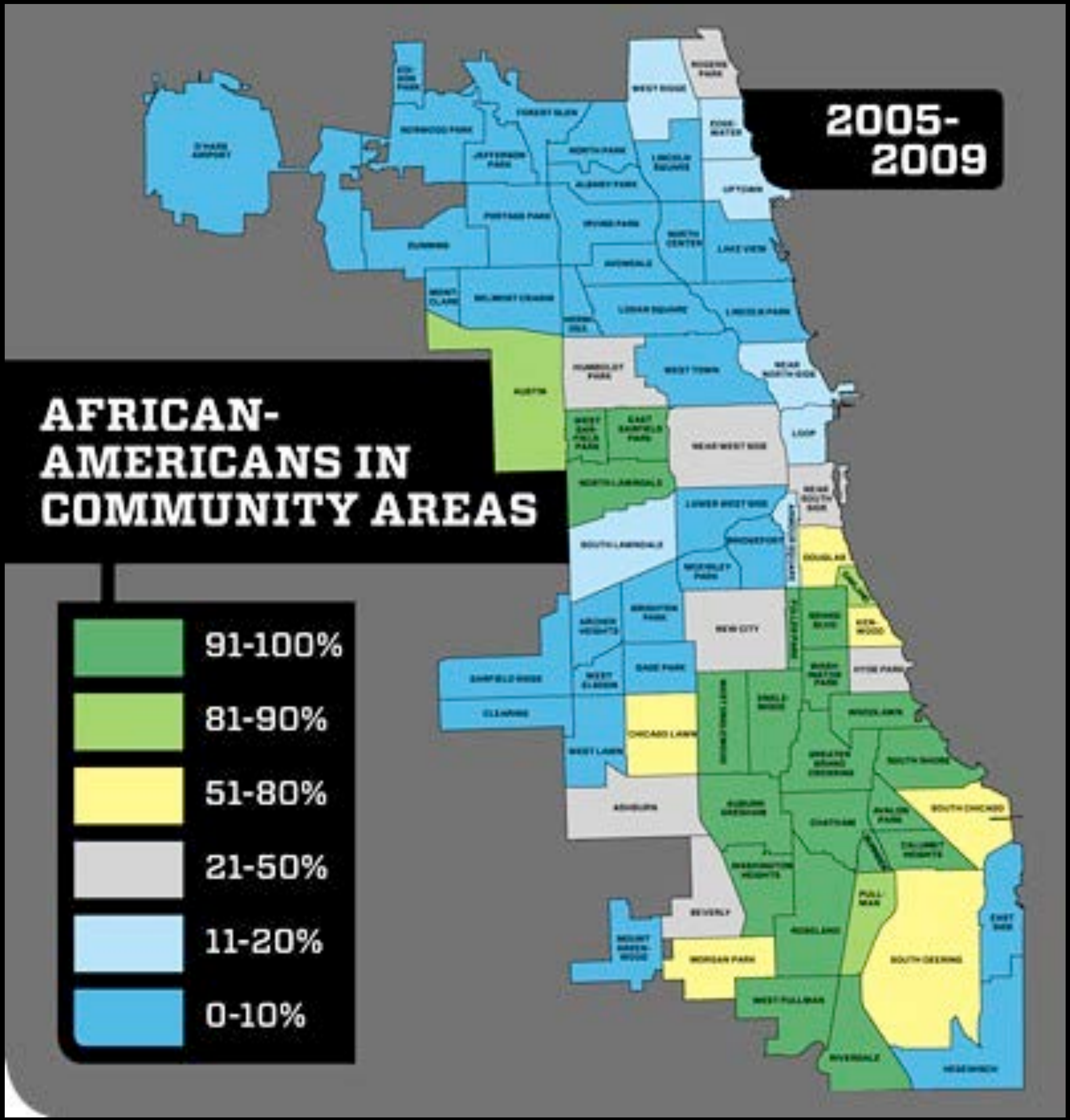
observed that humans are like liquids.

Some mix easily like wine and water, and some do not, like oil and water.

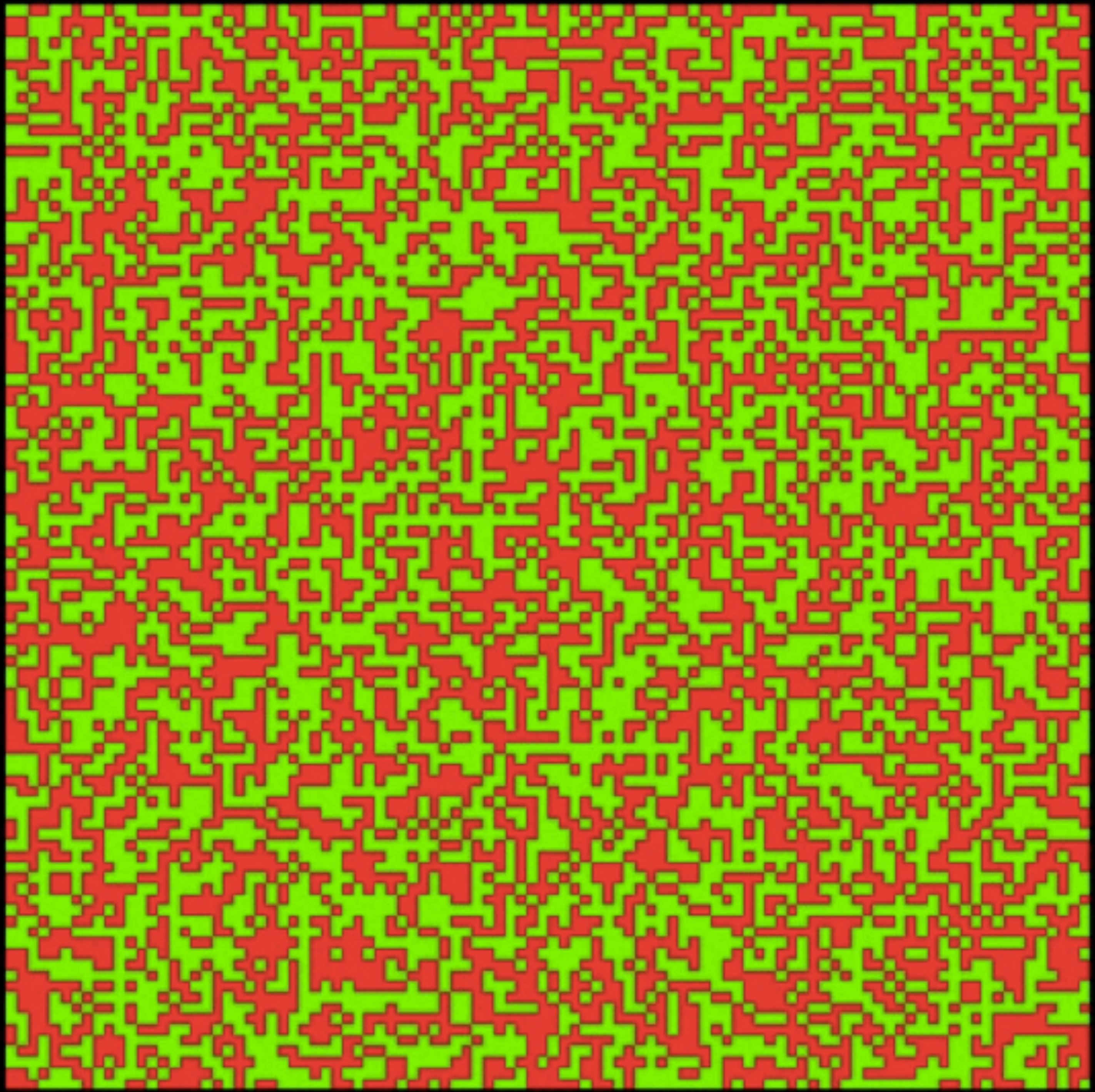
In the 60s **Thomas Schelling** transformed this idea into a quantitative model and studied it.

2005 Nobel prize for "having enhanced our understanding of conflict and cooperation through game theory analysis"

The Schelling model of segregation describes the formation of homogeneous communities in multicultural cities.



In the 2D Schelling model we start with a randomly colored grid

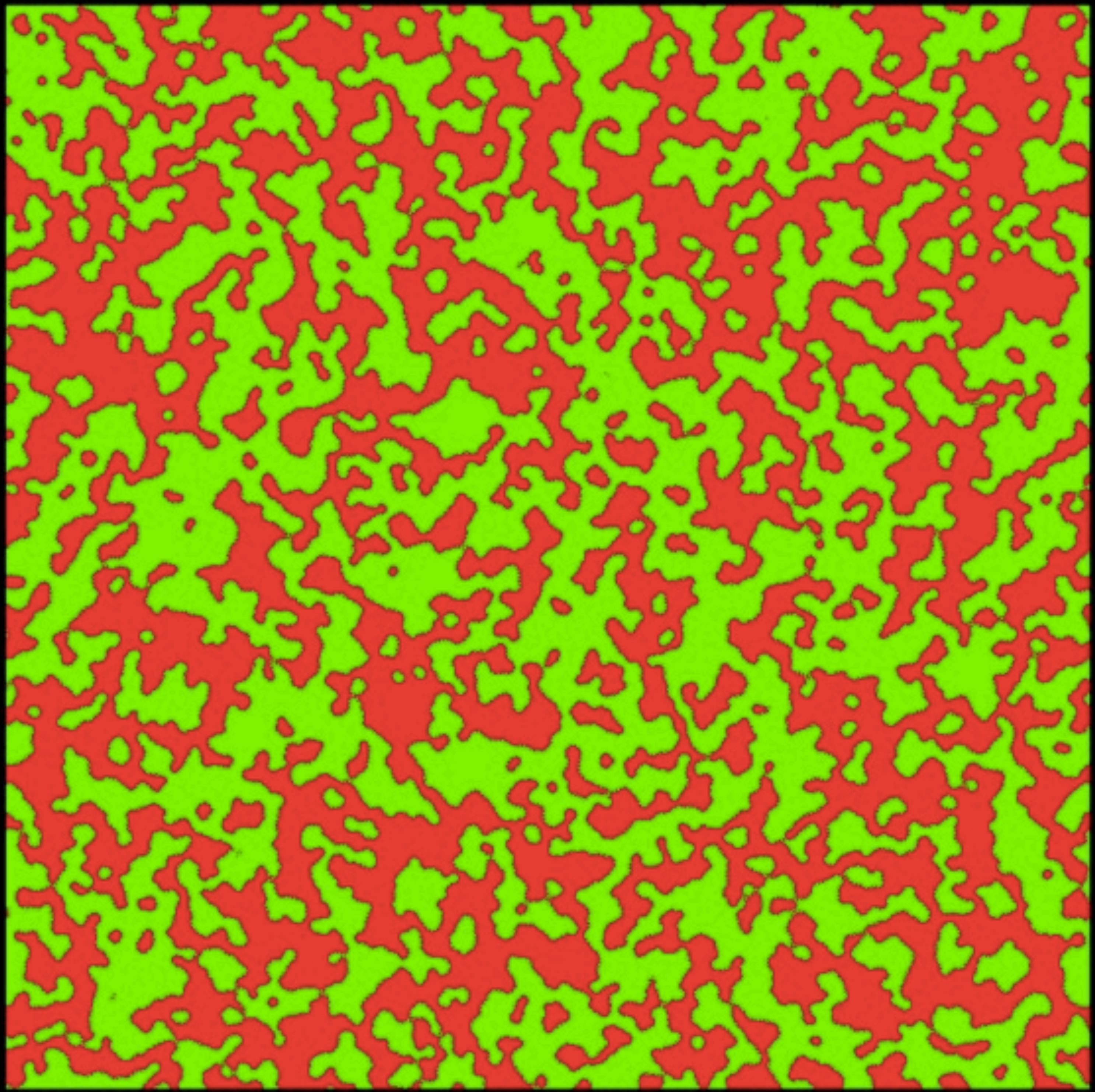


Parameters:

- Population n
- Neighborhood radius w
- Intolerance τ
- Distribution ρ

Happy/Unhappy

The swapping process often results in segregated regions.

