



**HAL**  
open science

## Diffusion in large networks

Michel Grabisch, Agnieszka Rusinowska, Xavier Venel

► **To cite this version:**

Michel Grabisch, Agnieszka Rusinowska, Xavier Venel. Diffusion in large networks. Journal of Economic Dynamics and Control, 2022, 139, pp.104439. 10.1016/j.jedc.2022.104439. halshs-03881455v1

**HAL Id: halshs-03881455**

**<https://shs.hal.science/halshs-03881455v1>**

Submitted on 5 Jun 2022 (v1), last revised 1 Dec 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Diffusion in large networks

Michel GRABISCH<sup>\*(a)(b)</sup>, Agnieszka RUSINOWSKA<sup>(a)(c)</sup>, Xavier VENEL<sup>(d)</sup>

<sup>(a)</sup> Centre d'Economie de la Sorbonne, 106-112 Bd de l'Hôpital  
75647 Paris, France

<sup>(b)</sup> Paris School of Economics, Université Paris I Panthéon-Sorbonne

<sup>(c)</sup> Paris School of Economics – CNRS

<sup>(d)</sup> Dipartimento di Economia e Finanza, LUISS

Viale Romania 32, 00197 Rome, Italy

{michel.grabisch,agnieszka.rusinowska}@univ-paris1.fr, xvenel@luiss.it

Version of May 27, 2022

## Abstract

We investigate the phenomenon of diffusion in a countably infinite society of individuals interacting with their neighbors in a network. At a given time, each individual is either active or inactive. The diffusion is driven by two characteristics: the network structure and the diffusion mechanism represented by an aggregation function. We distinguish between two diffusion mechanisms (probabilistic, deterministic) and focus on two types of aggregation functions (strict, Boolean). Under strict aggregation functions, polarization of the society cannot happen, and its state evolves towards a mixture of infinitely many active and infinitely many inactive agents, or towards a homogeneous society. Under Boolean aggregation functions, the diffusion process becomes deterministic and the contagion model of Morris (2000) becomes a particular case of our framework. Polarization can then happen. Our dynamics also allows for cycles in both cases. The network structure is not relevant for these questions, but is important for establishing irreducibility, at the price of a richness assumption: the network should contain at least one complex star and have enough space for storing local configurations. Our model can be given a game-theoretic interpretation via a local coordination game, where each player would apply a best-response strategy in a random neighborhood.

**JEL Classification:** C7, D7, D85

**Keywords:** diffusion, countable network, aggregation function, polarization, convergence, best-response

---

\*Corresponding author. Tel (+33)14407-8744.

# 1 Introduction

Studying diffusion in networks has a very strong established position in the literature. Adoption of a new technology, product purchasing and marketing, opinion formation, influence, social learning, information cascades and transmission, rumors, fashions, contagion and disease infection, financial contagion, these are only some examples of phenomena very relevant to network diffusion.

The present paper investigates diffusion in large networks. Our point of departure is the work of Morris (2000) on contagion, defined in his paper as a diffusion phenomenon occurring when one of two actions can spread from a finite set of individuals to the whole (countably infinite) population. The basic mechanism in Morris (2000) is a deterministic process, known under the name of *threshold model* (see, e.g., Granovetter (1978), Schelling (1978)), where an agent becomes infected (or *active*, as we will use this term throughout the paper) if the number of infected agents in its neighborhood exceeds a given threshold. To make the mechanism more realistic, we introduce some randomness and replace the threshold mechanism by the following: an agent becomes active with some probability, which is a function of the status (active or inactive) of every agent in its neighborhood. The probability is equal to 1 (respectively, 0) if all agents in the neighborhood are active (respectively, inactive), and it is required that the probability monotonically increases with the number of active agents. Such a function determining the probability to be active is called an *aggregation function*. Hence, the new model is in the vein of Gravner and Griffeath (1998), who study the growth of a cellular probabilistic automaton on  $\mathbb{Z}^2$ , although our perspective is different and we work on arbitrary countable networks, of which  $\mathbb{Z}^2$  is only a particular case. The approach based on aggregation functions also relates the present paper to previous works of two of the authors that study in depth models of opinion dynamics based on aggregation functions (see Grabisch and Rusinowska (2013); Förster et al. (2013); Grabisch et al. (2019)).

Similarly to Morris (2000); Grabisch and Rusinowska (2013); Förster et al. (2013), we assume that an agent is not a neighbor of itself. This is a reasonable assumption in an imitation model, for example when studying the diffusion of a technology. An agent tends to adopt or imitate the status/action of its neighbors. In the threshold model, Morris (2000) suggests a game-theoretic interpretation via a local coordination game, where each player would apply a best-response strategy in a random neighborhood. Since a player does not play against itself, it is then again natural in this interpretation to assume that a player is not a neighbor of itself.

There are two essential differences with our previous works: 1) First, we assume that the society of agents is not anymore finite but countably infinite, like in Morris (2000). Indeed, to accommodate with the large size of the network, there are basically two solutions. Either the network is considered to have an infinite number of agents, or it is considered to be periodic. Otherwise, taking a finite network of any size inevitably implies that border phenomena will occur. We have chosen the first option, as infinite networks are very common in the literature on opinion dynamics. For example, most of studies made in physics consider this assumption, in particular for random graphs, but also those in computer science (e.g., Gravner and Griffeath (1998); Moran (1995)). 2) Second, the underlying network is not anymore complete but is any undirected graph with a bounded degree for all agents, as in Morris (2000).

The general aim of the paper is to study the evolution in the long run of the diffusion process. More precisely, the fundamental question is: *How does the diffusion evolve from a finite set of active agents?* This question can be decomposed into more basic ones like: Is a cascade possible, leading to a state where all agents are active, or all inactive? Is polarization possible, i.e., a stable state where a part of the population is active while the other is not?

Are other remarkable phenomena, like cycles, possible? Our main aim is to disentangle which properties are induced by the aggregation function and which properties are induced by the network structure.

Our mechanism of diffusion is Markovian, and we investigate all its absorbing and transient classes, which permits to give a precise answer to the above questions. However, the assumption on the number of agents (countably infinite) makes the number of possible states of the process *uncountably* infinite. Therefore, we will have to use  $\sigma$ -fields and measure theory. Our study provides complete results for two cases of aggregation functions: strict, and Boolean. A strict aggregation function means that if any agent in the neighborhood of agent  $i$  becomes active (respectively, inactive), then the probability that  $i$  becomes active (respectively, inactive) is positive. A Boolean aggregation function simply means that the result is either 0 or 1. Then the model becomes deterministic and includes as particular case the initial model of Morris (2000).

We summarize our main findings. While in the infinite context it is typical to consider  $\mathbb{Z}^2$ , we provide more general results that concern arbitrary network structures. Most notably, we provide a condition called *richness* (see below) under which everything essentially works like in  $\mathbb{Z}^2$ .

In the case of strict aggregation functions, we distinguish the case of bipartite and non-bipartite networks. While the latter case is much more common for real networks, we prove that only for bipartite networks, a cycle of length two occurs, which therefore necessitates a particular study. We start our analysis with non-bipartite networks, and obtain as main result the exhaustive list of all possible absorbing and transient irreducible sets<sup>1</sup> (Theorems 1 and 2). Most notably, we find that: 1) The states where every agent is active or every agent is inactive are absorbing, which can result from a cascade effect; 2) Any state with a finite number of agents being active or inactive is necessarily transient, so that no polarization can occur, and the state of the society evolves either towards the states where all agents are active or all inactive, or towards a class of states which are a mixture of infinitely many active and infinitely many inactive agents, endlessly changing. Irreducibility is proved at the price of an additional hypothesis on the structure of the network (*not* on the aggregation function, interestingly), which we call *richness*: it says that the network has enough “space” to store a local configuration of active/inactive agents on its neighbors. This can be related to the existence of a matching in the graph, which was shown by Erdős and Renyi (1966) to be highly probable in connected random graphs.

The results for bipartite graphs are similar but more complicated due to the presence of a cycle (Theorems 3 and 4). Also, the richness condition is more demanding, as the network should contain a “complex star” (stars with 3 branches of length at least 2). Proposition 3 gives examples of classes of graph satisfying the richness assumption, while Proposition 4 characterizes those bipartite networks without complex stars.

As for Boolean aggregation functions, the results fairly differ. Indeed, while a cascade effect to the limit cases where all agents are active or all are inactive is still possible, there are many possibilities to reach other absorbing states (which is interpreted as polarization), or cycles, depending on the value of the threshold. A more detailed study depends essentially on the topology of the network, as we showed with  $\mathbb{Z}^2$  and the neighborhoods of 4 and 8 neighbors. Lastly, we show that in the grid  $\mathbb{Z}^d$ , cycles have length 2, while there are examples of networks, like infinite trees, where cycle may have arbitrary length.

---

<sup>1</sup>Since the Markov chain is uncountable, the formal definition will be the one of  $\phi$ -irreducibility where  $\phi$  is a reference measure. See definitions in Appendix A.

**Other related literature** Although the first works on diffusion do not refer to explicit network structure, certain early and widely used models (e.g., Bass (1969)) do incorporate some aspects of imitation. Since then, a variety of approaches and different ways of modeling diffusion have been used; for surveys see, e.g., two chapters in Jackson (2008) on diffusion through networks and learning, Acemoglu and Ozdaglar (2011) for an overview of Bayesian learning and imitation models, several chapters in Bramoullé et al. (2016) on diffusion, learning and contagion (Lamberson (2016), Golub and Sadler (2016), Cabrales et al. (2016), Acemoglu et al. (2016)), and Grabisch and Rusinowska (2020) for a survey on nonstrategic models of opinion dynamics; see also e.g., Rogers (2003) for overviews of the literature on diffusion and its applications, and Silverman (2001) for “word-of-mouth marketing”.

As the spread of ideas, products or rumours can be similar to that of infectious disease, many frameworks of diffusion in networks are indeed inspired by the standard epidemiological models such as SI, SIS, and SIR (Bailey (1975), Keeling and Rohani (2008)). These models differ, in particular, by the possibility of re-infection after recovering. Contrary to the SIS model, this option is not possible in the SIR model, while in the SI framework once individuals become infected, they remain infected forever. Different variants of the models are proposed, e.g., in the literature on network diffusion in economics (Jackson and Rogers (2007), López-Pintado (2008)).

There are two classic papers in the prior literature and a number of follow-up works which are particularly relevant, as they already consider the absorbing classes of states for finite graphs and show that the process always converges or the dynamic is two-periodic. These are Goles and Olivos (1980) who consider generalized threshold functions, and Poljak and Sura (1983) who analyze a weighted majority model. Also quite related is Berger (2001) who considers a majority model with finite graphs and dynamic monopoly being a subset of nodes whose opinion spreads everywhere; for an overview of dynamic monopolies, see e.g., Peleg (2002). Dreyer Jr. and Roberts (2009) consider several graph-theoretical threshold models of the spread of disease or opinion, and focus on an irreversible  $k$ -threshold process where a node, once getting infected, cannot become uninfected anymore. There are also other works that discuss diffusion in similar contexts, such as the linear threshold models of Yildiz et al. (2011) who study the dynamics of a diffusion process under deterministic topologies and deterministic thresholds values, and of Adam et al. (2012) who investigate a model of cascades over finite networks based on a deterministic binary linear threshold model. There is also a large literature on cascades under general threshold models. One example is Gao et al. (2016) who study stochastic attachment (preferential attachment being a particular case), with a growing network getting newcomers that connect stochastically to nodes already present in the network. Our model is also related to the extensive line of works on influence maximization originating from Kempe et al. (2003) who analyze mechanisms that are similar to the ones studied in this paper. However, the works along this line usually consider models in which an active node cannot become inactive again. When the random threshold is uniform, we get a particular case of our aggregation model.

The remainder of this paper is organized as follows. In Section 2 we introduce the main framework of countable networks and aggregation functions, and our diffusion process. We also present examples and some first results. In Section 3 we focus on strict aggregation functions and non-bipartite networks. Section 4 is dedicated to the case of strict aggregation functions and bipartite networks. Section 5 is devoted to the study of Boolean aggregation functions. We conclude in Section 6. Proofs of our results are presented in Appendices A till F.

## 2 The framework

### 2.1 Countable networks

A (*countable*) *network* is an undirected graph  $(\mathcal{X}, E)$ , where  $\mathcal{X}$  is the set of agents (society), and  $E$  the set of edges or links, satisfying the following properties:

1.  $\mathcal{X}$  is countably infinite;
2.  $(\mathcal{X}, E)$  is connected;
3. For every agent  $x \in \mathcal{X}$ , the neighborhood of  $x$ , denoted by  $\Gamma(x)$ , does not contain  $x$ ;
4. For every agent  $x \in \mathcal{X}$ , the size of  $\Gamma(x)$  (the degree of  $x$ ) is bounded by a fixed integer  $\gamma$ .

Moreover, we assume that for any  $x$ , we have an order on the neighbors in  $\Gamma(x)$ . We note that this framework is equivalent to the one used by Morris (2000), where the graph is induced by a neighborhood relation. When considering a subset  $X$  of  $\mathcal{X}$ , the corresponding subgraph is denoted by  $(X, E_X)$ .

We give several examples of networks below, borrowed from Morris (2000).

**Example 1 (2-dimensional grid).** Consider  $\mathcal{X} = \mathbb{Z}^2$ , and the neighborhood of  $x$  defined as the set containing all agents within a certain distance of  $x$ , excluding  $x$ . Taking for example the Euclidean distance, the neighborhood of  $x$  within distance 1 is formed by the 4 agents at north, south, east and west position, i.e.,  $(x_1, x_2 + 1)$ ,  $(x_1, x_2 - 1)$ ,  $(x_1 + 1, x_2)$  and  $(x_1 - 1, x_2)$  respectively. We call this the *1-neighborhood* of  $x$ . With distance  $\sqrt{2}$ , we get 8 neighbors (4 on the diagonals in addition to the previous ones). We may consider more general distances, not necessarily symmetric on both coordinates. Figure 1 shows the graphs corresponding to the 1-neighborhood and  $\sqrt{2}$ -neighborhood.

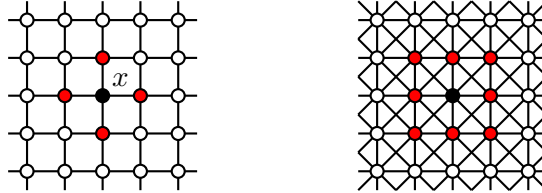


Figure 1: Graphs with  $\mathcal{X} = \mathbb{Z}^2$  and neighborhood induced by the Euclidean distance of 1 (left) and  $\sqrt{2}$  (right). The neighborhood of  $x$  (black node) is the set of red nodes.

**Example 2 ( $d$ -dimensional grid).** An immediate generalization of the previous example is to consider  $\mathcal{X} = \mathbb{Z}^d$ , with  $d$  any positive integer, and the neighborhood is again defined by the Euclidean distance.

**Example 3 (hexagonal pavement).** The hexagonal pavement is such that every node has 3 neighbors (see Figure 2).

**Example 4 (hierarchy).** A hierarchy is an infinite tree without leaf, i.e., a graph where each agent  $x$  has  $m(x) \geq 1$  subordinates, and one superior, except the root which has no superior (see Figure 3). The numbers  $m(x)$  may differ for each agent.

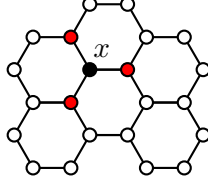


Figure 2: Graph of the hexagonal pavement

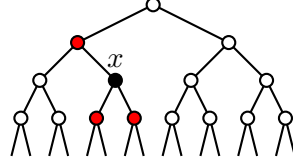


Figure 3: Graph of a hierarchy with  $m(x) = 2$  for each agent  $x$

We assume that each agent can have one of two statuses 0 and 1, which can be interpreted in various ways (opinion on a given subject, adoption of a new technology, infection by some disease, etc.). We will call this value the *status* of agent  $x$ . As we are interested in this paper by the spread over  $\mathcal{X}$  of one of the actions, we say that agent  $x$  is *active* if his status is 1 and that he is *inactive* if his status is 0.

At a given time, the society is composed of active and inactive agents. The *configuration* of the society describes who is active and who is inactive. Hence, a configuration can be represented either by a 0-1-valued function on  $\mathcal{X}$  (0 for inactive, 1 for active), or by the set of active agents. The set of all possible configurations is  $\Omega := \{0, 1\}^{\mathcal{X}} \cong 2^{\mathcal{X}}$ . In the sequel, we will freely use one or the other representation, whichever is most convenient. We elaborate on both representations below.

In the function representation, a configuration is a function  $\omega \in \Omega$ , where  $\omega(x) = 1$  if  $x$  is active, and  $\omega(x) = 0$  if  $x$  is inactive. For every  $X \subset \mathcal{X}$ , we denote by  $\pi_X$  the projection from  $\Omega$  to  $\{0, 1\}^X$ . Every element of  $\{0, 1\}^X$  will be called a *partial configuration* restricted to  $X$ . Given a subset  $X'$ , a *partial configuration in  $X'$*  is a partial configuration for a subset  $X$  of  $X'$ . A partial configuration  $\theta \in \{0, 1\}^X$  defines naturally the set  $\theta^+ = \{\omega \in \Omega, \pi_X(\omega) = \theta\}$  of all configurations compatible with  $\theta$ . This set will be called the *cylinder* generated by  $\theta$ . Two particular configurations are the constant functions  $\mathbf{1}, \mathbf{0}$ , which correspond respectively to the cases where all agents are active and no agent is active. This representation is particularly convenient for defining formally the dynamical system.

In the set representation, the state of the society is represented by the set of active agents. We have clearly a one-to-one correspondence between the two formulations since any configuration  $\omega$  is uniquely defined by the set  $X$  of active agents, with the correspondence  $1_X \equiv \omega$ . For example,  $\mathbf{1}, \mathbf{0}$  correspond respectively to  $\mathcal{X}$  and  $\emptyset$ . A partial configuration becomes a pair of disjoint sets  $(X, Y)$  such that agents in  $X$  are active and agents in  $Y$  are inactive. It naturally induces the cylinder  $(X, Y)^+$  of configurations such that agents in  $X$  are active, agents in  $Y$  are inactive, and the status of agents outside  $X \cup Y$  is not constrained. When  $X$  and  $Y$  are finite,

we will say that the cylinder is *finite*. This second representation is particularly convenient when using the notion of neighborhood.

## 2.2 Definition of the diffusion process

We now describe how the diffusion evolves in the society. Given a finite set  $X$ , we denote by  $|X|$  its cardinal, and for any natural number  $n$ , we denote by  $[n]$  the set  $\{1, \dots, n\}$ .

