**Università degli Studi di Padova**

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"
Corso di Laurea Magistrale in Matematica


TESI DI LAUREA MAGISTRALE

# The Axelrod model for cultural interaction


Candidato:

**Davide Rigon**
**Matricola 1141092**

Relatore:

**Prof. Paolo Dai Pra**



**Anno Accademico 2017-18**
**6 Luglio 2018**

# Contents

# Ringraziamenti

In questa pagina vorrei ringraziare tutti coloro che mi hanno sostenuto durante i vari anni di studi e durante la stesura di questo lavoro.
Un riconoscimento particolare va a Marta che è stata per me sempre di grandissimo aiuto e sostegno.
Ringrazio molto il professore Paolo Dai Pra che mi ha dato fiducia e sicurezza, e dal quale ho imparato molto.
Sono profondamente grato ai miei cari genitori che, nonostante tutto, mi hanno incoraggiato a studiare e mi hanno sostenuto in ogni mia attività da sempre.
Grazie a tutti gli amici e colleghi matematici per il tempo trascorso assieme e la preziosa collaborazione nello studio: Edoardo, Laura, Alessandro, Francesco, Sergej e molti altri.
Desidero poi ringraziare tutti gli amici per il loro affetto e supporto psicologico.

*Davide Rigon*

# Introduction

There has been a rapidly growing interest in agent-based models in the last years, in order to understand the behaviour of complex social systems. These models are characterized by agents, rules that govern the outcome of an interaction between two of them and a graphical structure that gives the pairs of agents that may interact due to friendship or geographical proximity. Research in this field has the purpose to deduce the macroscopic behaviour of the model from the microscopic rules, that depends also on the structure of the network of interactions.

From the mathematical point of view, these models are represented in a first simplified way by the interacting particle systems and in particular by spin systems, and in this category we can find for example the voter model or the contact process. This type of models consists in continuous time Feller processes on the configuration space $\{0,1\}^S$, where $S$ is the countable set of agents. The voter model, as in [6], accounts for social influence, that is the tendency of individuals to become more similar when they interact. Individuals are characterized by one of two competing opinions which they update at a constant rate by mimicking one of their neighbors, chosen uniformly at random.

This thesis is focused on one of the most popular models of social dynamics: the Axelrod model. This model describes the evolution of a simple interacting particle system, and it has been proposed by political scientist Robert Axelrod as a stochastic model for the dissemination of culture. In this model, individuals are represented as vertices of a connected graph $G = (V, E)$, where its set of edges $E$ refers to the interactions between them. The model differs from the voter model because it includes another important social factor: homophily, that is the tendency of individuals to interact more frequently with individuals that are more similar. To include this factor, individuals are characterized by a vector (a *culture*) of $F$ coordinates, called features, each of which can assume one of $q$ possible states, denoted as $1, ..., q$. Homophily can be modeled considering a cultural distance between individuals: pairs of neighbors interact at a rate equal to the fraction of features they have in common. Social influence is modeled by assuming that, after an interaction, one of the cultural features they do not share is chosen uniformly at random and the state of one of the two individuals is set equal to the state of the other individual for that cultural feature.
The macroscopic behaviour of the model when the graph is $G = \mathbb{Z}$ and the structure is linear is well known in some cases, as stated in [3],[4] and [5] and discussed in this work. After some preliminares on interacting particle systems and Chernoff bounds, arguments useful in the second part of the thesis, the second Chapter deals with the Axelrod model, in the case $G = \mathbb{Z}$ as mentioned above, and the behaviour with respect to different values

of the parameters. Finally, the last Chapter is a new evolution/modification of the model, in which it is introduced another individual, called a *media*, that has a fixed opinion during time and that interacts at a rate $\beta$ with all the other individuals. With a coupling with a contact process, when $F, q = 2$, we prove that the behaviour of the model depends strongly on $\beta$, and more precisely that if this rate is "small enough" then the model's behaviour is the same as in the case without media, while if it is "large enough" then all the individuals reply the media opinion.

# Chapter 1

# Preliminaries

## 1.1 Spin Systems

Interacting particle systems, and spin systems that are a special case of them, is a branch of probability theory that developed in the last fourty years, and it has rich connections with a great number of areas, as biology or physics, but also the social sciences. These processes are used to model spread of infection, tumor growth, economic systems and also magnetism.

We start with some definitions. Let $X$ be a compact metric space with measurable structure given by the Borel $\sigma$-algebra. Let $D[0,\infty)$ be the set of all functions $\eta_.$ on $[0,\infty)$ with values in $X$ that are right continuous and have left limits. For $s \in [0,\infty)$, the evaluation mapping $\pi_s$ from $D[0,\infty)$ to $X$ is defined by $\pi_s(\eta_.) = \eta_s$. Let $\mathscr{F}$ be the smallest $\sigma$-algebra on $D[0,\infty)$ relative to which all the mappings $\pi_s$ are measurable, and similarly, for $t \in [0,\infty)$, let $\mathscr{F}_t$ be the smallest $\sigma$-algebra on $D[0,\infty)$ relative to which all the mappings $\pi_s$, for $s \leq t$, are measurable.

**DEFINITION 1.1.** A *Markov process* on $X$ is a collection $\{\mathbb{P}^\eta, \eta \in X\}$ of probability measures on $D[0,\infty)$ indexed by $X$ with the following properties:

i) $\mathbb{P}^\eta[\xi_. \in D[0,\infty) : \xi_0 = \eta] = 1$ for all $\eta \in X$;

ii) the mapping $\eta \to \mathbb{P}^\eta(A)$ from $X$ to $[0,1]$ is measurable for every $A \in \mathscr{F}$;

iii) $\mathbb{P}^\eta[\eta_{s+.} \in A | \mathscr{F}_s] = \mathbb{P}^{\eta_s}(A)$ a.s. $(\mathbb{P}^\eta)$ for every $\eta \in X$ and $A \in \mathscr{F}$.

The expectation corresponding to $\mathbb{P}^\eta$ will be denoted by $\mathbb{E}^\eta$. Let $\mathcal{C}(X)$ be the set of continuous functions on $X$, regarded as a Banach space with

$$\|f\| = \sup_{\eta \in X} |f(\eta)|.$$

Define the operator $S(t) : \mathcal{C}(X) \to \mathcal{C}(X)$ by, for $\eta \in X$,

$$S(t)f(\eta) = \mathbb{E}^\eta f(\eta_t). \tag{1.1}$$

**DEFINITION 1.2.** A Markov process $\{\mathbb{P}^\eta, \eta \in X\}$ is said to be a *Feller process* if $S(t)f \in \mathcal{C}(X)$ for every $t \geq 0$ and $f \in \mathcal{C}(X)$.

The processes we will dicuss are spin systems, that are continuous time Feller processes $\eta_t$ on the compact configuration space $X = \{0,1\}^S$, where $S$ is a generic countable set, denoting the set of sites (or individuals).

**PROPOSITION 1.3.** *Let $(\eta_t : t \geq 0)$ be a Feller process on $X$. Then the family $(S(t) : t \geq 0)$ is a Markov semigroup, i.e.*

a) $S(0) = Id$;

b) $t \to S(t)$ *is right-continuous for all* $f \in \mathcal{C}(X)$;

c) $S(t+s)f = S(t)S(s)f$ *for all* $f \in \mathcal{C}(X)$ *and* $s, t \geq 0$;

d) $S(t)1 = 1$ *for all* $t \geq 0$;

e) $S(t)f \geq 0$ *for all non-negative* $f \in \mathcal{C}(X)$.

The dynamic of the process is usually described specifying the transition rates. In spin systems, these rates represent only a flip at a site $x$ from 0 to 1 or vice versa from 1 to 0. Indeed, states at different sites do no change simoultaneously, and this can be described by saying that

$$\mathbb{P}^\eta\big(\eta_t(x) \neq \eta(x), \eta_t(y) \neq \eta(y)\big) = o(t) \quad \text{as } t \to 0,$$

for each $x, y \in S$ with $x \neq y$, and for each $\eta \in X$. Here, as in (1.1), we use the notation $\mathbb{P}^\eta$ meaning the distribution of the process with initial configuration $\eta$.

The rates $c$ are functions of the site $x$ and the configuration $\eta$, and they are denoted by $c(x, \eta)$. In order to give the definition of $c(x, \eta)$, we have first to define the new configurations obtained from a given configuration following a single transition. In other words, for $\eta \in X$ and $x \in S$ define $\eta^x$ by

$$\eta^x(z) = \begin{cases} \eta(z) & \text{if } z \neq x, \\ 1 - \eta(x) & \text{if } z = x. \end{cases}$$

In this way, $\eta^x$ is obtained from $\eta$ by flipping the $x$-th state.

Defining the transition rates, we have to distinguish between the finite and the infinite case of the cardinality of $S$; indeed,

- if $S$ is finite, given a configuration $\eta$ and a vertex $x$, we say that the transition $\eta \to \eta^x$ occurs at rate $c(x, \eta)$ if

$$\mathbb{P}^\eta(\eta_t = \eta^x) = c(x, \eta)t + o(t) \text{ as } t \to 0. \tag{1.2}$$

- if $S$ is infinite, the intuitive meaning of rate is similar, replacing equation (1.2) with

$$\mathbb{P}^\eta(\eta_t = \eta^x \text{ on } A) = c(x, \eta)t + o(t) \text{ as } t \to 0 \tag{1.3}$$

  for large finite sets $A \subset S$.

We will assume that $c(x, \eta)$ is a uniformly bounded nonnegative function which is continuous in $\eta$ for each $x$ and which satisfies the condition

$$\sup_{x\in S}\sum_{u\in S}\sup_{\eta\in X}|c(x,\eta)-c(x,\eta^u)|<\infty. \tag{1.4}$$

The relation between transition rates and the process is provided by the infinitesimal generator $\Omega$. It is an operator defined on a (dense) subset of $\mathcal{C}(X)$ and it is determined by its values on cylinder functions, functions that depend only on a finite number of coordinates.

**DEFINITION 1.4.** A function $f \in \mathcal{C}(X)$ is a *cylinder function* if there exists a finite subset $\Delta \subset S$ such that $f(\eta^x) = f(\eta)$ for all $x \in \Delta$ and all $\eta \in X$, i.e. $f$ depends only on a finite set of coordinates of a configuration.

**DEFINITION 1.5.** The generator $\Omega : \mathscr{D} \to \mathcal{C}(X)$ for the process $(S(t) : t \geq 0)$ is given by

$$\Omega f := \lim_{t\to 0}\frac{S(t)f-f}{t} \tag{1.5}$$

where the domain $\mathscr{D} \subset \mathcal{C}(X)$ is the set of functions for which the limit exists.

In spin systems, with condition (1.4), given $f$ a cylinder function the form for the generator in the case of spin systems is the following:

$$\Omega f(\eta) = \sum_x c(x,\eta)\big[f(\eta^x) - f(\eta)\big]. \tag{1.6}$$

In fact, as stated in [1], under assumption (1.4), the closure of the operator defined in (1.6) is the generator of a unique Feller process $\eta_t$ on $X$, that satisfies also (1.5) and so, by uniqueness, it is the generator of the process.

**THEOREM 1.6.** *Given $\Omega$ as in (1.5) and $f \in \mathscr{D}$, the process $f(\eta_t) - \int_0^t \Omega f(\eta_s)ds$ is a martingale for the natural filtration.*

Much of the studies of interacting particle systems involves their invariant measures, and convergence to them.

**DEFINITION 1.7.** If $\mu$ is a probability measure on $X$, we denote with $\mu S(t)$ the distribution of $\eta_t$ when the initial distribution is $\mu$. It is defined by

$$\int_X f d\mu S(t) = \int_X S(t) f d\mu \quad \forall f \in \mathcal{C}(X) \tag{1.7}$$

with $S(t)f$ as in (1.1).

**REMARK 1.** *The fact that probability measures on $X$ can by characterised by all expected values of functions in $\mathcal{C}(X)$ is a direct consequence of Riesz representation theorem.*

**DEFINITION 1.8.** A probability measure $\mu$ on $X$ is said to be an *invariant measure* (or a *stationary distribution*) if $\mu S(t) = \mu$ for all $t > 0$. The set of invariant measures is denoted by $\mathcal{I}$.

The set $\mathcal{I}$, thanks to the Feller property and the compactness of $X$, is always nonempty. In some cases, there will be some stationary distributions that can be written down explicitely, and the task is to determine whether or not those exhaust all of $\mathcal{I}$. To check directly if a measure is invariant, we have the following

**THEOREM 1.9.** *A probability measure $\mu$ on $X$ is an invariant measure for the process if and only if*

$$\int_X \Omega f d\mu = 0$$

*for every $f$ cylinder function on $X$.*

One of the main questions one can ask is whether coexistence of types is possible in equilibrium, and in general what is the limiting behaviour of the system for large values of $t$.

**DEFINITION 1.10.** For a spin system $\eta_t$ we say that:

i) *clustering* occurs if

$$\lim_{t\to\infty} \mathbb{P}^{\eta}\big(\eta_t(x) \neq \eta_t(y)\big) = 0$$

for all $x, y \in S$ and all initial configurations $\eta$;

ii) *coexistence* occurs otherwise, i.e. if there exists a stationary distribution that concentrates on configurations with infinitely many 0's and 1's.

Note that, when there is clustering, the only stationary distributions are the ones supported on the set of configurations in which all sites share the same opinion.

### 1.1.1 Graphical Representation

Some interacting particle systems can be represented graphically using the so called graphical representation. This idea consists in drawing in the plan interactions between vertices, taking the vertical axis as the time (only the positive ray), and the horizontal axis as the graph (for this, $G$ must be one-dimensional, so $G = \mathbb{Z}$ or a subset of it). We will discuss it case by case.

### 1.1.2 The linear voter model

The linear voter model is a particular spin system, where $S = \mathbb{Z}^d$ and the transition rates are

$$c(x, \eta) = \sum_{y:\eta(y)\neq\eta(x)} p(x, y) \tag{1.8}$$

with $p(x, y)$ transition probabilities for a Markov chain in $S$. The interpretation is that sites are individuals that have one of two opinions (0 and 1) and, at exponential times of rate 1, an individual $x$ chooses another individual $y$ with probability $p(x, y)$, and he adopts $y$'s opinion.

There are two trivial stationary distributions, that are $\delta_0$ and $\delta_1$, respectively the pointmasses on $\eta \equiv 0$ and $\eta \equiv 1$.

For linear voter models it holds, for all $\eta \in X$,

$$\mathbb{P}^\eta\big(\eta_t(x) \neq \eta_t(y)\big) = \mathbb{P}\big(\eta(X_t) \neq \eta(Y_t)\big)$$

where $X_t$ and $Y_t$ are continuous time random walks on $S$ with $X_0 = x$ and $Y_0 = y$ and transition rates $p(x, y)$, and $\eta(X_t)$ is the position of the random walk at time $t$. $X_t$ is called the ancestor of $x$ at time $t$, and one can find it using a path in the graphical representation: the opinion $\eta_t(x)$ of the voter at $x$ at a large time $t$ came from some other voter at $x_1$ at some earlier time $t_1$; continuing backward until time $t_n = 0$ (for some $n \geq 1$) one can find its ancestor. Such a path is called a dual path. Note that $X_s$ and $Y_s$ are not independent, indeed they are independent only until the first time $\tau$ they meet, i.e. $X_\tau = Y_\tau$, and then they evolve together. For this, they are called coalescing random walks.

Called $\tilde{X}_t = X_t - Y_t$, we have the following

**THEOREM 1.11.** *The linear voter model clusters if $\tilde{X}_t$ is recurrent and coexists if $\tilde{X}_t$ is transient. In particular,*

i) *the process clusters if $d = 1$ and $\sum_x |x|p(0, x) < \infty$, or if $d = 2$ and $\sum_x |x|^2 p(0, x) < \infty$;*

ii) *the process coexists if $d \geq 3$.*

About the limiting behaviour of the system, we have another time to distinguish the recurrent case from the transient one.

**THEOREM 1.12.** *Suppose $\mu$ is any translation invariant probability measure on the state space $X$, and define $\rho := \mu\big(\{\eta : \eta(x) = 1\}\big)$. Then, denoting with $\Rightarrow$ the weak convergence:*

i) *if $\tilde{X}_t$ is recurrent, then $\mu S(t) \Rightarrow \rho\delta_1 + (1 - \rho)\delta_0$ as $t \to \infty$;*

ii) *if $\tilde{X}_t$ is transient, then the extremal invariant measures for $\eta_t$ form a one- parameter family $\{\mu_\rho, 0 \leq \rho \leq 1\}$; if $\mu$ is also ergodic, then $\mu S(t) \Rightarrow \mu_\rho$, as $t \to \infty$.*

### 1.1.3 A special linear voter model

A special case of the linear voter model is the basic voter model. In this case $S = \mathbb{Z}^d$ as above, and the probabilities are

$$p(x, y) = \begin{cases} 1/2d & \text{if } |x - y| = 1 \text{ and } \eta(x) \neq \eta(y); \\ 0 & \text{otherwise.} \end{cases}$$

In particular, the rate is

$$c(x, \eta) = \frac{1}{2d}\Big|\{y : |y - x| = 1, \eta_t(x) \neq \eta_t(y)\}\Big|.$$

Applying the results in the previous section, we have that

**THEOREM 1.13.** *The one dimensional basic voter model clusters.*

We will use this theorem in the next chapter to prove the clustering of the Axelrod model for particular values of its parameters, together with the following proposition.

**PROPOSITION 1.14.** *For the one dimensional basic voter model it holds*

$$\mathbb{P}\left(\eta_t(x) \neq \eta_s(x) \text{ for some } t > s\right) = 1 \qquad \text{for all } x \in \mathbb{Z} \text{ and } s > 0.$$

*Proof.* Let, for $y \in \mathbb{Z}$,

$$M_t(y) = \text{card}\{z \in \mathbb{Z} : y \text{ is the ancestor of } z \text{ at time } t\}$$

the process that keeps track of the descendants of $y$ at time $t$ for the process $\eta$.

This process is a martingale for the natural filtration: in fact, the generator of the dual process, that is the process that consider the dual paths and so the descendant, is the operator defined, for every function $f : \mathcal{F}(\mathbb{Z}) \to \mathbb{R}$ and for every $A \in \mathcal{F}(\mathbb{Z})$, that is $A \subset \mathbb{Z}$ finite, by

$$\mathcal{L}f(A) = \frac{1}{2} \sum_{x \in A} \sum_{y \sim x, y \notin A} \left[\left(f(A \cup \{y\}) - f(A)\right) + \left(f(A \setminus \{x\}) - f(A)\right)\right].$$

Taking $f$ equal to the cardinality function, one has $\mathcal{L}f(A) = 0$ for every $A \in \mathcal{F}(\mathbb{Z})$, and this means, thanks to theorem (1.6), that the cardinality of descendants, as a function of time $t$, is a martingale.