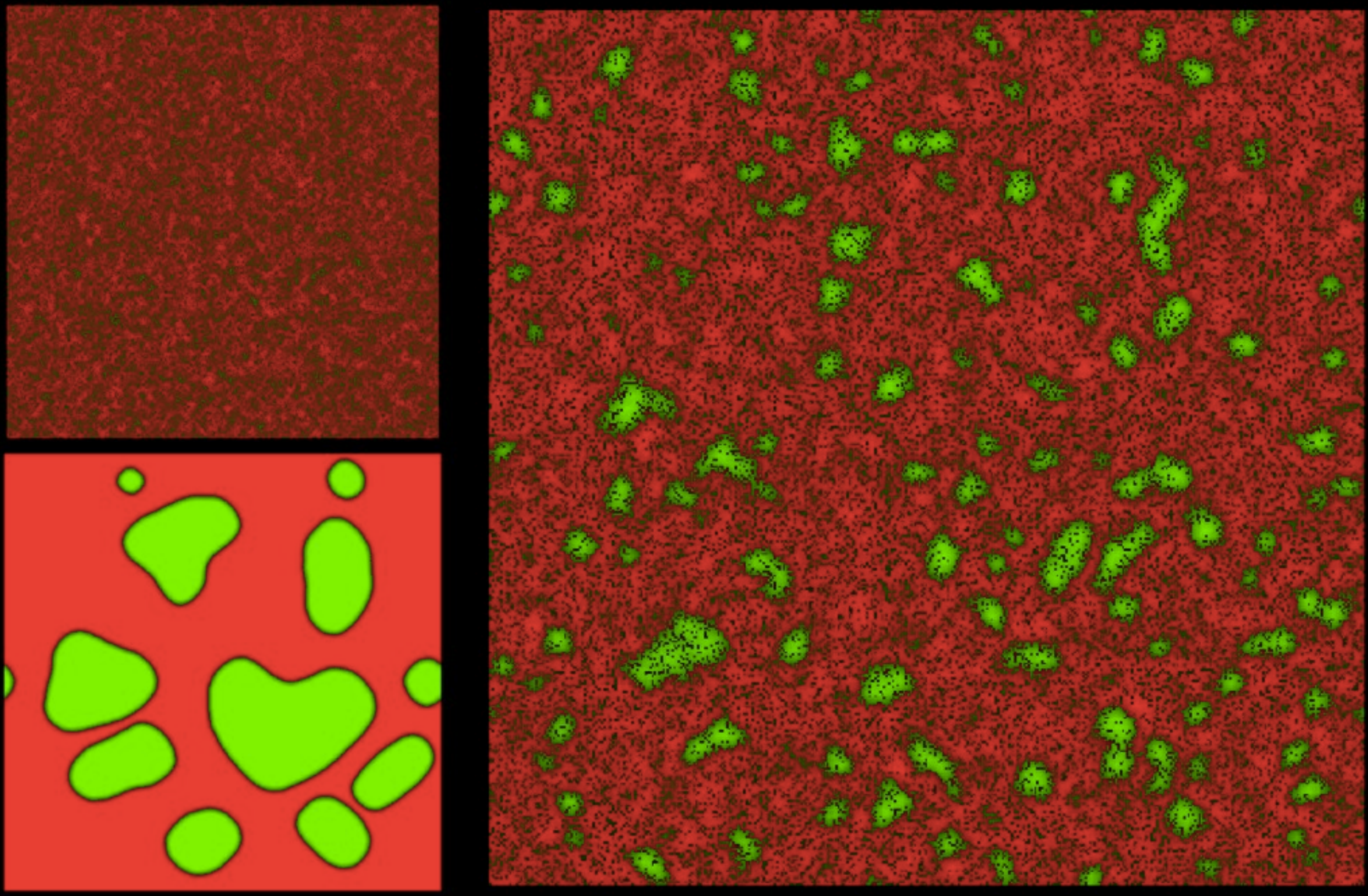


Problem:

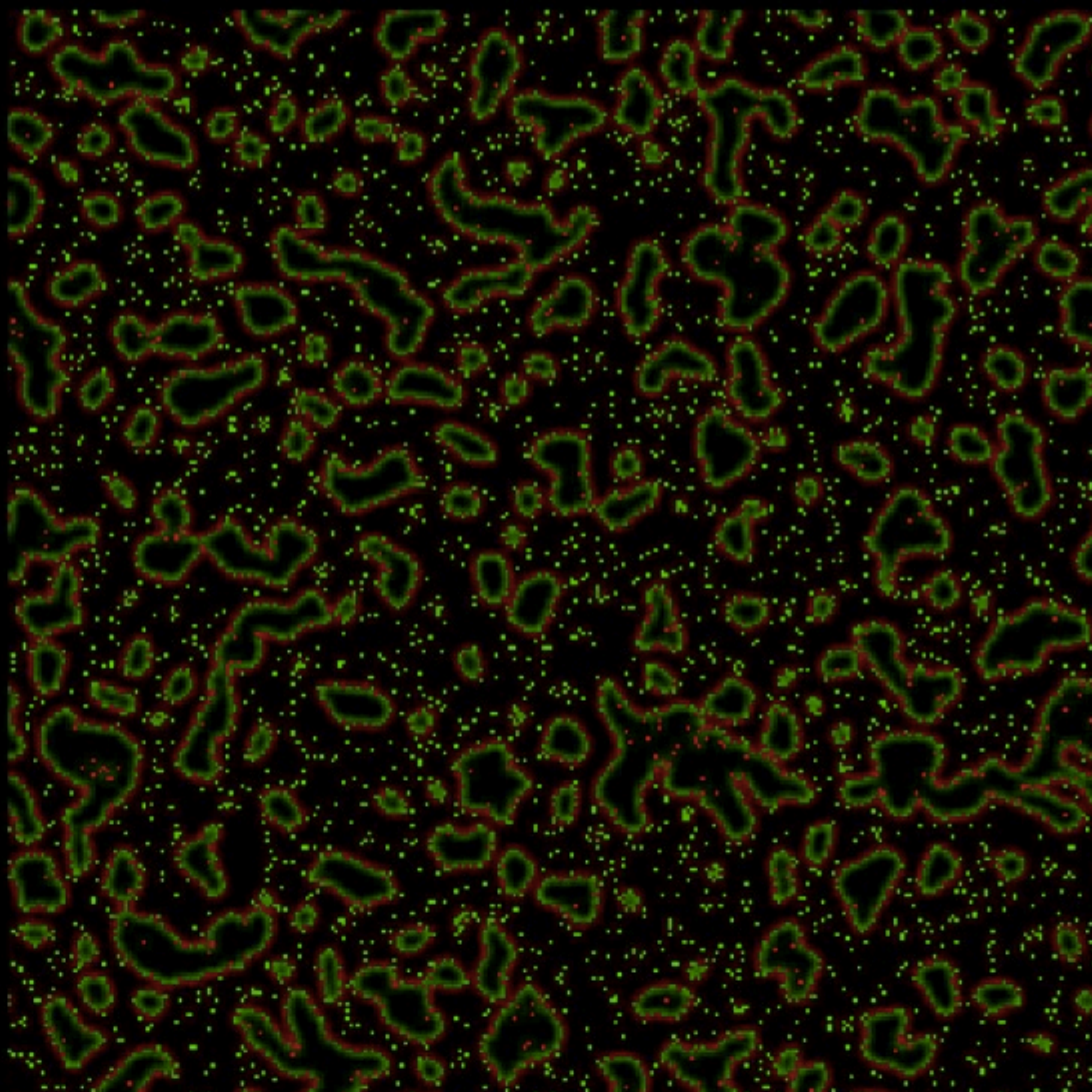
Given the parameters, predict:

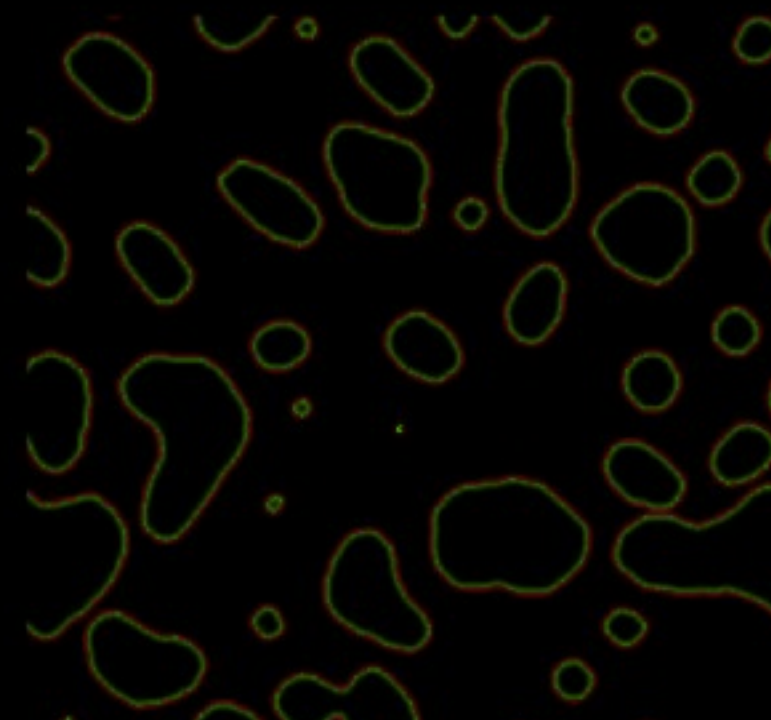
- ★ Extent of segregation
- ★ Expected time of process
- ★ Analyze the process