We assume that each agent's status is randomly updated as a function of the status of its neighbors. Given an integer  $l$ , an *aggregation function of size  $l$*  is a mapping  $A : [0, 1]^l \rightarrow [0, 1]$  which is nondecreasing w.r.t. each coordinate, and satisfies  $A(1, \dots, 1) = 1$ ,  $A(0, \dots, 0) = 0$ . It is *symmetric* or *anonymous* if  $A(z_1, \dots, z_l) = A(z_{\sigma(1)}, \dots, z_{\sigma(l)})$  for every permutation  $\sigma$  on  $[l]$ . Each agent  $x$  has an aggregation function  $A_x$  of size  $l = |\Gamma(x)|$ , i.e., the number of its neighbors. We denote by  $\mathbf{A} = (A_x)_{x \in \mathcal{X}}$  the vector of the aggregation functions of all agents.

We assume that the probability for an agent  $x$  to be active at time  $t + 1$ , given the configuration  $\omega$  at time  $t$ , is

$$P(x \mid \omega) = A_x(\pi_{\Gamma(x)}(\omega)). \quad (1)$$

In words, the probability for  $x$  to be active at the next stage is obtained by aggregating the vector of statuses of all the neighbors of  $x$  in the configuration  $\omega$ . By definition,  $\pi_{\Gamma(x)}(\omega)$  is a vector in  $[0, 1]^{\Gamma(x)}$  that we identify to  $[0, 1]^{|\Gamma(x)|}$  (following the order defined on  $\Gamma(x)$ ), hence the probability is well defined. The more active agents are in the neighborhood, the higher the probability. Moreover,  $x$  is active (respectively, inactive) for sure if all its neighbors (respectively, none of its neighbors) were active. We make the additional assumption that the probability for the agents to be active conditionally on  $\omega$  is independent across the agents. We will see in the next section how to define rigorously this process.

Our framework contains several models presented in the literature. For example, it may be seen as a sophistication of the voter model (Holley and Liggett, 1975), where an agent adopts the status of one of its neighbors, randomly chosen. When the aggregation function is Boolean, the updating process becomes deterministic, and contains the classical threshold model (see, e.g., Granovetter (1978)), as well as the framework of Morris (2000) as particular cases. This is exactly the framework of Section 5.

Moreover, as explained in the introduction, Morris (2000) suggests a game-theoretic interpretation of the threshold-model. The same interpretation can be provided here. Specifically, taking two players  $i$  and  $j$  with a coordination game payoff matrix yielding  $q \in [0, 1]$  for coordination on action 0,  $1 - q$  for coordination on action 1, and 0 otherwise, the best response strategy leads to the choice of action 1 for player  $i$  if this player assigns a probability at least  $q$  that player  $j$  chooses 1. Generalizing this to the neighborhood  $\Gamma(i)$  of player  $i$ , it is found that player  $i$ 's best response is to choose action 1 if and only if at least  $q\gamma$  neighbors choose action 1. Hence, the threshold model is recovered with threshold  $q\gamma$ . In order to get our model in its full generality, it remains to introduce some random device in the game. This can be done for example by proceeding like in the voter model: instead of meeting all of its neighbors, an agent meets only some of them at random, according to some probability distribution. Then the threshold will be exceeded with some probability, which causes the agent to become active (status 1) with some probability.

We define the diffusion process as a Markov process whose set of states is the set  $\Omega$  of configurations, based on Equation (1) and the independence across agents. As  $\Omega$  is uncountable, the definition of the Markov process requires to work on  $\sigma$ -fields (see, e.g., Hernández-Lerma



and Lasserre (2003)), and the probability of transition between two states is replaced by the *kernel*  $\mathcal{K}(\omega, \mathcal{A})$ , which gives the probability to go from the present state  $\omega$  into a state belonging to a measurable set of states  $\mathcal{A}$ . Formal definitions and details are given in Appendix A.

### 2.3 Interior and closure of a configuration

From (1) and this informal definition, we can already highlight one specific aspect of the dynamic process. Using the set representation, we define for any set  $X \subseteq \mathcal{X}$  its *closure*  $\text{clo}(X) = \overline{X}$  and *interior*  $\text{int}(X) = \overset{\circ}{X}$  by

$$\begin{aligned}\text{clo}(X) &= \overline{X} = \{x \in \mathcal{X} : \Gamma(x) \cap X \neq \emptyset\} \\ \text{int}(X) &= \overset{\circ}{X} = \{x \in \mathcal{X} : \Gamma(x) \subseteq X\}.\end{aligned}$$

It is sometimes convenient to iterate the operator several times. We will define  $\text{int}^n$  (respectively,  $\text{clo}^n$ ) to be the  $n$ th iteration of the operator  $\text{int}$  (respectively,  $\text{clo}$ ). Accordingly, we use the same notation for configurations:  $\overline{\omega}, \overset{\circ}{\omega}, \text{clo}(\omega), \text{int}(\omega)$ . Obviously,  $\overset{\circ}{X} \subseteq \overline{X}$ , but it is not true in general that  $\overset{\circ}{X} \subseteq X \subseteq \overline{X}$  (see Example 5 below). Also, it is easy to see that  $\text{clo}$  and  $\text{int}$ , viewed as mappings on  $(2^{\mathcal{X}}, \subseteq)$ , are monotone:

$$X \subset X' \Rightarrow \overset{\circ}{X} \subseteq \overset{\circ}{X'} \text{ and } \overline{X} \subseteq \overline{X'}.$$

With this new notation, we obtain:

$$x \in \overset{\circ}{X} \Rightarrow P(x | X) = 1, \quad x \notin \overline{X} \Rightarrow P(x | X) = 0. \quad (2)$$

This shows that, given that  $X$  is the set of active agents at time  $t$ , the set  $X'$  of active agents at time  $t + 1$  lies between  $\overset{\circ}{X}$  and  $\overline{X}$ . Formally, given  $X \subset Z$ , denote  $[X, Z] := \{Y \in 2^{\mathcal{X}} \mid X \subseteq Y \subseteq Z\}$ . Then the set  $X'$  of active agents at time  $t + 1$  lies in the interval  $[\overset{\circ}{X}, \overline{X}]$  with probability 1.

**Example 5.** Consider  $\mathcal{X} = \mathbb{Z}^2$  and the 1-neighborhood. Let  $\omega$  be the configuration defined by  $\omega(n, m) = 1$  iff  $|n| + |m| = 1$ , denoted by  $X$  in set representation (Figure 4 (left)).  $\overset{\circ}{X}$  and  $\overline{X}$  are given in Figure 4 (middle) and (right), respectively. One can see that  $\overset{\circ}{X} \not\subseteq X \not\subseteq \overline{X}$ , however  $\overset{\circ}{X} \subseteq \overline{X}$ .

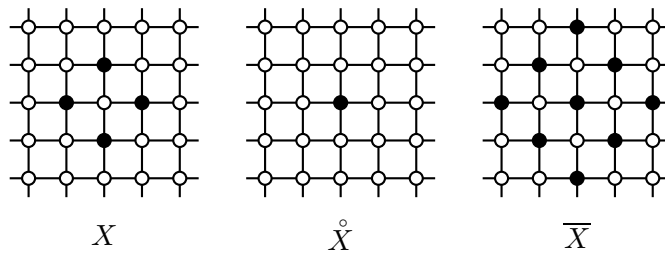


Figure 4: Interior and closure of a set  $X$

### 2.4 Examples and first results

Let us investigate some examples that will highlight three different aspects. The first one highlights the influence of the aggregation function, while the second and third ones highlight the influence of the structure of the graph.

**Example 6.** Consider the regular grid  $\mathbb{Z}^2$  with the  $\sqrt{2}$ -neighborhood and the following two aggregation functions

$$\forall z_1, \dots, z_8 \in \{0, 1\}, \quad A(z_1, \dots, z_8) = \frac{1}{8} \sum_{i=1}^8 z_i$$

and

$$A'(1_S) = \begin{cases} 0 & \text{iff } |S| \leq 4 \\ 1 & \text{iff } |S| \geq 5. \end{cases} \quad (3)$$

We suppose that all agents have the same aggregation function. Under both aggregation functions, the configuration where all agents have opinion 1 (resp. 0) are steady states (fixed points) of the dynamics. But, under the second case there exist many additional steady states, as illustrated on Figure 5.

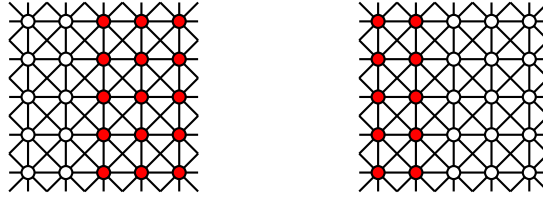


Figure 5: Two examples of steady states with aggregation function  $A'$

This leads us to distinguish 3 types of aggregation functions. For a given size  $l$ ,

1.  $A$  is *strict*<sup>2</sup>:  $A(z) = 0$  iff  $z$  is the 0 vector of size  $l$  and  $A(z) = 1$  iff  $z$  is the 1 vector of size  $l$ .
2.  $A$  is *Boolean*:  $A(z) = 0$  or  $1$  for all  $z \in [0, 1]^l$ .
3. None of the above applies:  $A$  is nonstrict and nonBoolean, i.e., there exists  $z \in [0, 1]^l$  s.t.  $0 < A(z) < 1$ , and there exists  $z' \neq 0$  s.t.  $A(z') = 0$  or  $z'' \neq 1$  s.t.  $A(z'') = 1$ .

We will focus on the two first types, that are the extreme cases. We assume that, although all agents may have a different aggregation function, *they all have an aggregation function of the same type*. Consequently, we will say that  $\mathbf{A} = (A_x)_{x \in \mathcal{X}}$  is of the strict type or of the Boolean type.

Thanks to this distinction, we can formalize the findings of Example 6.

**Proposition 1.** 1.  $X$  is a steady state for any vector of aggregation functions  $\mathbf{A}$  iff  $X = \mathcal{X}$  or  $X = \emptyset$ .

2. If  $\mathbf{A}$  is strict, then the only steady states are  $X = \mathcal{X}$  and  $X = \emptyset$ .

(see proof in Appendix A)

The second distinction concerns the type of networks. In the following Example 7, one can see the existence of a cycle of length two for the graph induced by the Euclidean distance 1, whereas it is not possible for the graph with distance  $\sqrt{2}$ .

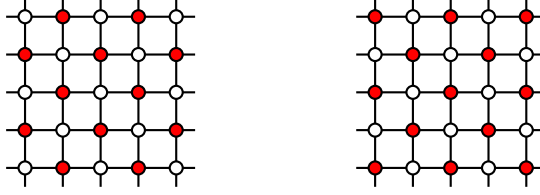


Figure 6: Deterministic cycle of length two

**Example 7 (Bipartite graph).** Consider  $\mathcal{X} = \mathbb{Z}^2$  and the 1-neighborhood. Then, there exists a cycle of length two represented on Figure 6.

The next result elucidates the existence of cycles in the dynamics, independently of the aggregation function, and shows that the key feature is the existence of a bipartition in the graph. This explains why a cycle of length 2 occurs in the grid  $\mathbb{Z}^2$  with the 1-neighborhood, which disappears when the  $\sqrt{2}$ -neighborhood is used, since then the graph is no more bipartite.

**Proposition 2.** Assume  $\mathbf{A}$  is not Boolean, i.e., at least one  $A_x$  is not Boolean. There exists a cycle if and only if  $(\mathcal{X}, E)$  is bipartite with bipartition  $(X, X^c)$ , in which case the unique cycle is  $(X, X^c)$  of length 2.

(see proof in Appendix A)

Apart from the grid  $\mathbb{Z}^2$  with the 1-neighborhood, many of the networks given in Examples 1 to 4 are bipartite. This is the case of  $\mathbb{Z}^d$  with the 1-neighborhood, for any positive integer  $d$ . The hexagonal pavement is also bipartite (1 node out of 2 on each hexagon; see Figure 7 (left)) and so is the hierarchy (all nodes of odd-numbered layers; see Figure 7 (right)).



Figure 7: Bipartite graphs: the hexagonal pavement (left), and the hierarchy (right)

**Remark 1.** The result on the cycle of length 2 is reminiscent of the classical result on finite graphs with the threshold model (recall that this is a particular case of our Boolean aggregation function model), saying that the process either reaches a fixed point or enters a cycle of length 2 (see Goles and Olivos (1980), Poljak and Sura (1983)). We mention that this classical result is no more valid for infinite graphs as shown by Moran (1995). We will go back to these results when studying Boolean aggregation functions.

We now provide a last example highlighting a difficulty when the graph is too “flat”. We will come back to this example with more details in the bipartite section.

---

<sup>2</sup>Note that if  $A$  is increasing (in the strict sense), then it is a strict aggregation function, but not the converse.

**Example 8.** Consider the network defined by  $\mathcal{X} = \mathbb{Z}$  and the 1-neighborhood:

$$\forall x \in \mathcal{X}, \Gamma(x) = \{x - 1, x + 1\},$$

together with the following strict aggregation function

$$A(1_S) = \begin{cases} 0 & \text{iff } |S| = 0 \\ 1/2 & \text{iff } |S| = 1 \\ 1 & \text{iff } |S| = 2. \end{cases} \quad (4)$$

It is clear that this graph is bipartite and is flat in the sense that it does not contain a node with three neighbors. Consider the configuration  $\omega_0$  in this set defined by

$$\omega_0(x) = \begin{cases} 0 & \text{if } x \text{ is odd,} \\ 0 & \text{if } x \leq 0 \text{ and } x \text{ is even,} \\ 1 & \text{if } x > 0 \text{ and } x \text{ is even.} \end{cases}$$

What are the possible configurations after one stage? We can see that there are only two possible configurations  $\omega_{-1}$  or  $\omega_1$  that are translations of  $\omega_0$  by respectively  $-1$  and  $+1$ . By iterating, it follows that from  $\omega_0$ , the Markov Chain induced by the aggregation function can only reach configurations which are translations of  $\omega_0$ . In particular, it is impossible to reach the cylinder  $\mathcal{A} = (X, Y)^+$  with  $X = \{0, 4\}$  and  $Y = \{2\}$ . Apart from the bipartite nature of the graph which forces the alternation of 0 and 1 on odd and even positions, the flatness of the graph (absence of nodes of degree larger than 2) prevents from moving patterns on the graph without erasing the current configuration. A sufficient condition to be able to do so will be formalized later as a complex star (see Definition 2).

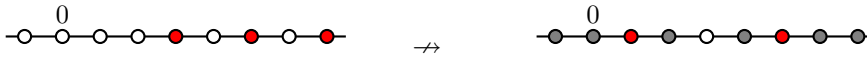


Figure 8: Initial configuration and Cylinder that can not be reached (gray nodes have either status 1 or 0)

Following the intuition provided by Examples 6 and 7, we split our analysis into three parts: non-bipartite graph with strict aggregation function  $\mathbf{A}$ , bipartite graph with strict aggregation function  $\mathbf{A}$ , and finally – Boolean aggregation function  $\mathbf{A}$ .

### 3 Strict aggregation functions and non-bipartite networks

In this section, we focus on the case of strict aggregation functions ( $\mathbf{A}$  is strict) and non-bipartite networks  $(\mathcal{X}, E)$ . A key element of the analysis is to define a partition of the set of configurations into blocks, according to the “number” of active/inactive nodes. Indeed, under a strict aggregation function and in an infinite network, there is some persistence of the number of active/inactive nodes: if there are no active nodes, there will never be and if there are infinitely many active (resp. inactive) nodes, there will always be an infinite number of active (resp. inactive) nodes. This leads to the definition of sets of configurations that are absorbing and transient. The next step is to study the irreducibility of these sets. We will see that it is necessary to impose some conditions on the graph for these sets to be irreducible and therefore to be classes.

### 3.1 Absorbing and transient sets

We consider the partition of the set of configurations where we count how many 1 and 0 exist in  $\omega$ . Formally, a block of the partition is a 2-uple  $(a, b)$  such that  $a, b \in \{0, F, \infty\}$ , where  $a$  is the “number” of positions with inactive status, and  $b$  is the number of positions with active status. The “number” of nodes is coded by either 0,  $F$  or  $\infty$ , where 0 means that there is no node,  $F$  means that the number of nodes is finite but positive, and  $\infty$  means that there are infinitely many nodes.

There are a priori  $3^2 = 9$  possible blocks in the partition, but actually only 5 are nonempty. Indeed, either 0 or 1 has to appear infinitely often, therefore only 5 possible cases remain:  $(0, \infty)$ ,  $(F, \infty)$ ,  $(\infty, 0)$ ,  $(\infty, F)$  and  $(\infty, \infty)$ . For example, with this notation we have:

- $(\infty, 0)$  is the singleton  $\{\mathbf{0}\} = \{\emptyset\}$ ,
- $(0, \infty)$  is the singleton  $\{\mathbf{1}\} = \{\mathcal{X}\}$ ,

Let us remind some standard definitions of the Markov chain theory. Let  $\mathcal{T}$  be a  $\sigma$ -field on  $\Omega$  (see Appendix A for details). A set  $\mathcal{A} \in \mathcal{T}$  is called *absorbing* if  $K(\omega, \mathcal{A}) = 1$  for every  $\omega \in \mathcal{A}$ . A set  $\mathcal{A} \in \mathcal{T}$  is called *transient* if for every configuration  $\omega \in \mathcal{A}$  there exists  $n \in \mathbb{N}$  such that  $K^n(\omega, \mathcal{A}) < 1$ . In words, a set of configurations is absorbing if from any configuration in this set, a transition yields a configuration still in this set with probability 1, while a set of configurations is transient if there is some probability to go outside after a certain number of steps. With these definitions, we can now formalize our introductory statement.

**Theorem 1.** The following sets of configurations are respectively:

- (i) finite absorbing sets:  $(\infty, 0)$  and  $(0, \infty)$ ;
- (ii) infinite uncountable absorbing sets:  $(\infty, \infty)$ ;
- (iii) infinite transient sets:  $(\infty, F)$  and  $(F, \infty)$ .

(see proof in Appendix B)

Theorem 1(i) recovers the result of Proposition 1, establishing that  $\mathbf{0}$  and  $\mathbf{1}$  are the only fixed points of the dynamics when  $\mathbf{A}$  is strict. Results (ii) and (iii), although important and interesting, are not conclusive for our study of convergence. Indeed, we have not yet proved that these sets are irreducible, therefore we cannot conclude that these sets are classes. We postpone their interpretation till Theorem 2, where this result will be established.