Furthermore, this martingale is absorbed at state 0: if $M_s(y) = 0$, then $M_t(y) = 0$ for every $t > s$, otherwise $M_s(y)$ should have been different from 0. Since it takes value in $\mathbb{N}$, and it converges almost surely, as the martingale convergence theorem states, the only possibility is that it converges to its adsorbing state 0.

Now, fix $x \in \mathbb{Z}$ and $s > 0$, and let $y = X_s(x)$ to be the ancestor, as introduced in the previous section, of $x$ at time $s$. Define

$$\phi(s) = \inf\{t > 0 : M_t(y) = 0\}.$$

that is a stopping time, larger than $s$ (there is a dual path from $(x, s)$ to $(y, 0)$, so $M_s(y) \geq 1$) and almost surely finite, because $M_t(y)$ is an integer valued process and it converges almost surely to 0.

Moreover, the spin at $(x, \phi(s))$ and at $(x, s)$ originates from different vertices at time 0, since there isn't a dual path from $(x, \phi(s))$ to $(y, 0)$. The two spins are so independent, because the initial distribution is a product measure. Inductively, define $s_0 = s$ and, for $i \geq 0$,

$$s_{i+1} = \phi(s_i) = \inf\{t > 0 : M_t(X_{s_i}(x)) = 0\}.$$

This is a sequence of (increasing) stopping times, all almost surely finite. Since all spins at $(x, s_i)$ are independent (they originate from different vertices at time 0) and each culture occurs initially with positive probability, one has

$$\lim_{n \to +\infty} \mathbb{P}(\eta_{s_0}(x) = \eta_{s_1}(x) = \cdots = \eta_{s_n}(x)) = 0$$

and so

$$\mathbb{P}\Big(\bigcup_{n=1}^{\infty} \left(\eta_{s_n}(x) \neq \eta_s(x)\right)\Big) = 1$$

and the proposition follows. □

### 1.1.4 The contact process

The contact process is another particular spin system, in which $S$ is thought as a connected undirected graph whose vertices have bounded degree, and the transition rates are

$$c(x,\eta) = \begin{cases} 1 & \text{if } \eta(x) = 1 \\ \lambda\big|\{y \sim x : \eta(y) = 1\}\big| & \text{if } \eta(x) = 0 \end{cases} \tag{1.9}$$

where $\lambda > 0$ is a (infection) parameter, and $y \sim x$ means that the two vertices are connected by an edge. The interpretation is that sites with $\eta(x) = 1$ are infected while sites with $\eta(x) = 0$ are healthy. If a site is infected, he recovers after an exponential time of rate 1, while if it is healthy it becomes infected at a rate equal to the number of its neighbors infected sites times $\lambda$. Configurations $\eta \in X$ will often by identified with subsets $A$ of $S$ via $A = \{x \in S : \eta(x) = 1\}$, so considering the infected vertices; $A_t$ so will denote the infected vertices at time $t$.

It is easy to see that the pointmass $\delta_0$ on the configuration $\eta \equiv 0$ is a stationary distribution. Another invariant measure is the so called upper invariant measure, and it is constructed by monotonicity. Indeed, taking $A_0 = S$, so when all vertices are initially infected, and $\mu_t$ the distribution of $A_t$, it exists

$$\bar{\nu} = \lim_{t\to\infty} \mu_t$$

and it is the biggest invariant measure of the process.

When $S$ is finite, $\delta_0$ is the only stationary distribution, and for any initial configuration $\eta_t$ is eventually $\equiv 0$ .

The general problem of determining when convergence of $\mu S(t)$ occurs is difficult, and the answer depends heavily on $S$ and $\lambda$. In the case $S = \mathbb{Z}^d$, if the initial distribution $\mu$ is homogeneous and it satisfies $\mu(\emptyset) = 0$, then $\mu S(t) \Rightarrow \bar{\nu}$ as $t \to \infty$, where $\Rightarrow$ denotes weak convergence. This means that in this case there are at most two extremal translation invariant measures in $\mathcal{I}$.

About convergence, there is another important concept:

**DEFINITION 1.15.** The process is said to *complete converge* if, for every initial configuration $A$,

$$A_t^A \Rightarrow \alpha_A \bar{\nu} + (1 - \alpha_A)\delta_0 \tag{1.10}$$

where

$$\alpha_A = \mathbb{P}^A\big(A_t \neq \emptyset \ \forall t \geq 0\big)$$

is called the survival probability.

The graphical representation in contact processes is as follows: for every vertex $x \in S$, assign a Poisson process $N_x$ of rate 1 and for every ordered pair of vertices $(x, y)$ connected by an edge assign a Poisson process $N_{x,y}$ of rate $\lambda$. All Poisson processes are independent. At each event time $t$ of $N_x$ place a symbol $\times$ at the point $(x,t) \in S \times [0,\infty)$ and for each event time $t$ of $N_{x,y}$ put an (infection) arrow $\leftarrow$ from $(x,t)$ to $(y,t)$. The case $S = \mathbb{Z}$ is shown in Figure 1.1.

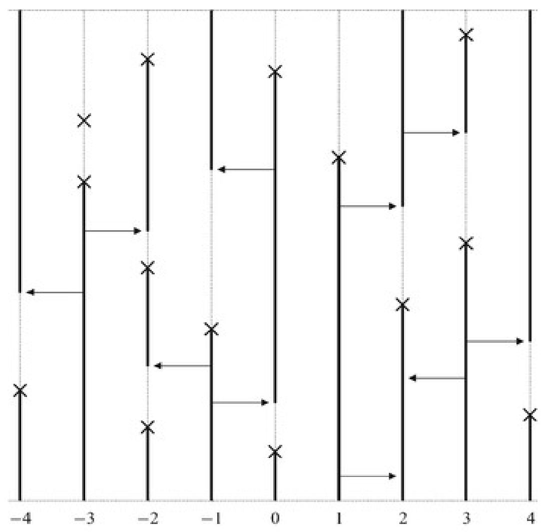


Figure 1.1: The graphical representation for the contact process.

One advantage of the graphical construction is that it provides a joint coupling of the process with arbitrary initial states, that is better known as *monotone coupling*:

$$A_0 \subset B_0 \Rightarrow A_t \subset B_t. \tag{1.11}$$

An active path in $S \times [0, \infty)$ is a connected path that moves along the time lines in the increasing $t$ direction and along arrows in their direction, but that do not pass through a recovery symbol. Define $A_t^A$ to be the set of infected sites at time $t$, with $A$ as initial set of infected sites; in other words, with the graphical representation,

$$A_t^A = \{y \in S : \exists x \in A \text{ s.t. there is an active path from } (x, 0) \text{ to } (y, t)\}.$$

The most important feature of the contact process is that survival and extinction can both occur, and this depends on the value of $\lambda$.

**DEFINITION 1.16.** The contact process is said to *die out* (or it becomes *extinct*) if

$$\mathbb{P}^{\{x\}}\big(A_t \neq \emptyset \ \forall t \geq 0\big) = 0$$

and it is said to *survive* otherwise. It is said to *survive strongly* if

$$\mathbb{P}^{\{x\}}\big(x \in A_t \text{ i.o.}\big) > 0$$

and to *survive weakly* if it survives but not strongly.

**REMARK 2.** *Note that all the previous properties do not depend on the choice of $x$, since $S$ is connected.*

For each graph $S$ as in the hypothesis there exist two critical values $0 \leq \lambda_1 \leq \lambda_2 \leq \infty$ such that

$$\begin{aligned} &A_t \text{ dies out if } \lambda < \lambda_1 \\ &A_t \text{ survives weakly if } \lambda_1 < \lambda < \lambda_2 \\ &A_t \text{ survives strongly if } \lambda > \lambda_2. \end{aligned} \tag{1.12}$$

When $S = \mathbb{Z}^d$ there are more information about the behaviour of the process.

**THEOREM 1.17.** *In the case $S = \mathbb{Z}^d$, it holds $\lambda_1 = \lambda_2$. Moreover, the process dies out at the common critical value, and complete convergence (1.10) holds for all $\lambda$.*

About some estimates of the common critical value $\lambda_c$, in the case $S = \mathbb{Z}^d$, we have:

$$\frac{1}{2d-1} \leq \lambda_c \leq \frac{2}{d}$$

and, in the case $d = 1$, $\lambda_c \geq 1.539$.

## 1.2 Laws of large numbers and Chernoff bounds

This section concerns some bounds in the case of a sequence of independent and identically distributed random variables. Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. random variables, and define the *sample mean* as

$$\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i.$$

The two following inequalities will be useful later.

**PROPOSITION 1.18.** 1. *(Markov inequality) Let $X$ be a nonnegative, integrable random variable. Then, for every $\epsilon > 0$*

$$\mathbb{P}(X \geq \epsilon) \leq \frac{\mathbb{E}[X]}{\epsilon}.$$

2. *(Chebischev inequality) Let $X$ be a square-integrable, real random variable, with mean $\mu$ and variance $\sigma^2$. Then, for every $\epsilon > 0$*

$$\mathbb{P}\big(|X - \mu| \geq \epsilon\big) \leq \frac{\sigma^2}{\epsilon^2}.$$

The Weak Law of Large Numbers states that the value of the sample mean can be predicted if $n$ is large, with a small error probability.

**THEOREM 1.19.** *Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. square integrable random variables with mean $\mu$ and variance $\sigma^2$. The, for every $\epsilon > 0$,*

$$\lim_{n\to+\infty} \mathbb{P}\big(|\bar{X}_n - \mu| \geq \epsilon\big) = 0. \tag{1.13}$$

In the proof of the previous theorem, the bound found for the probability in (1.13) is not very strong, indeed the order of estimate is $\frac{1}{n}$ and it comes from the Chebischev inequality applied to the sample mean. Chernoff estimates use more details of the distribution of the random variables $X_n$, but they require some more assumptions. Remember that the moment generating function of a random variable $X$ is

$$m_X(t) = \mathbb{E}[e^{tX}].$$

**THEOREM 1.20.** *Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. random variables, and assume that $m_{X_i}(t) < +\infty$ for all $t$ in an open interval $I$ containing zero. The, for every $\epsilon > 0$ there exist $c_1, c_2 > 0$ such that*

$$\mathbb{P}\big(\bar{X}_n \geq \mu + \epsilon\big) \leq e^{-c_1 n} \tag{1.14}$$

*and*

$$\mathbb{P}\big(\bar{X}_n \leq \mu - \epsilon\big) \leq e^{-c_2 n}. \tag{1.15}$$

*Proof.* Since all variables have the same distribution, they have the same moment generating function, so call $m(t) := m_{X_n}(t)$.

For every $t > 0$, it holds

$$\begin{aligned}
\mathbb{P}\big(\bar{X}_n \geq \mu + \epsilon\big) &= \mathbb{P}\big(X_1 + \cdots + X_n \geq n(\mu + \epsilon)\big) \\
&= \mathbb{P}\Big(e^{t(X_1+\cdots+X_n)} \geq e^{tn(\mu+\epsilon)}\Big) \\
&\leq \frac{\mathbb{E}\Big[e^{t(X_1+\cdots+X_n)}\Big]}{e^{tn(\mu+\epsilon)}}
\end{aligned}$$

where in the last step we have used Markov inequality for $e^{t(X_1+\cdots+X_n)}$. Rewriting the last term, we find

$$\begin{aligned}
\mathbb{P}\big(\bar{X}_n \geq \mu + \epsilon\big) &\leq \frac{m^n(t)}{e^{tn(\mu+\epsilon)}} \\
&= \exp[n \log m(t) - nt(\mu + \epsilon)] \\
&= \exp\big[-n\big(t(\mu+\epsilon) - \log m(t)\big)\big].
\end{aligned}$$

Now, define

$$g(t) := t(\mu + \epsilon) - \log m(t).$$

Since $m(0) = 1$, one has that $g(0) = 0$. Moreover, since $m'(0) = \mu$,

$$g'(0) = \mu + \epsilon - \frac{m'(0)}{m(0)} = \epsilon > 0.$$

This implies that for some $t > 0$ $g(t)$ is positive. The best possible estimate would be obtained by maximizing $g(t)$ over $t$, so define $c_1 := \max_I m(t)$. Then (1.14) follows.

For the second one, the proof is very similar. Infact, for every $t > 0$ one has

$$\begin{aligned}
\mathbb{P}\big(|\bar{X}_n \leq \mu - \epsilon\big) &= \mathbb{P}\big(X_1 + \cdots + X_n \leq n(\mu - \epsilon)\big) \\
&= \mathbb{P}\Big(e^{-t(X_1+\cdots+X_n)} \geq e^{-tn(\mu-\epsilon)}\Big) \\
&\leq \frac{\mathbb{E}\Big[e^{-t(X_1+\cdots+X_n)}\Big]}{e^{-tn(\mu-\epsilon)}} \\
&= \frac{m^n(-t)}{e^{-tn(\mu-\epsilon)}} \\
&= \exp\big[-n\big(-t(\mu-\epsilon) - \log m(-t)\big)\big].
\end{aligned}$$

As before, the function

$$h(t) := -t(\mu - \epsilon) - \log m(-t)$$

assume positive values for some $t > 0$, and defining $c_2 := \max_I h(t)$ it follows (1.15). □

# Chapter 2

# The Axelrod model

The Axelrod model is a stochastic process that includes social influence, as the voter model, but also homophily, both mechanism that are usually seen in the dynamic of populations. The first one is the tendency of individuals to become more similar when they interact, while the second is the tendency to interact more frequently with individuals that are more similar.

In this model, individuals are represented as vertices of a connected graph $G = (V, E)$, where its set of edges $E$ refers to the interactions between them. Each vertex $x \in V$ is characterized by a vector $X(x)$, its opinion, of $F$ cultural features, each of which can assume $q$ possible states. In other words,

$$X(x) = (X^1(x), ..., X^F(x))$$

and

$$X^i(x) \in \{1, ..., q\} \quad \forall i = 1, ..., F.$$

Note that the set of cultures, that is $\{1, ..., q\}^F$ and describe the state of a vertex at a time $t$, is equipped with a natural distance: the function that counts the number of disagreements between two cultures. This distance is essential to model both homophily and social influence: homophily is modeled by assuming that the rate at which two neighbors interact decreases with the distance between their cultures, and social influence by assuming that the result of an interaction is to decrease the cultural distance.
In particular, the dynamic is described as follows: at each time, a vertex $x$ is picked uniformly at random from $V$, and then one of its neighbors $y$. With a probability equal to the fraction of features $x$ and $y$ have in common, one of the features for which they disagree (if any) is chosen, and the state of $x$ corresponding to that feature is set equal to the $y$'s one. In the case the two vertices agree in all features nothing happens.

Assuming that the system evolves in continous time, with each pair of adjacent vertices interacting ar rate one, the process is actually a continuous time Markov process, whose state at time $t$ is

$$X_t : V \to \{1, ..., q\}^F$$

and the dynamic is described by the generator

$$\Omega_{ax} f(X) = \sum_{x \in V} \sum_{y \sim x} \sum_{i=1}^{F} \frac{1}{2F} \left[ \frac{F(x, y)}{1 - F(x, y)} \right] \mathbb{1}\{X^i(x) \neq X^i(y)\} \left[ f(X^i_{y \to x}) - f(X) \right] \quad (2.1)$$

where:

- $y \sim x$ means that $x$ and $y$ are connected by an edge, because interactions derive only from edges;
- $F(x,y) = \frac{1}{F}\sum_{i=1}^{F} \mathbb{1}\{X^i(x) = X^i(y)\}$ denotes the fraction of cultural features $x$ and $y$ share;
- $X^i_{y\to x}(x) = \big(X^1(x), ..., X^{i-1}(x), X^i(y), X^{i+1}(x), ..., X^F(x)\big)$ and $X^i_{y\to x}(z) = X(z)$ otherwise.

Note that
$$\frac{1}{2F}\left(\frac{F(x,y)}{1-F(x,y)}\right) = F(x,y) \times \frac{1}{2} \times \frac{1}{F(1-F(x,y))}$$
where the first term is the rate at which the two vertices interact, the second term is the probability that one rather than the other is chosen to be updated and the last term is the inverse of the number of disagreements.

**REMARK 3.** *The case $F = 1$ is not interesting because when there is only one feature, since Axelrod model accounts for homophily, nothing happens in every case. In fact, if two vertices agree nothing happens as prescribed, while if they disagree $F(x,y) = 0$ and so they are too discordant. The case $q = 1$ is trivial too, since every feature is forced to be equal to 1, the only possible state. From now, motivated by this remark, we will assume $F,\ q \geq 2$.*

We have now some important definitions, that reflect the asymptotic macroscopic behaviuor of the model.

**DEFINITION 2.1.** We say that there is *fluctuation* if
$$\mathbb{P}(X^i_t(x) \text{ changes infinitely often in } t) = 1 \qquad \forall x \in V, \quad \forall i = 1, ..., F.$$

**DEFINITION 2.2.** We say that there is *fixation* if $\forall x \in V$ and $\forall i = 1, ..., F$ $\exists \bar{q} \in \{1, ..., q\}$ such that
$$\mathbb{P}(X^i_t(x) = \bar{q} \text{ eventually in } t) = 1.$$

The previous two definitions reflect properties of vertices (individuals), more precisely how many times they change their opinion. Note that fixation or fluctuation depend not only on $F$ and $q$, but also on the initial condition: for example, if the system starts with a configuration in which all the individuals agree for a given cultural feature, whereas the states at the other ones are i.i.d., it always fluctuates, because individuals have at least one feature in common at any time, so there are infinitely many interactions and so there is fluctuation. On the other hand, take $G = \mathbb{Z}$ with every two consecutive vertices connected by an edge: if the system starts with a configuration in which all the even sites share the same culture, while the odd ones share another culture that is incompatible with the even ones (i.e. these two cultures are different at every level $i$), it always fixates. This is a consequence of the fact that there can not be interactions, because all adjacent vertices have no features in common, so in every case nothing happens.

As in spin systems, we recall the following definition.

**DEFINITION 2.3.** We say that for the system *clustering* occurs if

$$\lim_{t\to+\infty} \mathbb{P}(X_t^i(x) = X_t^i(y)) = 1 \qquad \forall x, y \in V, \quad \forall i = 1, ..., F$$

and we say the system *coexists* otherwise.

Clustering of the system means convergence to a global consensus, in the sense that, for large values of $t$, the probability that two individuals agree is approximately 1. Note that in general neither fluctuation implies clustering, nor fixation excludes clustering: the voter model in dimension at least three fluctuates but it does not cluster, in fact it coexists, while the biased voter model fixates and clusters. This disconnection derives from the fact that, while fixation/fluctuation consider the probability of an event eventually in time, clusterization refers to a limit of an event with time fixed.

In this chapter we will study the case $G = \mathbb{Z}$ and $E = \big\{\{x, x+1\} : x \in \mathbb{Z}\big\}$, the natural case. Almost all the results refer to the initial distribution $\pi_0$ in which the states of the cultural features within each vertex and among different vertices are independent and identically distributed. Sometimes, the request can be relaxed, as we will see in theorem's statements. We will always specify what are the minimal hypothesis in any theorem.