Unhappiness is incentive to move

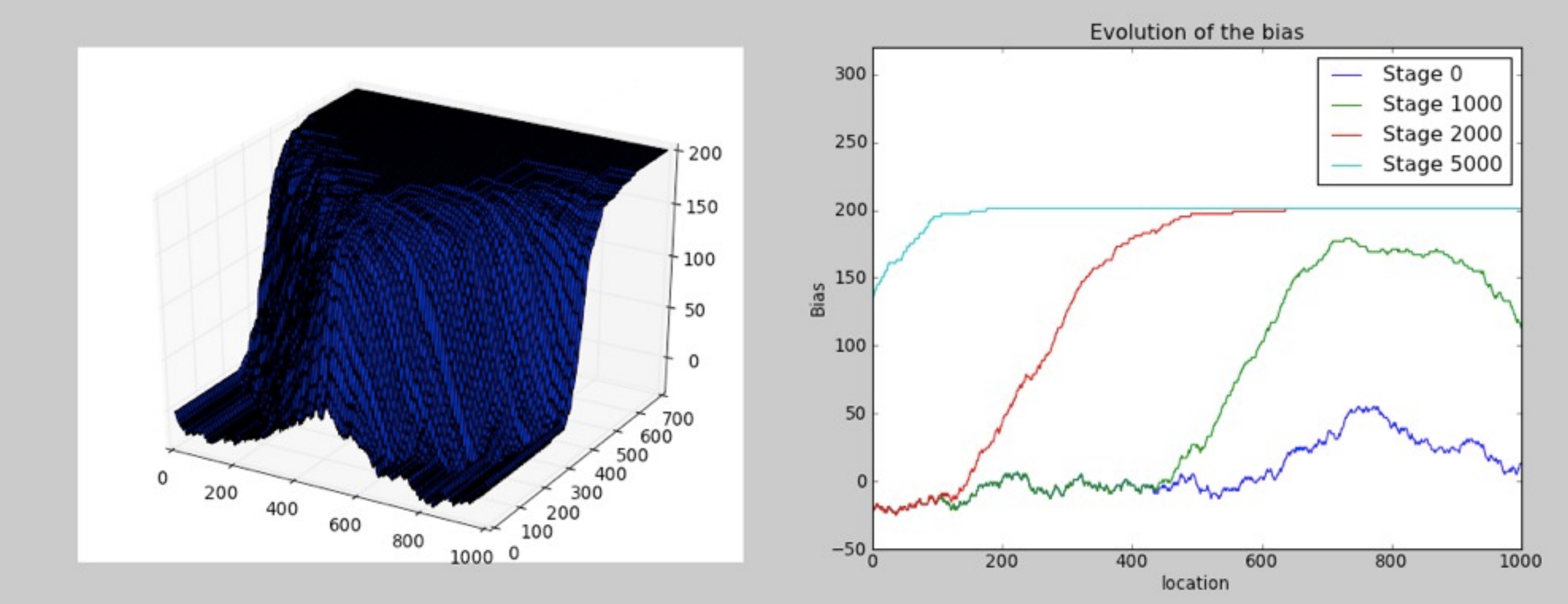


Unhappiness is incentive to move

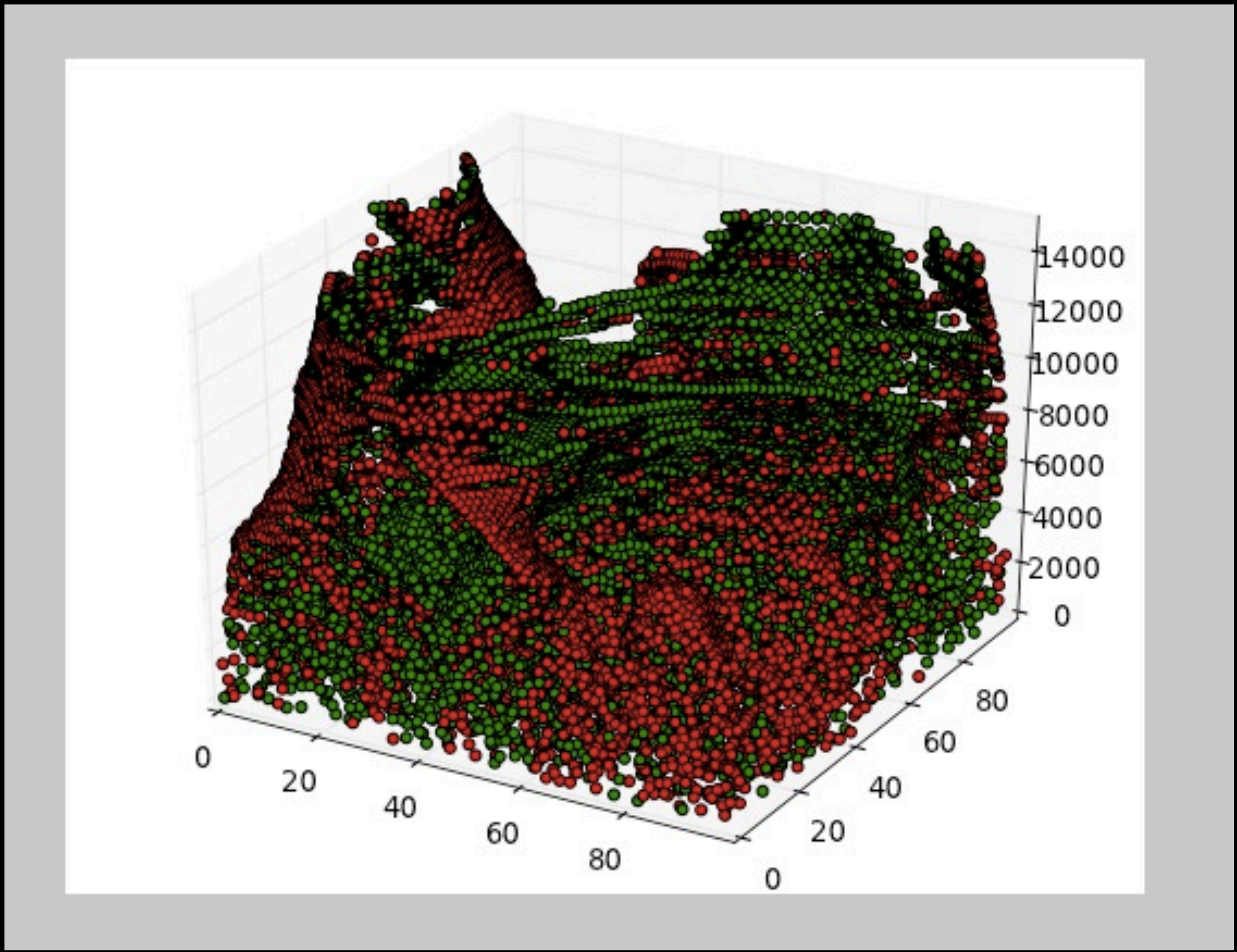




Unhappiness is incentive to move

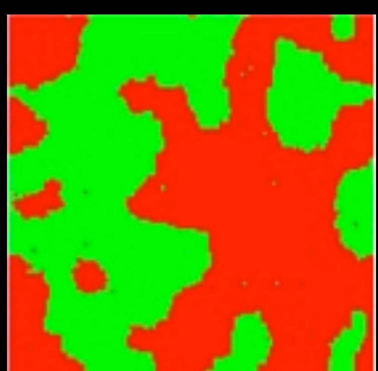


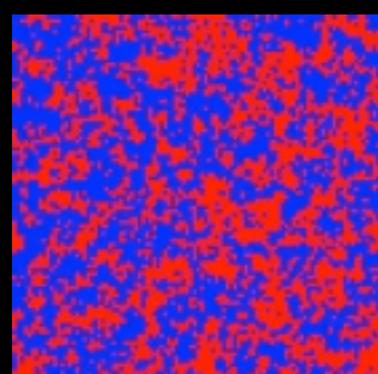
Unhappiness is incentive to move

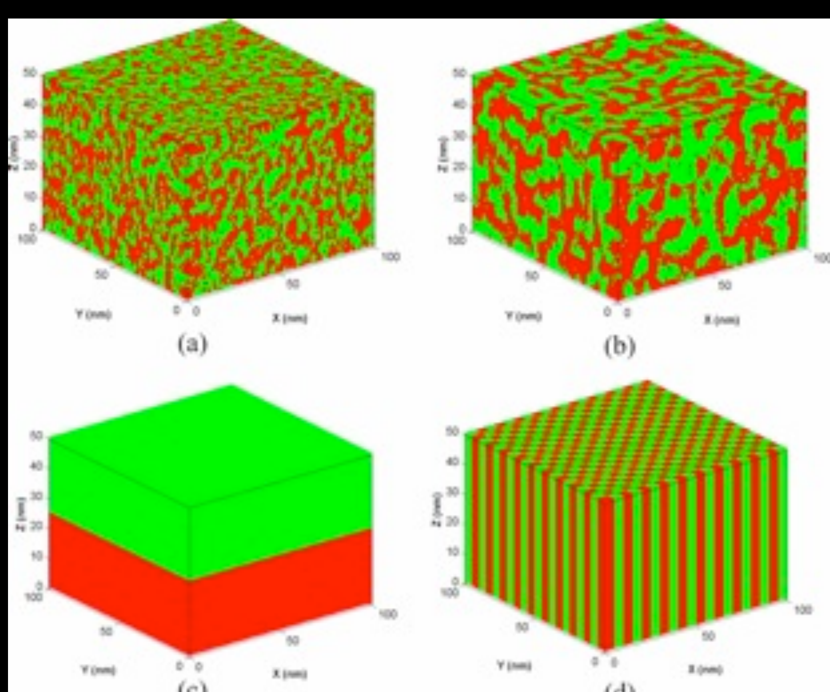


Schelling worked in a socio-economic context, unaware of the study of similar effects by

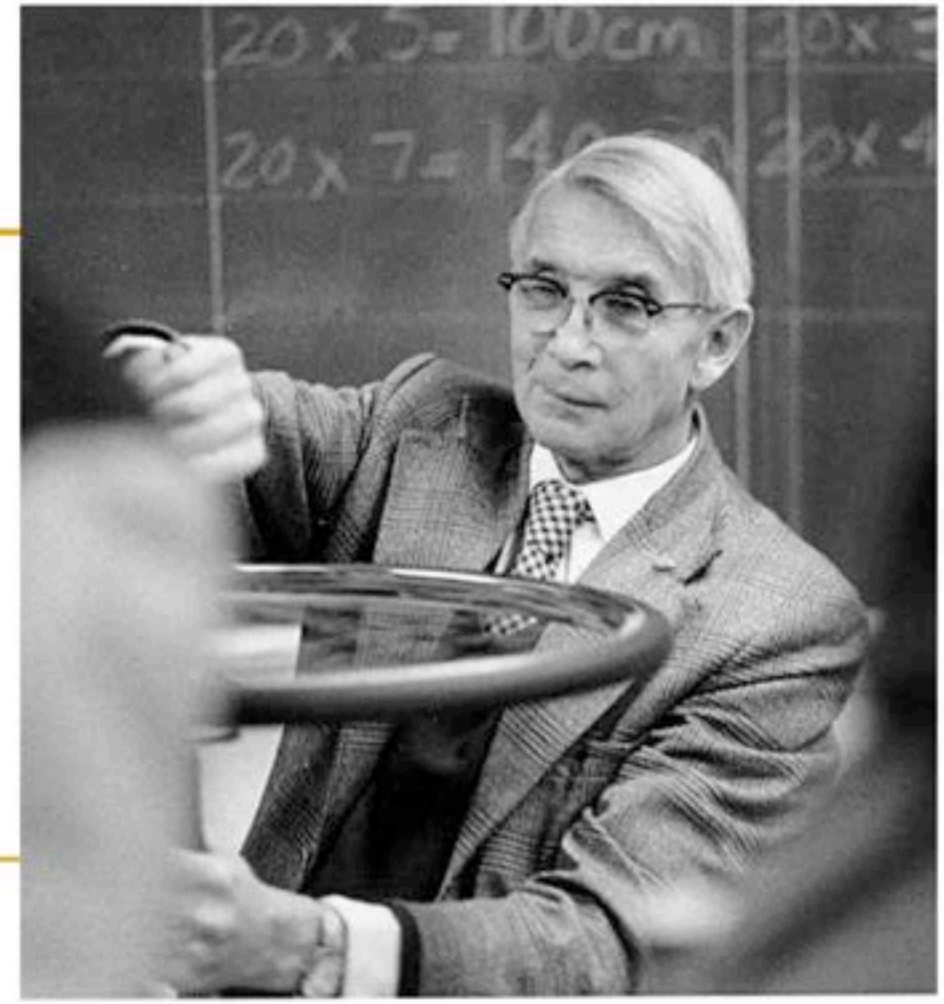
- Physicists (Ising model, 1925)
- Biologists (Morphogenesis)