### 3.2 Irreducibility

The aim of the rest of the section is to investigate whether the sets listed in Theorem 1 are irreducible. Since the Markov chain is uncountable, the definition of irreducible sets relies on the introduction of a measure of reference  $\phi$  on the set of configurations and is called  $\phi$ -irreducibility.

Contrarily to the previous results which hold without additional assumptions on the network, we need now some particular conditions to establish  $\phi$ -irreducibility. In this section, based on sufficient conditions on  $(\mathcal{X}, E)$ , we will prove that the sets highlighted in Theorem 1 are  $\phi$ -irreducible. Hence, we provide a decomposition of the set of configurations into three types: the sets in case (i) are finite absorbing  $\phi$ -irreducible sets, the set in case (ii) is an infinite absorbing  $\phi$ -irreducible set, and the ones in case (iii) are transient  $\phi$ -irreducible sets.

We first introduce the formal definition of  $\phi$ -irreducibility. Then, we provide two examples of non-bipartite graphs where the sets of Theorem 1 are not  $\phi$ -irreducible and finally sufficient conditions for them to be  $\phi$ -irreducible.

### 3.2.1 Formal definition of $\phi$ -irreducibility

We will consider as measure of reference the “uniform distribution”  $\phi$ -defined as follows: for every two finite disjoint sets  $X$  and  $Y$ ,

$$\phi((X, Y)^+) = \frac{1}{2^{|X|+|Y|}}.$$

For every  $X^*$  and  $Y^*$ , we impose the restriction of  $\phi$  conditionally on  $(X^*, Y^*)^+$  defined on the trace of  $\mathcal{T}$  on  $(X^*, Y^*)^+$  such that for every two finite disjoint sets  $X$  and  $Y$  disjoint with  $X^* \cup Y^*$ ,

$$\phi_{X^*, Y^*}((X^* \cup X, Y^* \cup Y)^+) = \frac{1}{2^{|X|+|Y|}}.$$

The Markov chain is  $\phi$ -irreducible if for all  $\mathcal{A} \in \mathcal{T}$  such that  $\phi(\mathcal{A}) > 0$ , for all  $\omega \in \Omega$ ,

$$\sum_{n=1}^{+\infty} K^n(\omega, \mathcal{A}) > 0.$$

In words, the Markov chain is  $\phi$ -irreducible if there is a positive probability that starting from any configuration, the process reaches after some step any set of configurations, provided this set has a positive measure (w.r.t.  $\phi$ ). In Theorem 1, we have seen that the singletons  $\{\mathbf{0}\}$  and  $\{\mathbf{1}\}$  are absorbing sets, hence the diffusion Markov process  $K$  is not irreducible.

Since  $K$  is not  $\phi$ -irreducible on  $\mathcal{T}$ , we may look for sub-fields of  $\mathcal{T}$  where  $K$  is  $\phi$ -irreducible. We introduce for any  $\mathcal{A} \in \mathcal{T}$  the *trace* of  $\mathcal{T}$  on  $\mathcal{A}$ , defined by  $\mathcal{T}_{|\mathcal{A}} = \{\mathcal{B} \in \mathcal{T} : \mathcal{B} \subseteq \mathcal{A}\}$ , and  $\phi_{\mathcal{A}}$  the regular conditional probability on  $\mathcal{A}$ . A set of configurations  $\mathcal{A} \in \mathcal{T}$  is a  $\phi$ -irreducible set if for every  $\omega \in \mathcal{A}$ , every  $\mathcal{B} \in \mathcal{T}_{|\mathcal{A}}$  such that  $\phi_{\mathcal{A}}(\mathcal{B}) > 0$ ,  $K^n(\omega, \mathcal{B}) > 0$  for some  $n$ . Combining both  $\phi$ -irreducibility and the property of being absorbing, we get the fundamental notion of class. A set of configurations  $\mathcal{A} \in \mathcal{T}$  is a  $\phi$ -irreducible class if it is both absorbing and a  $\phi$ -irreducible set. For example, it is clear by the previous results that  $\{\mathbf{1}\}$  and  $\{\mathbf{0}\}$  are finite  $\phi$ -irreducible classes for any  $\phi$ .

### 3.2.2 Counterexamples

In general, the sets in Theorem 1 are not  $\phi$ -irreducible for any  $(\mathcal{X}, E)$ . We present two counterexamples.

**Example 9.** Consider the plane graph completed by two additional agents linked to  $(0, 0)$ . Formally, let  $\mathcal{X} = \mathbb{Z}^2 \cup \{\alpha, \beta\}$ . We assume that the neighborhoods are given by the Euclidean distance with distance  $\sqrt{2}$ , except on  $\{0, \alpha, \beta\}$ :

- $\Gamma(\alpha) = \Gamma(\beta) = \{(0, 0)\}$ ,
- $\Gamma((0, 0)) = \{(-1, -1), (-1, 0), (-1, 1), (0, -1), (0, 1), (1, -1), (1, 0), (1, 1), \alpha, \beta\}$ ,

Then, this network is non-bipartite. Moreover, since  $\alpha$  and  $\beta$  have a unique neighbor, we know that for every dynamic of opinion, the status of  $\alpha$  (resp.  $\beta$ ) at stage  $t$  is the status of agent  $(0, 0)$  at stage  $t - 1$ . In particular, both statuses are equal and it is impossible to reach the cylinder  $(\{\alpha\}, \{\beta\})^+$  of configurations where  $\alpha$  is active whereas  $\beta$  is inactive.

In Example 9, as the nodes  $\alpha, \beta$  share the same neighbor, there is no possibility to “store” somewhere a partial configuration where  $\alpha$  and  $\beta$  would take different values.

### 3.2.3 Sufficient condition

Based on Example 9, we introduce the notion of *storing function*.

**Definition 1.** We say that a partial configuration  $(X, Y)$  of  $\mathcal{X}$  can be *stored* if there exists a mapping  $\theta$  from  $\mathcal{X}$  to  $\mathcal{X}$  such that

1. for every  $x \in X \cup Y$ ,  $\theta(x) \in \Gamma(x)$ ,
2. for every  $x \in X$  and every  $y \in Y$ ,  $\theta(x) \neq \theta(y)$ .

$\theta$  is called a *storing function*. Observe that  $\theta(X \cup Y) \subseteq \overline{X \cup Y}$ .

Informally, the existence of a storing function makes possible to “store” the configuration for one stage and therefore to come back in two stages (with positive probability) to the initial configuration. This will be fundamental in our procedure since we will fix the statuses of agent progressively and need to ensure that status that have been fixed previously are not lost.

Are there many non-bipartite graphs with a storing function? Let us remark that a particular case of storing function is provided by the notion of matching in a graph. Indeed, consider a finite partial configuration  $(X, Y)$  and suppose that for each  $x \in X \cup Y$ , there exists a unique  $y(x) \in \Gamma(x)$  such that the set of edges  $\{(x, y(x)) \mid x \in X \cup Y\}$  is a matching in the subgraph  $X \cup Y \cup \overline{X \cup Y}$ , i.e., no two edges have a common vertex. Then the mapping  $x \mapsto y(x)$  defines a storing function  $\theta$ , and moreover, it defines a perfect matching (i.e., covering all vertices) in the subgraph  $X \cup Y \cup \theta(X \cup Y)$ . About the existence of a perfect matching in a graph, Erdős and Renyi (1966) showed that for finite random graphs where the probabilities of each edge are i.i.d, if there exists a connected component, then it is essentially unique and with probability close to 1 there exists a perfect matching. Hence, in terms of random graphs, connectivity essentially implies that any configuration can be stored.

As highlighted by Example 8, it is also necessary to have some spaces to move values around. The assumption that the graph is not bipartite implies the existence of a cycle of odd length and in particular the existence of the following subgraph called a complex star. The existence of a complex star implies that the graph is not completely flat. Its importance will be clearer in the bipartite section.

**Definition 2.** A *complex star* is a 7-uple  $(s_*, s_1, s_2, s_3, s'_1, s'_2, s'_3) \in \mathcal{X}^7$  such that:

- $s_1, s_2, s_3$  are 3 distinct nodes;
- $\{s_1, s_2, s_3\} \subseteq \Gamma(s_*)$ ;
- $s'_1 \in \Gamma(s_1)$ ,  $s'_2 \in \Gamma(s_2)$  and  $s'_3 \in \Gamma(s_3)$ ;
- $s_* \notin \{s'_1, s'_2, s'_3\}$ .

Informally,  $s_*$  is the center of a star with three branches that have at least a depth of two.

Note that we do not assume that  $s'_1, s'_2$  and  $s'_3$  are distinct. However, it can be the case that  $s'_1 = s_2$  and  $s'_2 = s_1$  (or similarly with indices 1,3 or 2,3), in which case we say that the complex star is *degenerate*. Example 9 shows both types of complex star.

**Richness Assumption.**  $(\mathcal{X}, E)$  is said to be *rich* if:

1. There exists a complex star;



Figure 9: Degenerate complex star (on the left) and non-degenerate complex star (on the right).

2. Any partial configuration  $(X, Y)$  can be stored.

Recall that for non-bipartite graphs, the first condition is always satisfied. We obtain the following theorem.

**Theorem 2.** Assume that the graph  $(\mathcal{X}, E)$  is non-bipartite and satisfies the Richness Assumption. We have the following decomposition:

- (i)  $(\infty, 0)$  and  $(0, \infty)$  are finite  $\phi$ -irreducible classes,
- (ii)  $(\infty, \infty)$  is an infinite  $\phi$ -irreducible class,
- (iii)  $(\infty, F)$  and  $(F, \infty)$  are transient and  $\phi$ -irreducible sets.

Moreover, these are the only ones.

(see proof in Appendix C)

Let us present the sketch of the proof on a simple example. One considers the graph  $\mathcal{X}$  composed of an infinite ray with a terminal triangle depicted on Figure 10. We denote the nodes of the triangle by  $\{0, s_1, s'_1, s_2, s'_2\}$  and the elements of the infinite ray by natural numbers. Notice that 0 is the center of a complex star. Formally,  $\mathcal{X} = \mathbb{N} \cup \{s_1, s'_1, s_2, s'_2\}$  such that the neighborhoods are defined by

- for all  $n \in \mathbb{N} \setminus \{0\}$ ,  $\Gamma(n) = \{n - 1, n + 1\}$ ,
- $\Gamma(s'_1) = \{s_1, s'_2\}$ ,  $\Gamma(s'_2) = \{s_2, s'_1\}$ ,
- $\Gamma(s_1) = \{s'_1, 0\}$ ,  $\Gamma(s_2) = \{s'_2, 0\}$  and  $\Gamma(0) = \{s_1, s_2, 1\}$ .

We consider an initial configuration such that nodes 2 and 3 have status 0 (white color) and nodes 4 and 5 have status one (red color). Other statuses are left unconstrained (gray color). Our aim is to generate the statuses 0, 1, 1, 1, 0, 1, 0, 1 on nodes  $s'_1, s_1, s'_2, s_2, 0, 1, 2, 3$ , respectively.

The proof is decomposed into several steps:

- Step 1: Placing active and inactive statuses on the branches of the complex star.
- Step 2: Generating a well-chosen pattern on the ray.
- Step 3: Using this pattern to put the correct status at the correct node.

In the example, the first step is simply done by pushing the leftmost 0 to the suitable place in the complex star, then the second 0 and then the 1's. The second step is to reverse the process to generate the statuses by reverse order and place them temporarily on 0, 1, 2, 3, 4, 5, 6, 7, 8. When doing so, we exchange at every stage the opinions of  $s_1$  and  $s'_1$  (resp.,  $s_2$  and  $s'_2$ ). Finally, one can then place the statuses at their correct positions from the leftmost nodes by translating them.



Figure 10 shows the intermediate configuration at the end of each step. Figure 18 in Appendix C shows a possible trajectory between each of these intermediate configurations following the approach of the proof.<sup>3</sup>

Setting the status on the star:

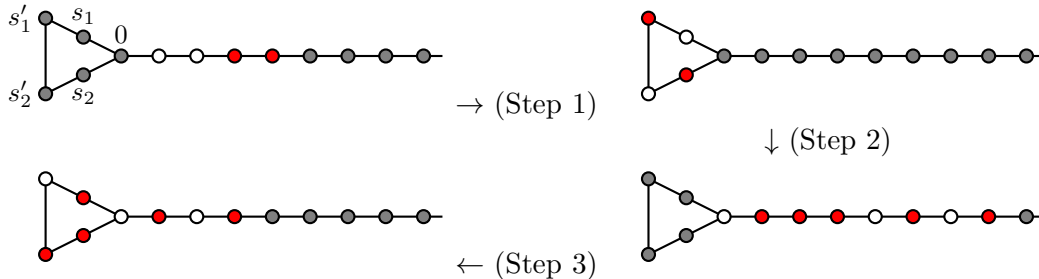


Figure 10: Intermediate configurations

We make some comments on the results given in Theorem 2.

Recall that our fundamental question was how the diffusion evolves from a finite set of active agents. Our results permit to answer this question. We have shown that any configuration with a finite number of active or of inactive agents belongs to a transient  $\phi$ -irreducible set. This means that the process cannot stay forever in this set, but must evolve and eventually reach one of the three  $\phi$ -irreducible classes we have enumerated: the class where every agent is active, the class where every agent is inactive, or the class  $(\infty, \infty)$ , which comprises all configurations with an infinite number of active and of inactive agents. It seems however very difficult to say to which of the three classes the process will converge and with which probability. It is not difficult to build examples of trajectories with  $\mathbb{Z}^2$  and the  $\sqrt{2}$ -neighborhood, where from a finite configuration the process converges to any of the three classes, but intuitively  $(\infty, \infty)$  should be much more probable than the two others.

Converging to the class  $(\infty, \infty)$  means that the diffusion in the network is “homogeneous”, in the sense that everywhere there are active and inactive agents. Any such configuration could happen, which means that there is no region in the network with specific properties.

## 4 Strict aggregation functions and bipartite graphs

We now investigate the case of bipartite graphs. Example 7 shows that there exists a cycle of length two inside  $(\infty, \infty)$ , implying that  $(\infty, \infty)$  is not  $\phi$ -irreducible. Therefore, the sets appearing in Theorem 1 have to be split in order to make them  $\phi$ -irreducible. We first exhibit relevant absorbing and transient sets. We then introduce the richness condition for bipartite graphs and then our main results.

### 4.1 Absorbing and transient sets

Let us denote by  $(\mathcal{X}_o, \mathcal{X}_e)$  the bipartition of the graph  $(\mathcal{X}, E)$ . We will call the nodes in  $\mathcal{X}_o$  the odd positions and the nodes in  $\mathcal{X}_e$  the even positions. We define a partition of  $\Omega$  as follows. A block of the partition of  $\Omega$  is characterized by a 4-uple  $(E_a, E_i, O_a, O_i)$ , where  $E_a, E_i, O_a, O_i$

<sup>3</sup>Notice that the general procedure on this example would lead to different trajectories (more complex) but with the same flavor.

are respectively the “numbers” of nodes in  $\mathcal{X}_e$  (even positions) with active status, in  $\mathcal{X}_e$  with inactive status, in  $\mathcal{X}_o$  with active status, and in  $\mathcal{X}_o$  with inactive status. As for the non-bipartite case, the “numbers” of nodes are coded by  $0, F, \infty$ .

There are a priori  $3^4 = 81$  possible blocks in the partition, but actually only 25 are nonempty. Indeed, if we restrict ourselves to even positions, either 0 or 1 has to appear infinitely often, therefore only 5 possible cases remain:  $(0, \infty)$ ,  $(F, \infty)$ ,  $(\infty, 0)$ ,  $(\infty, F)$  and  $(\infty, \infty)$ . Since what happens on odd and on even positions are independent, there are 25 nonempty blocks in the partition.

For example, with this notation we have:

- $(\infty, 0, \infty, 0)$  is the singleton  $\{\mathbf{0}\} = \{\emptyset\}$ ,
- $(0, \infty, 0, \infty)$  is the singleton  $\{\mathbf{1}\} = \{\mathcal{X}\}$ ,
- $(0, \infty, \infty, 0)$  is the singleton  $\{\mathcal{X}_e\}$  which is the configuration where even agents are active whereas odd agents are inactive,
- $(\infty, 0, 0, \infty)$  is the singleton  $\{\mathcal{X}_o\}$  which is the configuration where even agents are inactive whereas odd agents are active.

We obtain the following theorem.

**Theorem 3.** The following sets of configurations are respectively:

(i) finite  $\phi$ -irreducible classes:

- $(\infty, 0, \infty, 0)$ ,
- $(0, \infty, 0, \infty)$ ,
- $(0, \infty, \infty, 0) \cup (\infty, 0, 0, \infty)$ .

(ii) infinite uncountable absorbing sets:

- $(\infty, \infty, \infty, \infty)$ ,
- $(\infty, 0, \infty, \infty) \cup (\infty, \infty, \infty, 0)$ ,
- $(0, \infty, \infty, \infty) \cup (\infty, \infty, 0, \infty)$ .

(iii) infinite transient sets:

- $(\infty, F, \infty, F)$ ,
- $(F, \infty, F, \infty)$ ,
- $(F, \infty, \infty, F) \cup (\infty, F, F, \infty)$ ,
- $(\infty, F, \infty, \infty) \cup (\infty, \infty, \infty, F)$ ,
- $(F, \infty, \infty, \infty) \cup (\infty, \infty, F, \infty)$ ,
- $(\infty, F, \infty, 0) \cup (\infty, 0, \infty, F)$ ,
- $(F, \infty, 0, \infty) \cup (0, \infty, F, \infty)$ ,
- $(0, \infty, \infty, F) \cup (\infty, F, 0, \infty)$ ,
- $(F, \infty, \infty, 0) \cup (\infty, 0, F, \infty)$ .

Moreover, these are the only ones.

Compared to the results where the network is not bipartite (Theorem 1), we obtain results of the same nature but much more complicated. The introduction of the distinction between odd and even positions makes appear a cycle and periodic classes. The first two finite classes are the classes (i) in Theorem 1. The infinite class (ii) of this theorem has been split into a cycle (leading to the third finite class) and three infinite sets, while the transient set (iii) of

Theorem 1 is split into nine subsets. Our conclusion and interpretation given for the case of non-bipartite networks remain therefore identical, up to the appearance of the periodic classes.

The proof of this theorem is similar to the proof for non-bipartite graphs up to some small differences. Essentially, there is now a separation between the status of the agent as a function of the parity of the stage and the parity of the node:

- odd positions at odd stages and even positions at even stages
- odd positions at even stages and even positions at odd stages.

These statuses are never interacting. Further details are given in Appendix E.

We now turn to the study of  $\phi$ -irreducibility.