## 2.1 Constructions

### 2.1.1 System of random walks

Instead of considering cultural features of each vertex, we will consider the interfaces between vertices, introducing the set of sites. A site simply counts for disagreements between vertices that it refers to.

In particular, denote with

$$\mathbb{D} := \mathbb{Z} + 1/2$$

the set of sites, meaning that $u \in \mathbb{D} \Leftrightarrow \exists\, x \in \mathbb{Z}$ such that $u = x + 1/2$. To obtain a system of random walks, for every site $u \in \mathbb{D}$ and for every cultural feature $i = 1, ..., F$ let

$$\xi_t(u, i) = \mathbb{1}\{X_t^i(u - 1/2) \neq X_t^i(u + 1/2)\} \tag{2.2}$$

where $\mathbb{1}\{A\}$ is the indicator function of the set $A$, and put a particle at $(u, i)$ whenever $\xi_t(u, i) = 1$. A particle indicates if there is a disagreement between the two vertices.
To understand when there can be an interaction, so when there is at least one particle at a site, we let, for every $u \in \mathbb{D}$,

$$\xi_t(u) = \sum_{i=1}^{F} \xi_t(u, i). \tag{2.3}$$

**DEFINITION 2.4.** A site $u \in \mathbb{D}$ is called a $j$-site at time $t$ if $\xi_t(u) = j$.

This construction induces a system of $F$ non-independent symmetric random walks: in fact, at each level $i$, as specified by the rules, when a particle jumps it moves to the left or to the right with equal probability, unless another particle already occupies the site on which the first one tries to jump. It is a system of non-independent random walks because a particle jumps according to the number of disagreements between the two vertices, so its jump in general depends on all the other levels. These dependencies result actually from the inclusion of homophily in the model. The symmetry is due to the fact that, when two vertices interact, each of them is equally likely to be the one chosen to be updated.

The role of $q$ is in the result of collision between particles:

- when $q = 2$, if at time $t$ a particle at $(u, i)$ tries to jump to $(u + 1, i)$ and this is already occupied, there is an annihilation event: the presence of two particles means that
$$\begin{cases} X^i_{t-}(u - 1/2) \neq X^i_{t-}(u + 1/2) \\ X^i_{t-}(u + 1/2) \neq X^i_{t-}(u + 3/2) \end{cases}$$
so $X^i_{t-}(u - 1/2)$ must be equal to $X^i_{t-}(u + 3/2)$ because there are only two possible states. It follows that at time $t$ all three vertices agree on their $i$-th feature, so the previous particles kill each other and do not exist anymore at time $t$;

- when $q > 2$, in the same situation the result depends on the $i$-th cultural features of $u - 1/2$ and $u + 3/2$ at time $t-$: they can be equal, so there is an annihilation event as above, but they can also be different, in which case there is a coalescing event. This type of event is characterized by the fact that the particle that tries to jump is removed, because the two individuals $u - 1/2$ and $u + 1/2$ at time $t$ will agree on their $i$-th feature, but the other particle remains, since $u + 1/2$ and $u + 3/2$ still disagree at time $t$.

The following picture represents a collisions between two particles and a jump of a particle, with $F = 4$ and $q = 2$: on the left, we see that the particle at site $\frac{3}{2}$ wants to jump to its right, where there is another particle. So, after the jump, all three vertices 1,2,3 share the white dot, and an annihilation event has happened. In the second case, the particle at $\frac{5}{2}$ wants to jumps to its left, but the site to its left is not occupied by a particle, so simply the particle jumps.

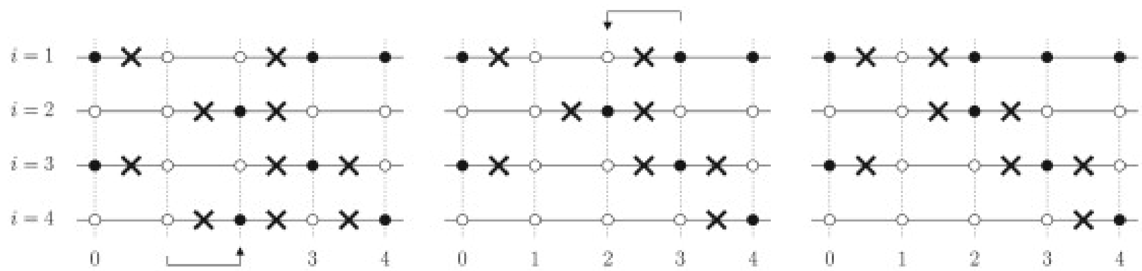


Figure 2.1: Random walks with F=4 and q=2.

This difference of collisions results is the reason for which the behaviour in the 2-states case is well known, as stated in theorems 2.7 and 2.8. In fact, when $q = 2$, knowing the configuration of the Axelrod model is unimportant in determining the evolution of the random walks, because every collision is an annihilating event, so in the proofs we will focus only on this particles, forgetting the configuration. In contrast, when $q > 2$, whether a collision results in an annihilating or a coalescing event depends on the configuration of the Axelrod model just before the time of collision. We will see that, in spite of this dependency, collisions result in either an annihilating or a coalescing event with some fixed probabilities: in other words, the outcome of a collision is independent of the past of the system of random walks, but it is not independent of the past of the model.

We end this section giving explicitly the rate of jump of particles. When $u$ is a $j$-site, with $j \neq 0$, the two vertices $u - 1/2$ and $u + 1/2$ disagree on exactly $j$ features, so they interact at a rate equal to $1 - j/F$, the fraction of features they share. This means that if $u$ is a $j$-site, each particle at site $u$ jumps at a rate

$$r(j) = \frac{1}{j}\left(1 - \frac{j}{F}\right) = \frac{1}{j} - \frac{1}{F} \quad \text{for } j \neq 0 \tag{2.4}$$

with $r(0) = 0$, because no features in common means total disagreements and so no interactions.

### 2.1.2 Graphical Representation

The Axelrod model is included into a general class of interacting particle systems that can be constructed using a graphical representation, as in the first chapter.
For all pairs of vertex-cultural feature $(x, i) \in \mathbb{Z} \times \{1, 2, ..., F\}$ we let:

- $(N_{x,i}(t) : t \geq 0)$ be independent rate one Poisson processes;
- $T_{x,i}(n) = \inf\{t \geq 0 : N_{x,i}(t) = n\}$;
- $(B_{x,i}(n) : n \geq 1)$ be collections of independent Bernoulli variables with parameter $p = \frac{1}{2}$ and with values in $\{1, -1\}$;
- $(U_{x,i}(n) : n \geq 1)$ be collections of independent Uniform random variables on the interval $(0, 1)$.

Poisson processes fix times of interactions between vertices, while Bernoulli variables choose the adjacent vertex that will update the one that the variable refers to (the value of $p$ reflects the symmetry). Finally, Uniform variables verify the correct rate, and split all interactions between potential and real ones.

More precisely, at each time $t = T_{x,i}(n)$, we draw an arrow from the pair

$$(y, i) := (x + B_{x,i}(n), i)$$

to the pair $(x, i)$, meaning that at time $t$ there is a potential interaction between vertices $x$ and $y$. We say that in this case there is an arrow from vertex $y$ to vertex $x$ at level $i$, at time $t$.

If in addition

$$X^i_{t-}(x) \neq X^i_{t-}(y) \quad \text{and} \quad U_{x,i}(n) \leq r(\xi_{t-}(x)) \tag{2.5}$$

than the arrow is said to be active. The meaning of active arrows is that they correspond exactly to interactions that updated the system. Note that, with these choices, the rate of interaction is exactly the one in (2.1).

### 2.1.3 Dual paths

Since Axelrod model is included into a general class of models in which one can apply the idea of graphical representation, one can apply also duality theory, searching dual paths from one vertex to another one. Duality is useful when it needs to keep track of the ancestry of each vertex going backward in time.

**DEFINITION 2.5.** Let $(z, s)$ and $(x, t)$ be two couples of vertices-times, and $i \in \{1, ..., F\}$. We say that there is an *active $i$-path* from $(z, s)$ to $(x, t)$ if there exist sequences of times and vertices

$$s_0 = s < s_1 < \cdots < s_{n+1} = t \quad \text{and} \quad x_0 = z, x_1, ..., x_n = x$$

such that the following two conditions hold:

(1) for $j = 1, ..., n$ there is an active $i$-arrow from $x_{j-1}$ to $x_j$ at time $s_j$;

(2) for $j = 0, ..., n$ there is no active $i$-arrow that points at $\{x_j\} \times (s_j, s_{j+1})$.

In this case, we use the notation $(z, s) \rightsquigarrow_i (x, t)$.

The meaning of active $i$-paths is the following: the first condition is necessary to have a connection between the two couples, that is vertex $x$ inherits the $i$-th feature of vertex $z$, but the second condition is essential to guarantee this connection, since otherwise there could exists an index $j$ for which another vertex different from $x_{j-1}$ interacts with $x_j$ in a time $\tau \in (s_j, s_{j+1})$.

Then, for $A \subset \mathbb{Z}$, the dual process starting at $(A, T)$ is the set valued process

$$\hat{Y}_s(A, T) = \{y \in \mathbb{Z} : \text{there is a dual path from } (x, T) \text{ to } (y, T - s) \text{ for some } x \in A\}. \tag{2.6}$$

Conditions (1) and (2) of the definition 2.5 implies that $\forall\ (x, t) \in \mathbb{Z} \times \mathbb{R}_+$ it exists a unique $z \in \mathbb{Z}$ such that $(z, 0) \rightsquigarrow_i (x, t)$: it is called the ancestor of $(x, t)$ for the $i$-th cultural feature.

**DEFINITION 2.6.** For every $(x, t) \in \mathbb{Z} \times \mathbb{R}_+$ we denote with $a_t(x, i)$ the *ancestor* of $(x, t)$ for the $i$-th cultural feature, that is the unique $z \in \mathbb{Z}$ such that $(z, 0) \rightsquigarrow_i (x, t)$.

Note that in particular it holds

$$X_t^i(x) = X_0^i(z) \quad \text{whenever} \quad (z, 0) \rightsquigarrow_i (x, t).$$

Actually, this relation holds in general for the dual process in (2.6): given $t > 0$, for all $s \in [0, t]$

$$X_t^i(x) = X_{t-s}^i(y) \qquad \forall y \in \hat{Y}_s(x, t).$$

Figure 2.2 is an example of dual paths, and going backward in time until time $s = 0$ one can find the ancestor.

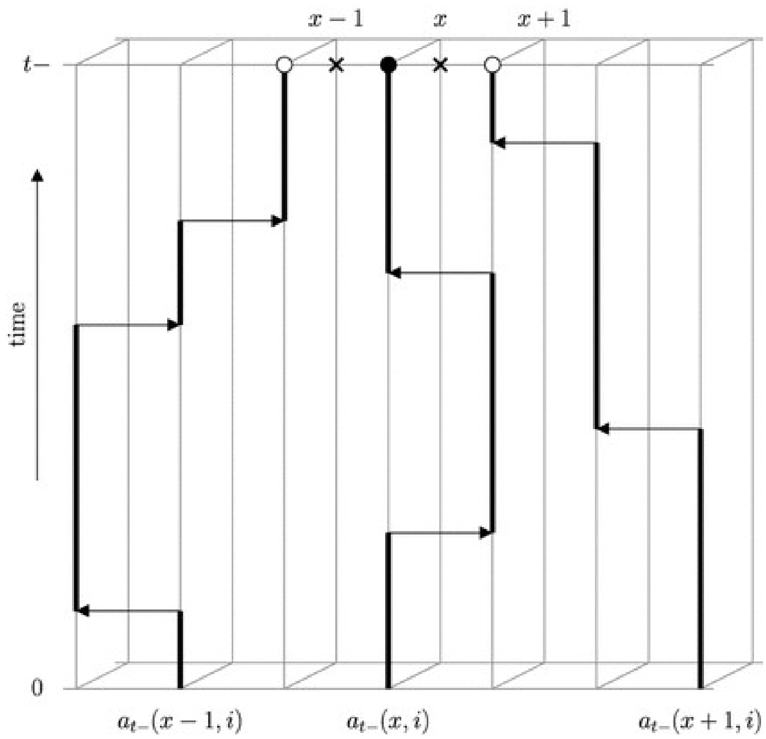


Figure 2.2: Graphical representation of the Axelrod model and ancestors.

## 2.2 Theorems

**THEOREM 2.7.** *Starting from a translation invariant product measure in which the cultures appear with positive probability, the 2-states 2-features Axelrod model clusters.*

We will see that, besides proving clustering, also fluctuation holds in this case, and it is a direct consequence of the site recurrence property proved in the first chapter.

Actually, there is a generalization of the previous result, when the number of states is 2:

**THEOREM 2.8.** *Starting from a translation invariant product measure in which the cultures at different individuals are i.i.d and appear with the same positive probability, the 2-states F-features Axelrod model clusters.*

Theorem 2.8 implies theorem 2.7, since the 2-2 case is a particular case of the $(2-F)$ one, but we want to prove it separately, to show that the 2-2 case can be deduced from properties of the basic voter model.

For the next results, we introduce $\omega : \mathbb{N}^2 \to \mathbb{Q}$, where

$$\omega(q,F) = q\left(1-\frac{1}{q}\right)^F - F\left(1-\frac{1}{q}\right).$$

This function has actually a particular meaning, that we will understand in the proof of the following theorem, where we exploit the positive condition on $\omega$. More precisely,

**THEOREM 2.9.** *Assume that $\omega(q, F) > 0$ and to start from the initial distribution $\pi_0$. Then, fixation occurs and clustering does not occur.*

Also in this case, there is an improvement of the previous result. Note that $\omega(3, 2) = 0$, so theorem 2.9 can not be applied. In spite of this,

**THEOREM 2.10.** *The system with $F = 2$ and $q = 3$ fixates and coexists.*

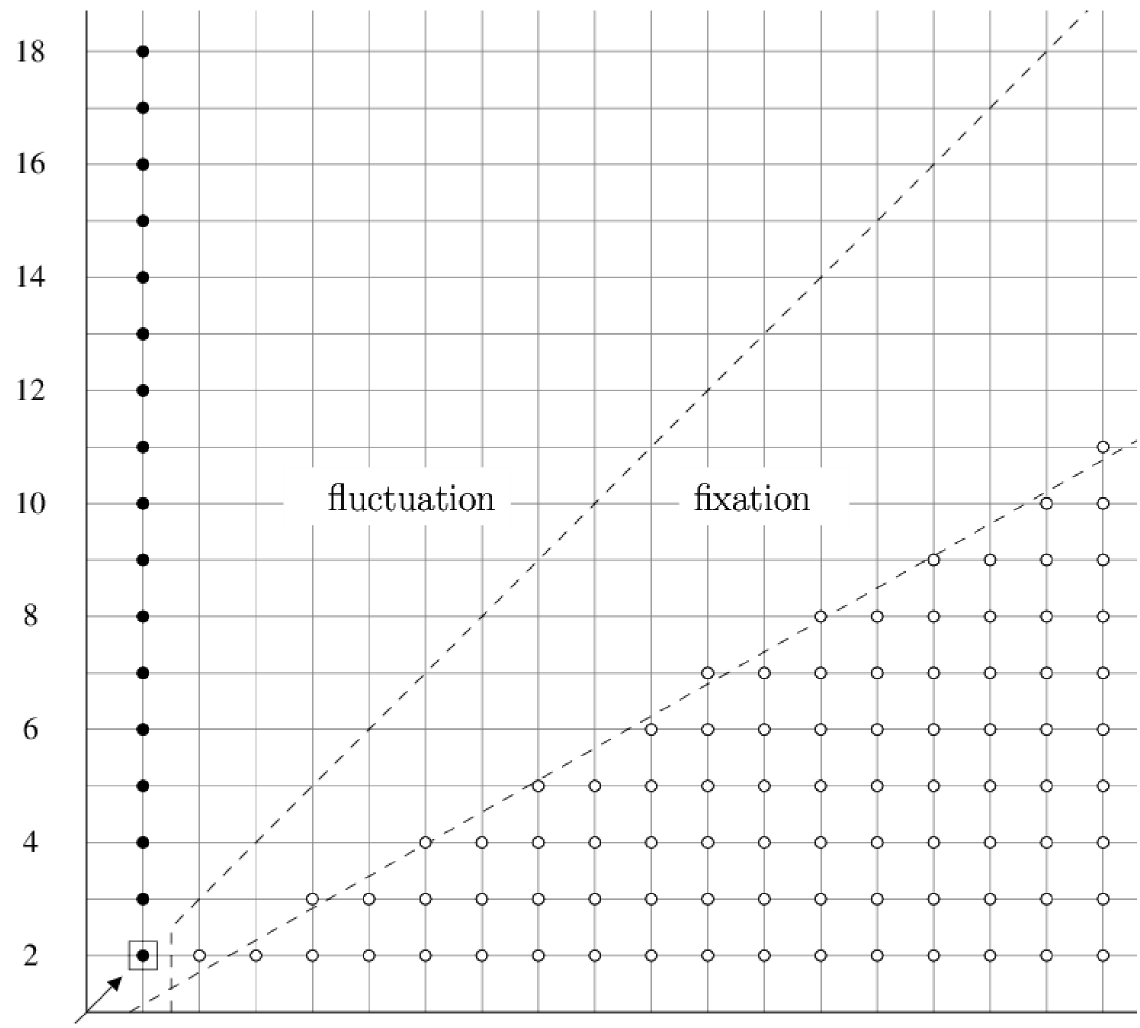


Figure 2.3: Parameter region of the $F - q$ plane that satisfies $\omega(q, F) > 0$.

As the picture suggets, the region identified by the condition in theorem 2.9 is almost equal to the set of parameters below a certain straight line. Since it goes through the origin, to identify it it remains to find its slope $c$. Replacing $F = cq$,

$$\frac{\omega(q, F)}{q} = \left(1 - \frac{1}{q}\right)^{cq} - c\left(1 - \frac{1}{q}\right) \xrightarrow[q\to\infty]{} e^{-c} - c.$$

If $e^{-c} = c$, we find

$$\begin{aligned}
(cq-1)\ln\left(1-\frac{1}{q}\right) - \ln(c) &= (1-cq)\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{1}{q}\right)^n + c \\
&= \sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{1}{q}\right)^n - \sum_{n=0}^{\infty}\frac{c}{n+1}\left(\frac{1}{q}\right)^n + c \\
&= \sum_{n=1}^{\infty}\left(\frac{n(1-c)+1}{n(n+1)}\right)\left(\frac{1}{q}\right)^n > 0.
\end{aligned}$$

This means that

$$cq\ln\left(1-\frac{1}{q}\right) > \ln(c) + \ln\left(1-\frac{1}{q}\right) \Leftrightarrow \left(1-\frac{1}{q}\right)^{cq} > c\left(1-\frac{1}{q}\right) \Leftrightarrow \omega(q,F) > 0$$

so $c$ satisfying $e^{-c} = c$, ($c \approx 0,567$), is the slope.

The next sections are dedicated to the proofs of the previous theorems.