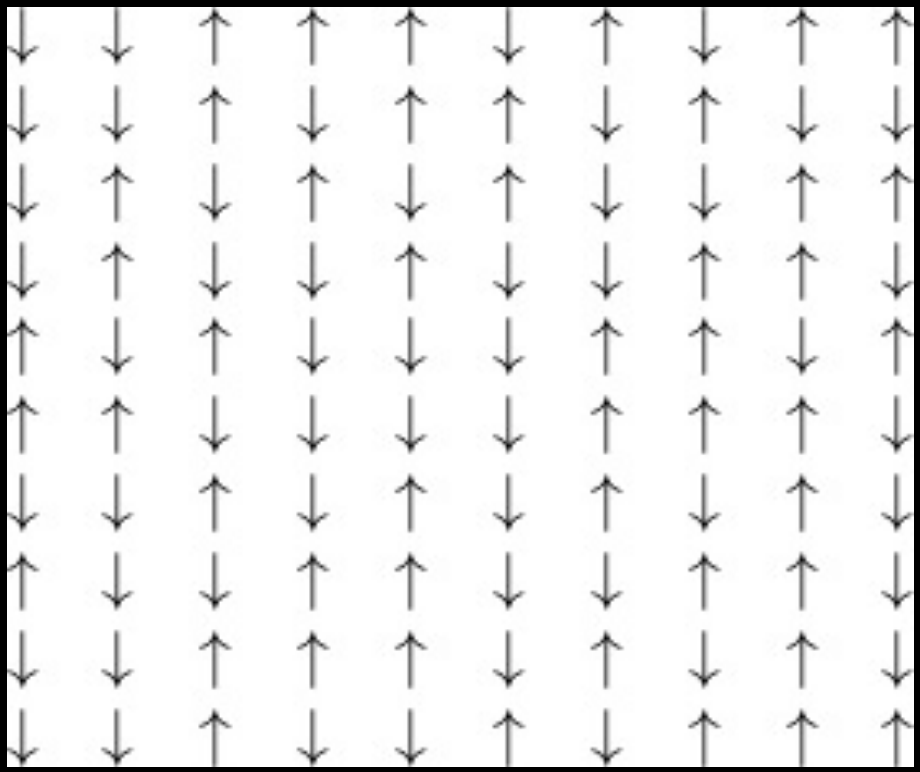
Ising Model



Dr. Ernst Ising

May 10, 1900 – May 11, 1998

- Studied ferromagnetism with his model
- Looked for phase transitions when varying temperature
- Concluded that in the 1D case no phase transitions exist
- Wrongly argued that same is true in higher dimensions



- 1924: PhD thesis with the Ising model
- 1933: Barred from teaching and research
- 1934: Teacher at a Jewish school
- 1938: School destroyed by Nazis
- 1939: Fled to Luxembourg
- 1947: Moved to the US

Today the Ising model is used to address problems in

- Statistical mechanics
- Ferromagnetism, phase transitions
- Neural networks
- Protein folding
- Biological membranes
- Social behavior

About 800 papers on the Ising model are published every year

Schelling's work is regarded as the

ARCHETYPE OF AGENT-BASED MODELING

in economics

Since the 60s numerous works have been produced acknowledging the interdisciplinarity of the model.

- Physics: simulations and statistical mechanics (Boltzmann distribution)
- Computer/Network science: Dynamical systems, combinatorics
- Social science: Evolutionary game theory

As a Dynamical system

- It is an irregular Markov chain
- Not reversible, doesn't satisfy "detailed balance"
- It has many stationary distributions (state explosion problem)

All studies up to recently introduced noise to the system in order to overcome these problems.

Occasionally, agents make decisions that are detrimental to their utility function (with small probability).

Some recent articles

An Analysis of One-Dimensional Schelling Segregation (Brandt, Immorlica, Kamath, Kleinberg: STOC 2012

Ising, Schelling and Self-Organising Segregation (Stauffer, Solomon: Eur. Phys. J. 2007)

Individual strategy and social structure (Young, Monograph 1998)

Self-organizing, two-temperature Ising model describing human segregation (Odor, Int. J. Modern Phys. 2008)

A physical analogue of the Schelling model (Vincovic, Kirman: PNAS 2006)

Some recent articles

An Analysis of One-Dimensional Schelling Segregation (Brandt, Immorlica, Kamath, Kleinberg: STOC 2012

Ising, Schelling and Self-Organising Segregation (Stauffer, Solomon: Eur. Phys. J. 2007)

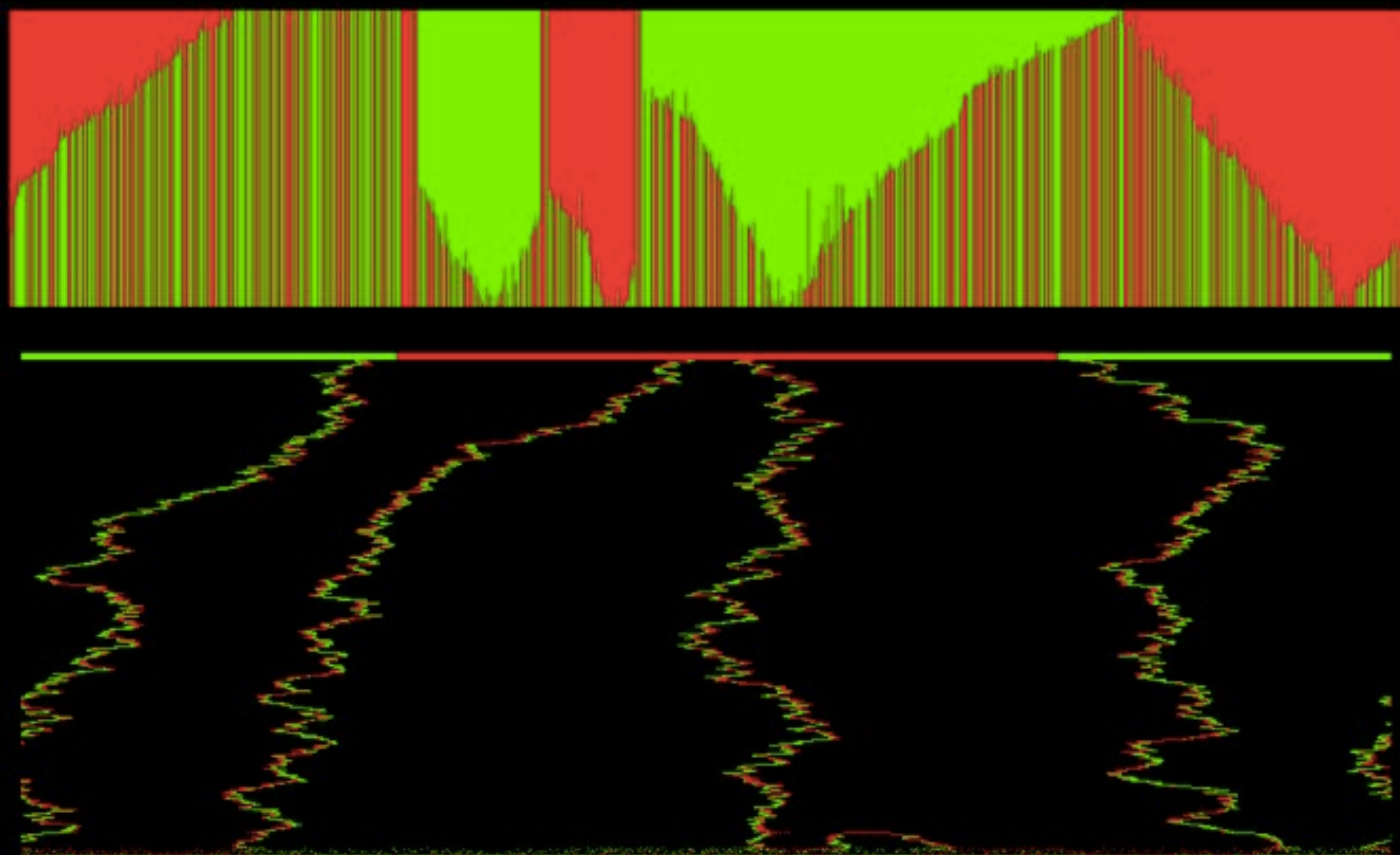
Individual strategy and social structure (Young, Monograph 1998)

Self-organizing, two-temperature Ising model describing human segregation (Odor, Int. J. Modern Phys. 2008)

A physical analogue of the Schelling model (Vincovic, Kirman: PNAS 2006)

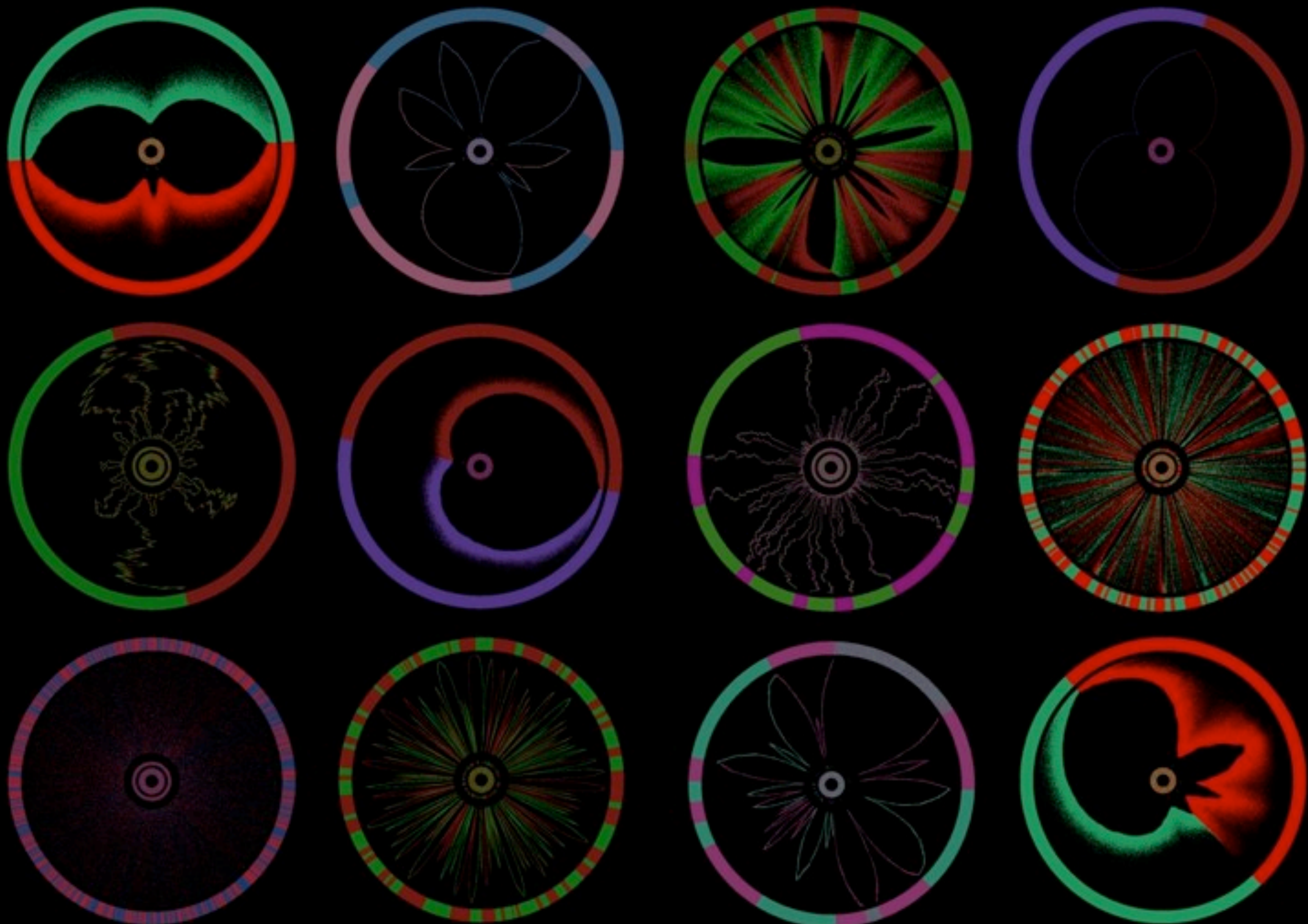
1D Schelling model

Individuals are arranged on a line



or a circle....