## 4.2 Richness condition for bipartite graph

First, Example 9 can be immediately adapted to a bipartite framework by simply changing the graph on  $\mathbb{Z}^2$ .

**Example 10.** Consider the plane graph completed by two additional agents linked to  $(0,0)$ . Formally, let  $\mathcal{X} = \mathbb{Z}^2 \cup \{\alpha, \beta\}$ . We assume that the neighborhoods are given by

- $\Gamma(\alpha) = \Gamma(\beta) = \{(0,0)\}$ ,
- $\Gamma((0,0)) = \{(-1,0), (0,-1), (1,0), (0,1), \alpha, \beta\}$ ,
- for every  $(n,m) \neq (0,0)$ ,  $\Gamma((n,m)) = \{(n-1,m), (n,m-1), (n+1,m), (n,m+1)\}$ .

Then, this network is bipartite with  $\mathcal{X}_e = \{(n,m) \text{ s.t. } n+m \text{ is even}\}$  and  $\mathcal{X}_o = \{(n,m) \text{ s.t. } n+m \text{ is odd}\} \cup \{\alpha, \beta\}$ . Moreover, since  $\alpha$  and  $\beta$  have a unique neighbor, we know that for every dynamic of opinion: the status of  $\alpha$  (resp.  $\beta$ ) at stage  $t$  is the status of agent  $(0,0)$  at stage  $t-1$ . In particular, both status are equal and it is impossible to reach the cylinder  $(\{\alpha\}, \{\beta\})^+$  of configurations where  $\alpha$  is active whereas  $\beta$  is inactive.

As the nodes  $\alpha, \beta$  share the same neighbor, there is no possibility to store somewhere a partial configuration where  $\alpha$  and  $\beta$  would take different values. We still need to rely on the existence of a storing function. Observe that any storing function  $\theta$  has now the property

$$\theta(\mathcal{X}_o) \subset \mathcal{X}_e \text{ and } \theta(\mathcal{X}_e) \subset \mathcal{X}_o.$$

Hence, it is easy to check that if a partial configuration  $(X_1, Y_1)$  of  $\mathcal{X}_o$  can be stored, and a partial configuration  $(X_2, Y_2)$  of  $\mathcal{X}_e$  can be stored, then the partial configuration  $(X_1 \cup X_2, Y_1 \cup Y_2)$  can also be stored. Hence, one can separate the problem of odd positions and even positions.

We now present classes of bipartite graphs where a storing function exists. As for the case of non-bipartite graphs, storing functions can be related to matching in graphs, but here we can be more specific. Consider a partial configuration  $(X, Y)$  in  $\mathcal{X}_o$ , storable by  $\theta$  and suppose in addition that  $\theta$  is an injection from  $X \cup Y$  to  $\overline{X \cup Y} \subseteq \mathcal{X}_e$ . In terms of graph theory, we may rephrase this by saying that the bipartite graph  $(X \cup Y, \overline{X \cup Y}, E)$  admits a matching of  $X \cup Y$  ( $E$  is of course restricted to the nodes in the bipartite graph).

There are classes of graphs  $(\mathcal{X}, E)$  which admit a matching of  $X \cup Y$  for any partial configuration  $(X, Y)$  in  $\mathcal{X}_o$  and in  $\mathcal{X}_e$ : these are the  $k$ -regular graphs and the hierarchies (see Example 4). A graph is  $k$ -regular ( $k \geq 2$ , integer) if each node has exactly  $k$  neighbors.

**Proposition 3.** If  $(\mathcal{X}, E)$  is  $k$ -regular and bipartite, or is a hierarchy, then any configuration can be stored.

(see proof in Appendix D)

Concerning the second assumption (existence of a complex star), the situation is different. Indeed, a bipartite graph may not admit a complex star. This observation leads to a simple counterexample.

**Example 11.** Consider the network defined by  $\mathcal{X} = \mathbb{Z}$  and the 1-neighborhood:

$$\forall x \in \mathcal{X}, \Gamma(x) = \{x - 1, x + 1\}.$$

It is clear that this graph is bipartite and it does not admit a complex star. We want to show that the set of configurations  $(\infty, \infty, \infty, 0) \cup (\infty, 0, \infty, \infty)$  is not  $\phi$ -irreducible. Consider the configuration  $\omega$  in this set defined by

$$\omega(x) = \begin{cases} 0 & \text{if } x \text{ is odd,} \\ 0 & \text{if } x \leq 0 \text{ and } x \text{ is even,} \\ 1 & \text{if } x > 0 \text{ and } x \text{ is even.} \end{cases}$$

and the cylinder  $\mathcal{A} = (X, Y)^+$  with  $X = \{0, 4\}$  and  $Y = \{2\}$ . We show that the cylinder  $\mathcal{A}$  is not reachable from  $\omega$ . The reason is that at any stage, the current configuration  $\omega'$  has the form  $\omega'(x) = \omega(x - z)$ , for some  $z \in \mathbb{Z}$ , so that the succession 1-0-1 (when looking only at even positions in an even stage or only at odd positions in an odd stage) is not possible. Let us show this for the first transition from  $\omega$ . The new configuration  $\omega'$  is such that  $\omega'(2k + 1) = 1$  for any  $k \geq 1$ ,  $\omega'(1)$  can be either 1 or 0, and  $\omega'(x) = 0$  otherwise. If  $\omega'(1) = 0$ , then the claim is true with  $z = 1$ , and if  $\omega'(1) = 1$ , then the claim is true with  $z = -1$ . As  $\omega'$  is equal to  $\omega$  up to a shift, the same reasoning applies at every stage.

Intuitively, in Example 11, the set of configurations is not  $\phi$ -irreducible because the graph is a line, and there is not enough room to move 0 and 1 without erasing patterns that we want to preserve. We can see in a bipartite graph more clearly why the existence of a complex star is important. It is a device permitting to store an inactive status or an active status along time. Its existence prevents the graph to be similar to a line.

In the following proposition, which characterizes networks having no complex stars, we identify a sequence with the set of its values.

**Proposition 4.** The network  $(\mathcal{X}, E)$  does not contain a complex star if and only if it has one of the following two forms:

- First case:

1. there exists a sequence  $(x_m)_{m \in \mathbb{N}}$  in  $\mathcal{X}^{\mathbb{N}}$  such that all nodes are different,  $\Gamma(x_0) = \{x_1\}$  and for every  $n \geq 1$ ,  $\Gamma(x_n) \cap (x_m)_{m \in \mathbb{N}} = \{x_{n-1}, x_{n+1}\}$ ,
2. for all  $x \in \mathcal{X}$ , if  $x \notin (x_m)_{m \in \mathbb{N}}$  then there exists  $n \in \mathbb{N}$  such that  $\Gamma(x) = \{x_n\}$ .

- Second case:

1. there exists a sequence  $(x_m)_{m \in \mathbb{Z}}$  in  $\mathcal{X}^{\mathbb{Z}}$  such that all nodes are different and for every  $n \in \mathbb{Z}$ ,  $\Gamma(x_n) \cap (x_m)_{m \in \mathbb{Z}} = \{x_{n-1}, x_{n+1}\}$ ,

2. for all  $x \in \mathcal{X}$ , if  $x \notin (x_m)_{m \in \mathbb{Z}}$  then there exists  $n \in \mathbb{Z}$  such that  $\Gamma(x) = \{x_n\}$ .

(see proof in Appendix D)

Observe that all the networks introduced in Examples 1 to 4 satisfy the richness assumption (except  $\mathbb{Z}^d$  with  $d = 1$  which has no complex star), since they are all  $k$ -regular or a hierarchy, and contain a complex star.

### 4.3 $\phi$ -irreducibility

We now turn to the study of  $\phi$ -irreducibility. We show that any set of configurations of Theorem 3 is  $\phi$ -irreducible, hence they are all classes or transient  $\phi$ -irreducible sets.

**Theorem 4.** Assume that  $(\mathcal{X}, E)$  satisfies the Richness Assumption. The following sets are

(i) Finite  $\phi$ -irreducible classes:

**(1)**  $(0, \infty, 0, \infty)$  (this is  $\mathcal{X}$ );

(2)  $(\infty, 0, \infty, 0)$  (this is  $\emptyset$ );

**(3)**  $(0, \infty, \infty, 0) \cup (\infty, 0, 0, \infty)$  (this is the 2-cycle on the bipartition);

(ii) Infinite (uncountable)  $\phi$ -irreducible classes:

**(4)**  $(\infty, \infty, \infty, \infty)$ ;

**(5)**  $(\infty, 0, \infty, \infty) \cup (\infty, \infty, \infty, 0)$ ;

(6)  $(0, \infty, \infty, \infty) \cup (\infty, \infty, 0, \infty)$ ;

(iii) Transient and  $\phi$ -irreducible sets:

**(a)**  $(\infty, F, \infty, F)$ ;

(b)  $(F, \infty, F, \infty)$ ;

**(c)**  $(F, \infty, \infty, F) \cup (\infty, F, F, \infty)$ ;

**(d)**  $(\infty, F, \infty, \infty) \cup (\infty, \infty, \infty, F)$ ;

(e)  $(F, \infty, \infty, \infty) \cup (\infty, \infty, F, \infty)$ ;

**(f)**  $(\infty, F, \infty, 0) \cup (\infty, 0, \infty, F)$ ;

(g)  $(F, \infty, 0, \infty) \cup (0, \infty, F, \infty)$ ;

**(h)**  $(0, \infty, \infty, F) \cup (\infty, F, 0, \infty)$ ;

(i)  $(F, \infty, \infty, 0) \cup (\infty, 0, F, \infty)$ .

Moreover, these are the only ones.

(see proof in Appendix F)

We make some comments on these results.

- As explained for the results of Theorem 3, the results of the non-bipartite case are split due to the introduction of the distinction odd/even nodes, which also provokes the appearance of a cycle and periodic classes. However, the comments we made for Theorem 2 (non-bipartite case) remain essentially valid. Specific comments to the bipartite case follow.
- We have found 3 finite classes, 3 infinite (uncountable) classes and 9 transient sets. We observe that these numbers can be reduced by taking into account symmetries. Indeed, the collection of classes and the collection of transient sets are invariant by interchanging “active” and “inactive” (coordinates 1 and 2, and 3 and 4). This yields respectively 2 types of finite classes ((1) and (3)), 2 types of infinite classes ((4) and (5)), and 5 types of transient sets ((a), (c), (d), (f) and (h)) (we have put these labels in bold). It is therefore enough to restrict our analysis to these classes and sets.

We also observe that every class and every transient set is invariant to interchanging “odd” and “even” (that is, coordinates 1 and 3, and 2 and 4).

- Among the classes we have 2 singletons ((1) and (2)), one cycle of length 2 ((3)), two periodic classes of length 2 ((5) and (6)), and one aperiodic class ((4)). Classes (5) and (6) are periodic, because taking for instance (5) and starting from a configuration in  $(\infty, 0, \infty, \infty)$ , we have at next stage a configuration in  $(\infty, \infty, \infty, 0)$ , and conversely. This is because if there are no active agents on even nodes, all odd nodes have all their neighbors inactive, so all the odd nodes must become inactive at next stage. This is similar to a 2-cycle, up to the difference that the statuses of agents of one block (odd or even, alternatively) is not determined.

## 5 Boolean aggregation functions

The aggregation functions being Boolean, the diffusion model (1) becomes deterministic. In the previous sections, we used heavily the fact that the aggregation functions were strict. Since this is not true anymore, we obtain very different results.

As we have mentioned, the Boolean case comprises as particular case the classical threshold model, which has been studied in depth, together with particular cases like the majority model. Let us mention also that other models, like the voter model and all its variants, use as basic mechanism a simple count in the neighborhood of the number of active agents. This shows how central is the mechanism studied in this section. We start our analysis by imposing additional conditions on the aggregation functions, and by taking as underlying network the grid  $\mathbb{Z}^2$ . We show that a detailed analysis of all possible classes becomes very combinatorial and unfeasible. In a second step, we establish general results for the Boolean case.

### 5.1 Particular cases and examples

Let us start by assuming that the aggregation functions  $A_x$  are anonymous (symmetric) for all  $x \in \mathcal{X}$ . Then for each agent  $x$  there exists a threshold  $0 \leq \tilde{q}_x \leq \gamma$  such that

$$A_x(1_S) = \begin{cases} 1 & \text{if } |S| \geq \tilde{q}_x \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

This corresponds to the classical threshold model introduced, e.g., by Granovetter (1978). In this model, the rule of diffusion with threshold  $0 \leq q_x \leq 1$  for agent  $x$  is the following. Given a configuration  $X(t)$  at time  $t$ , next configuration  $X(t+1)$  is the set of agents  $x$  having a proportion of neighbors in  $X(t)$  at least equal to  $q_x$ :

$$X(t+1) = \left\{ x \in \mathcal{X} : \frac{|\Gamma(x) \cap X(t)|}{|\Gamma(x)|} \geq q_x \right\}. \quad (6)$$

Comparing (6) with (1) and (5) yields the equivalence between the two models.

Assuming in addition that all aggregation functions have the same threshold  $q$  for all agents, we obtain the contagion model of Morris (2000). Consequently, all aggregation functions  $A_x$  are identical, up to the size of the neighborhood, and we denote them by  $A$ . The *contagion threshold*  $\xi$  is the largest  $q$  such that infection spreads over  $\mathcal{X}$  from some finite group  $X(0)$ . Morris (2000) has shown that for any network (in the sense of Section 2.1),  $\xi \leq \frac{1}{2}$ .

By definition, if  $q$  is below the contagion threshold, then the absorbing states are the trivial states  $\emptyset$  and  $\mathcal{X}$ . Otherwise, other nontrivial absorbing classes may occur, as will be shown.

Let us adopt the assumptions of Morris' model and try first to find nontrivial absorbing states. As we noted earlier,  $X$  is an absorbing state if it is a fixed point, i.e.,  $X(t+1) = X(t)$ . We introduce the following convenient notation. The *frontier* of  $X$  is the set

$$\partial X = \overline{X} \setminus \overset{\circ}{X} = \{x \in \mathcal{X} : \Gamma(x) \cap X^c \neq \emptyset \text{ and } \Gamma(x) \cap X \neq \emptyset\}.$$

We also introduce  $\partial^+ X = \partial X \cap X$  the *inner frontier* of  $X$ , and  $\partial^- X = \partial X \cap X^c$  the *outer frontier* of  $X$ . The following results are immediate.

**Proposition 5.** Suppose  $A$  is an anonymous Boolean aggregation function.  $X$  is an absorbing state if and only if:

1. Each inner frontier node  $x \in \partial^+ X$  has at least  $\lceil \gamma q \rceil$  neighbors in  $X$ , and
2. Each outer frontier node  $x \in \partial^- X$  has at least  $\lfloor \gamma(1-q) \rfloor + 1$  neighbors in  $X^c$ .

*Proof.*  $X$  will not lose an element if each element of  $X$  has at least  $\gamma q$  neighbors in  $X$ , and  $X$  will not attract a new element if each element of  $\mathcal{X} \setminus X$  has strictly less than  $\gamma q$  neighbors in  $X$ , hence at least  $\lfloor \gamma(1-q) \rfloor + 1$  neighbors in  $\mathcal{X} \setminus X$ .  $\square$

We examine two particular networks from Example 1 for illustration.

**The case  $\mathbb{Z}^2$  with the 1-neighborhood** Let us consider various values of  $q$ . The following table gives the minimal number of neighbors in  $X$  for frontier nodes, so that  $X$  is an absorbing state.

$q$	1/4	1/2	3/4	1
minimal nb of interior neighbors for nodes in $X$	1	2	3	4
minimal nb of interior neighbors for nodes outside $X$	4	3	2	1

Let us describe the shape of the frontier of a connected set  $X$  with  $|X| > 1$  satisfying the above requirements. To this end, we consider a frontier node  $x$  of  $X$  and the number of its interior neighbors  $\eta(x) := |\Gamma(x) \cap X|$ .

1.  $\eta(x) \geq 1$  for every frontier node  $x$ . This is true for any  $X$  such that  $|X| \geq 2$ .
2.  $\eta(x) \geq 2$  for every frontier node  $x$ . Only the situation  $\eta(x) = 1$  is forbidden, which corresponds to “antennas” (see Figure 11, left). Therefore,  $X$  is any shape without antennas, with  $|X| \geq 3$ .
3.  $\eta(x) \geq 3$  for every frontier node  $x$ . Then additional forbidden situations are “convex corners” and “isthms” (see Figure 11). Then  $X$  is any shape without antennas, convex corners and isthms, and  $|X| \geq 4$ . Note that since  $X$  may be infinite, one can have shapes with concave corners (which are allowed) but without convex corners (see Figure 12).
4. Evidently,  $\eta(x) \geq 4$  is impossible by definition of frontier nodes.

From the previous analysis, it follows that only  $q = \frac{1}{2}$  or  $\frac{3}{4}$  may lead to nontrivial absorbing classes. Also, these two cases are exact complements of each other, in the sense that  $X$  is a possible absorbing state for  $q = \frac{1}{2}$  if and only if  $\mathcal{X} \setminus X$  is a possible absorbing state for  $q = \frac{3}{4}$ .

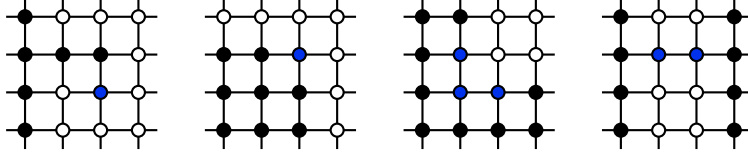


Figure 11:  $4 \times 4$  portion of  $\mathbb{Z}^2$ , where  $X$  is in black or blue. By convention, the color of nodes extends infinitely in any direction. From left to right: antenna ( $x$  such that  $\eta(x) = 1$  in blue), convex corner ( $x$  such that  $\eta(x) = 2$  in blue), concave corner in blue, isthm ( $x$  such that  $\eta(x) = 2$  in blue)

Taking for example the latter, each connected component of  $X$  should be of size at least 4, and should have no convex corner, no antenna and no isthm, while each connected component of the complement set should be of size at least 3 and have no antennas. Obvious examples are half-planes, concave corners extending infinitely in both directions, and infinite shapes with squared “holes” of width greater than 1 (see the example in Figure 12).