## 2.3 Proof of theorem 2.7

In the case $F = q = 2$ the expression of (2.1) becomes

$$\Omega_{ax} f(X) = \frac{1}{4}\sum_{x\in\mathbb{Z}}\sum_{x\sim y}\sum_{i\neq j}\mathbb{1}\{X^i(x) \neq X^i(y)\}\mathbb{1}\{X^j(x) = X^j(y)\}[f(X^i_{y\to x}) - f(X)]$$

because, given that $x$ and $y$ disagree at least in one feature, the only case in which $F(x,y) \neq 0$ is when they agree exactly in one feature (the other one), and in this case $F(x,y) = \frac{1}{2}$, so $\frac{F(x,y)}{1-F(x,y)} = 1$.

Motivated by the fact that there is a product of two indicator functions, we let, for every $x \in \mathbb{Z}$,

$$Y(x) := |X^1(x) - X^2(x)|$$

and we note that the event $\{Y(x) \neq Y(y)\}$ corresponds to exactly one agreement between them. In other words,

$$\{Y(x) \neq Y(y)\} = \bigcup_{i,j\in\{1,2\}:i\neq j}\{X^i(x) \neq X^i(y)\} \cap \{X^j(x) = X^j(y)\}.$$

Infact, if $X(x) = X(y)$ obviously we have that $Y(x) = Y(y)$, but also when both the coordinates $X^1$ and $X^2$ are different we have the same conclusion: since $q = 2$, letting the two states $q_1, q_2$, $|q_1 - q_2| = |q_2 - q_1|$.

The reason for using $Y$ is that the generator becomes

$$\Omega_{vm} f(Y) = \frac{1}{4}\sum_{x\in\mathbb{Z}}\sum_{x\sim y}\mathbb{1}\{Y(x) \neq Y(y)\}[f(Y_{y\to x}) - f(Y)]$$

where as usual

$$Y_{y\to x}(x) = Y(y) \quad \text{and} \quad Y_{y\to x}(z) = Y(z) \quad \text{otherwise.}$$

This means that $\{Y_t : t \geq 0\}$ is a time change of the special linear one-dimension voter model, that run at rate $\frac{1}{2}$. Since, as in the statement of the theorem, we are only interested in the limiting distribution of the Axelrod model, we can speed up $Y$ by a factor two and consider it as an usual voter model.

The proof consists in three parts: we first prove that there is almost sure extinction of 1-sites, then we prove the same for 2-sites and finally we collect together these two results to obtain the theorem.

**LEMMA 2.11.** *There is almost sure extinction of 1-sites, that is*

$$\lim_{t\to+\infty} \mathbb{P}(\xi_t(u) = 1) = 0 \quad \textit{for all } u \in \mathbb{D}.$$

*Proof.* We fix a site $u = x + \frac{1}{2} \in \mathbb{D}$ and we recall that

$$\{\xi_t(u) = 1\} = \{Y_t(x) \neq Y_t(x+1)\}.$$

It follows from the clustering of the special linear one-dimensional voter model that

$$\lim_{t\to+\infty} \mathbb{P}\left(\xi_t(u) = 1\right) = \lim_{t\to+\infty} \mathbb{P}\left(Y_t(x) \neq Y_t(x+1)\right) = 0.$$

□

To prove the almost sure extinction of 2-sites, we use the fact that the one-dimensional voter model is site recurrent, as stated in proposition 1.14.

**LEMMA 2.12.** *There is almost sure extinction of 2-sites, that is*

$$\lim_{t\to+\infty} \mathbb{P}\left(\xi_t(u) = 2\right) = 0 \quad \textit{for all } u \in \mathbb{D}.$$

*Proof.* First of all, since the initial configuration is translation invariant in space, and the evolution rules too, the probability in the statement does not depend on the choice of $u$.

Given a vertex $x \in \mathbb{Z}$ and the pair of sites $x - \frac{1}{2}$ and $x + \frac{1}{2}$, we have that the culture of $x$ flips at a positive rate if and only if at least one of both sites is a 1-site, and when there is an update only one feature changes at that time. So, there are only four possibilities for the pair of sites (obviously, less than symmetry):

$$(1,2) \to (0,1), \quad (1,1) \to (0,2), \quad (1,1) \to (0,0), \quad (1,0) \to (0,1).$$

From this, it follows that the probability of $u$ being a 1-site is nonincreasing, because their number is nonincreasing as function of time: infact, as result of an "interaction" of two 1-sites there is their annihilation, and in the other cases the number of them remains the same. An other consequence is that 2-sites can only be generated by the annihilation of a pair of 1-sites, as in the second possibility above.

Now, given $0 < s < t < +\infty$, we denote with $\Omega_-^{s,t}$ and $\Omega_+^{s,t}$ the following sets:

- $\Omega_-^{s,t} = \{u \in \mathbb{D} : \xi_\tau(u) = 2 \text{ for all } \tau \in (s,t]\}$;
- $\Omega_+^{s,t} = \{u \in \mathbb{D} : \xi_t(u) = 2, \xi_\tau(u) = 1 \text{ for some } \tau \in (s,t)\}$.

In other words, the first set denotes sites that have been lately updated by time s, while the second corresponds to sites that have been lately updated after time s.

In order to estimate the probability that a site belongs to $\Omega_+^{s,t}$, we note that, in the model in which the graph is the torus (with a finite number of vertices, precisely $2N$) it holds

$$\begin{aligned}
p_N := \mathbb{P}(\exists\tau \in (s,t] : \xi_\tau(u) = 1) &= \frac{1}{2N}\sum_{i=-N}^{N-1} \mathbb{P}\Big(\exists\tau \in (s,t] : \xi_\tau(i+\frac{1}{2}) = 1\Big) \\
&= \frac{1}{2N}\,\mathbb{E}\Big[\sup_{\tau\in(s,t]}\sum_{i=-N}^{N-1} \mathbb{1}_{\{\xi_\tau(i+1/2)=1\}}\Big] \\
&\le \frac{1}{2N}\,\mathbb{E}\Big[\sum_{i=-N}^{N-1} \mathbb{1}_{\{\xi_s(i+1/2)=1\}}\Big] \\
&= \mathbb{P}(\xi_s(u) = 1)
\end{aligned}$$

where we use the invariance in translation, thanks to the initial distribution and the evolution rules, and the fact that, as anticipated above, the number of 1-sites is decreasing in time.

Now, if $u \in \Omega_+^{s,t}$, then for some time $\tau \in (s,t]$ site $u$ must verify $\xi_\tau(u) = 1$ and so, by an approximating argument,

$$\mathbb{P}(u \in \Omega_+^{s,t}) \le \lim_{N\to\infty} p_N \le \mathbb{P}(\xi_s(u) = 1)$$

and this, together with lemma 2.11, implies the existence of a time $s$ large such that

$$\mathbb{P}(u \in \Omega_+^{s,t}) \le \epsilon.$$

For $\Omega_-^{s,t}$, we note that there exists $t > s$ such that

$$\begin{aligned}
\mathbb{P}(u = x + \frac{1}{2} \in \Omega_-^{s,t}) &\le \mathbb{P}\big(Y_\tau(x) = Y_\tau(x+1) \text{ for all } \tau \in (s,t)\big) \\
&\le \mathbb{P}\big(Y_\tau(x) = Y_s(x) \text{ for all } \tau \in (s,t)\big) \\
&\le \epsilon
\end{aligned}$$

Infact, the first inequality holds because, since $q = 2$, both $X_\tau^1(x) \ne X_\tau^1(x+1)$ and $X_\tau^2(x) \ne X_\tau^2(x+1)$ implies that $Y_\tau(x) = Y_\tau(x+1)$; the second one holds because if $Y_\tau(x) = Y_\tau(x+1)$ $\forall\tau \in (s,t)$, then $Y_\tau(x)$ must be equal to its initial value $Y_s(x)$, since no interactions can be happened in the time interval $(s,t)$; finally, the last one is a direct consequence of 1.14. Combining the previous estimates, for all $\epsilon > 0$ there exists a time $s > 0$ and a time $t > s$ such that

$$\mathbb{P}(\xi_t(u) = 2) = \mathbb{P}(u \in \Omega_+^{s,t}) + \mathbb{P}(u \in \Omega_-^{s,t}) \le 2\epsilon$$

and this establishes that the limit is zero. □

We now simply collect together the previous two results: fix $x < y$ in $\mathbb{Z}$, and let $k := |x-y|$, $z_i = x + i$ and $u_i = z_i + \frac{1}{2}$ for every $i = 0, ..., k-1$. We have:

$$\begin{aligned} \lim_{t\to+\infty} \mathbb{P}\big(X_t(x) \neq X_t(y)\big) &\leq \lim_{t\to+\infty} \mathbb{P}\big(X_t(z_i) \neq X_t(z_{i+1}) \text{ for some } i = 0, ..., k-1\big) \\ &\leq \lim_{t\to+\infty} \sum_{i=0}^{k-1} \mathbb{P}\big(\xi_t(u_i) = 2 \text{ or } \xi_t(u_i) = 1\big) = 0 \end{aligned} \tag{2.7}$$

because the sum is a finite sum.

**REMARK 4.** *Note that the site recurrence property given by the proposition 1.14 is equivalent to fluctuation of the system, so in this case we have proved that the system clusters and fluctuates.*

## 2.4 Proof of theorem 2.8

The approach in this section is the same of the previous: keeping track of the disagreements between adjacent vertices rather than their culture; infact, consensus (clustering) of the Axelrod model is equivalent to extinction of interfaces.

The proof consists as usual in more parts, divided in subsections. The strategy is to prove almost surely extinction of $j$-sites, for all $j \geq 1$, and then use the same argument as in (2.7). To do this, we need some preliminary results.

### The process cannot become frozen

The target of this section is to prove the following proposition, that states that no site can remain an $F$-site forever.

**PROPOSITION 2.13.** *Let $u \in \mathbb{D}$ be an $F$-site at time $t$. Then, defined $T := \inf\{s > t : \xi_s(u) < F\}$, it holds*

$$T < \infty \quad a.s.$$

The idea is to consider only the sites to the right of $u$, showing that eventually one particle from the right jumps into $u$, unless another left-particle has jumped before. For this reason, let $G$ be the graph induced by $\mathbb{N}$, and suppose the left-most site $\frac{1}{2}$ be an $F$-site at time 0, while every level of every other edge has initially a particle with probability $1/2$, independently from the others. In this context, let

$$p := \mathbb{P}\big(\xi_s(\{0,1\}) = F \quad \forall s > 0\big).$$

The aim is to prove that $p = 0$, deducing then the result of the proposition. For this, we need two preliminary lemmas, the first of which is similar to the proposition, but it requires a stronger assumption, as in the following definition.

**DEFINITION 2.14.** Given $u, v \in \mathbb{D}$ with $u < v$, the interval $\{u, u+1, ..., v\}$ is said to be *active* at time $t$ if the numbers of particles at two different levels are different in parity, that is $\exists\ i, j \in \{1, ..., F\}$ with $i \neq j$ such that

$$\sum_{z=u}^{v} \xi_t(z, i) \neq \sum_{z=u}^{v} \xi_t(z, j) \bmod 2.$$

**LEMMA 2.15.** *Assume that the interval $\{u, ..., v\}$ is active at time $t$, and $\xi_t(u) = F = \xi_t(v)$. Defined $\tau := \inf\{s > t : \xi_s(u) + \xi_s(v) < 2F\}$, it holds*

$$\tau < \infty \ a.s.$$

*Proof.* Assume, by contradiction, that $\xi_s(u) = F = \xi_s(v)$ for every time $s > t$. Since, under this assumption, at the boundary there are no events involving particles, annihilation events happen only in the interior of the interval, and in these cases particles annihilate in pairs. It follows that the parity of the number of particles at each level is the same at every later time, and this means that the interval $\{u, ..., v\}$ is active at every time $s > t$.
This implies that, at every time $s > t$, the interval contains at least one site that is neither a 0-site nor an $F$-site: indeed, if all sites are 0-sites or $F$-sites, the number of particles at each level is the same, and this contraddicts the fact that the interval is active. Therefore, for every $s > t$, the interval contains at least one active particle at time $s$. The contradiction is now reached: this particle will hit one of the boundaries in an almost surely finite time, because, since it is active, it jumps at a positive rate. □

**LEMMA 2.16.** *There exists a sequence of random times $t_0 < t_1 < \cdots < \infty$ with $t_0 = t$ such that, denoting $A_k$ as the event that a particle at site $u$ is annihilated at some time $s \in (t_{k-1}, t_k]$, it holds*

$$\mathbb{P}\left(A_k \middle| \bigcap_{n=1}^{k-1} A_n^c\right) \geq \frac{p}{2} \tag{2.8}$$

*for every $k > 0$.*

*Proof.* The idea is to prove (2.8) by induction on $k$, but to do this we have to introduce some other helpful objects.

For all $w \in \mathbb{D}$ with $w \geq u$ and for all $s \geq 0$, let $\mathcal{F}_w(s)$ be the $\sigma$-algebra generated by the graphical representation through time $s$ and over the spacial interval $\{u, ..., w\}$, that is

$$\begin{aligned}\mathcal{F}_w(s) \quad = \quad & \sigma(\xi_0(v,i), (N_{v,i}(\tau) : 0 \leq \tau \leq s), (B_{v,i}(n) : 1 \leq n \leq N_{v,i}(s)), \\ & (U_{v,i}(n) : 1 \leq n \leq N_{v,i}(s)) : u \leq v \leq w, i = 1, ..., F).\end{aligned}$$

We define a *stopping pair* as a pair of random variables $(V, T)$ that satisfies:

i) $V$ is $\{u, u+1, ...\}$-valued;

ii) $T$ is $[0, +\infty)$-valued;

iii) for every $w \in \mathbb{D}$ with $w \geq u$ and for every $s \geq 0$, $\{V \leq w\} \cap \{T \leq s\} \in \mathcal{F}_w(s)$.

Given a stopping pair $(V, T)$, define $\mathcal{F}_V(T)$ to be the $\sigma$-algebra representing the information of the graphical representation through time $T$ and over the spatial interval $\{u, ..., V\}$, that is the $\sigma$-algebra of all events $A$ such that

$$A \cap \{V \leq w\} \cap \{T \leq s\} \in \mathcal{F}_w(s) \quad \text{for all } w \in \mathbb{D}, w \leq u \text{ and for all } s \geq 0.$$

In order to prove (2.8) by induction, we want to prove that for every $k \geq 1$ there exist a random site $v_k$ and a random time $t_{k-1}$ such that:

h1) $N_{v_k,i}(t_{k-1}) = N_{v_k+1,i}(t_{k-1}) = N_{v_k+2,i}(t_{k-1}) = 0$ for $i = 1, ..., F$;

h2) $(v_k + 2, t_{k-1})$ is a stopping pair and $A_1, ..., A_{k-1} \in \mathcal{F}_{v_k+2}(t_{k-1})$

and then to prove that

$$\mathbb{P}(A_k|\mathcal{F}_{v_k+2}(t_{k-1})) \geq \frac{p}{2}. \tag{2.9}$$

It will follow from (2.9) that

$$\mathbb{P}\left(A_n \middle| \bigcap_{k=1}^{n-1} A_k^c\right) \geq \frac{p}{2} \quad \text{because} \bigcap_{k=1}^{n-1} A_k^c \in \mathcal{F}_{v_k+2}(t_{k-1}).$$

Certainly, for $k = 0$ it is true: take $t_0 = t$ and

$$v_1 = \min\{w > u : N_{w,i}(t) = N_{w+1,i}(t) = N_{w+2,i}(t) = 0 \ \forall i = 1, ..., F\}.$$

Note that, by Borel-Cantelli, $v_1$ is almost surely finite. The pair $(v_1 + 2, t_0)$ is a stopping pair, and conditions h1 and h2 are satisfied.

Fixed $k$, we want to prove the existence of $v_{k+1}$ and $t_k$ as above. Condition h1 implies that there is no arrow starting at either site $v_k$ or $v_k + 1$ or $v_k + 2$ by time $t_{k-1}$, so in particular particles starting to the right of $v_k$ do not reach $v_k$ by time $t_{k-1}$ and particles starting in $\{u, ..., v_k\}$ do not reach $v_k + 1$ by time $t_{k-1}$.

Partition out the infinite interval $\{u, u + 1, ...\}$ into subintervals of the same length $d = v_k - u + 3$, as $\{u, ..., v_k + 2\}$. Precisely, for $j \geq 0$ we let

$$B_{j,k} = \{jv_k + (1 - j)u + 3j, ..., (j + 1)v_k - ju + 3j + 2\}$$

because the initial site is $u + jd = jv_k + (1 - j)u + 3j$ and the final one is $u + (j + 1)d - 1 = (j + 1)v_k - ju + 3j + 2$.

Now, let $J_k$ be the smallest positive integer $j$ that satisfies:

j1) $N_{w,i}(t_{k-1}) = 0$ for all $w \in B_{j,k}$ and $i = 1, ..., F$;

j2) for $m = 0, ..., v_k - u + 1$ and $i = 1, ..., F$ it holds translation or reflection, where

- translation means $\xi_0(jv_k - (j - 1)u + 3j + m, i) = \xi_{t_{k-1}}(u + m, i)$;
- reflection means $\xi_0((j + 1)v_k - ju + 3j + 2 - m, i) = \xi_{t_{k-1}}(u + m, i)$.

Condition j2) relates the initial configuration of $B_{J_k,k}$ and configuration of $B_{0,k}$ at time $t_{k-1}$: in the first case, the configurations are the same, while in the second they are one the mirror image of the other, but in both cases the right-most site of $B_{J_k,k}$ is excluded, as the left-most one (resp. the right-most one) in the translation case (resp. in the reflection case). Note that $J_k$ is almost surely finite.

The first step is to prove that the probability of a reflection is at least one half. Translation or reflection events depend only on the initial configuration of the sites in $B_{J_k,k}$, and recall that the initial configuration is the one in which all cultures at different sites are independent and they appear with the same probability. Because Poisson processes are independent of the initial configuration, knowing that j1) holds for a particular $j$ provides no information about the initial configuration in $B_{j,k}$. Moreover, h2) implies that the initial configuration to the right of $v_k + 2$ is independent of the site $v_k$ and time $t_{k-1}$. So, conditional on the values of $v_k$ and $t_{k-1}$ and on the event that j1) holds for a particular $j$, all possible values for $\xi_0(w, i)$ are equally likely, for $w \in B_{j,k}$ and $i = 1, ..., F$. In particular, the probabilities of translation and reflection are the same, because being the same configuration as the one in the interval $B_{0,k}$ at time $t_{k-1}$ excluding the right-most site is as likely as being the mirror image of the configuration in the interval $B_{0,k}$ at time $t_{k-1}$, escluding the right-most site. Because either translation or reflection must occur, and their intersection has a positive probability (indeed, both will occur if the configuration at time $t_{k-1}$ between $u$ and $v_k + 1$ is a mirror image of itself), the probability of a reflection must be larger than one half.