Flowers of Segregation



Winner of the info-graphics category in this year's
Picturing Science competition of the Royal Society

A word about simulations...

Schelling's model features in many agent-based modeling tools

Repast, Net-logo, online java applets, ...

However these are very slow.

Fast simulations require good algorithms and low level coding.

We did our simulations in C++

- Fast graphics with OpenGL
- Dynamic arrays cost time
- Static arrays require handling empty entries
- Compiler optimization makes a huge difference!

Some recent articles

An Analysis of One-Dimensional Schelling Segregation (Brandt, Immorlica, Kamath, Kleinberg: STOC 2012

Ising, Schelling and Self-Organising Segregation (Stauffer, Solomon: Eur. Phys. J. 2007)

Individual strategy and social structure (Young, Monograph 1998)

Self-organizing, two-temperature Ising model describing human segregation (Odor, Int. J. Modern Phys. 2008)

A physical analogue of the Schelling model (Vincovic, Kirman: PNAS 2006)

Brandt, Immorlica, Kamath, Kleinberg

They did the first study of the unperturbed model

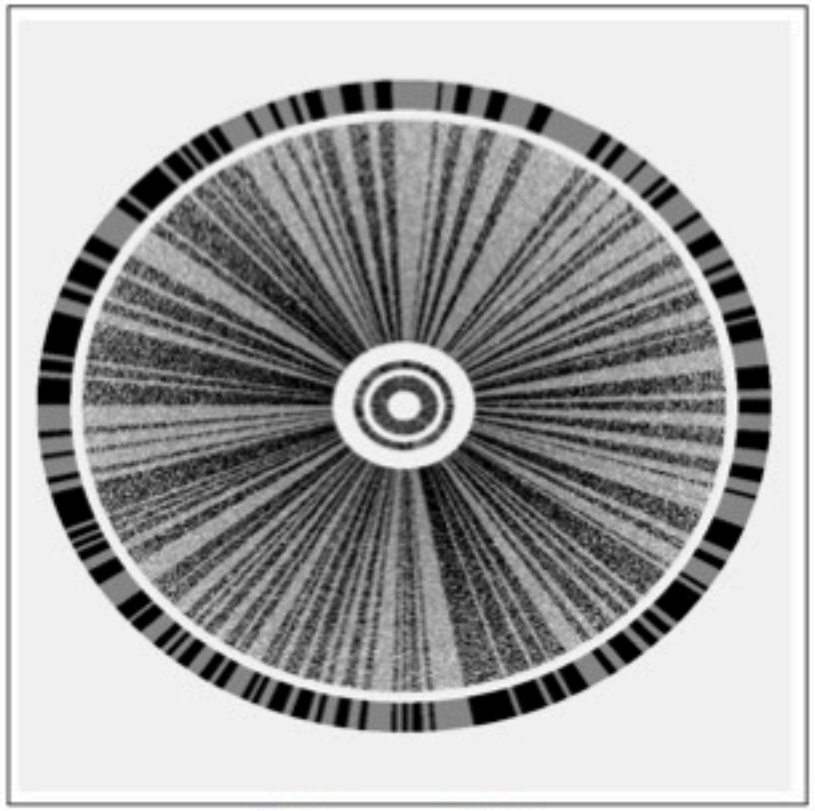
Result:

If tolerance is 1/2 and initial distribution is uniform, segregated blocks have length polynomial in the neighborhood size

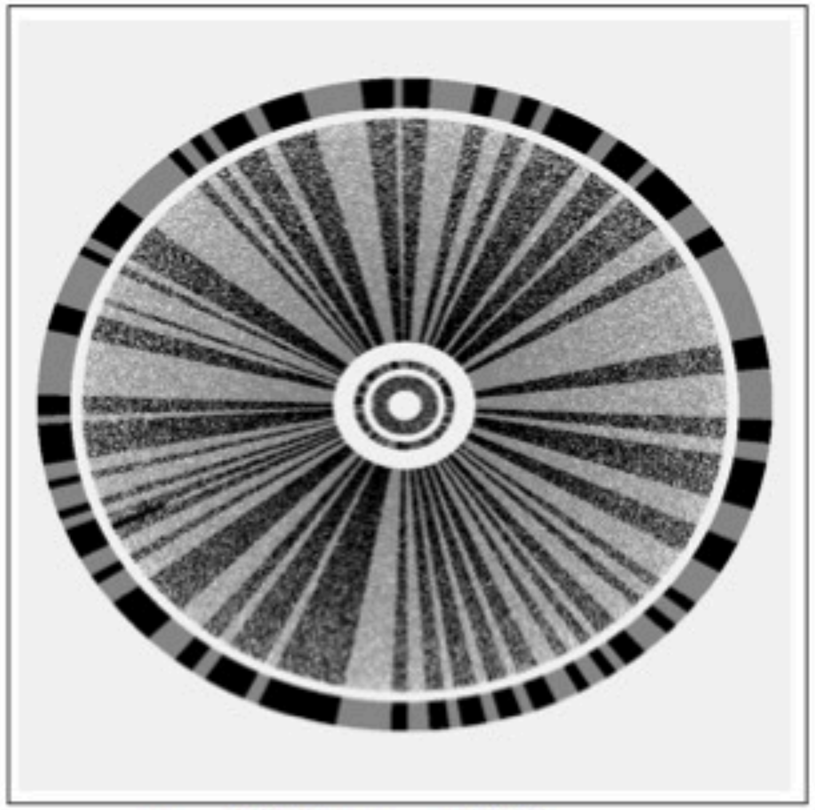
- The average run length in the final configuration is $O(w^2)$
- There is $c>0$ such that the probability that a random node belongs to a run of length $>kw^2$ is less than c^k.

Tools and methods

- ✓ Central limit theorem
- ✓ Wormald differential equation technique
- ✓ Symmetry arguments
- ✓ Combinatorics

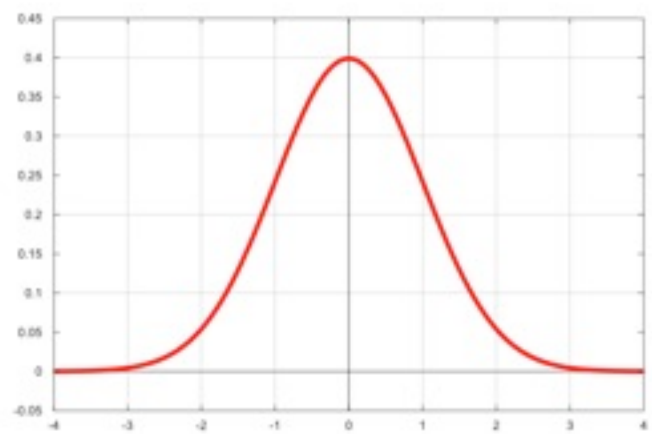


$w = 1500, \; \tau = 0.50$

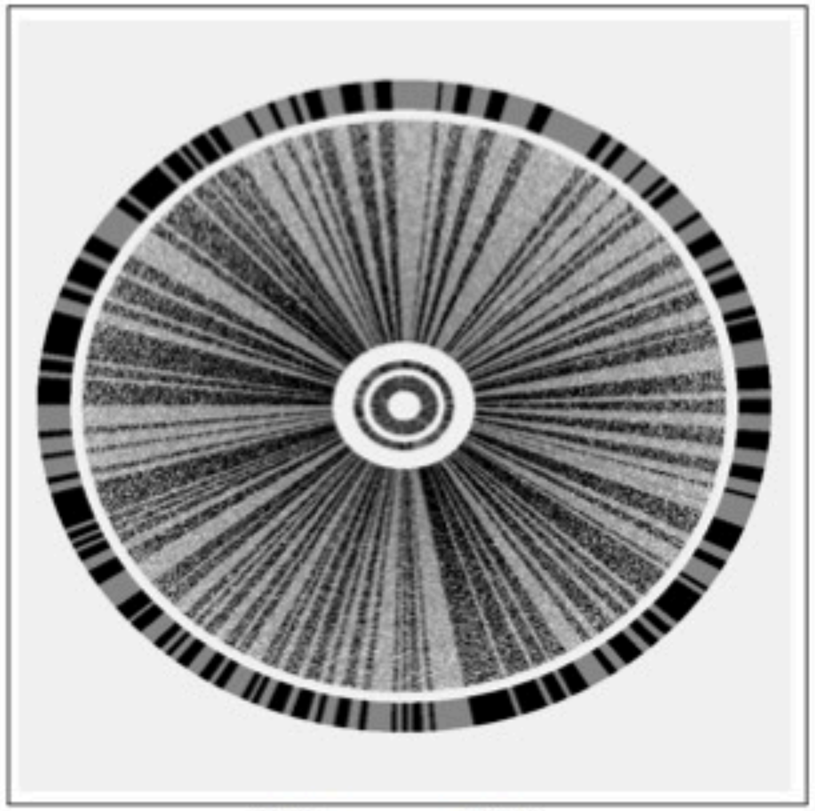


$w = 3000, \; \tau = 0.50$

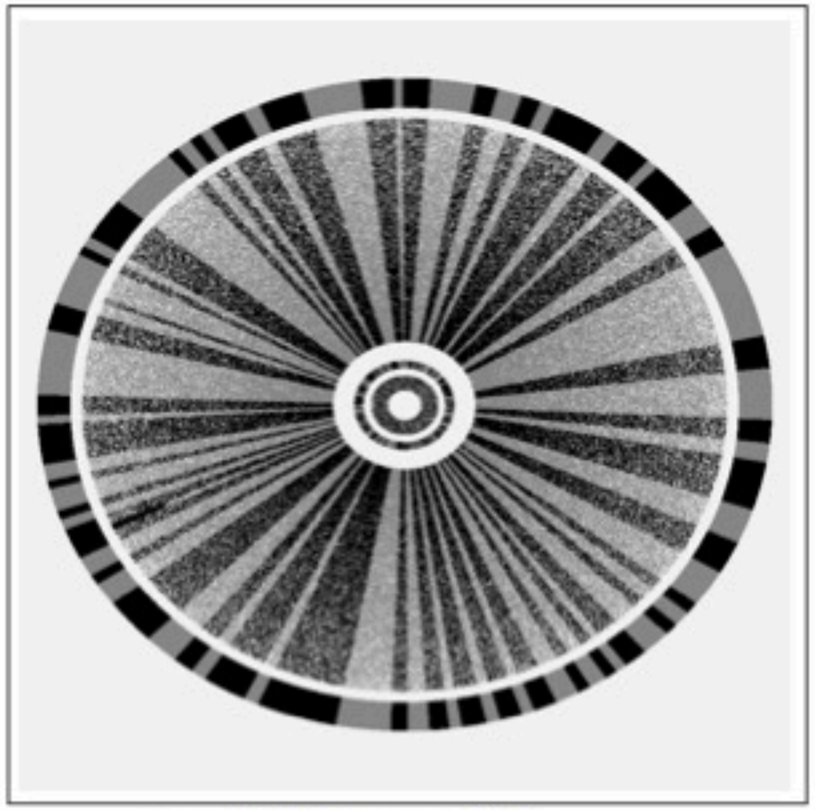
$$\sqrt{n}\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \mu\right) \xrightarrow{d} N(0,\ \sigma^2).$$



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$



$w = 1500, \ \tau = 0.50$



$w = 3000, \ \tau = 0.50$

Process

Many unhappy of both colors near to each other

↓

Incubators in every block of length O(w)

↓

Incubators to firewalls with positive probability

Irregularities

Different numbers of unhappy green and red nodes

↓

Not every unhappy node is equally likely to be chosen as part of a swapping pair

↓

Simple model: switching instead of swapping

Simple model is closer to the Ising model and other network cascading processes that model spread of viruses etc.