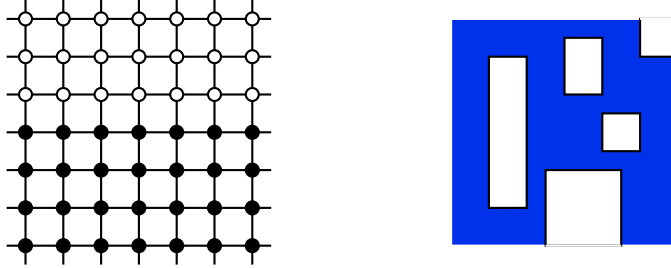


Figure 12: Examples of absorbing states with  $q = \frac{3}{4}$ . Left: horizontal half-plane. Right: shape with squared holes (in blue; the grid is not represented)

**The case of  $\mathbb{Z}^2$  with the  $\sqrt{2}$ -neighborhood** We obtain the following table.

$q$	1/8	1/4	3/8	1/2	5/8	3/4	7/8
minimal nb of interior neighbors for nodes in $X$	1	2	3	4	5	6	7
minimal nb of interior neighbors for nodes outside $X$	8	7	6	5	4	3	2

The analysis of possible shapes in this topology reveals to be much more complex and we will not provide a full analysis as in the previous example. Obviously,  $q = \frac{1}{8}$  is ruled out. If  $\eta(x) = 7$ , then the only possibility is that  $\mathcal{X} \setminus X$  is a singleton. Since then  $\mathcal{X} \setminus X$  cannot satisfy the condition  $\eta(x) \geq 2$ , this rules out  $q = \frac{1}{4}$  and  $q = \frac{7}{8}$ . If  $\eta(x) \geq 3$ , as before, antennas and isthms of length greater than 2 are forbidden, but not convex corners. For  $\eta(x) \geq 4$ , squared convex corners are no more allowed (but  $45^\circ$  corners are possible). For  $\eta(x) \geq 5$ , no corners are allowed except concave ones. For  $\eta(x) \geq 6$ , a remarkable situation is where  $\mathcal{X} \setminus X$  is a square with edges of length 1. Therefore, as possible absorbing states (we do not pretend to be exhaustive) we find:

1.  $q = 3/8$ : a connected component of  $X$  is a square with edge of length 1 (see Figure 13, left)



2.  $q = 1/2$  or  $q = 5/8$ : the two cases are complements of each other. Taking the latter, a connected component of  $X$  is any shape of size at least 6 without convex corners, antennas nor isthms, and each connected component of the complement set should be of size at least 5 and have no antennas, isthms and squared convex corners. For example, half-planes are allowed, as well as infinite shapes with squared holes whose corners are “rounded” (see Figure 13, right).
3.  $q = 3/4$ : a connected component of  $\mathcal{X} \setminus S$  is a square with edge of length 1.

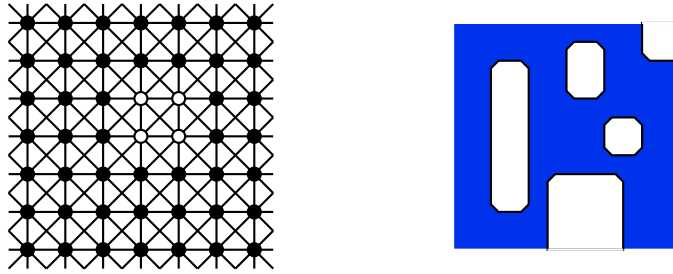


Figure 13: Left: example of absorbing state with  $q = \frac{3}{4}$ . Right: example with  $q = \frac{5}{8}$  (in blue; the grid is not represented)

**Other absorbing classes** Let us reconsider the notion of trajectory. As the process is deterministic, a (deterministic) trajectory is merely a sequence of states  $X_1, X_2, X_3, \dots$ . We show that cycles (periodic trajectories) and infinite (aperiodic) trajectories may exist. Figure 14 presents an example of a cycle, and Figure 15 an example of an infinite trajectory.

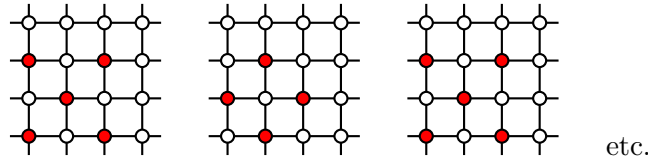


Figure 14: Example of a cycle with  $\mathbb{Z}^2$  and the 1-neighborhood ( $q = \frac{1}{2}$ )

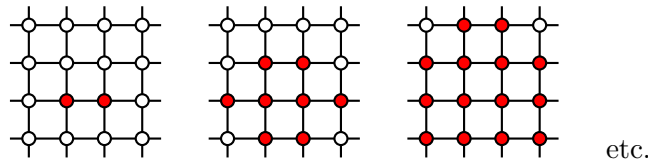


Figure 15: Example of an infinite trajectory with  $\mathbb{Z}^2$  and the 1-neighborhood ( $q = \frac{1}{4}$ ). The initial  $X$  constantly grows.

## 5.2 General results

We now assume that  $\mathbf{A}$  is Boolean, without any additional assumption, and  $(\mathcal{X}, E)$  is arbitrary. The next proposition establishes the different types of possible absorbing classes.

**Proposition 6.** Suppose  $\mathbf{A}$  is Boolean. Then absorbing classes are either singletons  $\{X\}$ , where  $X \in 2^{\mathcal{X}}$ , or cycles (periodic trajectories) of nonempty sets  $\{X_1, \dots, X_k\}$  with the condition that all sets are pairwise incomparable by inclusion.

*Proof.* Every transition being with probability 1, the evolution is a deterministic trajectory, hence a sequence of configurations  $X_0, X_1, \dots, X_k, \dots$ . Note that if a repetition occurs, then the trajectory is periodic. The following cases are exhaustive and exclusive:

1. Fixed points,  $X_{k+1} = X_k$  for  $k \geq K$  exist, and are absorbing states.
2. Cycles (periodic trajectories) exist, and are periodic absorbing classes.
3. Otherwise, the trajectory does not have any repetition and is infinite. But this cannot be a class, because there would be no transition from state  $X_{k+1}$  to  $X_k$ .

□

The results show that, contrarily to the probabilistic model with strict aggregation functions, polarization might arise: there could exist finite or infinite sets of agents being active or inactive forever, provided these sets satisfy some properties regarding their “shape”. Cycles may also occur.

As exemplified by the results in Section 5.1, it seems very difficult to give more precise results on absorbing states in the general case. However, something general can be said on cycles, thanks to the results obtained by Moran (1995) for the threshold model. Moran proved that for infinite graphs using a majority model (threshold model with  $q = \frac{1}{2}$ ), the classical result saying that for the threshold model in its general form (non-anonymous aggregation functions), cycles have length 2 (see Goles and Olivos (1980)), is no more valid, as longer cycles may exist. Specifically, Moran proved that cycles have length 2 if two conditions are satisfied: 1) the degree of a node is bounded 2) the growth of any closed ball is subexponential.

As the first condition is satisfied in our model ( $\gamma$  is the upper bound), we elaborate on the second condition. First, we give a precise definition. The *distance* between two nodes  $x, y$  is the length of a shortest path between  $x$  and  $y$ . We define the closed ball  $B(x, n)$  centered at  $x \in \mathcal{X}$  of radius  $n \in \mathbb{N} \cup \{0\}$  as the set of nodes lying within distance  $n$  to  $x$ . We denote by  $b(x, n)$  its cardinality. The *growth* of  $(\mathcal{X}, E)$  is defined by

$$g(\mathcal{X}, E) = \limsup_{n \in \mathbb{N}} (b(x, n))^{1/n}$$

for any node  $x$  (if  $(\mathcal{X}, E)$  is locally finite and connected, which is our case, the choice of  $x$  is unimportant). The growth is *subexponential* if  $g(\mathcal{X}, E) = 1$ .

Let us compute the growth in the grid  $\mathbb{Z}^2$ , with the  $\sqrt{2}$ -neighborhood (8 neighbors). We have  $b(x, 1) = 8 + 1$ , then  $b(x, 2) = b(x, 1)$  plus the additional neighbors which form a  $5 \times 5$  square (16 nodes), etc. Therefore

$$b(x, n) = b(x, n - 1) + 8n = 8(n + (n - 1) + \dots + 1) + 1 = 4n(n + 1) + 1.$$

More generally, for the  $\mathbb{Z}^d$  grid,  $d \geq 2$ , and the neighborhood given by the all nodes in the hypercube of side 2 centered at  $x$  (all nodes at Euclidean distance  $\sqrt{d}$ ), we obtain

$$b(x, n) = b(x, n - 1) + (2n)2^d = 2^d n(n + 1) + 1.$$

In any case, we get

$$\lim_{n \rightarrow \infty} 2^{\frac{d}{n}} (n(n + 1))^{\frac{1}{n}} = 1,$$

hence the growth is subexponential. It can be noticed that augmenting the Euclidean distance for defining the neighborhood does not change the conclusion. We have shown:

**Proposition 7.** Consider the grid  $\mathbb{Z}^d$  and the  $k$ -neighborhood, where  $k$  is the Euclidean distance and  $A_x$  is the majority aggregation function for all  $x \in \mathcal{X}$ . Then no cycle of length greater than 2 exists.

We end this section by giving an example of a graph admitting a cycle of length greater than 2. The typical example of a graph with an exponential growth is the hierarchy (Example 4), under the condition that  $m(x) > 1$  for all  $x \in \mathcal{X}$ . Moran (1995) defines a shift system as a generic example to produce cycles of arbitrary length. It consists in partitioning the graph in blocks of increasing size, and such that each node of block  $k$  has its neighborhood contained in the two neighbor blocks  $k-1, k+1$ , with more neighbors in block  $k+1$ . Considering a hierarchy  $(\mathcal{X}, E)$ , we define the layer  $L_k$  as the set of nodes at distance  $k$  from the root, for some  $k \geq 1$ . Then the partition of  $(\mathcal{X}, E)$  into the layers  $1, 2, \dots$ , is a shift system, as it is easy to check. Suppose now that the initial configuration is:

$$x \in L_k \text{ is active} \Leftrightarrow k = 0 \pmod{\ell}$$

for some fixed integer  $\ell > 2$ . Then the system enters a cycle of length  $\ell$ .

## 6 Concluding remarks

Our main question was how the diffusion evolves from a finite set of active agents and more generally from any set of active agents. The main finding is that the answer depends on the aggregation function for some properties and on the structure of the network for others. We first showed that the transience/persistence of a state relies only on the type of diffusion mechanism, which in our model amounts to the mechanism of aggregating the statuses of the neighbors, without any further restriction on the network. On the contrary, the possibility to go from one configuration to another one inside a class (irreducibility) is closely related to the structure of the graph. We provide a mild sufficient condition on the structure, called richness, which permits to obtain irreducibility.

Among diffusion mechanisms, we clearly establish a distinction between the probabilistic and the deterministic mechanism, the latter being nothing other than the threshold model, studied by Morris (2000). In the former, we have supposed in our analysis that the probability of being active (resp., inactive) becomes positive as soon as one neighbor is active (resp., inactive). Under this assumption, we have shown that no polarization can occur, even in the weak sense: no set of active agents can remain stable, even if we allow some variation around this set. On the contrary, the diffusion is erratic and homogeneous on the whole network, and does not fix on some peculiar region of it. By contrast, the deterministic model allows the appearance of stable finite or infinite sets of active/inactive agents, that is, polarization can appear, and under many different forms.

The next step is to study nonstrict nonBoolean aggregation function, which reveals to be pretty tricky. To highlight the difficulties, let us assume for simplicity that all aggregation functions  $A_x$  are equal to some aggregation function  $A$ , which is anonymous, nonstrict and nonBoolean. This implies that there exist  $\ell, r \in \{0, \dots, \gamma - 1\}$  such that

$$A(1_S) = \begin{cases} 0 & \text{iff } |S| \leq \ell \\ 1 & \text{iff } |S| \geq \gamma - r, \end{cases} \quad (7)$$

with  $\ell \vee r > 0$  and  $\ell + r < \gamma$ . This case is more complex than the two previous ones. Indeed, it has the aspect both of the Boolean case and of the strict aggregation case. For example, take  $\mathbb{Z}^2$  with the 1-neighborhood and the following anonymous aggregation function

$$A(1_S) = \begin{cases} 0 & \text{iff } |S| = 0 \\ 1/2 & \text{iff } |S| = 1, \\ 1 & \text{iff } |S| \geq 2. \end{cases} \quad (8)$$

We see an asymmetry between the role of 0 and 1.

Proposition 8 that focuses on the persistence of the number of elements is only partially true here: the two first bullet points concerning the case with no active agents or the case with a finite number of active agents are still valid. On the contrary, the third one concerning the case with an infinite number of agents is not true anymore, as shown by the following configuration. Let  $\omega$  be such that

$$\forall (n, m) \in \mathbb{Z}^2, \omega(n, m) = 0 \text{ if and only if } n + m \in 10\mathbb{Z}.$$

The transition is deterministic and leads to  $\mathbf{1}$  since the inactive agents are all isolated. Proposition 9 that proves the possibility for a finite number of active agents to disappear is not true anymore, since a square of active agents will never disappear. On the contrary, with positive probability a finite set of inactive agents may disappear.

This makes the study of nonstrict nonBoolean aggregation function pretty tricky, and we leave it for further research. It will need to disentangle the condition for each of the lemmas developed for strict aggregation functions and boolean aggregation functions.

## Acknowledgment

The authors thank Matthew Jackson, Ron Peretz and Marcus Pivato for fruitful discussions, and three anonymous referees for careful reading and constructive suggestions.

## Funding

M. Grabisch and A. Rusinowska acknowledge the financial support by the National Agency for Research, Project DynaMITE (ANR-13-BSH1-0010-01). X. Venel acknowledges the financial support by the National Agency for Research, Project CIGNE (ANR-15-CE38-0007-01).

## Conflict of interest

The authors declare that they have no conflict of interest.

## References

- D. Acemoglu and A. Ozdaglar. Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1:3–49, 2011.
- D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi. Networks, shocks, and systemic risk. In Y. Bramoullé, A. Galeotti, and B. W. Rogers, editors, *The Oxford Handbook of the Economics of Networks*, pages 569–610. Oxford University Press, 2016.

- E. Adam, M. Dahleh, and A. Ozdaglar. On the behavior of threshold models over finite networks. In *Proceedings of the 2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, 2012.
- C. Aliprantis and K. Border. *Infinite Dimensional Analysis*. Springer, 2006. 3rd edition.
- N. Bailey. *The Mathematical Theory of Infectious Diseases and its Applications*. London: Charles Griffin and Company Ltd, 1975.
- F. M. Bass. A new product growth model for consumer durables. *Management Science*, 15: 215–227, 1969.
- E. Berger. Dynamic monopolies of constant size. *Journal of Combinatorial Theory*, 83:191–200, 2001.
- Y. Bramoullé, A. Galeotti, and B. W. Rogers, editors. *The Oxford Handbook of the Economics of Networks*. Oxford University Press, 2016.
- A. Cabrales, D. Gale, and P. Gottardi. Financial contagion in networks. In Y. Bramoullé, A. Galeotti, and B. W. Rogers, editors, *The Oxford Handbook of the Economics of Networks*, pages 543–568. Oxford University Press, 2016.
- R. Diestel. *Graph Theory (Graduate Texts in Mathematics)*. Springer, 2005.
- P. A. Dreyer Jr. and F. S. Roberts. Irreversible  $k$ -threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion. *Discrete Applied Mathematics*, 157:1615–1627, 2009.
- P. Erdős and A. Renyi. On the existence of a factor of degree one of a connected random graph. *Acta Mathematica Hungarica*, 17(3-4):359–368, 1966.
- M. Förster, M. Grabisch, and A. Rusinowska. Anonymous social influence. *Games and Economic Behavior*, 82:621–635, 2013.
- J. Gao, G. Ghasemiesfeh, G. Schoenebeck, and F.-Y. Yu. General threshold model for social cascades: Analysis and simulatons. In *EC 2016 - Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 617–634. Association for Computing Machinery, Inc, 2016.
- E. Goles and J. Olivos. Periodic behaviour of generalized threshold functions. *Discrete Mathematics*, 30:187–189, 1980.
- B. Golub and E. Sadler. Learning in social networks. In Y. Bramoullé, A. Galeotti, and B. W. Rogers, editors, *The Oxford Handbook of the Economics of Networks*, chapter 19, pages 504–542. Oxford University Press, 2016.
- M. Grabisch and A. Rusinowska. A model of influence based on aggregation functions. *Mathematical Social Sciences*, 66:316–330, 2013.
- M. Grabisch and A. Rusinowska. A survey on nonstrategic models of opinion dynamics. *Games*, 11(4):65, 2020.
- M. Grabisch, A. Poindron, and A. Rusinowska. A model of anonymous influence with anti-conformist agents. *Journal of Economic Dynamics and Control*, 109:103773, 2019.

- M. Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83: 1420–1443, 1978.
- J. Gravner and D. Griffeath. Cellular automaton growth on  $\mathbb{Z}^2$ : Theorems, examples and problems. *Advances in Applied Mathematics*, 21:241–304, 1998.
- O. Hernández-Lerma and J. B. Lasserre. *Markov Chains and Invariant Probabilities*. Springer-Verlag, Basel, 2003.
- R. A. Holley and T. M. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter models. *Annals of Probability*, 3:643–663, 1975.
- M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.
- M. O. Jackson and B. W. Rogers. Relating network structure to diffusion properties through stochastic dominance. *The B.E. Journal of Theoretical Economics (Advances)*, 7(1):1–13, 2007.
- M. Keeling and P. Rohani. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2008.
- D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. *Proceedings of 9th International Conference on Knowledge Discovery and Data Mining*, pages 137–146, 2003.
- P. J. Lamberson. Diffusion in networks. In Y. Bramoullé, A. Galeotti, and B. W. Rogers, editors, *The Oxford Handbook of the Economics of Networks*, pages 479–503. Oxford University Press, 2016.
- D. López-Pintado. Diffusion in complex social networks. *Games and Economic Behavior*, 62(2): 573–590, 2008.
- G. Moran. On the period-two-property of the majority operator in infinite graphs. *Transactions of the American Mathematical Society*, 347(5):1649–1667, 1995.
- S. Morris. Contagion. *Review of Economic Studies*, 67:57–78, 2000.
- D. Peleg. Local majorities, coalitions and monopolies in graphs: a review. *Theoretical Computer Science*, 282:231–257, 2002.
- S. Poljak and M. Sura. On periodical behaviour in societies with symmetric influences. *Combinatorica*, 3(1):119–121, 1983.
- E. M. Rogers. *Diffusion of Innovations*. Simon and Schuster, 5th edition, 2003.
- T. Schelling. *Micromotives and Macrobehavior*. Norton, 1978.
- G. Silverman. *The Secrets of Word-of-Mouth Marketing: How to Trigger Exponential Sales through Runaway Word of Mouth*. New York: Amacom Books, 2001.
- E. Yildiz, D. Acemoglu, A. Ozdaglar, and A. Scaglione. Diffusion of innovations on deterministic topologies. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

## A Diffusion process and proofs of the first results

**Formal definition of the diffusion process** We define the diffusion process as a Markov process whose set of states is the set  $\Omega$  of configurations, based on Equation (1) and the independence across agents. As  $\Omega$  is uncountable, the definition of the Markov process requires to work on  $\sigma$ -fields (see, e.g., Hernández-Lerma and Lasserre (2003)).