Now, consider the process conditioned to the event $J_k = j$ and the fact that a reflection occurs. Furthermore, let

$$r_k := (j + 1)v_k - ju + 3j + 2$$

be the right-most site of $B_{j,k}$ and

$$l_k := jv_k - (j - 1)u + 3j$$

be the left-most site of $B_{j,k}$, and consider the truncated process, that is the process in the subinterval $\Gamma_k = \{u, ..., r_k\}$. By condition h1) and by reflection, particles at $v_k + 1$ and at $l_k + 1$ have no interactions with other particles before time $t_{k-1}$, so the configuration of particles at these two sites is the same at time $t_{k-1}$ and at time 0. So, the probability that the numbers of particles at two given levels of $v_k + 1$ and $l_k + 1$ have the same parity is exactly one half, another time due to the initial configuration. In other words, the interval $\Gamma_k$ is an active interval with probability at least one half, since in definition 2.4 it is required only the existence of a pair of levels $i$ and $j$. It follows that the probability that both a reflection occurs and the interval $\Gamma_k$ is active is at least $1/2 \times 1/2 = 1/4$.

In case a reflection does not occur, or the interval is not active, we can repartition $\{u, u + 1, ...\}$ into subintervals of length $(|J_k| + 1)d$, starting with $B_{0,k,2} = \Gamma_k$ instead of $B_{0,k}$. We define $J_{k,2}$ to be the smallest integer satisfying conditions j1) and j2), with respect to this new partition. Applying another time the previous arguments, after a finite number of steps $n$ we will find a partition such that reflection occurs and the interval $\Gamma_{k,n-1} = \{u, ..., r_{k,n-1}\}$ is active. Actually, what we prove with this construction is that $n$, the number of steps that we have to wait for a realization, is stochastically smaller than a geometric random variable of parameter $1/4$. Since the last part of the proof does not depend on the value of $n$, we drop the subscript that refers to it.

Now, we want to prove that the law of the truncated process is the same as the law of its mirror image. First, note that, since no particles in the interval $B_{j,k}$ can jump before time $t_{k-1}$, as required in condition j1), the configuration in this interval is the same at time $t_{k-1}$ and at time zero. In addition, reflection implies that

$$\xi_{t_{k-1}}(u + m, i) = \xi_{t_{k-1}}(r_k - m, i) \ \forall m = 0, ..., v_k - u + 1, \ \forall i = 1, ..., F. \tag{2.10}$$

Since there is no arrow starting or pointing at either $v_k + 1$ or $l_k + 1$ by time $t_{k-1}$, the particles at $I_k := \{v_k + 1, ..., l_k + 1\}$ at time zero evolve independently of the particles outside $I_k$, at time zero. Since the graphical representation is independent of the initial configuration, the law of the graphical representation in $I_k$ is equal to the law of its mirror image by time $t_{k-1}$. Furthermore, conditional on $J_k = j$, all the initial configurations in the intervals $B_{i,k}$ for $i = 1, ..., j-1$, except the two configurations described in condition j2) that have probability zero, have all the same probability. This means that each possible initial configuration in $B_{i,k}$ is as likely as its mirror image, so the law of the initial configuration in $I_k$ is the same as the law of its mirror image. And this also implies that the law of the configuration in $I_k$ at all times until $t_{k-1}$ is equal to the law of its mirror image, since the they depend only on the initial configuration and the graphical representation in this interval. The remaining part of the interval $\Gamma_k$ not in $I_k$ is symmetric, as stated in (2.10), so it also follows that the law of the configuration in $\Gamma_k$ at time $t_{k-1}$ is the same as the law of its mirror image. But also the law of the graphical representation after time $t_{k-1}$ is the same as its mirror image, so the law of the truncated process after time $t_{k-1}$ is the same as the law of its mirror image. In particular, denoting as $D_k$ the analog of $A_k$ for the truncated process restricted to $\Gamma_k$, i.e. the event that eventually in time there is a change to $u$ before a change to $r_k$, it holds

$$\mathbb{P}(D_k|\mathcal{F}_{v_k+2}(t_{k-1})) = \frac{1}{2}. \tag{2.11}$$

To complete, we define $t_k$ to be the smallest time at which there is a change either to site $u$ or to site $r_k$.

Now return to the scenario in which only sites to the left of $u$ are removed, so that some particle could jump onto $r_k$ from its right. Thinking of $r_k$ as the left-most site in the definition of $p$ (particles of $r_k$ are frozen until at least $t_k$), there is a probability larger than $p$ that the site $r_k$ does not change (so, no particle from $r_k + 1$ jumps onto $r_k$) until time $t_k$, because $p$ is the probability that the left-most site does not change at every time. This, together with (2.11), means that there is a probability of at least $p/2$ that site $u$ will change at time $t_k$, that is exactly (2.9). We choice $v_{k+1}$ respecting the property h1), so

$$v_{k+1} = \min\{w > r_k : N_{w,i}(t_k) = N_{w+1,i}(t_k) = N_{w+2,i}(t_k) = 0 \ \forall i = 1, ..., F\}$$

that is a.s. finite. Since no particle to the right of $v_{k+1}$ can reach $v_{k+1}$ by time $t_k$, the event $\{v_{k+1}+2 \leq w\} \cap \{t_k \leq s\}$ depends only on the initial configuration and the graphical representation through time $s$ and over the interval $\{u, ..., w\}$, so it belongs to $\mathcal{F}_w(s)$, that is $(v_{k+1} + 2, t_k)$ is a stopping pair. Finally, $A_k \in \mathcal{F}_{v_{k+1}+2}(t_k)$, because again no particle to the right of $v_{k+1}$ can reach $v_{k+1}$ by time $t_k$.

□

**Proof of Proposition 2.13.** As mentioned above, consider the graph in which all sites to the left of $u$ are removed. Seeking again a contradiction, as in lemma 2.15, note that the thesis of the proposition is actually equivalent to

$$\mathbb{P}\big(\xi_s(u) = F \text{ for all } s > t\big) = 0,$$

so assume that there is a positive probability that the site $u$ never changes after time $t$. This implies that $p > 0$, in fact:

- if $t = 0$ then the probability that site $u$ never changes after time $t$ is exactly $p$, so there is nothing to prove;

- if $t > 0$ it is possible to reduce to the previous case. We first claim that with probability one there are infinitely many $v > u$ such that, for every $i = 1, ..., F$ it holds
$$N_{v,i}(t) = 0 = N_{v+1,i}(t).$$
Indeed, divide the sites to the right of $u$ into groups of 2. Every event of the form $\{N_{v,i}(t) = 0 \ \forall i = 1.,,,F\}$ has positive probability, equal to $\mathbb{P}(\{X > t\})^F$, where $X$ is an esponential random variable of parameter 1. It follows that, for $n \geq 1$, the events
$$E_n := \{N_{u+2n-1,i}(t) = 0 = N_{u+2n,i}(t) \ \forall i = 1, .., F\}$$
have all positive probability, and $(E_n)_{n\geq 1}$ is a family of independent events. By Borel-Cantelli, with probability 1, infinitely many $E_n$ occurs. Call $V$ the set of such sites. Letting
$$B_v = \big\{(\xi_s(u) = F \ \forall s > t \text{ and } N_{v,i}(t) = 0 = N_{v+1,i}(t) \ \forall i = 1, ..., F)\big\}$$
by intersection with the event in the statement one has
$$\mathbb{P}\left(\bigcup_{v\in V} B_v\right) > 0$$
and since it is a countable union, there must exists a site $\bar{v} > u$ such that $\mathbb{P}(B_{\bar{v}}) > 0$. The fact that $N_{\bar{v},i}(t) = 0 = N_{\bar{v}+1,i}(t)$ means exactly that there is no arrow connecting them until time $t$, so the evolution of the process on the interval $\{u, ..., \bar{v}\}$ under this condition is independent of its evolution on $\{\bar{v} + 1, ...\}$, until time $t$. Since there are only finitely many possible configurations for the sites in the first interval (indeed they are in finite number), with an argument similar to the one used before there must exist numbers $c_{z,i} \in \{0, 1\}$ for $z = u + 1, ..., \bar{v}$ and $i = 1, ..., F$ such that
$$\mathbb{P}\big(B_{\bar{v}} \cap \{\xi_t(z, i) = c_{z,i} \ \forall z = u + 1, ..., \bar{v}, \forall i = 1, ..., F\}\big) > 0.$$
For $z = u$ let $c_{u,i} = 1$ for every $i = 1, ..., F$, because the assumption is that $u$ is an $F$-site at time $t$. For the initial distribution and Poisson process rules, there is a positive probability that
$$\xi_0(z, i) = c_{z,i} \text{ and } N_{z,i}(t) = 0 \ \forall z = u, ..., \bar{v}, \ \forall i = 1, ..., F$$
so
$$\begin{aligned} q := \mathbb{P}(\xi_s(u) = F \ \forall s > t, \xi_0(z, i) = c_{z,i} \ \forall u \leq z \leq \bar{v} \text{ and } i = 1, ..., F,\\ N_{z,i}(t) = 0 \ \forall u \leq z \leq \bar{v} + 1) > 0.\end{aligned}$$
But
$$q \leq \mathbb{P}(\xi_s(u) = F \ \forall s > 0, \xi_0(z, i) = 1 \ \forall i = 1, ..., F) \leq p$$
and so $p$ must be positive.

So in all cases $p > 0$. But this gives a contradiction: by (2.8), one has

$$\frac{P\left(\bigcap_{k=1}^{n} A_k^c\right)}{P\left(\bigcap_{k=1}^{n-1} A_k^c\right)} \leq 1 - \frac{p}{2}$$

so, letting $a_n = P\left(\bigcap_{k=1}^{n} A_k^c\right)$,

$$a_n \leq a_0 \left(1 - \frac{p}{2}\right)^n \xrightarrow[n\to\infty]{} 0.$$

This means that

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n^c\right) = 0$$

that is equivalent to

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = 1$$

which said that, with probability one, eventually some particle at site $u$ will be annihilated. This contradiction implies the result. □

**Extinction of particles**

In this section we prove that each site is eventually a 0-site, that is equivalent to almost sure extinction of particles. In other words, for all $u \in \mathbb{D}$,

$$\lim_{t\to\infty} \mathbb{P}(\xi_t(u) = 0) = 1.$$

As always, the proof is divided into steps. The first lemma is preliminary to the two other ones, which separate the active and the frozen case.

**LEMMA 2.17.** *The following limits exist and do not depend on the choice of* $u \in \mathbb{D}$*:*

1. $\lim_{t\to\infty} \mathbb{E}[\xi_t(u)\mathbb{1}\{\xi_t(u) < F\}]$;
2. $\lim_{t\to\infty} \mathbb{E}[\xi_t(u)\mathbb{1}\{\xi_t(u) = F\}]$.

*Proof.* Since the initial configuration, in which the cultures at different sites are i.i.d. and at a given site every possible culture appears with the same probability, is invariant in space, and the evolution rules are invariant in space too, one has that, $\forall\ u, v \in \mathbb{D}$ and $\forall\ j = 1, ..., F$,

$$\mathbb{P}(\xi_t(u) = j) = \mathbb{P}(\xi_t(v) = j)$$

and so the upper and lower limits of the expected values in the statement do not depend on the choice of the site $u$. Fix $u \in \mathbb{D}$; given times $s < t$, note that

$$\xi_t(u) - \xi_s(u) = J_1(u) - J_2(u) - A(u)$$

where:

- $J_1(u)$ is the number of particles that jump onto the site $u$ in the time window $(s, t]$;

- $J_2(u)$ is the number of particles that jump away from the site $u$ in the time window $(s,t]$;
- $A(u)$ is the number of particles that annihilate at site $u$ in the time window $(s,t]$.

Exploiting another time the fact that the dynamic is symmetrical, i.e. the fact that particles jump to their left or to their right with same probability $1/2$, taking the expected value of $J_1$ and $J_2$ it holds

$$\mathbb{E}[J_1(u)] = \mathbb{E}[J_2(u)]$$

and so, by linearity of the expected value,

$$\mathbb{E}[\xi_t(u)] - \mathbb{E}[\xi_s(u)] = -\mathbb{E}[A(u)] \leq 0. \tag{2.12}$$

It follows that the expected value of particles at site $u$ is a non-increasing function of time, and so it exists

$$L := \lim_{t\to\infty} \mathbb{E}[\xi_t(u)].$$

This allows us to prove only that the first of the two limits exists and does not depend on the choice of $u$, because

$$\mathbb{E}[\xi_t(u)\mathbb{1}\{\xi_t(u) = F\}] = \mathbb{E}[\xi_t(u)] - \mathbb{E}[\xi_t(u)\mathbb{1}\{\xi_t(u) < F\}]$$

and from the first one we deduce the second one.

To show that the first limit exists, we proceed by contradiction assuming that it doesn't exist. This means exactly that $L_- < L_+$, where $L_-$ and $L_+$ are respectively the limit inferior and the limit superior of the same function.
Given an arbitrary $\epsilon > 0$, by limit definition there exists $t_0 > 0$ such that, for all $t \geq t_0$,

$$L - \epsilon < \mathbb{E}[\xi_t(u)] < L + \epsilon, \tag{2.13}$$

and by definition of limit superior/inferior, there exist two increasing sequences of times $(t_n)_{n\in\mathbb{N}}$ and $(s_n)_{n\in\mathbb{N}}$ tending to $+\infty$ such that:

- $\mathbb{E}[\xi_{s_n}(u)\mathbb{1}\{\xi_{s_n}(u) < F\}] < L_- + \epsilon$;
- $\mathbb{E}[\xi_{t_n}(u)\mathbb{1}\{\xi_{t_n}(u) < F\}] > L_+ - \epsilon$.

Clearly the two sequence can be chosen in such a way that

$$t_0 < s_1 < t_1 < s_2 < \cdots < s_n < t_n < \cdots < \infty \tag{2.14}$$

because the sequences are constructed choosing an element in an arbitrary neighbourhood of $+\infty$, so that's enough to change neighbourhood restricting it only to the right of the previous element in (2.14).
Collecting together the previous extimates, one has

$$\begin{aligned} & \mathbb{E}[\xi_{t_n}(u)\mathbb{1}\{\xi_{t_n}(u) = F\}] - \mathbb{E}[\xi_{s_n}(u)\mathbb{1}\{\xi_{s_n}(u) = F\}] \\ = \; & \mathbb{E}[\xi_{t_n}(u)] - \mathbb{E}[\xi_{s_n}(u)] - \mathbb{E}[\xi_{t_n}(u)\mathbb{1}\{\xi_{t_n}(u) < F\}] + \mathbb{E}[\xi_{s_n}(u)\mathbb{1}\{\xi_{s_n}(u) < F\}] \\ \leq \; & (L+\epsilon) - (L-\epsilon) - (L_+ - \epsilon) + (L_- + \epsilon) = L_- - L_+ + 4\epsilon \end{aligned}$$

Since until now $\epsilon$ was arbitrary, we fix $\epsilon = \frac{L_+ - L_-}{5}$ so that

$$\mathbb{E}[\xi_{t_n}(u)\mathbb{1}\{\xi_{t_n}(u) = F\}] - \mathbb{E}[\xi_{s_n}(u)\mathbb{1}\{\xi_{s_n}(u) = F\}] < -\epsilon.$$

Moreover, by the fact that in those events the value of $\xi$ is obliged to be equal to $F$, also the following inequality holds:

$$\mathbb{P}(\xi_{t_n}(u) = F) - \mathbb{P}(\xi_{s_n}(u) = F) < -\frac{\epsilon}{F}.$$

Now remember that $q = 2$, so when two particles interact, they cannot create a new particle in an other site (there is only annihilation, nor coalescing events). This means that new active particles can only result from the annihilation of two particles, one active and one frozen, corresponding to the event in which an $F$-site becomes an $(F-1)$-site. Therefore, in a finite interval, the number of particles annihilated between times $s$ and $t$ is at least equal to twice the number of times an $F$-site becomes an $(F-1)$-site between times $s$ and $t$. The last quantity is in turn at least equal to twice the number of $F$-sites at time $s$ minus twice the number of $F$-sites at time $t$, because surely these ones during the time interval $(s,t]$ have become $(F-1)$-sites. Using (2.12) and the last extimates, one found, for every $s < t$,

$$\mathbb{E}[\xi_t(u)] - \mathbb{E}[\xi_s(u)] \leq 2(\mathbb{P}(\xi_t(u) = F) - \mathbb{P}(\xi_s(u) = F))$$

that, with $s = s_n$ and $t = t_n$ becomes

$$\mathbb{E}[\xi_{t_n}(u)] - \mathbb{E}[\xi_{s_n}(u)] \leq -2\frac{\epsilon}{F}. \tag{2.15}$$

Taking now $t = t_F$ and $s = s_1$, applying (2.15) $F$ times we get

$$\begin{aligned}
\mathbb{E}[\xi_{t_F}(u)] - \mathbb{E}[\xi_{s_1}(u)] &= \sum_{n=1}^{F} (\mathbb{E}[\xi_{t_n}(u)] - \mathbb{E}[\xi_{s_n}(u)]) + \sum_{n=1}^{F-1} \left(\mathbb{E}[\xi_{s_{n+1}}(u)] - \mathbb{E}[\xi_{t_n}(u)]\right) \\
&\leq \sum_{n=1}^{F} (\mathbb{E}[\xi_{t_n}(u)] - \mathbb{E}[\xi_{s_n}(u)]) \leq -2\epsilon
\end{aligned}$$

because for the second sum, all terms are non positive, as stated in (2.12). But applying (2.13), since $t_F, s_1 > t_0$ one has

$$\mathbb{E}[\xi_{t_F}(u)] - \mathbb{E}[\xi_{s_1}(u)] > (L - \epsilon) - (L + \epsilon) = 2\epsilon$$

and this leads to a contradiction. □

**REMARK 5.** *Note that, in the first limit of the previous lemma, the expected value can be written also as*

$$\mathbb{E}\left[\xi_t(u)\mathbb{1}\{0 < \xi_t(u) < F\}\right]$$

*because when $\xi_t(u) = 0$, there is nothing to add. This equivalent way will be useful in the next result.*

The next lemma deals with active particles. More precisely, it states almost all we want to prove clustering, in fact it rules out the frozen case, that is treated below.

**LEMMA 2.18.** *There is extinction of active particles, i.e.*

$$\lim_{t\to\infty} \mathbb{P}(0 < \xi_t(u) < F) = 0 \quad \textit{for all } u \in \mathbb{D}.$$

*Proof.* By the previous lemma and remark, we know that it exists

$$L := \lim_{t\to\infty} \mathbb{E}\left[\xi_t(u)\mathbb{1}\{0 < \xi_t(u) < F\}\right].$$

In order to seek a contradiction, we assume that this limit is strictly positive: $L > 0$. Then, for all $t$ large enough it holds

$$\mathbb{E}\left[\xi_t(u)\mathbb{1}\{0 < \xi_t(u) < F\}\right] \geq \frac{L}{2}. \tag{2.16}$$

Now, remember that active particles evolve according to symmetric random walks that run at positive rate, and in one dimension they are recurrent. For this, given any two active particles at time $s$ at the same level, there exists an almost surely finite random time $t > s$ such that at time $t$ one of the following conditions is verified:

- one of these two particles annihilates, due to a collision with a third one;
- one of these two particles becomes frozen;
- both the particles annihilate each other.