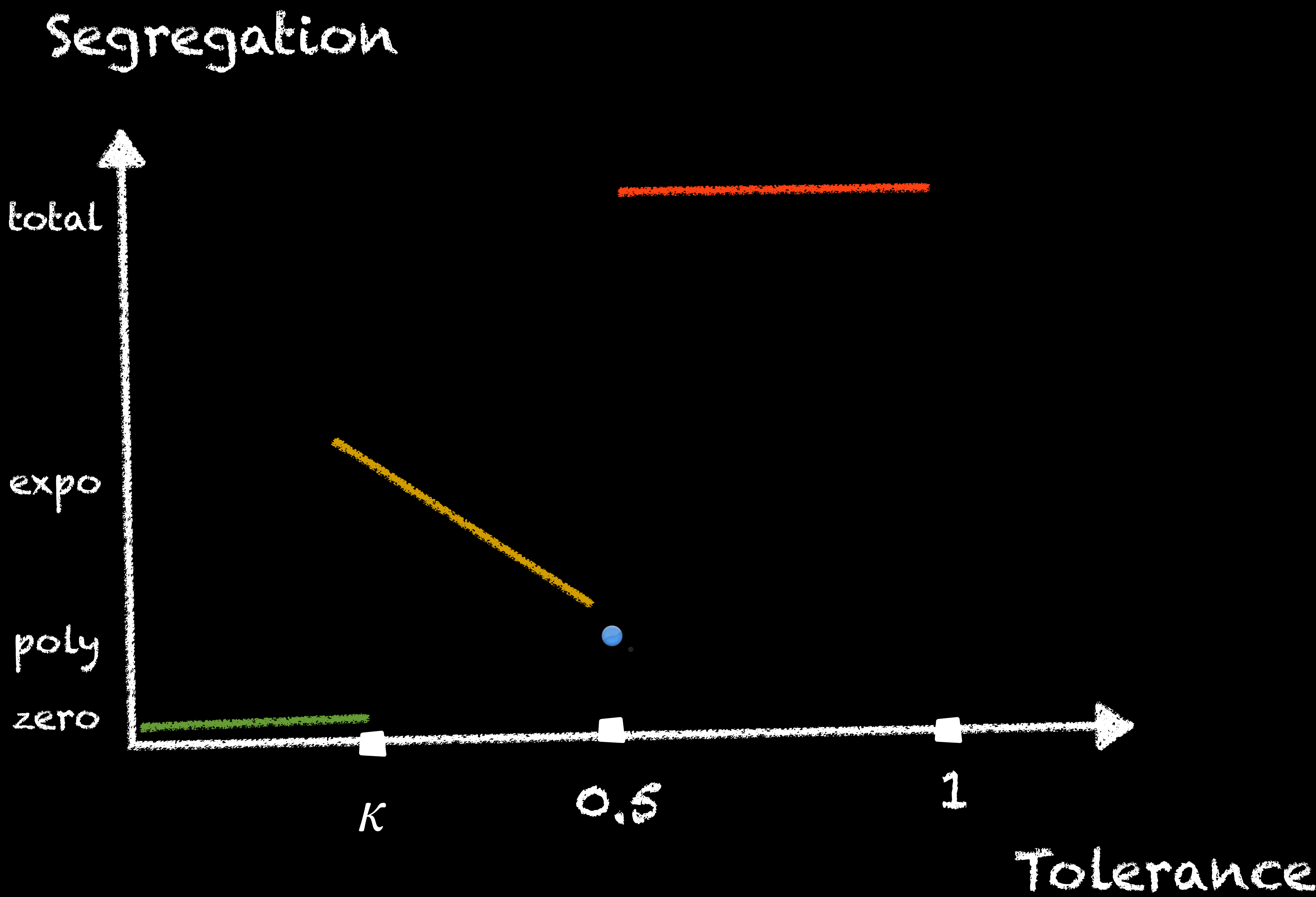
We provide an analysis of the 1D model for any tolerance

Summary

- If $\tau < 0.353$ no segregation
- If $0.353 < \tau < 1/2$ exponential segregation
- If $\tau = 1/2$ polynomial segregation (Brant et. al.)
- If $\tau > 1/2$ total segregation

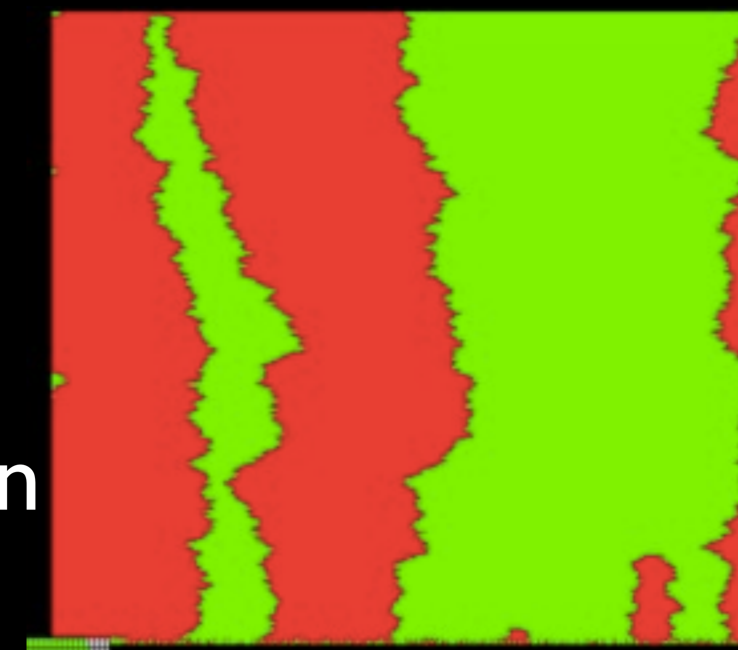
Paradox: in the interval [0.353, 1/2] increased tolerance leads to increased segregation

(Schelling: Micromotives and Macrobehaviour)


Segregation
total
expo
poly
zero
κ
0.5
1
Tolerance

Total segregation for τ>1/2 almost surely

- Unhappiness increases
- Markov chain with an absorbing state
- From any configuration to total segregation
- Unhappy of both colors at any stage



Skewed initial distribution

- Unhappiness unbalanced
- Many absorbing states
- Whp Unhappy of both colors at any stage
- Total segregation whp

Intolerance <1/2

- More tolerance
- More happiness

Analysis:

Stable intervals: length w with bias> 2wt
(probability goes to 0 as w goes to infinity)

Unhappy nodes
(probability goes to 0 as w goes to infinity)