We consider on  $\Omega$  the product topology which is generated by the finite cylinders and we denote by  $\mathcal{T}$  the Borelian  $\sigma$ -field associated to this topology. It is also the product  $\sigma$ -field.

Let us now define a Markov Kernel on  $\Omega$ , i.e., a mapping  $K$  from  $\Omega \times \mathcal{T}$  to  $[0, 1]$  such that:

- For every  $\omega \in \Omega$ ,  $K(\omega, \cdot)$  is a probability on  $\mathcal{T}$ ,
- For every  $\mathcal{A} \in \mathcal{T}$ ,  $K(\cdot, \mathcal{A})$  is measurable.

$K(\omega, \mathcal{A})$  can be interpreted as the probability that from configuration  $\omega$  the process jumps at the next time step into a configuration belonging to  $\mathcal{A}$ . Similarly,  $K^n(\omega, \mathcal{A})$  is the probability that in  $n$  steps the process jumps from  $\omega$  into a configuration in  $\mathcal{A}$ .

Fix  $\omega \in \Omega$ . In order to define  $K(\omega, \cdot)$ , we first construct a family of probability distributions  $(\mu_{Y,\omega})_{Y \subset \mathcal{X}, Y \text{ finite}}$  on finite partial configurations. Consider  $h \in \{0, 1\}^Y$ , we set

$$\mu_{Y,\omega}(\{h\}) = \prod_{y \in Y} (P(y | \omega)h(y) + (1 - P(y | \omega))(1 - h(y))), \quad (9)$$

where  $P(y | \omega) = A(\pi_{\Gamma(y)}(\omega))$  (this is Eq. (1)).  $\mu_{Y,\omega}(\{h\})$  is the probability that, given the present configuration  $\omega$ , the next (partial) configuration in  $Y$  is  $h$ . Notice that this probability does not depend on the entire  $\omega$  but on its projection on  $\bar{Y}$ . For every  $Y$ ,  $\mu_{Y,\omega}$  is a probability measure on the Borel  $\sigma$ -field of  $Y$  (finite set). Moreover, this family of probability measures satisfies the assumption of Kolmogorov's Extension Theorem (Aliprantis and Border, 2006, Corollary 15.27). It follows that there exists  $K(\omega, \cdot)$ , a probability distribution over  $\Omega$  and the infinite product  $\sigma$ -field, for every  $\omega \in \Omega$ .

**Lemma 1.** The function  $K(\cdot, \cdot)$  constructed above is a Markov kernel on  $\Omega$ .

*Proof.* By construction,  $K(\omega, \cdot)$  satisfies the first hypothesis of a kernel. We now check that given  $\mathcal{A} \in \mathcal{T}$ , the function  $K(\cdot, \mathcal{A})$  is measurable. We first consider a set  $\mathcal{A}$  in the basis of  $\mathcal{T}$ . Hence, it is a finite cylinder, i.e., there exist two finite disjoint sets  $X, Y$  such that  $\mathcal{A} = (X, Y)^+$ . Then we know that  $K(\omega, (X, Y)^+)$  depends only on the restriction of  $\omega$  to agents in  $\bar{X} \cup Y$ , hence on a finite set of agents. By definition of the product topology, the function is continuous and therefore measurable. Let

$$\mathcal{H} = \{\mathcal{A} \in \mathcal{T}, \omega \rightarrow K(\omega, \mathcal{A}) \text{ is measurable}\} \subset \mathcal{T}$$

be the set of sets in  $\mathcal{T}$  such that the mapping is measurable. Then  $\mathcal{H}$  is a  $\sigma$ -field since:

- it is nonempty:  $\emptyset \in \mathcal{H}$ ;
- it is stable under complementation since  $K(\omega, \bar{\mathcal{A}}) = 1 - K(\omega, \mathcal{A})$ ;
- it is stable under countable unions: Let  $(\mathcal{A}_n)_{n \in \mathbb{N}}$ , then

$$K(\omega, \cup_{n \in \mathbb{N}} \mathcal{A}_n) = \sum_{n \in \mathbb{N}} K(\omega, \mathcal{A}_n).$$

The countable sum of measurable mappings is also measurable.

Since  $\mathcal{H}$  contains the finite cylinders, it follows that  $\mathcal{T} \subset \mathcal{H}$  and we have the equality: for every  $\mathcal{A} \in \mathcal{T}$ , the function  $K(\cdot, \mathcal{A})$  is measurable.  $\square$

We can now rewrite Equations (2) in terms of the Markov kernel. Switching to the set notation, let  $X$  and  $Y$  be two disjoint sets,  $(X, Y)^+$  is the interval  $[X, Y^c]$ , where  $Y^c$  is the complement of  $Y$ , and with some abuse of notation we keep the same symbol for the  $\sigma$ -field  $\mathcal{T}$ . Given that  $\overset{\circ}{X}$  is the present configuration, the configuration at the next time step lies in the interval  $[\overset{\circ}{X}, \overline{X}]$  with probability 1:

$$K(X, \mathcal{A}) = 1 \text{ if } [\overset{\circ}{X}, \overline{X}] \subseteq \mathcal{A} \quad (X \subseteq \mathcal{X}). \quad (10)$$

Observe that if  $X$  is finite,  $[\overset{\circ}{X}, \overline{X}]$  can easily be seen to be an element of  $\mathcal{T}$  since

$$[\overset{\circ}{X}, \overline{X}] = \bigcap_{x \notin \overline{X}} (\overset{\circ}{X}, \{x\})^+.$$

The same conclusion holds with infinite sets as well, so that any interval  $[X, Y]$  is an element of  $\mathcal{T}$ , and so are the singletons  $\{X\}$  for any  $X \subseteq \mathcal{X}$ .

Formula (9) permits to compute the kernel when everything is finite. Specifically, with  $X, Y$  finite, we have  $K(X, \{Y\}) = \mu_{\overline{X}, X}(\{Y\})$ . We can also compute the kernel for every element of the basis. Supposing  $\mathcal{A} = (X, Y)^+$ , we have simply  $K(Z, \mathcal{A}) = \mu_{X \cup Y, Z}(\{X\})$ . By  $\sigma$ -additivity, we can deduce that  $K(X, \mathcal{A}) = 0$  if  $|X|$  is infinite and  $|\mathcal{A} \cap [\overset{\circ}{X}, \overline{X}]|$  is finite.

**Proof of Proposition 1** Observe that  $X$  is a steady state (fixed point) iff  $\overset{\circ}{X} = X = \overline{X}$ .

(1) Let  $X$  satisfy  $\overset{\circ}{X} = \overline{X} = X$  and suppose that  $X \neq \emptyset, \mathcal{X}$ . Then there exists  $x \in X$  and  $y \notin X$ . By connectedness, there exists a finite path  $x_1 = x, x_2, \dots, x_k = y$  connecting  $x$  to  $y$ . However, as  $X = \overset{\circ}{X}$ ,  $\Gamma(x_1) \subseteq X$ , and consequently  $\Gamma(x_2) \subseteq X$ ,  $\Gamma(x_3) \subseteq X$ , etc., till  $\Gamma(x_{k-1}) \subseteq X$ . However,  $y = x_k \in \Gamma(x_{k-1})$ , a contradiction.

The converse statement is obvious.

(2) We only have to prove that if  $X \neq \emptyset, \mathcal{X}$ ,  $X$  cannot be a fixed point. The above reasoning can be used without change because when  $A$  is strict, the implications in (2) become equivalences.

**Proof of Proposition 2.** Observe that a cycle is a sequence of deterministic transitions with no fixed points. Assume that  $\mathbf{A}$  is not Boolean, i.e., at least one  $A_x$  is not Boolean. Then a transition from  $X$  is deterministic iff  $\overset{\circ}{X} = \overline{X}$ , and we must have  $\overset{\circ}{X} \neq X$  in order to avoid fixed points. We claim that if  $X$  is such that  $\overset{\circ}{X} = \overline{X} \neq X$ , then (1)  $X^c$  has the same property; (2)  $\overset{\circ}{X} = X^c$ .

*Proof of the claim:* We first observe that  $\overset{\circ}{X} = \overline{X} \neq X$  is equivalent to the two conditions:

- Any  $x \in \mathcal{X}$  has all its neighbors in  $X$  or none of them. (\*)
- Either there exists  $x \notin X$  which has (all) neighbors in  $X$  or there exists  $x \in X$  with no neighbor in  $X$ . (\*\*)

1. The two conditions (\*) and (\*\*) are invariant to the change  $X \rightarrow X^c$ .



2. Assume by (\*\*) that there exists  $x \in X$  with no neighbor in  $X$  (the other case works by (\*)). We claim that  $X \cap \overline{X} = \emptyset$ . Indeed, suppose not and consider  $y \in X \cap \overline{X}$ . By connectedness, there exists a finite path  $x_1 = x, x_2, \dots, x_k = y$  connecting  $x$  to  $y$ . However,  $x_2 \in \Gamma(x_1)$ , therefore  $x_2 \in X^c$ . Since the graph is undirected,  $\Gamma(x_2) \ni x_1$ , which by (\*) implies that  $\Gamma(x_2) \subseteq X$ , i.e.,  $x_2 \in \overline{X}$ . Therefore,  $x_3 \in X$ , but  $x_3 \notin \overline{X}$  because by symmetry  $X^c \ni x_2 \in \Gamma(x_3)$ . Hence,  $x_\ell$  with odd index  $\ell$  belongs to  $X \cap (\overline{X})^c$  while those with even indices belong to  $X^c \cap \overline{X}$ . Therefore,  $y$  can never be reached, a contradiction.

By the same argument, it follows that  $X^c \cap (\overline{X})^c = \emptyset$ . Consequently,  $X^c = \overline{X}$ .

The claim is proved, which shows that the only possible deterministic transition which is not a fixed point is  $X \rightarrow X^c$ , which yields a cycle of length 2.

It remains to show that such a transition exists iff the graph  $(\mathcal{X}, E)$  is bipartite with bipartition  $(X, X^c)$ . Suppose that  $X$  with the above property exists and take  $x \in X$ . Then  $\Gamma(x) \cap X = \emptyset$ . Indeed, if  $y \in \Gamma(x) \cap X$ , then  $y \in \overline{X} = \overset{\circ}{X} = X^c$  by Claim (2), a contradiction. The same property holds for  $X^c$  by Claim (1). Therefore,  $(X, X^c)$  is a bipartite graph. In addition it is well-known that the bipartition of a bipartite graph is unique.

The converse statement is obvious.

## B Proof of Theorem 1

We have only to prove (ii) and (iii). In Section B.1, we focus on the “number” of inactive/active nodes on even/odd positions and prove that the sets containing only a null or infinite number of each element are absorbing. Then, in Section B.2, we prove that the sets containing finitely many elements are transient.

### B.1 Persistence of the “number” of elements

The following proposition states that the “number” of active agents stay constant along the dynamics.

**Proposition 8.** For every  $\omega \in \Omega$ , let  $\omega'$  be a configuration randomly chosen according to the kernel  $K(\omega, \cdot)$ . Then:

- if  $\{x \in \mathcal{X} \mid \omega(x) = 1\}$  is empty then  $\omega'$  has almost surely no active agents;
- if  $\{x \in \mathcal{X} \mid \omega(x) = 1\}$  is finite then  $\omega'$  has almost surely a finite (maybe zero) number of active agents;
- if  $\{x \in \mathcal{X} \mid \omega(x) = 1\}$  is infinite then  $\omega'$  has almost surely an infinite number of active agents .

*Proof.* Denote by  $\Theta = \{x \in \mathcal{X} \mid \omega(x) = 1\}$  the set of active agents.

First, it is clear that if this set is empty, then every position is surrounded by only inactive agents and therefore is inactive with probability 1 at the next stage. Hence, there is almost surely no active node in  $\omega'$ .

Secondly, let  $x$  be a position. Since the aggregation function is strict, the probability that  $\omega'(x)$  is equal to 1 is strictly positive, if and only if, there exists a neighbor  $x'$  of  $x$  such that  $\omega(x') = 1$ . Each agent has at most  $\gamma$  neighbors, therefore almost surely

$$|\{x \in \mathcal{X} \mid \omega'(x) = 1\}| \leq \gamma |\{x \in \mathcal{X} \mid \omega(x) = 1\}|.$$

Hence, if the right-hand side is finite, so is the left-hand side.

Finally, we assume that  $\{x \in \mathcal{X} \mid \omega(x) = 1\}$  is infinite. We compute the probability that  $\omega'$  has a finite number of active nodes on odd states. If this is the case, it implies that there exists an infinite number of elements  $x \in \Theta$  such that

$$\forall x' \in \Gamma(x), \omega'(x') = 0.$$

But by assumption,  $x'$  has at least one neighbor which is active, it follows that the probability that  $\omega'(x') = 0$  is strictly smaller than 1. Since the probabilities are independent, it follows that the total probability of the event is 0.  $\square$

Theorem 1.(ii) is then an immediate consequence of the above proposition.

## B.2 Transient set

We now prove that if there is a strictly positive but finite number of active agents, then there is a positive probability to reach a configuration where there are no active agents anymore. Similar results hold for inactive agents. This result can be expressed in terms of reachability.

**Lemma 2.** Let  $X \subset \mathcal{X}$ .

- If  $X$  is a finite set, then  $\text{int}(\text{int}(X)) \subset X$ .
- If  $X$  is a co-finite set, then  $X \subset \text{clo}(\text{clo}(X))$ .

*Proof.* We prove the first result. The second one can be deduced immediately by complementation.

We start by proving the inclusion, then in a second step we prove that the inclusion is strict. Assume that  $\text{int}(\text{int}(X)) \neq \emptyset$ , otherwise we are done, and take  $x \in \text{int}(\text{int}(X))$ . Then  $\Gamma(x) \subseteq \text{int}(X)$ . Take  $y \in \Gamma(x)$ . Then  $\Gamma(y) \subseteq X$ , hence  $x \in X$  since  $x \in \Gamma(y)$ .

Suppose by contradiction that for all  $x \in X$ ,  $x \in \text{int}(\text{int}(X))$ . It follows that for all  $x \in X$ ,  $\Gamma(x) \subseteq \text{int}(X)$ , hence

$$\forall x \in X, \forall y \in \Gamma(x), \forall z \in \Gamma(y), z \in X. \quad (11)$$

Since  $X$  is finite, pick some  $t \in X^c$ . By connectedness of the graph, there should exist a finite path from  $x \in X$  to  $t \notin X$ . However, no such path can exist by Equation (11).

Therefore, there exists  $x \in X$  such that  $x \notin \text{int}(\text{int}(X))$ . This proves that the inclusion is strict.  $\square$

We can immediately deduce from Lemma 2 the following proposition.

**Proposition 9.** For every  $\omega \in \Omega$ , if  $\{x \in \mathcal{X} \mid \omega(x) = 1\}$  is finite then the configuration  $\mathbf{0}$  is reachable.

*Proof.* As  $X := \{x \in \mathcal{X} \mid \omega(x) = 1\}$  is finite, Lemma 2 implies that the sequence  $\text{int}^2(X), \text{int}^4(X), \dots$  converges to  $\emptyset$  in a finite number of steps. Since  $A$  is a strict aggregation rule, for every  $n \geq 0$ , the probability to reach  $\text{int}^{2n+2}(X)$  from  $\text{int}^{2n}(X)$  is positive, hence we constructed a trajectory between  $X$  and  $\emptyset$  with even stages: hence a trajectory from a configuration in  $(\cdot, F)$  to  $\mathbf{0}$ .  $\square$

This yields the proof of Theorem 1.(iii).

## C Proof of Theorem 2

In order to prove Theorem 2, we establish the following proposition. The theorem is then an immediate consequence.

**Definition 3.** Let  $\omega$  be a configuration and  $\mathcal{A} \in \mathcal{T}$  be a set of configurations. We say that  $\mathcal{A}$  is *reachable* from  $\omega$  if after a certain number of transitions from  $\omega$ , we may obtain a configuration in  $\mathcal{A}$ :

$$\sum_{n \in \mathbb{N}} K^n(\omega, \mathcal{A}) > 0.$$

**Proposition 10.** Assume that  $(\mathcal{X}, E)$  satisfies the Richness Assumption. Let  $\omega$  be a configuration such that there exist both 0 and 1 on some positions. Let  $X \subset \mathcal{X}$  and  $Y \subset \mathcal{X}$  be two finite sets, then the cylinder  $(X, Y)^+$  is reachable from  $\omega$ .

In order to prove that a set is reachable, we will mainly rely on the following definition of a trajectory. Indeed, the existence of a trajectory from  $\omega$  to  $\mathcal{A}$  is equivalent to the fact that  $\mathcal{A}$  is reachable from  $\omega$ .

**Definition 4.** A *trajectory* from  $\omega$  to  $\mathcal{A}$  is a finite sequence of sets of configurations  $(\mathcal{A}_1, \dots, \mathcal{A}_n)$  such that

- $\mathcal{A}_1 = \{\omega\}$ ,
- $\mathcal{A}_n = \mathcal{A}$ ,
- for every  $1 \leq l < n$ ,  $\forall m \in \mathcal{A}_l$ ,  $K(m, \mathcal{A}_{l+1}) > 0$ .

We will often apply Definition 4 such that every set  $\mathcal{A}_i$  is a cylinder generated by a partial configuration  $(X_i, Y_i)$ . When doing so, it will be easier to speak directly about a *trajectory of partial configurations*.