This suggests to consider the pair of annihilating and freezing events. We say that at site $u$ at time $t$ there is:

1. an Annihilating event if $\xi_t(u) + \xi_t(u+1) = \xi_{t-}(u) + \xi_{t-}(u+1) - 2$;
2. a Freezing event if $\xi_t(u) = F$ and $\xi_{t-}(u) = F - 1$.

In annihilating events two particles destroy one each other, and so the number of them decreases by 2, while in freezing events site $u$, being at $t-$ an $(F-1)$-site, becomes an $F$-site. Denote with $\mathcal{A}_t(u)$ and $\mathcal{F}_t(u)$ the number of annihilating and freezing events that occur at site $u$ by time $t$.
Since, as in (2.16), for large times $t$ the expected value of active particles at a site is strictly positive, bounded away from zero, and since every active particle will be sooner or later part of an annihilating or a freezing event, one has that the number of annihilating or freezing events by time $t$ tends to infinity, when $t$ tends to infinity, that is

$$\lim_{t\to\infty} \mathbb{E}[\mathcal{A}_t(u) + \mathcal{F}_t(u)] = +\infty. \tag{2.17}$$

Now, we prove that (2.17) is not true, obtaining a contradiction.
For the annihilating events, note that every particle can get annihilated only once, and annihilation happens with pairs of particles. This means that, fixed two times and a finite interval, the number of particles destroyed in this space-time window is exactly twice the number of annihilation events, occuring in this space-time window. Taking the expected

value per site, and remembering that every site has at most $F$ particles, we obtain, for all $t \geq 0$,

$$\frac{1}{2N}\,\mathbb{E}\Big[\sum_{i=-N}^{N-1} \mathcal{A}_t\big(i+\frac{1}{2}\big)\Big] \leq \frac{1}{2N}\frac{2N\cdot F}{2} = \frac{F}{2}$$

so, by an approximation argument and translation invariance,

$$\mathbb{E}[\mathcal{A}_t(u)] = \lim_{N\to\infty} \frac{1}{2N}\,\mathbb{E}\Big[\sum_{i=-N}^{N-1} \mathcal{A}_t\big(i+\frac{1}{2}\big)\Big] \leq \frac{F}{2}.$$

For freezing events, we separate once-freezing case from more than two freezing case. In fact, when there is only once freezing event, clearly the expected value is 1; when there are at least two freezing events, it means that, after the first one, the site from an $F$-site becomes again an $(F-1)$-site, and so there is an annihilating event. This event can occurs not only at $u$, but also at $u-1$: it depends on which direction (left or right) the other particle arrives. In other words,

$$\mathbb{E}[\mathcal{F}_t(u)] \leq 1 + \mathbb{E}[\mathcal{A}_t(u)] + \mathbb{E}[\mathcal{A}_t(u-1)] \leq 1 + F.$$

Collecting the two last extimates,

$$\mathbb{E}[\mathcal{A}_t(u) + \mathcal{F}_t(u)] \leq 1 + \frac{3}{2}F < +\infty$$

that contradicts (2.17). In particular, $L = 0$ and

$$\lim_{t\to\infty} \mathbb{P}(0 < \xi_t(u) < F) \leq \lim_{t\to\infty} \mathbb{E}[\xi_t(u)\mathbb{1}\{0 < \xi_t(u) < F\}] = 0.$$

□

The next result deals with frozen particles. In particular,

**LEMMA 2.19.** *There is extinction of frozen particles, i.e.*

$$\lim_{t\to\infty} \mathbb{P}(\xi_t(u) = F) = 0 \quad \text{for all } u \in \mathbb{D}.$$

*Proof.* By lemma 2.17, it exists

$$\lim_{t\to\infty} \mathbb{E}[\xi_t(u)\mathbb{1}\{\xi_t(u) = F\}]$$

and it does not depend on the choice of $u \in \mathbb{D}$. But, since the value of $\xi_t(u)$ is forced to be equal to $F$, the previous limit is actually equal to

$$\lim_{t\to\infty} \mathbb{E}[\xi_t(u)\mathbb{1}\{\xi_t(u) = F\}] = F \times \lim_{t\to\infty} \mathbb{P}(\xi_t(u) = F)$$

so the limit in the statement exists and does not depend on the choice of $u$. Furthermore, by lemma 2.18, for every $\epsilon > 0$ it exists a time $t_0 > 0$ such that

$$\mathbb{E}[\xi_t(u)\mathbb{1}\{0 < \xi_t(u) < F\}] < \epsilon/2 \quad \text{for all } t \geq t_0. \tag{2.18}$$

Recall now proposition 2.13 in an equivalent way: if $u$ is an $F$-site at time $t$, then

$$\mathbb{P}(\xi_s(u) = F \ \forall s > t) = 0$$

that is, by the continuity of probability, for any increasing sequence of times $(s_n)_{n\in\mathbb{N}}$ that tends to infinity,

$$\lim_{n\to\infty} \mathbb{P}(\xi_s(u) = F \ \forall s \in (t, s_n)) = 0. \tag{2.19}$$

Applying (2.19) inductively, given $t_0$ as in (2.18), it exists an increasing sequence of times $(t_n)_{n\in\mathbb{N}}$ such that, for every $n \geq 0$,

$$\mathbb{P}(\xi_s(u) \neq F \ \exists s \in (t_n, t_{n+1}) | \xi_{t_n}(u) = F) > \frac{1}{2}. \tag{2.20}$$

In order to estimate the expected number of particles killed per site $u$ between times $t_n$ and $t_{n+1}$, we divide the case in which $u$ at time $t_n$ is an $F$-site from the case in which it is a $j$-site, with $0 \leq j < F$. In the first one, we note that, as in (2.20), with probability at least one half the site will be part of an annihilation event of two particles, so in this case (with this probability) the number of killed particles is at least 2. In the other case, simply remember that the number of killed particles is non-negative (for example, see (2.12)). This implies that, for all $n \geq 0$,

$$\mathbb{E}[\xi_{t_n}(u)] - \mathbb{E}[\xi_{t_{n+1}}(u)] \geq 2 \times \frac{1}{2} \times \mathbb{P}(\xi_{t_n}(u) = F) = \mathbb{P}(\xi_{t_n}(u) = F). \tag{2.21}$$

Collecting (2.21) and (2.18), we have the following chain of inequalities:

$$\begin{aligned}
F \cdot \mathbb{P}(\xi_{t_{n+1}}(u) = F) &\leq \sum_{i=0}^{F} i \cdot \mathbb{P}(\xi_{t_{n+1}}(u) = i) \\
&= \mathbb{E}[\xi_{t_{n+1}}(u)] \\
&\leq \mathbb{E}[\xi_{t_n}(u)] - \mathbb{P}(\xi_{t_n}(u) = F) \\
&= F \cdot \mathbb{P}(\xi_{t_n}(u) = F) + \mathbb{E}[\xi_{t_n}(u) \mathbb{1}\{0 < \xi_{t_n}(u) < F\}] \\
&\quad - \mathbb{P}(\xi_{t_n}(u) = F) \\
&\leq F \cdot \mathbb{P}(\xi_{t_n}(u) = F) + \frac{\epsilon}{2} - \mathbb{P}(\xi_{t_n}(u) = F) \\
&= (F-1) \cdot \mathbb{P}(\xi_{t_n}(u) = F) + \frac{\epsilon}{2}.
\end{aligned}$$

Calling $x_n = \mathbb{P}(\xi_{t_n}(u) = F)$, what is stated before is

$$x_{n+1} \leq \left(1 - \frac{1}{F}\right) x_n + \frac{\epsilon}{2F},$$

and by induction

$$\begin{aligned}
x_n &\leq \left(1-\frac{1}{F}\right)^n x_0 + \frac{\epsilon}{2F}\sum_{i=0}^{n-1}\left(1-\frac{1}{F}\right)^i \\
&\leq \left(1-\frac{1}{F}\right)^n + \frac{\epsilon}{2F}\frac{1-\left(1-\frac{1}{F}\right)^n}{1-\left(1-\frac{1}{F}\right)} \\
&\leq \left(1-\frac{1}{F}\right)^n + \frac{\epsilon}{2F}\frac{1}{1-\left(1-\frac{1}{F}\right)} \\
&= \left(1-\frac{1}{F}\right)^n + \frac{\epsilon}{2}.
\end{aligned}$$

where in the second inequality we use the fact that $x_0 \leq 1$ because it is a probability.

Since $F \geq 2$, $\left(1-\frac{1}{F}\right)^n \to 0$ and so $x_n \leq \epsilon$ for all $n$ sufficiently large. And this is exactly the thesis: since the limit in the statement exists, it is equal to the limit in a particular sequence of times going to $+\infty$. Taking $(t_n)_n$ as this sequence,

$$\lim_{t\to\infty}\mathbb{P}(\xi_t(u)=F) = \lim_{n\to\infty}\mathbb{P}(\xi_{t_n}(u)=F) \leq \epsilon$$

that is what we want. □

Now, as in the last part of the proof of the first theorem, it is sufficent to collect together what we have proved until now, that is

$$\lim_{t\to\infty}\mathbb{P}(\xi_t(u)=0)=1 \quad \forall u \in \mathbb{D}. \tag{2.22}$$

In fact, to see that (2.22) implies Theorem 2, note that, for every $x, y \in \mathbb{Z}$ with $x < y$ it holds

$$\begin{aligned}
\lim_{t\to\infty}\mathbb{P}(X_t(x)\neq X_t(y)) &\leq \lim_{t\to\infty}\mathbb{P}(X_t(x+n-1)\neq X_t(x+n)\ \exists\ 1\leq n\leq y-x) \\
&= \lim_{t\to\infty}\mathbb{P}(\xi_t(x+n-\frac{1}{2})\neq 0\ \exists\ 1\leq n\leq y-x) \\
&\leq \sum_{n=1}^{y-x}\lim_{t\to\infty}\mathbb{P}(\xi_t(x+n-\frac{1}{2})\neq 0) = 0
\end{aligned}$$

because, if vertices $x$ and $y$ have different cultural opinions at time $t$, there must exists an $n = 1, ..., y-x$ such that vertices $x+n-1$ and $x+n$ have different cultural opinions. This complete the proof.

## 2.5 Proof of theorem 2.9

We start this section by proving a preliminary result, introduced in section 2.1. We use the notation

$$(u,i) \longrightarrow (u+1,i) \quad \text{at time } t$$

meaning that $(u,i)$ and $(u+1,i)$ are occupied at time $t-$ and the particle at $(u,i)$ jumps one unit to the right at time $t$, so that there is a collision at $(u+1,i)$ at time $t$.

**LEMMA 2.20.** *Conditionally on the realization of the system of random walks until time $t-$, and the event $(u,i) \longrightarrow (u+1,i)$, it holds:*

i) *the collision results in an annihilation event with probability $(q-1)^{-1}$;*

ii) *the collision results in a coalescing event with probability $(q-2)\cdot(q-1)^{-1}$.*

**REMARK 6.** *Note that this fact holds in the case $q=2$, too. Indeed, as we know, when there are only two possible states, a collision of two particles can only annihilate them. Obviously, the interest of the lemma is when $q>2$.*

*Proof.* First of all, since a collision can result either in a coalescing or an annihilating event, it's enough to prove the first statement, because the second probability in that case must be equal to $1-(q-1)^{-1}=(q-2)\cdot(q-1)^{-1}$.

To prove the first, let $x := u+1/2 \in \mathbb{Z}$. For the second property that $i$-paths have to satisfy, they cannot cross each other, otherwise there will be an active $i$-arrow that points at some intermediate point. So

$$a_s(x-1,i) \le a_s(x,i) \le a_s(x+1,i) \quad \forall s \ge 0. \tag{2.23}$$

where $a_s(x,i)$ is the ancestor of $x$ for the $i$-th feature, as in definition (2.6). Moreover, since there is a collision $(u,i) \longrightarrow (u+1,i)$, there are two particles at time $t-$, one at $(u,i)$ and one at $(u+1,i)$, so

$$\begin{cases} X_0^i\big(a_{t-}(x-1,i)\big) = X_{t-}^i(x-1) \neq X_{t-}^i(x) = X_0^i\big(a_{t-}(x,i)\big) \\ X_0^i\big(a_{t-}(x+1,i)\big) = X_{t-}^i(x+1) \neq X_{t-}^i(x) = X_0^i\big(a_{t-}(x,i)\big). \end{cases} \tag{2.24}$$

From (2.24), we deduce that, conditional on the event $(u,i) \longrightarrow (u+1,i)$ at time $t$, the three ancestors in (2.23) must be different, i.e.

$$a_s(x-1,i) < a_s(x,i) < a_s(x+1,i) \quad \forall s < t.$$

This means that

$$(u,i) \xrightarrow{annih} (u+1,i) \Leftrightarrow \begin{cases} (u,i) \longrightarrow (u+1,i) \\ X_0^i\big(a_{t-}(x-1,i)\big) = X_0^i\big(a_{t-}(x+1,i)\big). \end{cases}$$

Therefore, the conditional probability of being an annihilating collision depends only on the ancestors, and so only on the initial condition. Since we take $\pi_0$ as the initial configuration, this probability is

$$p := \mathbb{P}(U_1 = U_3 | U_1 \neq U_2 \cap U_3 \neq U_2)$$

where we denote with $U_1, U_2, U_3$ three independent uniform variables over $\{1,...,q\}$. We have

$$\begin{aligned} p &= \frac{\mathbb{P}(U_1 = U_3, U_1 \neq U_2, U_3 \neq U_2)}{\mathbb{P}(U_1 \neq U_2, U_3 \neq U_2)} \\ &= \frac{\sum_{j=1}^q \mathbb{P}(U_1 = U_3, U_1 \neq j, U_3 \neq j)\mathbb{P}(U_2 = j)}{\sum_{j=1}^q \mathbb{P}(U_1 \neq j, U_3 \neq j)\mathbb{P}(U_2 = j)} \\ &= \frac{\frac{1}{q}\sum_{j=1}^q \mathbb{P}(U_1 = U_3, U_1 \neq j, U_3 \neq j)}{\frac{1}{q}\sum_{j=1}^q \mathbb{P}(U_1 \neq j, U_3 \neq j)} \\ &= \frac{(q-1)/q^2}{(q-1)^2/q^2} = \frac{1}{q-1}. \end{aligned}$$

□

The next lemma gives a sufficient condition for fixation, based on properties of active $i$-paths. We will apply this later to prove fixation. Define, for all $(z,i) \in \mathbb{Z} \times \{1, ..., F\}$,

$$T(z,i) := \inf\{t \geq 0 : (z,0) \rightsquigarrow_i (0,t)\}$$

that is the first time an active $i$-path that originates from the $z$ hits the origin.

**LEMMA 2.21.** *Consider the Axelrod model with initial distribution $\pi_0$. If in addition*

$$\lim_{n\to\infty} \mathbb{P}\big(T(z,i) < \infty \textit{ for some } z < -n \textit{ and some } i = 1, ..., F\big) = 0 \tag{2.25}$$

*then the system fixates.*

*Proof.* In order to prove fixation, we prove that the probability that one vertex (we choose $x = 0$, and this is the reason for the definition of $T$ as above) changes infinitely often her culture is zero. This implies fixation, since the graph is connected. In order to estimate that probability, we have to define stopping times that consider when the vertex at the origin changes the state of her $i$-th cultural feature. So, we let, for every $i = 1, ..., F$ and $j \geq 1$,

$$\tau_{i,j} = \inf\{t > \tau_{i,j-1} : X_t^i(0) \neq X_{\tau_{i,j-1}}^i(0)\}$$

with obviously $\tau_{i,0} = 0$, and

$$B_i := \{\tau_{i,j} < \infty \text{ for all } j \geq 1\}.$$

The individual at the origin changes her culture infinitely often if and only if at least one of the events $B_i$ occurs, so fixation is equivalent to prove that every event $B_i$ has probability zero.

We have also to consider $a_{i,j}$, the ancestor of vertex 0 at this stopping times $\tau_{i,j}$ for the $i$-th feature, because we want to use the hypothesis. Indeed, defined

$$G_{i,n} := \{|a_{i,j}| < n \text{ for all } j \geq 1\}$$

we have that

$$\mathbb{P}\left(\bigcup_n G_{i,n}\right) = 1.$$

This follows from the fact that (2.25) together with reflection symmetry implies that

$$\lim_{n\to\infty} \mathbb{P}\big(T(z,i) < \infty \text{ for some } z > n \text{ and some } i = 1, ..., F\big) = 0$$

because there is a natural bijection between $(-\infty, n)$ and $(n, +\infty)$, and the initial distribution is symmetrical.

From this,

$$\begin{aligned}\mathbb{P}\left(\bigcup_{i=1}^{F} B_i\right) &\leq \sum_{i=1}^{F} \mathbb{P}(B_i)\\ &= \sum_{i=1}^{F} \mathbb{P}\left(B_i \cap \left(\bigcup_{n} G_{i,n}\right)\right)\\ &= \sum_{i=1}^{F} \mathbb{P}\left(\bigcup_{n} (B_i \cap G_{i,n})\right)\\ &\leq \sum_{i=1}^{F}\sum_{n=1}^{\infty} \mathbb{P}(B_i \cap G_{i,n}).\end{aligned}$$

It remains to prove that, for every $i = 1, ..., F$ and $n \geq 1$

$$\mathbb{P}(B_i \cap G_{i,n}) = 0. \tag{2.26}$$

To prove this, we repeat the construction as in Proposition 1.14. We denote with

$$I_t(x,i) := \{z \in \mathbb{Z} : (x,0) \rightsquigarrow_i (z,t)\}$$

the set of descendants of $x$ at time $t$ for the $i$-th feature, and with $M_t(x,i)$ we denote its cardinality. $M_t$ is a martingale, with constantly expected value one, so by the martingale convergence theorem it exists $M_\infty$ such that

$$\mathbb{P}\big(M_t(x,i) \xrightarrow[t\to\infty]{} M_\infty(x,i)\big) = 1,$$

where $\mathbb{E}[|M_\infty(x,i)|] < \infty$. Remembering that the process $M_t$ is an integer valued process,

$$\tau_M(x,i) := \inf\{t > 0 : M_t(x,i) = M_\infty(x,i)\} < \infty \quad \text{with probability 1.}$$

Since the process is governed by Poisson processes, simultaneous updates occur with probability zero. This implies that also $I_t$ enjoys the same properties of $M_t$, that is, with probability one:

- $\mathbb{P}\big(I_t(x,i) \xrightarrow[t\to\infty]{} I_\infty(x,i)\big) = 1$, where $I_\infty(x,i)$ is a random interval almost surely finite;
- $\tau_I(x,i) := \inf\{t > 0 : I_t(x,i) = I_\infty(x,i)\} < \infty$.