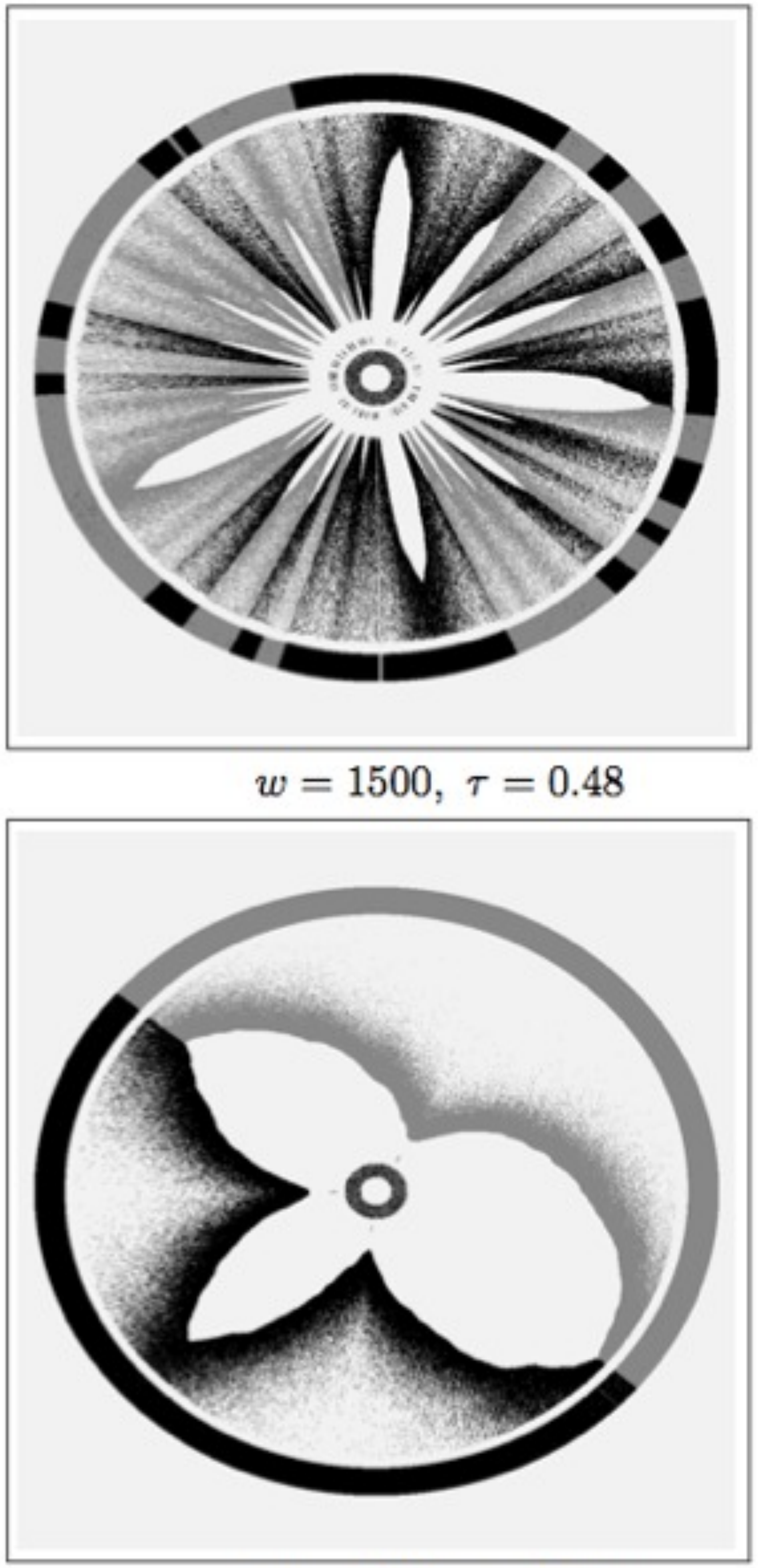
Compare the two probabilities

Stable intervals are more likely

Likely that sites don't change

Unhappy nodes more likely

Unhappy initiate cascades



$w = 1500,\ \tau = 0.48$

$w = 3000,\ \tau = 0.48$

Spreading firewalls

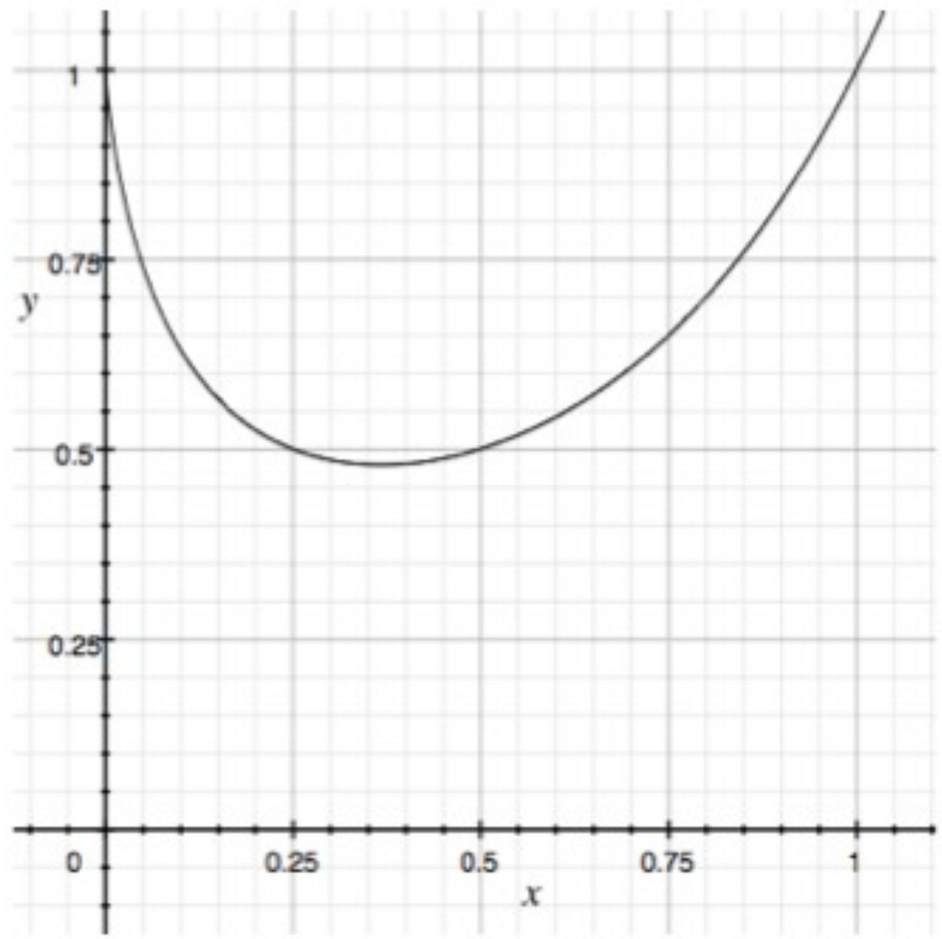
Compare binomial distributions

B(w, 2wτ) and B(2w, 2w(1-τ))

(stable and unhappy events)

By a powerful approximation result of the **binomial** by **normal**

$f(x) = x^{2x}$



the threshold is the solution to

$$f(\frac{1}{2} - \tau) = f(1 - \tau).$$

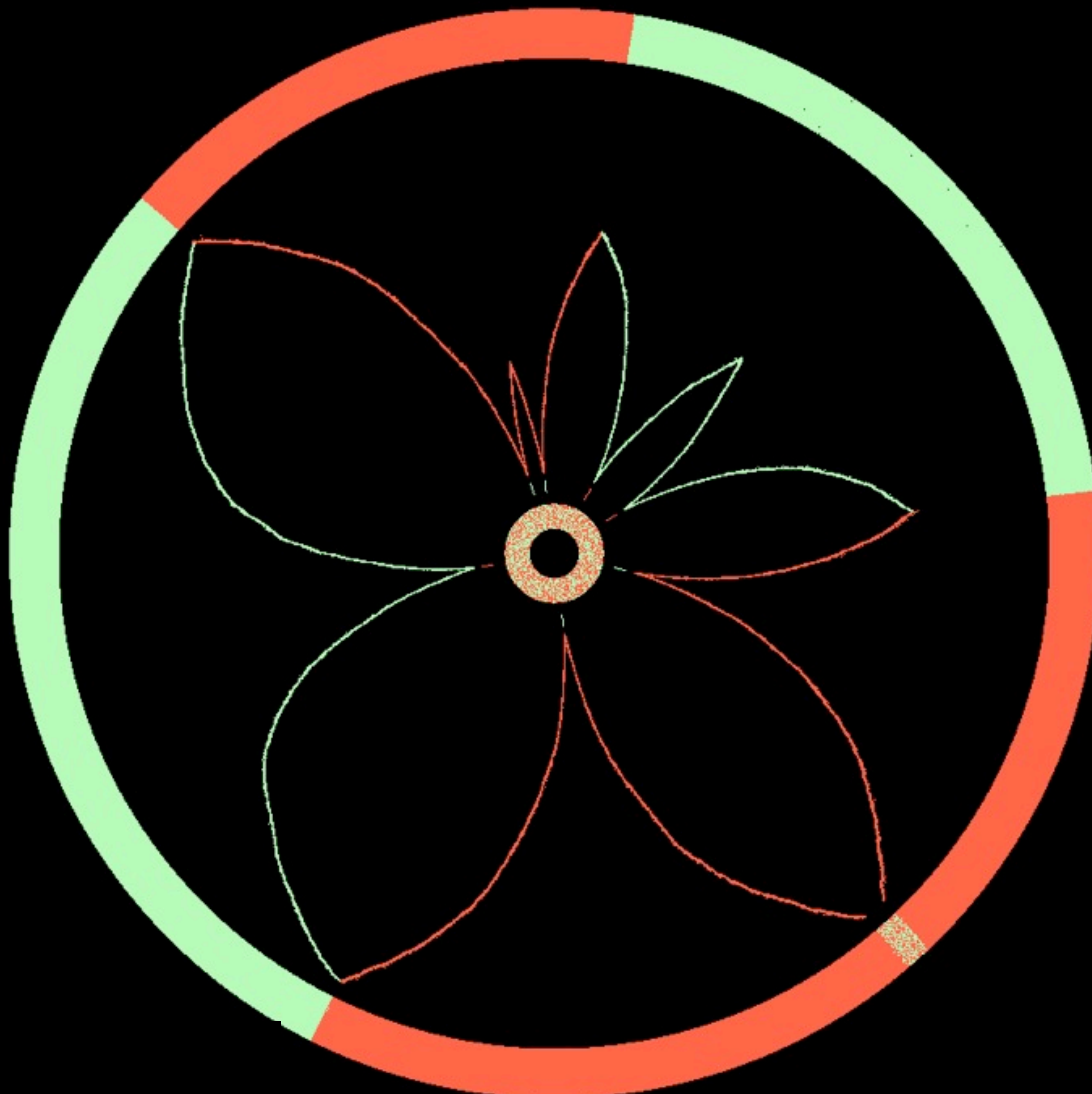
Thus κ is unique, and numerical analysis gives $\kappa \approx 0.353092313$ (which is just slightly less than $\sqrt{2}/4$)

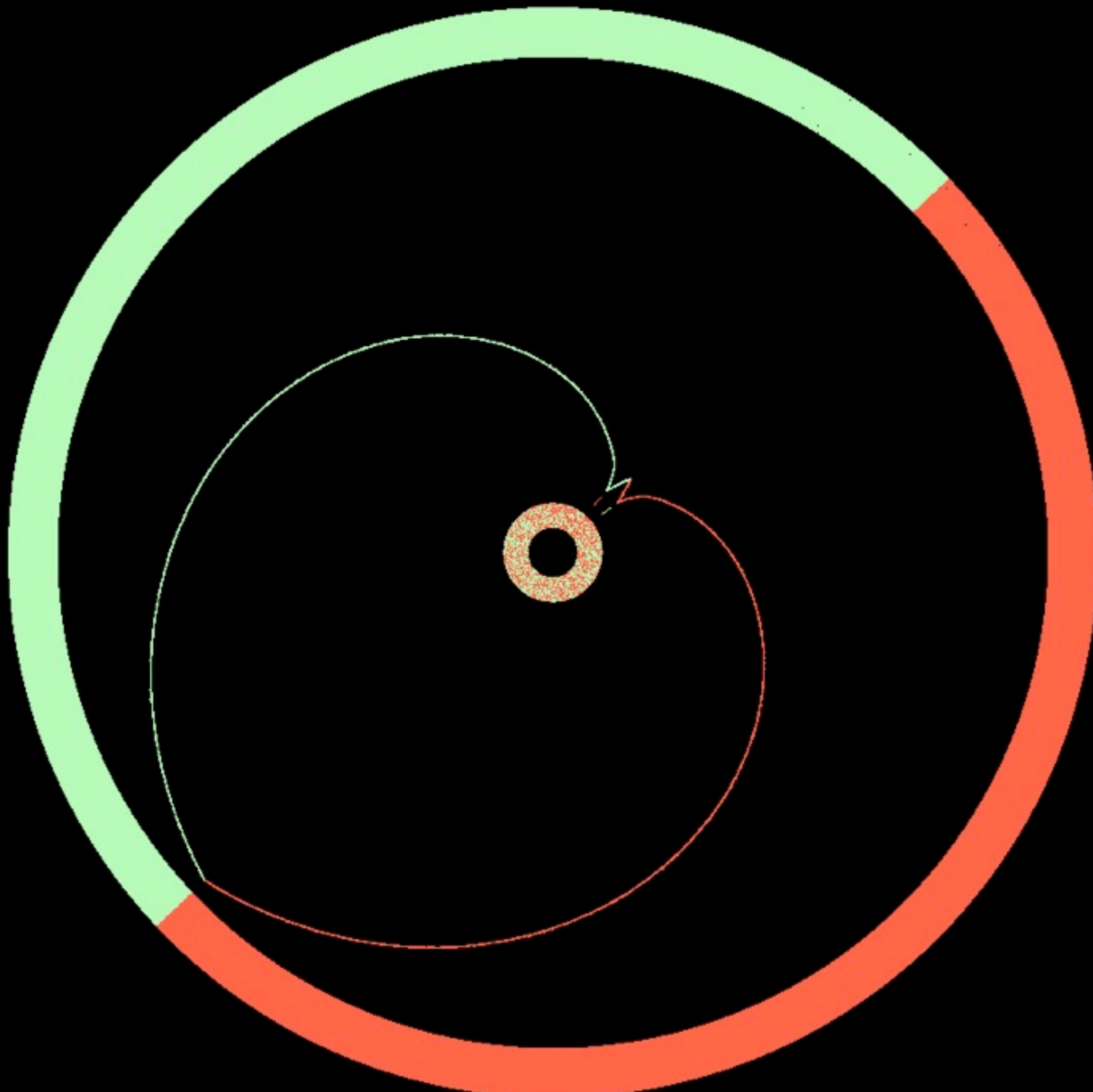
Theorems

Theorem 1.1. *Suppose $\tau < \kappa$ and $\epsilon > 0$. For all sufficiently large w, if u is chosen uniformly at random, then the probability that any node in $\mathcal{N}(u)$ is ever involved in a swap is $< \epsilon$. Thus there exists a constant d such that, for sufficiently large w, the probability u belongs to a run of length $> d$ in the final configuration is $< \epsilon$.*