In Proposition 10, we restricted ourselves to finite cylinders. We can extend these results by using the fact that  $\Omega$  is a polish set. Let  $\mathcal{A}$  be such that  $\phi(\mathcal{A}) > 0$ .  $\phi$  is a Borel measure and therefore is tight:

$$\forall \mathcal{A} \in \mathcal{T}, \phi(\mathcal{A}) = \sup_{K \text{ compact } \subset \mathcal{A}} \phi(K).$$

It follows that there exists a closed set  $K$  such that  $\phi(K) > 0$ . Any open set can be written as a countable union of finite cylinders, and therefore any closed set can be written as a countable intersection of complements of finite cylinders. However, the complement of a finite cylinder is a union of finite cylinders:

$$((X, Y)^+)^c = \left( \bigcup_{x \in X} (\emptyset, \{x\})^+ \right) \cup \left( \bigcup_{y \in Y} (\{y\}, \emptyset)^+ \right).$$

Hence, a closed set can be written as a countable union of countable intersections of finite cylinders, and therefore one of these intersections has a non-zero measure under  $\phi$ : there exists  $(X_n, Y_n)_{n \in \mathbb{N}}$  a sequence of finite cylinders such that

$$\phi \left( \bigcap_{n \in \mathbb{N}} (X_n, Y_n)^+ \right) > 0.$$

This is only possible if  $X_\infty = \bigcup_{n \in \mathbb{N}} X_n$  and  $Y_\infty = \bigcup_{n \in \mathbb{N}} Y_n$  are finite and disjoint. In this case, we have

$$\bigcap_{n \in \mathbb{N}} (X_n, Y_n)^+ = (X_\infty, Y_\infty)^+.$$

$(X_\infty, Y_\infty)^+$  is reachable from  $\omega$  and therefore  $\mathcal{A}$  is reachable from  $\omega$ .

In the next section, we will prove Proposition 10. The proof relies on constructing an explicit trajectory between the initial partial configuration  $\omega$  to any other partial configuration  $\omega'$  compatible with the condition in the lemma. In order to do so, the intuition is to fix the status of each agent in the network one by one. In doing so, one needs to be careful about not erasing a status that has been fixed before in the procedure.

## C.1 Proof of Proposition 10

The exposition of the proof is divided into several parts. We first present two partial mechanisms of transitions. The first one is a storing mechanism of partial configuration conditionally on the existence of a storing function. The second one is the propagation of an active status or the propagation of an inactive status in the network. These two procedures are defined locally, i.e., on partial configurations.

Then, we will combine these partial mechanisms into a global well-defined procedure. Hence, it will induce a trajectory. The key point is to ensure that these elementary mechanisms are not conflicting each other by using the same nodes at the same time.

### C.1.1 Storing and propagation

We begin by remarking that one can store the status of a node on one of its neighbor. Let  $\omega \in \Omega$  be a configuration such that agent  $x$  in  $\mathcal{X}$  is active, and  $x'$  be one of its neighbors. The following sequence of cylinders does the job:

- for every  $l \in 2\mathbb{N}$ ,  $\mathcal{A}_l = (\{x\}, \emptyset)^+$ ,
- for every  $l \in 2\mathbb{N} + 1$ ,  $\mathcal{A}_l = (\{x'\}, \emptyset)^+$ .

The next mechanism for storing partial configurations is merely a generalization of the previous idea.

**Lemma 3.** If a partial configuration  $(X, Y)$  of  $\mathcal{X}$  can be stored then for every  $n \in 2\mathbb{N}$ ,

$$\forall \omega \in (X, Y)^+, K^n(\omega, (X, Y)^+) > 0.$$

The procedure used in the proof will be called “storing a partial configuration in the global procedure”.

*Proof.* We assume that the finite configuration  $(X, Y)$  of  $\mathcal{X}$  can be stored. Hence, there exists  $\theta$  such that for all  $x \in X \cup Y$ ,  $\theta(x) \in \Gamma(x)$  and for all  $x \in X$ , for all  $y \in Y$ ,  $\theta(x) \neq \theta(y)$ . Define  $\theta(X) = \{\theta(x), x \in X\}$  and  $\theta(Y) = \{\theta(y), y \in Y\}$ . We consider the following sequence of cylinders:

- for every  $l \in 2\mathbb{N}$ ,  $\mathcal{A}_l = (X, Y)^+$ ,
- for every  $l \in 2\mathbb{N} + 1$ ,  $\mathcal{A}_l = (\theta(X), \theta(Y))^+$ .

By construction,  $\theta(X) \subseteq \overline{X}$  and  $\theta(Y) \subseteq \overline{Y}$ . Moreover, since the graph is undirected,  $X \subseteq \overline{\theta(X)}$  and  $Y \subseteq \overline{\theta(Y)}$ . By definition of  $\theta$  we know that  $\theta(X) \cap \theta(Y) = \emptyset$ . Hence, this sequence is well defined and

$$\forall m \in \mathcal{A}_l, K(m, \mathcal{A}_{l+1}) > 0.$$

Moreover, this can be ensured by only imposing restriction on the status of agents in  $X \cup Y$  at even stages and in  $\overline{X \cup Y}$  at odd stages.  $\square$

We now turn to the problem of propagating the status of a node in a partial configuration to any node in the network. This procedure will be called “propagating in the global procedure”.

**Lemma 4.** Let  $(X, Y)$  be a partial configuration of  $\mathcal{X}$  such that  $X$  is not empty. Then, for any  $\omega \in (X, Y)^+$  and for every  $x' \in \mathcal{X}$ , there exists  $n \in \mathbb{N}$ , which can be chosen odd or even, such that

$$K^n(\omega, (\{x'\}, \emptyset)^+) > 0,$$

Moreover,  $n$  is the length of a path between an active node in  $\omega$  and  $x'$ .

*Proof.* Let  $(X, Y)$  be a partial configuration of  $\mathcal{X}$  such that  $X$  is not empty. Let  $x \in X$  be an active agent in  $\mathcal{X}$  and  $x'$  be another agent. Since the graph is not bipartite, we know that there exists a path of both even and odd length. Consider a path  $x_0, \dots, x_n$  between  $x = x_0$  and  $x' = x_n$ . Informally, we propagate the status of  $x$  to  $x'$  along this path. More formally, consider the following sequence of cylinders:

$$\forall l \in \{0, \dots, n\}, \mathcal{A}_l = (\{x_l\}, \emptyset)^+.$$

This sequence is well defined, since for every  $l \in \{0, \dots, n-1\}$ ,  $x_{l+1} \in \Gamma(x_l)$ , and we have by construction

$$\forall m \in \mathcal{A}_l, K(m, \mathcal{A}_{l+1}) > 0.$$

We obtain that a configuration where  $x'$  is active is reachable.  $\square$

A similar procedure can be described for inactive agents.

A key point in the previous construction is that the propagation requires a constraint at stage  $l$  only on the status of  $x_l$ . All other agents can be indifferently active or not. This will allow us to combine the procedure with other procedures without having them interfering with each other.

### C.1.2 Global procedure

In order to prove this result, we decompose the construction of the trajectory into several steps. First, we will focus on placing active and inactive statuses on the complex star from the existence of an active and of an inactive agents. We provide a simple proof if the star is degenerate and then a long proof for the case where no complex star is degenerate. Second, we show that one can reach any partial configuration on a ray adjacent to a complex star. Finally, we will extend the result to any partial configuration.

Let  $\mathcal{I}$  be the set of configurations such that  $s_1, s'_2$  have values 0 and  $s_2, s'_1$  have values 1 or  $s_1, s'_2$  have values 1 and  $s_2, s'_1$  have values 0. Formally, we have

$$\mathcal{I} = (\{s_1, s'_2\}, \{s_2, s'_1\})^+ \cup (\{s_2, s'_1\}, \{s_1, s'_2\})^+.$$

Let us highlight that even in the case where  $s'_2 = s_1$  and/or  $s'_1 = s_2$ ,  $\mathcal{I}$  is well defined. Configurations are depicted in Figure 16.

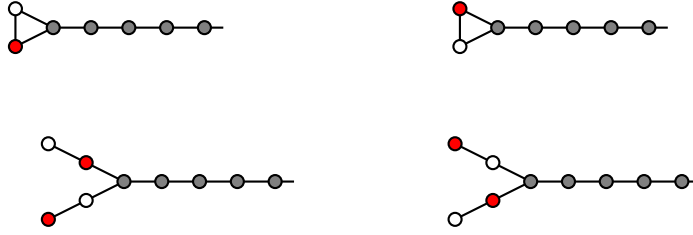


Figure 16: Configurations in  $\mathcal{I}$ .

**From  $\omega$  to a complex star.** Let us denote by  $(s_*, s_1, s_2, s_3, s'_1, s'_2, s'_3)$  the nodes of a complex star. Let  $\omega$  be a configuration with at least one active agent denoted  $x$  and one inactive agent denoted  $y$ . We show that there exists a trajectory from  $\omega$  to  $\mathcal{I}$ .

**Lemma 5.** Let  $\omega \in \Omega$ . Assume that there exist  $x, y \in \mathcal{X}$  such that  $\omega(x) = 1$  and  $\omega(y) = 0$ . Then there exists  $n \in \mathbb{N}$  such that

$$K^n(\omega, \mathcal{I}) > 0.$$

We split the proof of this first step in two cases depending if there exists a degenerate complex star or not. The proof for the degenerate complex star is easy, whereas the proof for the non-degenerate case is very technical since there are many cases to investigate.

Case of a degenerate complex star

*Proof.* There are three cases depending on the statuses of  $s_1$  and  $s_2$  under  $\omega$ .

- $s_1$  and  $s_2$  have different statuses. Then  $\omega$  is already in  $\mathcal{I}$ .
- Both statuses are equal to 0. Consider then a shortest path from  $x$  to  $s_*$ .
  - If this shortest path goes through  $s_1$ , propagate 1 until  $s_1$ .
  - If not and if this shortest path goes through  $s_2$ , propagate 1 until  $s_2$ .
  - Otherwise propagate to  $s_*$  and then to  $s_1$

Keep both nodes  $s_1, s_2$  equal to 0 until the last step where the status 1 is propagated to one of them.

- Both values are equal to 1. The proof is similar to the previous case by considering a shortest path from  $y$  to  $s_*$ .

□

Case of a non-degenerate complex star

The situation is more complicated if  $s_1$  and  $s_2$  are not directly connected. Assume that no complex star is degenerated. In this case, a complex star is essentially bipartite and one needs to use the property that the graph is not bipartite by considering other nodes that those composing the complex star.

The proof is done in three steps. We first introduce a technical lemma showing that any status on the complex star can be move around (on the star). Then, we show that depending

on the parity of the shortest paths from  $x$  and  $y$  to  $s_*$ , one can reach a partial configuration. Finally, it is then possible to use the fact that the graph is non-bipartite to reach the set  $\mathcal{I}$ .

Let  $\mathcal{J}$  be the set of configurations such that at least one node in  $\{s_1, s_2, s_3\}$  has value 0 and at least one node in  $\{s_1, s_2, s_3\}$  has value 1. Formally, we have

$$\mathcal{J} = (\{s_1\}, \{s_2\})^+ \cup (\{s_2\}, \{s_1\})^+ \cup (\{s_1\}, \{s_3\})^+ \cup (\{s_3\}, \{s_1\})^+ \cup (\{s_2\}, \{s_3\})^+ \cup (\{s_3\}, \{s_2\})^+.$$

**Lemma 6.** For every pair of partial configurations  $(X', Y')$  and  $(X'', Y'')$  in  $\mathcal{J}$  there exists a trajectory of even length between  $(X', Y')$  and  $(X'', Y'')$ .

*Proof.* Consider a restricted configuration on  $\{s_1, s_2, s_3\}$ . We assume without loss of generality that  $s_1$  is active whereas  $s_2$  is inactive. We want to prove that we can reach any restricted configuration on  $\{s_1, s_2, s_3\}$ .

First, let us prove that we can generate both the configuration where  $s_3$  is inactive or active without using any node outside the complex star  $\{s_*, s_1, s_2, s_3, s'_1, s'_2, s'_3\}$ . Let us show that we can reach the case where  $s_3$  is active. We consider the following sequence of restricted configurations:

- $s_1$  is active and  $s_2$  is inactive,
- $s_*$  is active and  $s'_2$  is inactive,
- $s_1, s_3$  are active and  $s_2$  is inactive.

We constructed a trajectory between our original configuration and the one we wanted. Notice that in fact, we only use as auxiliary state  $s_*$  and  $s'_2$ , hence there are no constraints on the statuses of  $s'_1$  and  $s'_3$  (if they are different from  $s'_2$ ). By exchanging the role of  $s_1$  and  $s_2$ , we obtain a procedure such that  $s_3$  is inactive.

Once  $s_3$  is fixed, it remains to fix by the same argument the statuses of  $s_1, s_2$ : if  $s_3$  is active, by considering  $s_2$  as the inactive, fix the status of  $s_1$  as it should be and then finally fix the status of  $s_2$ ; if  $s_3$  is inactive, exchange  $s_1$  with  $s_2$ .  $\square$

One can then deduce the following result.

**Lemma 7.** Let  $\omega \in \Omega$ . Assume that there exist  $x$  and  $y$  such that  $\omega(x) = 1$  and  $\omega(y) = 0$ . We distinguish three cases depending on the parity of the distances of  $x, y$  to  $s_*$

1. If the distances have different parity, then there exists  $n \in 2\mathbb{N}$  such that

$$K^n(\omega, (\{s_1\}, \{s'_1\})^+) > 0.$$

2. If both distances are odd, then there exists  $n \in 2\mathbb{N}$  such that

$$K^n(\omega, (\{s_1\}, \{s_2\})^+) > 0.$$

3. If both distances are even, then there exists  $n \in 2\mathbb{N} + 1$  such that

$$K^n(\omega, (\{s_1\}, \{s_2\})^+) > 0.$$

These intermediate configurations are depicted in Figure 17.

*Proof.* Let  $x, y \in \mathcal{X}$  such that  $x$  is active and  $y$  is inactive in configuration  $\omega$ .



Figure 17: Intermediate configuration: Different parity (left)/Same parity (right)

1. Assume that  $x$  is at an odd distance whereas  $y$  is at an even distance. Then, one can propagate both statuses simultaneously along their shortest path until  $s_*$  and then to the star. Once a status is fixed, at  $s'_1$  (resp.  $s_1$ ), it is stored at  $s_1$  (resp.  $s'_1$ ). The difference of parity ensures that the two paths (including the storing) are always using different nodes at the same time yielding no incompatibility.
2. Assume that  $x$  and  $y$  are at an odd distance of  $s_*$ . The situation is more complicated since the two protocols to propagate the status could a priori use the same node at the same time. We will distinguish by which one is closer to the center of the star  $s_*$ . Notice that since  $x$  and  $y$  are at an odd distance, their distance to  $s_*$  differs by an even number. We choose a shortest path  $p$  from  $x$  to  $s_*$ .
  - If  $d(x, s_*) < d(y, s_*)$  then  $p$  does not go through  $y$ . Consider a shortest path from  $y$  to  $s_*$ . By definition of a complex star, there exists  $i \in \{1, 2, 3\}$  such that  $s_i$  is not on the shortest path between  $y$  and  $s_*$ . Let us assume without loss of generality that it is  $s_1$ . We propagate the active status to  $s_1$  through  $p$  while storing  $y$  on any of its neighbors. By the distance assumption, none of the neighbors of  $y$  can be on  $p$  or it will contradict the assumption on distances. We distinguish two cases:
    - the shortest path from  $y$  to  $s_*$  goes through  $s_2$  (resp.  $s_3$ ): one propagates the inactive status to  $s_2$  (resp.  $s_3$ ) through the shortest path while storing the status of  $s_1$  in  $s_*$ .
    - the shortest path from  $y$  to  $s_*$  does not go through  $\{s_1, s_2, s_3\}$ . Therefore, there exists  $s_4 \in \Gamma(s_*)$  such that  $s_4 \notin \{s_1, s_2, s_3\}$ . One can propagate the inactive status to  $s_4$  through the shortest path while storing the status of  $s_1$  in  $s_*$ . To conclude, by applying Lemma 6 to the complex star  $\{s_*, s_1, s'_1, s_2, s'_2, s_4, s'_4\}$ , one can place the inactive status in  $s_2$  while preserving the active status at  $s_1$ .
  - If  $d(x, s_*) > d(y, s_*)$  then the proof is similar to the first case by inverting the role of active and inactive agents.
  - If  $d(x, s_*) = d(y, s_*)$  then we need to distinguish three cases depending on the neighbors of  $y$  and  $x$ :
    - If  $y$  has only one neighbor, not on  $p$ : We follow the proof of the first case by storing the status of  $y$  on this neighbor.
    - If  $y$  has more than one neighbor: Then, one of them has to be outside of  $p$ . Indeed, if two of them are on the path  $p$ , then it means than one of these neighbors is at distance less than  $d(x, s_*) - 3$  of  $s_*$  which would put  $y$  at a distance less than  $d(x, s_*) - 2$ , a contradiction. We follow the proof of the first case by storing the status of  $y$  on this neighbor.
    - If  $y$  has only one neighbor which is on  $p$ : Then by assumption on the distances, there exists  $z \in \mathcal{X}$  in the neighborhoods of  $x$  and  $y$ . The case  $\Gamma(x) = \Gamma(y) = \{z\}$



is impossible as the status of  $x, y$  cannot be stored in the same node  $z$ . Hence,  $|\Gamma(x)| > 1$  and we can treat this case by exchanging the role of  $x$  and  $y$ : first propagate the status of  $y$  while storing  $x$  on its other neighbor.

3. Assume that  $x$  and  $y$  are at an even distance. The proof is similar to the previous one.

□

Finally, we show that one can reach  $\mathcal{I}$ . Since the graph is not bipartite, there exists a cycle of odd length from  $s_*$  to  $s_*$ . Consider the smallest such cycle. Let us state some facts on the relation between this shortest cycle and the complex star.

- $s_*$  only appears as origin and end of the cycle. Indeed, otherwise our shortest odd cycle would be composed of two smaller cycles and one of them would be of odd length.
- There exists  $s_i$  such that the shortest cycle does not use  $s_i$ . Indeed, suppose the cycle, denoted by  $c$ , contains  $s_1, s_2, s_3$  in this order and denote by  $p_1$  the part of the cycle from  $s_*$  to  $s_2$  containing  $s_1$  and by  $p_2$  the remaining part of the cycle from  $s_2$  back to  $s_*$  via  $s_3$ . As  $s_2$  is a neighbor of  $s_*$ , we have two cycles  $c_1, c_2$ , with  $c_1$  formed by  $p_1$  and the link  $s_2$  to  $s_*$ , and  $c_2$  formed by  $p_2$  and the link  $s_2$  to  $s_*$ . As the cycle  $c$  is odd, one of the cycles  $c_1, c_2$  must also be odd, say  $c_1$ . Since  $c_1$  is shorter than  $c$  and odd, this contradicts the definition of  $c$ .
- If the shortest cycle does not use  $s_i$  it also does not use  $s'_i$  since we assumed that there was no triangle. Hence  $s'_i$  is at least at distance 2. If the shortest path from  $s_*$  uses  $s'_i$  then one of the two paths  $s_*$  to  $s'_i$  or  $s'_i$  to  $s_*$  has an odd length and the other part can be replaced by the path of length two which is the shortest possible path back to  $s_*$ .