And from this it follows the conclusion: indeed, fixed $i$ and $n$, if the individual at the origin changes infinitely often her $i$-th cultural feature, and all its ancestors belong to $(-n,n)$, then at least one of its ancestors $\bar{x}$ has to satisfy $\tau_I(\bar{x},i) = \infty$, otherwise there is a finite number of ancestors, all with a finite $\tau_I$, and so that individual at the origin should change only finitely many times her $i$-th cultural feature. So,

$$\mathbb{P}(B_i \cap G_{i,n}) = \mathbb{P}(\tau_I(x,i) = \infty \text{ for some } x \in (-n,n)) = 0$$

as stated above, and this is exactly (2.26). □

**Proof of Theorem 2.9.** In order to prove fixation, recalling the previous lemma, it suffices to prove that the probability of the event as in (2.25) tends to zero as $n \to \infty$. We denote that event $H_n$. Note that $H_n$ can be described in an equivalent way: denoted by $\tau_n$ the infimum of $T(z,i)$, with conditions as in (2.25), namely

$$\tau_n := \inf\{T(z,i) : z < -n, i = 1, ..., F\},$$

one has

$$H_n = \{\tau_n < \infty\}.$$

Indeed, it exists a $z < -n$ and an index $i = 1, ..., F$ for which $T(z,i)$ is finite if and only if the infimum of $T(z,i)$, with $z$ and $i$ satisfying the same conditions, is finite. $\tau_n$ is exactly the first time an active $i$-path, that originates from $\bar{z} \in (-\infty, -n)$, hits the origin. We have to consider the minimum and the maximum of vertices connected by a generalized active path to the origin, that are

$$\begin{aligned} z_- &:= \min\{z \in \mathbb{Z} : (z,0) \rightsquigarrow (0,\tau_n)\} \leq \bar{z} < -n, \\ z_+ &:= \max\{z \in \mathbb{Z} : (z,0) \rightsquigarrow (0,\sigma) \text{ for some } \sigma < \tau_n\} \geq 0. \end{aligned} \tag{2.27}$$

In general, $z_- < \bar{z}$, since the former is defined from active $i$-paths, while the latter is defined from generalized active paths, that are concatenations of active $i$-paths, with different values of $i$. Define $I = (z_-, z_+)$.

Note that each blockade that is initially in $I$ must have been destroyed by time $\tau_n$, since surely a particle that derives from a generalized active path creates an annihilating event. This particle, however, can not be initially outside the interval $I$, because active particles can not jump into the space-time region given in (2.27): if one particle came in, then $z_-$ or $z_+$ would not be the minimum or the maximum, as they are defined. So, on the event $H_n$, all the blockades initially in $I$ must have been destroyed by either active particles initially in $I$ or active particles that results from these blockade destructions.

In order to estimate the probability of $H_n$, we define a comparison function $\phi$ in this way:

$$\phi(u) = \begin{cases} -i & \text{if } \xi_0(u) = i \neq F \\ \psi(u) - (F-1) & \text{if } \xi_0(u) = F. \end{cases} \tag{2.28}$$

where $\psi(u)$ are independent geometric random variables with mean $q-1$. The meaning is the following: since the number of collisions to break a blockade is bounded by the number of active particles initially in $I$ or created from destruction of blockades initially in $I$, we define $\phi(u) = -i$ whenever $\xi_0(u) = i \neq F$ to count the former ones. For the second case, we note that the number of collisions required to break a blockade is geometric (by the definition of geometric random variables) and, by lemma 2.20, its mean is $q-1$: in fact, the probability that the blockade breaks is exactly the probability that there is an annihilation event given that there is a collision, and this probability is $(q-1)^{-1}$. The definition in the case $\xi_0(u) = F$ serves to count active particles created from destruction of blockades.

Using the function $\phi$, as in (2.28), and using what said until now,

$$H_n \subseteq \left\{ \sum_{u \in I} \phi(u) \le 0 \right\}.$$

Furthermore, since $(-n, 0) \subset I$,

$$H_n \subseteq \left\{ \sum_{u \in I} \phi(u) \le 0 \right\} \subseteq \bigcup_{\substack{r \ge 0 \\ l < -n}} \left\{ \sum_{u=l}^{r} \phi(u) \le 0 \right\}. \tag{2.29}$$

Remember that there is initially a particle at site $u$ at level $i$ if and only if $u - 1/2$ and $u + 1/2$ have differents $i$-th features, so with probability $1 - \frac{q}{q^2} = 1 - \frac{1}{q}$. Moreover, particles at different levels are independent, so the distribution of $\xi_0(u)$ is Binomial with parameters $F$, the number of cultural features, and $1 - \frac{1}{q}$, the probability that two vertices disagree on a specifical feature. This means that

$$\begin{aligned} \mathbb{E}[\phi(u)] &= (\mathbb{E}[\psi(u)] - (F-1))\mathbb{P}(\xi_0(u) = F) - \sum_{i=0}^{F-1} i \cdot \mathbb{P}(\xi_0(u) = i) \\ &= (\mathbb{E}[\psi(u)] + 1)\mathbb{P}(\xi_0(u) = F) - \sum_{i=0}^{F} i \cdot \mathbb{P}(\xi_0(u) = i) \\ &= q \cdot \mathbb{P}(\xi_0(u) = F) - \mathbb{E}[\xi_0(u)] \\ &= q\left(1 - \frac{1}{q}\right)^F - F\left(1 - \frac{1}{q}\right) = \omega(q, F) \end{aligned}$$

and the reason for the expression of $\omega(q, F)$ is now clear: one expects fixation when $\mathbb{E}[\phi(u)] > 0$, since all $\phi(u)$ are independent.

To make rigorously this idea, we show that the number of collisions required to break all the blockades in a large interval does not deviate too much from its expected value, and for this we need a lemma.

**LEMMA 2.22.** *Let $I_n := (-n, 0) \cap \mathbb{D}$ and assume that $\omega(q, F) > 0$. Then,*

$$\mathbb{P}\left( \sum_{u \in I_n} \phi(u) \le 0 \right) \le \exp(-nc_3)$$

*for a suitable constant $c_3 > 0$, and all $n$ sufficiently large.*

*Proof.* We divide two cases, conditioning to a particular event. For $i = 0, ..., F$, denote with

$$N_i = \sum_{u \in I_n} \mathbb{1}\{\xi_0(u) = i\}$$

the number of sites that are initially $i$-sites. $N_i$ is a random variable with mean $n\mu_i$, where $\mu_i = \mathbb{P}(\xi_0(u) = i)$, because in the interval $I_n$ there are exactly $n$ sites. By large deviation estimates, for all $\epsilon > 0$ it exists $c_4 > 0$ such that for all $i$

$$\mathbb{P}(N_i \notin (\mu_i - \epsilon)n, (\mu_i + \epsilon)n)) \le \exp(-nc_4). \tag{2.30}$$

Let $\Omega = \{(\mu_i - \epsilon)n \leq N_i \leq (\mu_i + \epsilon)n$ for all $i = 0, ..., F\}$. On $\Omega$, it holds

$$\frac{1}{n}\sum_{i=0}^{F-1} iN_i \leq \sum_{i=0}^{F-1} i(\mu_i + \epsilon) = \mathbb{E}[\xi_0(u)] - F\mu_F + C\epsilon$$

where $C = \frac{(F-1)F}{2}$. Since initially there are at least $(\mu_F - \epsilon)n$ $F$-sites, the sum of $\phi(u)$ restricted to $F$-sites is at least the sum of $\phi(u)$ for $K$ sites, where $K = \lfloor(\mu_F - \epsilon)n\rfloor$. So,

$$\begin{aligned}
\mathbb{P}\left(\sum_{u\in I_n}\phi(u) \leq 0 \Big| \Omega\right) &\leq \mathbb{P}\left(\sum_{u\in I_K}(\psi(u) - (F-1)) \leq (\mathbb{E}[\xi_0(u)] - F\mu_F + C\epsilon)n\right) \\
&\leq \mathbb{P}(\sum_{u\in I_K}\psi(u) \leq (\mathbb{E}[\xi_0(u)] - \mu_F + (C - F + 1)\epsilon)n) \qquad (2.31)
\end{aligned}$$

In the second inequality we use the fact that $K \leq (\mu_F - \epsilon)n$, since it is its integer part, to obtain

$$(-F\mu_F + C\epsilon)n + (F-1)(\mu_F - \epsilon)n = (-\mu_F + (C - F + 1)\epsilon)n$$

as in (2.31). The left hand side part of (2.31) looks like the estimate in theorem 1.20 (precisely, the second one), so we have to estimate another time the right hand side, and for this we now use the hypothesis $\omega(q, F) > 0$. In fact, taking $\epsilon \leq \omega(q, F)/(C - F + \mu_F + q)$, one has

$$\begin{aligned}
\mathbb{E}[\xi_0(u)] - \mu_F + (C - F + 1)\epsilon &= (q-1)\mu_F - \omega(q, F) + (C - F + 1)\epsilon \\
&\leq (q-1)\mu_F - \epsilon(\mu_F + q - 1) \\
&\leq (q-1)\mu_F - \epsilon(\mu_F + q - 1) + \epsilon^2 \\
&= (q - 1 - \epsilon)(\mu_F - \epsilon)
\end{aligned}$$

so, using theorem 1.20 and (2.31), we get

$$\mathbb{P}\left(\sum_{u\in I_n}\phi(u) \leq 0 \Big| \Omega\right) \leq \mathbb{P}\left(\sum_{u\in I_K}\psi(u) \leq (q - 1 - \epsilon)K\right) \leq \exp(-Kc_2) \qquad (2.32)$$

for all $K$ sufficiently large, because another time we stress the fact that $K$ is the integer part of $(\mu_F - \epsilon)n$. The result follows intersecting with $\Omega$ and its complementary: called $A = \{\sum_{u\in I_n}\phi(u) \leq 0\}$, using (2.30) and (2.32),

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(A \cap \Omega) + \mathbb{P}(A \cap \Omega^c) \\
&\leq \mathbb{P}(A|\Omega) + \mathbb{P}(\Omega^c) \\
&\leq \exp(-c_2(\mu_F - \epsilon)n) + \sum_{i=0}^{F}\mathbb{P}(N_i \notin ((\mu_i - \epsilon)n, (\mu_i + \epsilon)n) \\
&= \exp(-c_2(\mu_F - \epsilon)n) + (F+1)\exp(-nc_4)
\end{aligned}$$

for all $N$ sufficiently large. $\square$

The conclusion now follows applying in inclusion (2.29) what said in Lemma 2.22:

$$
\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(H_n) &\leq \lim_{n\to\infty} \sum_{l<-n} \sum_{r\geq 0} \mathbb{P}\left(\sum_{u=l}^{r} \phi(u) \leq 0\right) \\
&\leq \lim_{n\to\infty} \sum_{l<-n} \sum_{r\geq 0} e^{-c_3(r-l)} \\
&\leq \left(\sum_{r\geq 0} e^{-c_3 r}\right) \lim_{n\to\infty} \sum_{l<-n} e^{c_3 l} \\
&= \frac{1}{1-e^{-c_3}} \cdot \lim_{n\to\infty} \frac{e^{-c_3(n+1)}}{1-e^{-c_3}} = 0.
\end{aligned}
$$

□

**REMARK 7.** *Note that we prove also coexistence for these values of parameters. Indeed, with notation as at the beginning of this section, the time $\tau_n$ is almost surely infinite, and with a positive probability some of the blockades in the interval $I$ is not destroyed before time $\tau_n$, so there is survival of blockades. And this fact excludes clustering, getting the system trapped in a highly fragmented configuration.*

## 2.6 Proof of theorem 2.10

The idea is to exploit lemma 2.21 another time, as in the proof of theorem 2.9, but the comparison function defined in that contest is now not very useful, since $\mathbb{E}[\phi(u)] = \omega(3,2) = 0$. Indeed, with the previous comparison function, all particles initially active have a weight of -1, that corresponds to the worst case scenario in which the active particle hits a blockade. Actually, an active particle can also form a new blockade or hit another active particle. The possible outcomes are four:

1. if an active particle hits a blockade, it is assigned, as before, a weight of -1;

2. if an active particle annihilates with another active particle, no collisions with a blockade can be generated, because the particle (actually, the pair of particles) is destroyed, so it is assigned a weight of 0;

3. if an active particle coalesces with another active particle, then as a result of the collision from two particles only one remains, and this one can generated at most only one collision with a blockade, so each particle of this pair is assigned a weight of -1/2;

4. finally, if an active particle forms a blockade with another active particle, then, as before, the pair is assigned a weight of a geometric random variable with mean $q-1$, plus -1.

In view of what said until now, we define a new comparison function $\phi$ in this way:

$$\phi(u) = \begin{cases} 0 & \text{if } \xi_0(u) = 0, \\ \psi(u) - 1 & \text{if } \xi_0(u) = 2, \\ -1 & \text{if } \xi_0(u) = 1 \text{ and the particle at } u \text{ hits a blockade}, \\ -1/2 & \text{otherwise}, \end{cases} \tag{2.33}$$

with $\psi(u)$ again independent geometric random variables with parameter $(q-1)^{-1} = 1/2$ and mean 2. Note that the last case is exactly the case in which $\xi_0(u) = 1$ and the active particle initially at site $u$ either collides with another active particle or forms a blockade with another active particle. As in (2.29), we have

$$H_n \subseteq \bigcup_{\substack{r \geq 0 \\ l < -n}} \left\{ \sum_{u=l}^{r} \phi(u) \leq 0 \right\}. \tag{2.34}$$

To prove that the probability of $H_n$ converges to zero as $n \to \infty$, we imitate the same strategy as in lemma 2.22, with some more details. Actually, the results does not hold only for the $(3-2)$-case but also in general for $(q-2)$-case, with $q \geq 3$, so we propose the proof in the general case.

**LEMMA 2.23.** *Let $I_n = (-n, 0) \cap \mathbb{D}$, $F = 2$ and $q \geq 3$. Then, there exists $c_5 > 0$ such that*

$$\mathbb{P}\left( \sum_{u \in I_n} \phi(u) \leq 0 \right) \leq \exp(-c_5 n)$$

*for all $n$ sufficiently large.*

*Proof.* First, we need to find a lower bound for the initial number of active particles that will collide with another active particle or form a blockade. This because the new comparison function $\phi$ distinguishes between these active particles and the active ones that collides with a blockade. To do so, we divide the lattice $\mathbb{D}$ into countably many pairs of adjacent sites, and we said that an active particle initially at site $u = 2n - 1/2$ is a *good particle* if

$$\xi_0(2n - 1/2) = \xi_0(2n + 1/2) = 1. \tag{2.35}$$

An active particle that is not a good particle is called *bad particle*. Recall that the variables $\xi_0(u)$ are independent binomial random variables with parameters $F = 2$ and $1 - 1/q$, the probability that at a specific level there is initially a particle. We introduce, for $u$ as above, and $v = u + 1$:

- $\nu_0 = \mathbb{P}(\{u, v\}$ is initially occupied by a pair of good particles);
- $\nu_1 = \mathbb{P}(u$ is initially occupied by a bad particle);
- $\nu_2 = \mathbb{P}(u$ is initially a blockade).

Obviously, these probabilities do not depend on the choice of $u$ (and $v$ in the first case). Moreover, the events that pairs of adjacent sites that are non overlapping are occupied by two good particles, or one bad particle, or a blockade, or one bad particle and a blockade, or finally two blockades are independent, we have, with a standard large deviation argument, for $i = 0, 1, 2$,

$$\mathbb{P}(N_i \notin (\nu_i - \epsilon)n) \leq \exp(-c_6 n) \tag{2.36}$$

where $N_0, N_1, N_2$ denote respectively the initial number of good particles, the initial number of bad particles and the initial number of blockades in $I_n$.

Now, note that, given a pair of sites $\{u, u+1\} \subset \mathbb{D}$, in the graphical representation there are exactly six possible arrows that may affect the system of random walks at this pair, that are:

i) $u - 1/2 \to u + 1/2$, that is the individual in the middle imitates her left neighbor;

ii) $u + 3/2 \to u + 1/2$ that is the individual in the middle imitates her right neighbor;

iii) $u + 1/2 \to u - 1/2$ that is the individual on the left imitates her right neighbor, so the individual in the middle;

iv) $u + 1/2 \to u + 3/2$ that is the individual on the right imitates her left neighbor, so the individual in the middle;

v) $u - 3/2 \to u - 1/2$, that is the individual on the left imitates her left neighbor;

vi) $u + 5/2 \to u + 3/2$, that is the individual on the right imitates her right neighbor.

If a pair of sites $\{u, u+1\}$ has initially a pair of good particles, then either collision or the making of a blockade from these particles occurs only when the middle individual imitates one of her neighbors, that is only in the first two cases of the list. Moreover, parts of the graphical representation associated with nonadjacent pairs do not intersect, so the events that the first two arrows in the list appear before any of the other ones are independent for nonadjacent pairs. In particular, the initial number of good particles that either collide or form a blockade, that we denote with $J$, is stochastically larger than a binomial random variable with parameters $n\nu_0/2$ and one third: the first parameter is $n/4$, the number of nonadjacent pairs in the interval $I_n$, times $\nu_0$, the probability that a pair contains a pair of good particles, times 2, the number of such good particles, while the second parameter is 2/6=1/3, the probability that one of the two first arrows in the previous list appear before any of the four other ones. Calling $X$ this binomial variable, it holds $\mathbb{E}[X] = n\nu_0/6$. We want to use again large deviation estimates for binomial variables: note that, for $a > 0$,

$$\mathbb{P}(J \leq a) \leq \mathbb{P}(X \leq a)$$

because if $J \leq a$, then $X \leq J \leq a$. Conditioning to the event that there are "enough" initially good particles, i.e. $N_0 > (\nu_0 - \epsilon)n$, we find

$$\mathbb{P}\left(J \leq (\frac{1}{6} - \epsilon)(\nu_0 - \epsilon)n \middle| N_0 > (\nu_0 - \epsilon)n\right) \leq \exp(-c_7 n) \tag{2.37}$$

for a suitable constant $c_7 > 0$. In analogy with lemma 2.22, let

$$\Omega = \Big\{ J > (\frac{1}{6} - \epsilon)(\nu_0 - \epsilon)n, (\nu_i - \epsilon)n < N_i < (\nu_i + \epsilon)n \ \forall i = 0, 1, 2 \Big\}.$$

Conditioning to $\Omega$, letting $K = \lfloor (\nu_2 - \epsilon)n \rfloor$,

$$\mathbb{P}\left( \sum_{u \in I_n} \phi(u) \Big| \Omega \right) \leq \mathbb{P}\left( \sum_{u \in I_K} (\phi(u) - 1) \leq \frac{1}{2} J + (N_0 + N_1 - J) \right) \tag{2.38}$$

because initially there are $J$ good particles that either collide or form a blockade, and these have weight $1/2$. Moreover, on $\Omega$ it holds

$$\begin{aligned} \frac{1}{2} J + (N_0 + N_1 - J) &= N_0 + N_1 - \frac{1}{2} J \\ &< (\nu_0 + \nu_1 + 2\epsilon)n - \frac{1}{2}(\frac{1}{6} - \epsilon)(\nu_0 - \epsilon)n \\ &= (\frac{11}{12}\nu_0 + \nu_1 + C\epsilon)n \end{aligned}$$

with $C = \frac{25}{12} + \frac{\nu_0}{2} - \frac{\epsilon}{2}$.