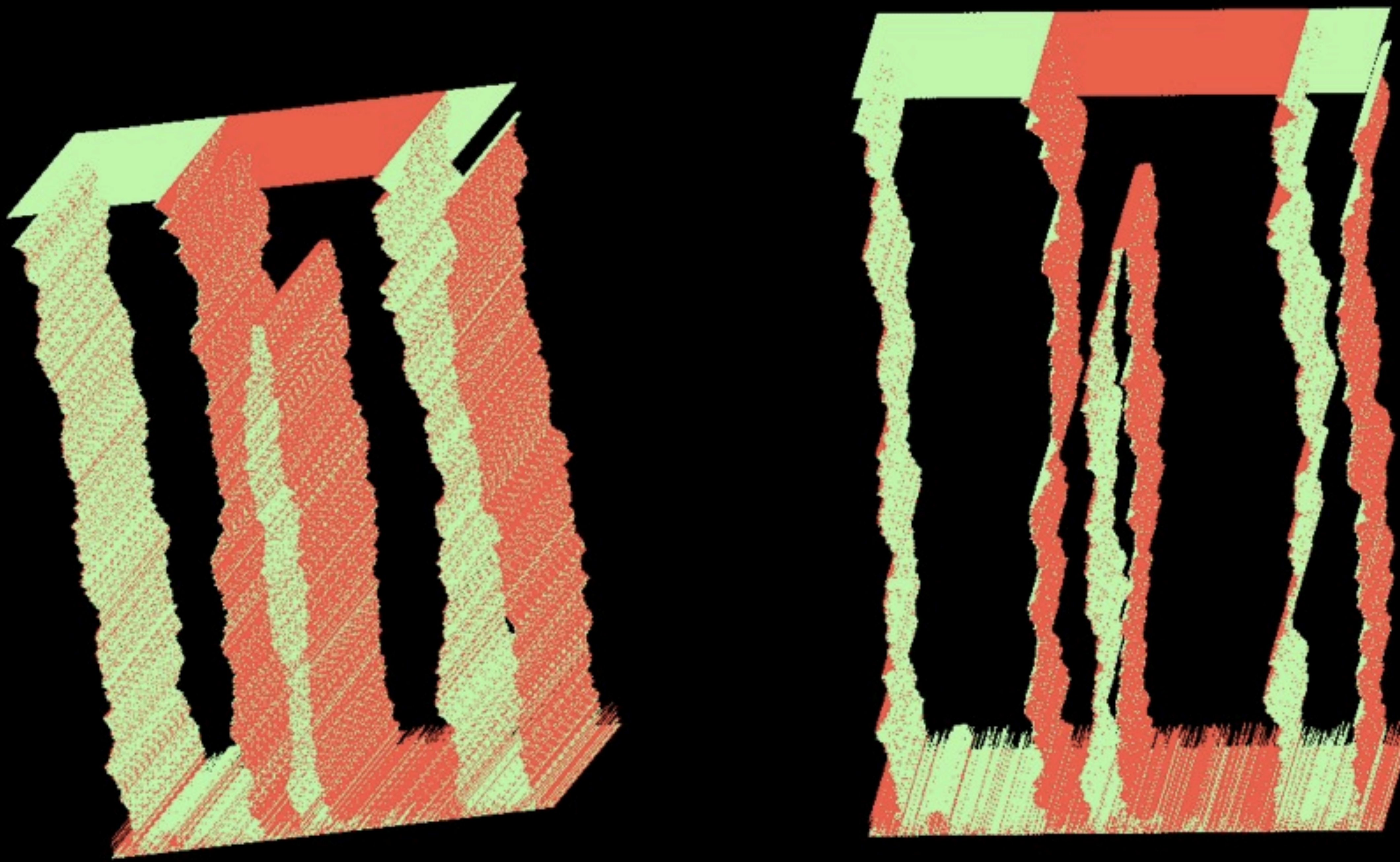
Theorem 1.2. *Suppose $\tau \in (\kappa, \frac{1}{2})$ and $\epsilon > 0$. There exists a constant d such that (for all w and $n \gg w$) the probability that u chosen uniformly at random will belong to a run of length $\geq e^{w/d}$ in the final configuration, is greater than $1 - \epsilon$.*

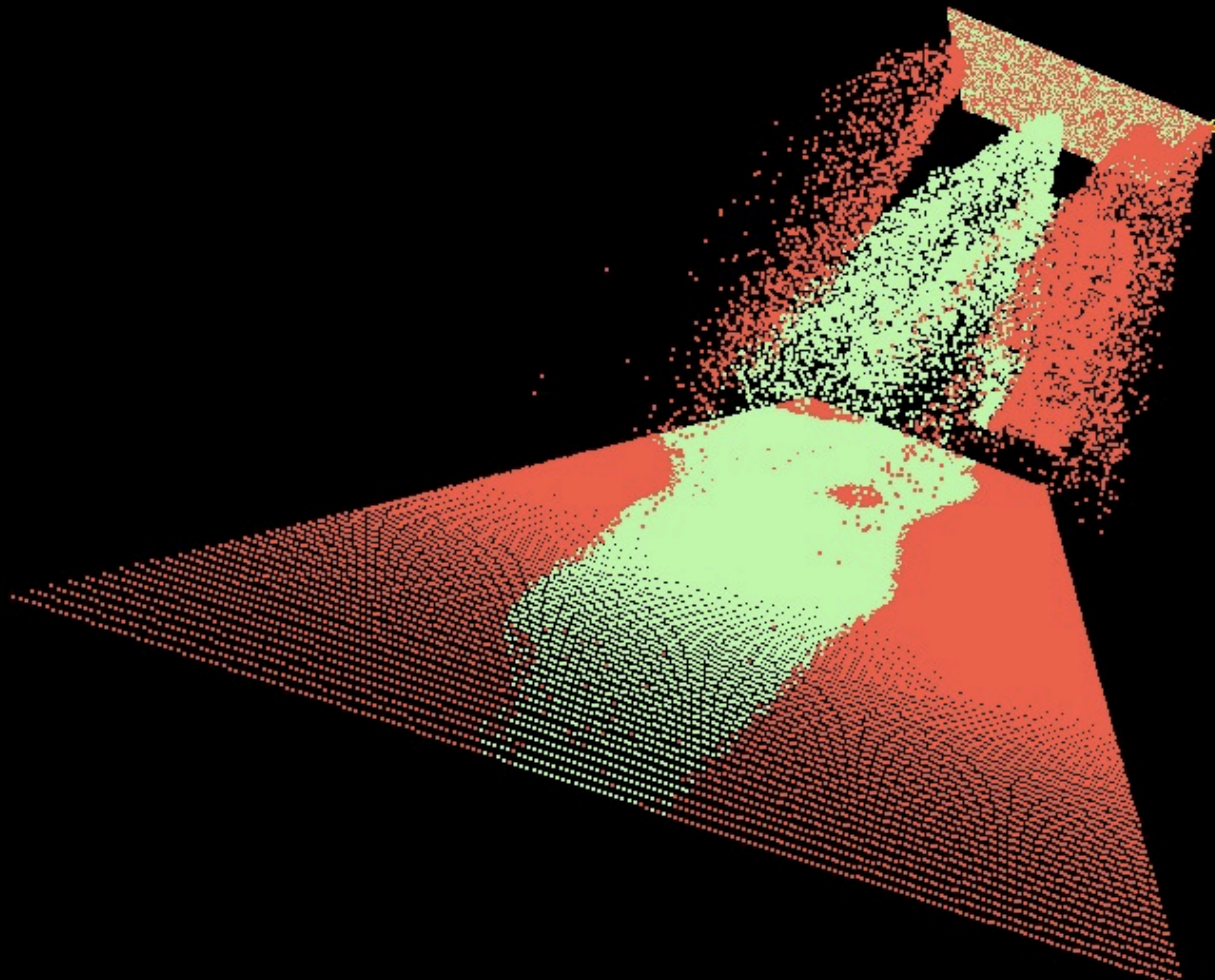
Theorem 1.4. *Suppose that $\tau > \frac{1}{2}$, and that w is sufficiently large that $\tau > \frac{w+1}{2w+1}$ (so that the process is not identical to that for $\tau = \frac{1}{2}$). Then, with probability tending to 1 as $n \to \infty$, the initial configuration is such that complete segregation eventually occurs with probability 1.*

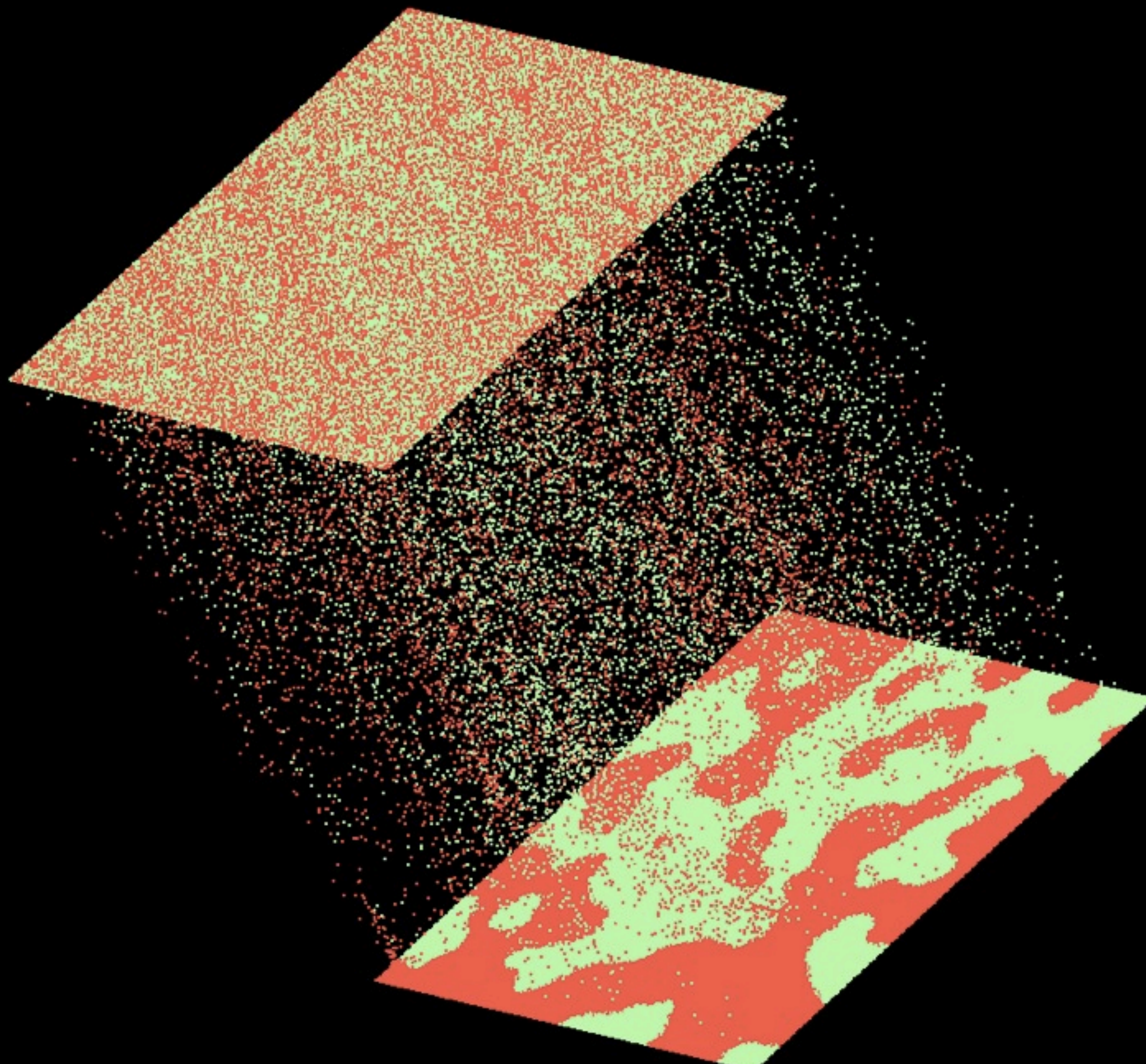




3D representations







Preprints

Barmpalias/Elwes/Lewis

- Digital morphogenesis via Schelling segregation
- Analysis of the skewed 1D Schelling model
- Tipping points in Schelling segregation

Arxiv

barmpalias.net / richardelwes.co.uk / aemlewis.co.uk