Combining (2.38) and the previous estimates, we find

$$\begin{aligned} \mathbb{P}\left( \sum_{u \in I_n} \phi(u) \leq 0 \Big| \right) &\leq \mathbb{P}\left( \sum_{u \in I_K} (\phi(u) - 1) \leq (\frac{11}{12}\nu_0 + \nu_1 + C\epsilon)n \right) \\ &\leq \mathbb{P}\left( \sum_{u \in I_K} \phi(u) \leq \Big( \frac{11}{12}\nu_0 + \nu_1 + \nu_2 + (C-1)\epsilon \Big) n \right) \end{aligned} \tag{2.39}$$

where as usual we use $K \leq (\nu_2 - \epsilon)n$.

We want to use theorem 1.20 for the last term in (2.39), but we have to find a further upper bound for its right hand side: note that, if $Y$ denotes a binomial random variable of parameters 2 and $(1 - 1/q)$,

$$\begin{aligned} (q-2)\nu_2 - \nu_1 - \frac{11}{12}\nu_0 &= (q-2)\mathbb{P}(Y=2) - \mathbb{P}(Y=1)\mathbb{P}(Y \neq 1) - \frac{11}{12}\mathbb{P}(Y=1)^2 \\ &= (q-2)\mathbb{P}(Y=2) - \mathbb{P}(Y=1) + \frac{1}{12}\mathbb{P}(Y=1)^2 \\ &= (q-2)\left(1 - \frac{1}{q}\right)^2 - \frac{2}{q}\left(1 - \frac{1}{q}\right) + \frac{1}{12}\left(\frac{2}{q}\left(1 - \frac{1}{q}\right)\right)^2 \\ &= (q-3)\left(1 - \frac{1}{q}\right) + \frac{1}{3}\left(\frac{1}{q}\left(1 - \frac{1}{q}\right)\right)^2 \\ &\geq \frac{1}{3}\left(\frac{1}{3}\left(1 - \frac{1}{3}\right)\right)^2 = \frac{4}{243} > 0 \end{aligned}$$

for all $q \geq 3$, since the function in the fourth line is increasing in $q$, when $q \geq 3$. Furthermore, since $C > 2$, $\nu_2 + q + C - 2 > 0$, so there exists $\epsilon > 0$ small such that

$$\begin{aligned}
\frac{11}{12}\nu_0 + \nu_1 + \nu_2 + (C-1)\epsilon &= (q-1)\nu_2 - \Big((q-2)\nu_2 - \nu_1 - \frac{11}{12}\nu_0\Big) + (C-1)\epsilon \\
&\leq (q-1)\nu_2 - (\nu_2 + q + C - 2)\epsilon + (C-1)\epsilon \\
&= (q-1)\nu_2 - (\nu_2 + q - 1)\epsilon \\
&\leq (q-1)\nu_2 - (\nu_2 + q - 1)\epsilon + \epsilon^2 = (q-1-\epsilon)(\nu_2 - \epsilon).
\end{aligned}$$

Since $\mathbb{E}[\phi(u)] = q - 1$, theorem 1.20 and (2.39) imply that

$$\mathbb{P}\left(\sum_{u \in I_n} \phi(u) \leq 0 \Big| \Omega\right) \leq \mathbb{P}\left(\sum_{u \in I_K} \phi(u) \leq (q-1-\epsilon)K\right) \leq \exp(-c_2 n) \tag{2.40}$$

for all $K$ sufficiently large, because we use the fact that $K$ is the integer part of $(\nu_2 - \epsilon)n$.

We denote with $A := \{\sum_{u \in I_n} \phi(u) \leq 0\}$; then, combining (2.36), (2.37) and (2.40) we have, as in lemma 2.22,

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(A \cap \Omega) + \mathbb{P}(A \cap \Omega^c) \\
&\leq \mathbb{P}(A|\Omega) + \mathbb{P}(\Omega^c) \\
&\leq \exp\big(-c_2(\nu_2 - \epsilon)n\big) + \exp(-c_7 n) + 3\exp(-c_6 n)
\end{aligned}$$

where, since $\Omega^c$ is a union of two events, one that is $\{J \leq (1/6 - \epsilon)(\nu_0 - \epsilon)n\}$ and the other one that is $\{N_i \notin ((\nu_i - \epsilon)n, (\nu_i + \epsilon)n)\ \exists i = 0, 1, 2\}$. Then we use conditional probability as in (2.37) and the fact that the second term is a union of three events, all with probability bounded as in (2.36), from which it derives the factor three. This gives the thesis. □

As in the previous section, (2.34) and lemma 2.23 imply

$$\lim_{n \to \infty} \mathbb{P}(H_n) \leq \lim_{n \to \infty} \sum_{l < -n} \sum_{r \geq 0} e^{-c_5(r-l)} = 0$$

and this is fixation in the case $F = 2$ and $q = 3$.

**REMARK 8.** *Note that the coexistence derives from the same argument as in remark 7.*

## 2.7 Missing proof of fluctuation in the (F-2) case

In the article [5], it is written that, when $q = 2$, the annihilating symmetric random walks constructed in the previous section have a certain site recurrence property, which is equivalent to fluctuation of the Axelrod model, when starting from $\pi_0$. The difference between the $(2-2)$ case and the $(F-2)$ case, and this is the reason why we divide this easier case from the more general one, is that the $(2-2)$ case can be coupled with the voter model obtained identifying cultures with no features in common, while the more general $(F-2)$ case cannot be coupled with a voter model, since there are more than two features. The site recurrence property, that is the proposition (1.14), is exactly fixation, because as we wrote in remark 4, with probability one, fixed a vertex, there is a finite time such that the configuration on that vertex is different from the current one.

In conclusion, the fluctuation of the Axelrod model when $q = 2$ is not clear, even because in [4] the theorem states only clustering but not fluctuation.

# Chapter 3

# The Axelrod model with one media

In this chapter we want to add an individual to the Axelrod model, that has a function of *media*. This is an individual $v$ that does not change his opinion during the time, and that influences all the other individuals. The case treated is the one in which $V = \mathbb{Z}$ and $E$ are nearest neighbors edges as in the previous chapter. The parameter $\beta$ reflects the intensity of the media influence. In this case, the generator is

$$\begin{aligned} \Omega_m f(X) \quad = \quad & \sum_{x\in V}\sum_{y\sim x}\sum_{i=1}^{F}\frac{1}{2F}\left[\frac{F(x,y)}{1-F(x,y)}\right]\mathbb{1}\{X^i(x)\neq X^i(y)\}\big[f(X^i_{y\to x})-f(X)\big] \\ +\beta \quad & \sum_{x\in V}\sum_{i=1}^{F}\frac{1}{2F}\left[\frac{F(x,v)}{1-F(x,v)}\right]\mathbb{1}\{X^i(x)\neq X^i(v)\}\big[f(X^i_{v\to x})-f(X)\big] \end{aligned} \tag{3.1}$$

The interest is to understand how the parameter $\beta$ influences the dynamics of the model, and if for some value there is a difference of behaviour compared to the one of the Axelrod model without media. We treat the case $F, q = 2$.

## 3.1 Graphical representation

First of all, we have to describe how the graphical representation changes. For each $x \in \mathbb{Z}$, we consider the following Poisson processes:

i) $N_{x,r}$ and $N_{x,l}$ of intensity 1;

ii) $N_{x,m}$ of intensity $\beta$.

Define, for $n \geq 1$, $T_{x,r}(n) := \inf\{t \geq 0 : N_{x,r}(t) = n\}$, $T_{x,l}(n) := \inf\{t \geq 0 : N_{x,l}(t) = n\}$ and $T_{x,m}(n) := \inf\{t \geq 0 : N_{x,m}(t) = n\}$. For every $s = T_{x,r}(n)$ (resp. $s = T_{x,l}(n)$) draw an arrow on the plane $\mathbb{Z} \times \mathbb{R}^+$ from $(x-1, s)$ to $(x, s)$ (resp. from $(x+1, s)$ to $(x, s)$). For every $s = T_{x,m}(n)$ draw a cross on $(x, s)$.

This construction can be used in order to define the trajectories of the process. If there is an arrow from $(x-1, s)$ to $(x, s)$, set $X_s(x) = X_{s-}(x-1)$ if $d\big(X_{s-}(x), X_{s-}(x-1)\big) = 1$, where $d$ is the *Hamming distance* defined by

$$d\big(X_t(x), X_t(y)\big) := \sum_{i=1}^{2} \mathbb{1}\{X_t^i(x) \neq X_t^i(y)\}.$$

Otherwise, set $X_s(x) = X_{s-}(x)$ because the arrow is not actually active, with notation as in the previous chapter. The opinion of the media is set equal to $(1, 2)$, so, for any cross at $(x, s)$ set $X_s(x) = (1, 2)$ if $d\big(X_{s-}(x), (1, 2)\big) = 1$, otherwise $X_s(x) = X_{s-}(x)$. If we are given the initial condition $X_0(x)$, for every $x \in \mathbb{Z}$, then these rules define the above process up to time scaling $\frac{1}{2F}$.

## 3.2 Fixation/fluctuation cases

**THEOREM 3.1.** *Let $\lambda$ be the critical infection rate of the one-dimension contact process, and suppose that the initial condition for the Axelrod model with media is chosen i.i.d. with each of the states having positive probability. Then,*

1. *for $\beta \geq \lambda^{-1}$ the model fixates;*
2. *for $\beta < \lambda^{-1}$ the model fluctuates.*

*Proof.* Define $Y_t(x) := |X_t^1(x) - X_t^2(x)|$. The process $Y$ is Markovian and its evolution can be described using the graphical representation of $X$ as above. Indeed, starting from $(x, t)$ and going backward in time following the arrows (reversed), we run across a unique path and we stop as we meet a cross or we reach time $s = 0$. Such a path is called *dual path*, starting at $(x, t)$. If this dual path ends at $(y, 0)$ then we set $Y_t(x) = Y_0(y)$, otherwise $Y_t(x) = 1$. The last setting derives from the fact that $Y(v) = 1$ and that a cross is an interaction with the media $v$, so the value of $Y$ will be the same of the $v$'s one.

Now, the dynamics of $Y$ can be coupled with that of a contact process $Z$ for which the arrows denote infections and the crosses denote recovery, but the recovery state is 1, instead as usual the zero state. The process $Z$ has recovery rate $\beta$ and infection rate 1, because by coupling the recovery corresponds to a cross, and so to an event of $N_m$, while infection corresponds to an arrow, and so to an event of $N_l$ or $N_r$. It follows that, if $Z_0 = Y_0$, then $Z_t \leq Y_t$, where as usual the order is by components. Infact, if $Z_t(x) = 1$ for some $x \in \mathbb{Z}$ and some $t \geq 0$, then $x$ at time $t$ is healthy, and so his dual path (in the sense of the contact process) ends with a cross or it ends at time 0 with an individual that is initially healthy; this means that $Y_t(x) = 1$ because of the initial condition $Z_0 = Y_0$ and the rules explained above. Note that it can be possible that $Z_t(x) = 0$ and $Y_t(x) = 1$ for some $x \in \mathbb{Z}$ and $t \geq 0$; indeed, suppose that $Y_0(x) = Z_0(x) = 0$ and $Y_0(x+1) = Z_0(x+1) = 1$. If there are no other interactions before they interact, their interaction at time $\tau$ results in changing $Y_\tau(x) = 1$, but $Z_\tau(x) = 0$ because in the contact process they do not interact.

Since we are considering the infection rate fixed (and equal to 1) and the recovery rate variable (and equal to $\beta$), we have that the process $Z$ dies out almost surely if $\beta^{-1} \leq \lambda$, that is if $\beta \geq \lambda^{-1}$. In this case, $Z_t(x) = 1$ definitely (remember that state 1 is healthy) and so $Y_t(x) = 1$ definitely. This further implies that also $X$ fixates: indeed, $Y_t(x) = 1$ means that $X_t(x) = (1, 2)$ or $X_t(x) = (2, 1)$, but $X(x)$ can not jump from $(1, 2)$ to $(2, 1)$ due to the fact that in Axelrod model, as in spin systems, a jump changes only one coordinate.

In the case $\beta < \lambda^{-1}$, consider the process $Z$ as above but with initial condition all zero (all vertices are infected). In this case, with probability 1, it holds the following fact: for every $(x, s) \in \mathbb{Z} \times \mathbb{R}^+$ there exists a time $t > s$ such that $Z_t(x) = 0$. Indeed, note that

$$\mathbb{P}\big(\exists s > 0 : \xi_t(x) = 1 \ \forall t \geq s\big) = 0 \Leftrightarrow \mathbb{P}\big(\exists n > 0 : \xi_t(x) = 1 \ \forall t \geq n\big) = 0$$

and so it's enough to prove that $\mathbb{P}(A_{n,x}) = 0$ for all $x \in \mathbb{Z}$ and $n \in \mathbb{N}$, where $A_{n,x} := \{\exists n > 0 : \xi_t(x) = 1 \ \forall t \geq n\}$. By contradiction, assume that $\mathbb{P}(A_{n,x}) = \epsilon > 0$ and let $y \in \mathbb{Z}, y < x$ such that $\mathbb{P}(B_{n,x,y}) \geq 1 - \delta$, where $B_{n,x,y} := \{r_n(y) < x\}$. For semplicity, we call $A = A_{n,x}$ and $B = B_{n,x,y}$. Such a $y$ exists because $\mathbb{P}(B^c) \leq \mathbb{P}(W \leq n)^{x-y}$, where $W$ is an exponential variable of parameter $\lambda$, and this probability goes to zero when $y \to -\infty$. Choose the initial condition $\xi^-$ in such a way that $\xi_0^-(z) = 0 \ \forall z \leq y$ and $\xi_0^-(z) = 1 \ \forall z > y$. Let, as in [1] pag. 276, $r_t := \max\{x : \eta_t^-(x) = 0\}$. With the choice of $\delta < \epsilon$ one has

$$\mathbb{P}(A \cap B) \geq \epsilon + (1 - \delta) - 1 = \epsilon - \delta > 0.$$

From [1] pag. 283, we know that in the supercritical case $\frac{r_t}{t} \to \alpha > 0$ a.s., that implies $r_t \to +\infty$ a.s. In particular, for almost every $\omega \in A \cap B$, $r_t(y)(\omega) \to +\infty$. By the fact that $r_t$ can move to the right only to one step, it exists a $t > n$ such that $r_t(y)(\omega) = x$. This means that $\xi_t(x)(\omega) = 0$ and in particular that $\omega \notin A$. Since it holds for almost every $\omega \in A \cap B$, this is a contradiction, that gives $\mathbb{P}(A) = 0$. Finally, since the contact process is monotone, as stated in (1.11), the same conclusion holds starting from an initial configuration in which all vertices are infected.

Fix $(x, s)$ and let $t_1$ such that $Z_{t_1}(x) = 0$. Since this vertex is infected, there must exists a dual path from $(x, t_1)$ to $(y_1, 0)$ for some $y_1 \in \mathbb{Z}$, and $Y_{t_1}(x) = Y_0(y_1)$. Now, consider the process that counts the descendants of an individual, that is, for $y \in \mathbb{Z}$

$$M_t(y) := \big|\{x \in \mathbb{Z} : \text{ there exists a dual path from } (x, t) \text{ to } (y, 0)\}\big|.$$

$M_t(y)$ is a positive supermartingale with adsorbing state 0: if $M_s(y) = 0$, then $M_t(y) = 0$ for every $t > s$, otherwise $M_s(y)$ should have been different from zero. To motivate the fact that $M_t$ is a supermartingale note that the generator of the dual process is the operator defined, for every function $f : \mathbb{Z} \to \mathbb{R}$ and for every $A \subset \mathbb{Z}$ finite by

$$\begin{aligned} \mathcal{L}f(A) &= \sum_{x \in A} \sum_{y \sim x, y \notin A} \Bigg( \frac{1}{2+\beta} \big[ \big(f(A \cup \{y\}) - f(A)\big) + \big(f(A \setminus \{x\}) - f(A)\big) \big] \\ &\quad + \frac{\beta}{2+\beta} \big[ f(A \setminus \{x\}) - f(A) \big] \Bigg). \end{aligned}$$

Taking $f$ equal to the cardinality function, one has $\mathcal{L}f(A) \leq 0$ for every $A$, and this means that the cardinality of descendants, as a function of time $t$, is a supermartingale. Since it takes values in $\mathbb{N}$ and it converges almost surely, as the supermartingale convergence theorem states, the only possibility is that it converges almost surely to its adsorbing state zero.

Now, let for $y \in \mathbb{Z}$,

$$\phi(y) := \inf\{t > 0 : M_t(y) = 0\}.$$

By the convergence above, $\phi(y)$ is a stopping time almost surely finite, so let $\bar{t}_2$ such that $M_t(y_1) = 0$ for all $t > \bar{t}_2$. Again, as before, we can find almost surely a time $t_2 > \bar{t}_2$ such that there is a dual path from $(x, t_2)$ to $(y_2, 0)$, for some $y_2 \in \mathbb{Z}$. Note that necessarily $y_2 \neq y_1$, because $M_{t_2}(y_1) = 0$ and so $y_1$ can not be the ancestor of $x$ at time $t_2$.
Iterating this procedure, we find a sequence of increasing times $t_n$ and a sequence of distinct integers $y_n$ such that

$$Y_{t_n}(x) = Y_0(y_n).$$

The initial condition is chosen in such a way that all the states have positive probability, so $Y_0(y_n) \neq Y_s(x)$ for infinitely many $n$'s. This is exactly fluctuation for $Y$, that implies that also $X$ fluctuates.

$\square$

# Bibliography

[1] T.M. Liggett, Interacting Particle Systems, 1999.

[2] T.M. Liggett, Stochastic Interacting Systems: Contact, Voter and Exclusion Processes, 1999.

[3] N. Lanchier, The Axelrod model for the dissemination of culture revisited, *The Annals of Applied Probability*, 2012.

[4] N. Lanchier, J. Schweinsberg, Consensus in the two-state Axelrod model, *Stochastic Processes and their Applications*, 2012.

[5] N. Lanchier, S. Scarlatos, Fixation in the one-dimensional Axelrod model, *The Annals of Applied Probability*, 2013.

[6] R.A. Holley, T.M. Liggett, Ergodic theorems for weakly interacting particle systems and the voter model, *The Annals of Applied Probability*, 1975.

[7] R. Arratia, Site recurrence for annihilating random walks on $\mathbb{Z}^d$, *The annals of Applied Probability*, 1983.