Let us assume that the configuration is in  $(\{s_1\}, \{s'_1\})^+$ . We propagate the configuration to an alternative branch that is not on the shortest odd cycle from  $s_*$  to  $s_*$ . It is then possible to propagate the statuses as follows:

- from the branch  $\{s_1, s'_1\}$  to  $s_*$ ,
- along the cycle to  $s_*$  again,
- then to another branch.

Since the cycle is odd, we obtain the configuration inverted to the original one, hence we have reached the cylinder  $\mathcal{I}$ .

Let us now assume that the configuration is in  $(\{s_1\}, \{s_2\})^+$ . We use Lemma 6 to set the status 0 on one branch outside of the shortest odd cycle. This allows us to propagate the status 1 until  $s_*$  along the cycle while ensuring that the status 0 is never lost. Since the cycle has an odd length, we can then propagate 1 to the same branch as 0 obtaining a configuration in  $(\{s_1\}, \{s'_1\})^+$ . We are back to the previous case.

**From the star to a partial configuration  $(X, Y)$  on a ray** Let us consider a ray originating from the center of the star  $s_*$ . We can assume without loss of generality that the ray is extending one of the branches of the complex star (up to replace one of the branches by it). We denote each node of the ray by a nonnegative natural number that is equal to its distance (along the ray) to  $s_*$ . One can reach the partial configuration  $(X, Y)$  from a configuration in  $\mathcal{I}$  by simply setting progressively the statuses from decreasing distance to  $s_*$  (along the ray).

**Lemma 8.** Let  $\omega \in \mathcal{I}$ . Let  $X \subset \mathcal{X}$  and  $Y \subset \mathcal{X}$  be two finite subsets of the ray from  $s_*$ . Then the cylinder  $(X, Y)^+$  is reachable from  $\omega$ .

*Proof.* Denote the nodes of the ray by an integer representing their distance to  $s_*$  in the restricted graph where only the ray remains. Notice that with this notation,  $s_*$  coincides with node 0.

Denote by  $N$  the maximal index in  $X \cup Y$ . Without loss of generality, one can assume that  $X \cup Y = \{0, \dots, N\}$  up to fixing arbitrarily the status in  $\{0, \dots, N\} \setminus (X \cup Y)$  to 1. Reaching the new cylinder implies in particular reaching  $(X, Y)^+$  which is larger. As remarked in the introduction, a partial configuration  $(X, Y)$  is a finite sequence of values  $\omega' = \{0, 1\}^N$  such that  $\omega'(n)$  is the status of node  $n$ . Restricted to the ray, there exists a total order on the indices.

One considers the following sequence of partial configurations:

- the statuses on  $\{s_1, s'_1, s_2, s'_2\}$  alternate between the two partial configurations allowed in  $\mathcal{I}$ :

$$\omega_t(s_1) = \omega_t(s'_1) = 1 \text{ and } \omega_t(s_2) = \omega_t(s'_2) = 0 \text{ if } t \text{ is even,}$$

and

$$\omega_t(s_2) = \omega_t(s'_2) = 1 \text{ and } \omega_t(s_1) = \omega_t(s'_1) = 0 \text{ if } t \text{ is odd,}$$

- the status of  $n \geq 1$  is equal to the status of its predecessor at the preceding stage:

$$\omega_t(n) = \omega_{t-1}(n-1),$$

- the status of  $s^* = 0$  at  $t$  is equal to the final expected status of node  $N - t$ ,

$$\omega_t(0) = \omega'(N - t).$$

By construction, the status at stage  $l$  of node  $m$  is equal to

$$\omega_l(m) = \omega_{l-m}(0) = \omega'(N - l + m).$$

In particular at stage  $N$ ,

$$\omega_N(m) = \omega_{m-N}(0) = \omega'(N - N + m) = \omega'(m).$$

□

**From a partial configuration on a ray to a generic partial configuration.** In this section, we show that the existence of a trajectory from a configuration in  $\mathcal{I}$  to a partial configuration on a ray (Lemma 8) implies that any configuration  $(X, Y)$  is reachable.

First, without loss of generality one can assume that the graph restricted to  $X \cup Y$  is a connected subgraph. Indeed, consider  $H$  to be a set of nodes such that  $X \cup Y$  is a subset of  $H$  and  $H$  is connected.

One can construct  $H$  by taking the union of all shortest paths between each pair of nodes in  $X \cup Y$ . One can now replace the condition to reach the cylinder  $(X, Y)^+$  by reaching the cylinder  $(H \setminus Y, Y)^+$ . By construction, this is subset of  $(X, Y)^+$ , hence reaching  $(H \setminus Y, Y)^+$  implies reaching  $(X, Y)^+$ .

The proof relies on the following lemma.

**Lemma 9.** Given a partial configuration  $(X, Y)$ , there exists a ray and a partial configuration  $(X', Y')$  restricted to the ray such that  $(X, Y)$  is reachable from  $(X', Y')$ .

*Proof.* Let us fix a ray  $r$  of a complex star and adopt the numbering system of Lemma 8 for its nodes:  $0 = s_*$ , 1, 2, etc. We first prove that there exists a ray  $r'$  (possibly different from the original ray) with origin  $o$  such that  $r' \setminus \{o\}$  is disjoint from  $X \cup Y \cup \overline{X} \cup \overline{Y}$ .

Since  $X \cup Y \cup \overline{X} \cup \overline{Y}$  is finite whereas  $r$  is infinite, there exists a natural number  $n_0$  such that for every  $n \geq n_0$ ,  $n$  (on  $r$ ) does not belong to  $X \cup Y \cup \overline{X} \cup \overline{Y}$ . Moreover, let us define the following classical distance in the graph

$$\forall x, x' \in \mathcal{X}, d_1(x, x') = \inf\{n \in \mathbb{N}, \exists (x_l)_{1 \leq l \leq n} \text{ a path between } x \text{ and } x'\}.$$

Then the following quantity is well defined

$$l = \min_{x \in X \cup Y \cup \overline{X} \cup \overline{Y}} \min_{n \geq n_0} d(x, n).$$

Let us denote by  $(n_1, x_1)$  a pair of minimizers. By definition, the path between  $x_1$  and  $n_1$  is disjoint (up to the extremities) of both the ray  $r$  and  $X \cup Y \cup \overline{X} \cup \overline{Y}$ . We obtain our result by considering  $x_1$  as the new origin (node 0) and the concatenation of the path between  $x_1$  and  $n_1$  together with the ray  $n_1, n_1 + 1, \dots$  as the new ray  $r'$ .

Recall that there exists a storing function  $\theta$ . We will in fact place the status of  $x$  at  $\theta(x)$  first and then propagate to  $x$  only at the last stage.

We introduce a second distance  $d_2$  where we consider shortest paths but restricted to the subgraph induced by  $X \cup Y \cup \overline{X} \cup \overline{Y}$ : given  $x, x' \in X \cup Y \cup \overline{X} \cup \overline{Y}$ ,  $d_2(x, x')$  is the length of a shortest path included into  $X \cup Y \cup \overline{X} \cup \overline{Y}$  from  $x$  to  $x'$ .

Consider now a spanning tree of  $X \cup Y \cup \overline{X} \cup \overline{Y}$  rooted at 0 induced by the distance  $d_2$ , and consider an injection function  $\phi$  from  $X \cup Y$  to  $\mathbb{N}$  such that

$$d_2(\theta(x'), 0) > d_2(\theta(x), 0) \Rightarrow \phi(x) > \phi(x')$$

and

$$\forall x \in X \cup Y, d_2(\theta(x), \phi(x)) \text{ is even.}$$

We define the partial configuration  $\omega'$  on  $r'$  by

$$\omega'(\phi(x)) = \begin{cases} 1 & \text{if } x \in X, \\ 0 & \text{if } x \in Y. \end{cases}$$

From the partial configuration  $\omega'$ , it is possible to reach the partial configuration  $\omega$ . One propagates each value to its target position by increasing value of  $\phi(x)$ . When placing the status of agent  $x$ , one stores the statuses of the other agents as follows:

- if  $\phi(x') < \phi(x)$ , then the status of  $x'$  has already being set. The status is stored at  $x'$  for odd stages and  $\theta(x')$  for even stages,
- if  $\phi(x') > \phi(x)$ , then the status of  $x'$  has not been set. The status is stored at  $\phi(x') + 1$  for odd stages and  $\phi(x')$  for even stages.

Once all of the statuses are fixed, then one can reach the partial configuration  $(X, Y)$  in one stage.  $\square$

**Conclusion** We obtain immediately Proposition 10 as a corollary of the previous lemma. By Lemma 5, one constructs a path from the original configuration  $\omega$  to  $\mathcal{I}$  on a complex star with an infinite ray. By Lemma 9, one can associate to  $(X, Y)$  an auxiliary configuration  $(X', Y')$  on this ray. By Lemma 8, it is possible to reach this cylinder on the ray and therefore by Lemma 9, it is also possible to reach the partial configuration  $(X, Y)$ . Figure 18 illustrates the main steps of the whole proof.

## D Complex star and bipartite graphs

**Proof of Proposition 3.** Due to the previous discussion, it suffices to prove that for any partial configuration  $(X, Y)$  on  $\mathcal{X}_o$  (or  $\mathcal{X}_e$ ), the bipartite graph  $(X \cup Y, \overline{X \cup Y}, \sim_{X \cup Y \cup \overline{X \cup Y}})$  admits a matching of  $X \cup Y$ . We consider first the case of a  $k$ -regular graph. Observe that by assumption, the degree of each node of  $X \cup Y$  in the bipartite graph is  $k$ , while the degree of each node of  $\overline{X \cup Y}$  in the bipartite graph is at most  $k$  (some of these nodes have necessarily neighbors outside  $X \cup Y$ ). Consider  $S \subseteq X \cup Y$  and the set of neighbors of  $S$ , denoted by  $\Gamma(S) := \bigcup_{x \in S} \Gamma(x)$ . By the above observation, the number of edges from  $S$  to  $\Gamma(S)$  is  $k|S|$ , and the total number of edges arriving at  $\Gamma(S)$  is at most  $k|\Gamma(S)|$ . Therefore  $|\Gamma(S)| \geq |S|$ . Then, by Hall's Theorem (see, e.g., (Diestel, 2005, Th. 2.1.2)), a matching of  $X \cup Y$  exists.

Suppose now that  $(\mathcal{X}, E)$  is a hierarchy. We know that  $\mathcal{X}_o$  is the set of nodes in the odd layers of the hierarchy. Take  $x \in X \cup Y$ . It suffices to define  $\theta(x)$  as one of the subordinates of  $x$ , which lies on the next even layer. This will form a matching of  $X \cup Y$  as two distinct nodes in  $X \cup Y$  have necessarily disjoint sets of subordinates.

**Proof of Proposition 4.** The “if” part is obvious. As for the “only if” part, let us first assume that there exists a double ray in the graph, i.e., there exists a sequence  $(x_m)_{m \in \mathbb{Z}}$  in  $\mathcal{X}^{\mathbb{Z}}$  such that all nodes are different and for every  $n \in \mathbb{Z}$ ,  $\{x_{n-1}, x_{n+1}\} \subseteq \Gamma(x_n)$ .

Moreover, since any node on the double ray admits already two branches of size 2, the absence of complex stars implies that:

- any additional node linked to the ray has no other neighbor. We will say that it forms an antenna.
- for every distinct  $i, j \in \mathbb{Z}$  which are not consecutive there is no link between  $x_i$  and  $x_j$ , since similarly this would generate a complex star with  $x_i$  as a center. Therefore,  $\Gamma(x_n) \cap (x_m)_{m \in \mathbb{Z}} = \{x_{n-1}, x_{n+1}\}$ .

Hence, we obtain the second case in our characterization of networks without complex stars.

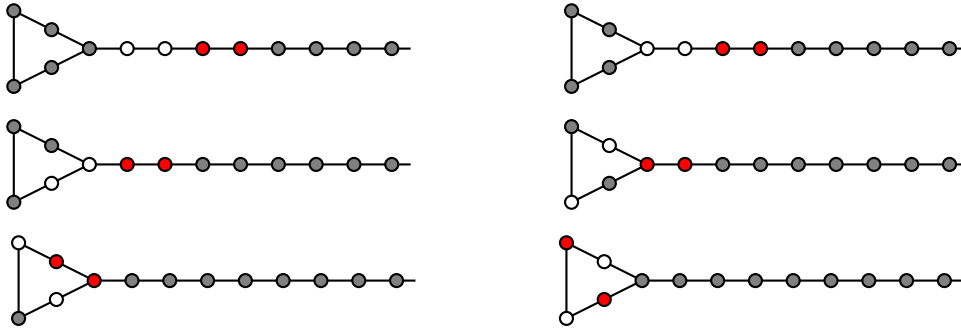
Let us suppose now that there exists no double ray. Since  $(\mathcal{X}, E)$  is an infinite connected graph with bounded degree, we know by Proposition 8.2.1. from Diestel (2005) that there exists a ray, i.e., there exists  $(x_i)_{i \geq 0}$  such that  $x_1 \in \Gamma(x_0)$  and

$$\forall i \geq 1, \{x_{i-1}, x_{i+1}\} \subseteq \Gamma(x_i).$$

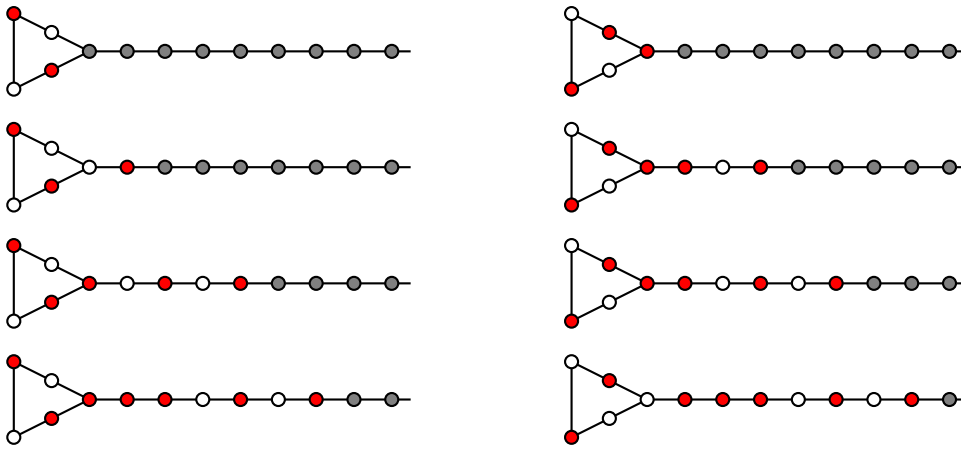
The preceding reasoning on additional nodes can be applied for every  $i \geq 2$  (it is unknown if  $x_0$  or  $x_1$  have more than one neighbors): any node not in  $(x_i)_{i \geq 0}$  and connected to  $(x_i)_{i \geq 2}$  has only one neighbor.

Denote by  $\mathcal{X}_1 = \bigcup_{i \geq 2} (\{x_i\} \cup \Gamma(x_i))$  and consider the graph with the set of nodes  $\mathcal{X}_2 = (\mathcal{X} \setminus \mathcal{X}_1) \cup \{x_1\}$  and the set of edges  $E_{\mathcal{X}_2}$ . This new graph is still connected. Moreover, it

Setting the statuses on the star:



Creating a generating sequence on the line:



Placing the status at their right place:

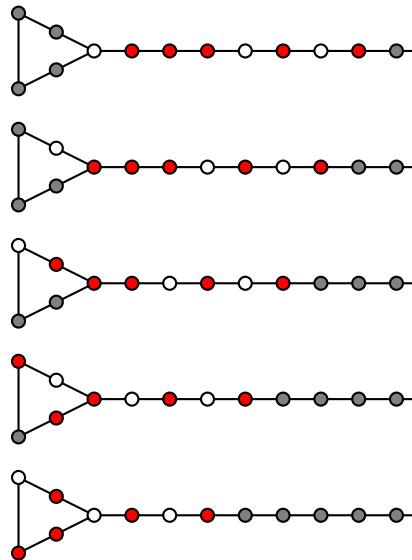


Figure 18: Sketch of the proof on an example

is finite, otherwise there would exist a ray which could be connected to the previous ray to generate a double ray. Let us consider a node  $x'$  at maximal distance from  $x_1$  in the subgraph  $(\mathcal{X}_2, E_{\mathcal{X}_2})$ . We define the new sequence  $(x'_n)_{n \geq 0}$  with  $x'_0 = x'$ , made from the path from  $x'$  to  $x_1$  concatenated with the ray  $(x_n)_{n \geq 1}$ .

- The reasoning on antennas can now be applied to any node in  $\{x'_n, n \geq 2\}$ .
- Moreover, every neighbor of  $x'_1$  (except  $x'_2$ ) has only  $x'_1$  as neighbor, otherwise  $x'$  would not be a farthest node from  $x_1$ .
- Finally, there cannot be any internal link in the ray  $(x'_n)_{n \geq 0}$ , since any such link has at least one extremity in  $\{x'_n, n \geq 2\}$  and would create a complex star.

## E Proof of Theorem 3

We have to prove only (ii) and (iii). The proof is similar to the proof of Theorem 1 but one needs to go more into details. In Section E.1, we focus on the “number” of inactive/active nodes on even/odd positions and prove that the sets containing only a null or infinite number of each element are absorbing. Then, in Section E.2, we prove that the sets containing finitely many elements are transient.

### E.1 Persistence of the “number” of elements

The following proposition states that the “number” of active agents on even (respectively odd) positions yields the same “number” of active agents on odd (respectively even) positions at the next stage. We establish the result for the case of active agents on even positions. The same result holds if active is replaced by inactive and/or  $\mathcal{X}_e$  by  $\mathcal{X}_o$ .

**Proposition 11.** For every  $\omega \in \Omega$ , let  $\omega'$  be a configuration randomly chosen according to the kernel  $K(\omega, \cdot)$ . Then:

- if  $\{x \in \mathcal{X}_e \mid \omega(x) = 1\}$  is empty then  $\omega'$  has almost surely no active agents on odd positions;
- if  $\{x \in \mathcal{X}_e \mid \omega(x) = 1\}$  is finite then  $\omega'$  has almost surely a finite (maybe zero) number of active agents on odd positions;
- if  $\{x \in \mathcal{X}_e \mid \omega(x) = 1\}$  is infinite then  $\omega'$  has almost surely an infinite number of active agents on odd positions.

*Proof.* Denote by  $\Theta = \{x \in \mathcal{X}_e \mid \omega(x) = 1\}$  the set of active agents on an even position.

First, it is clear that if this set is empty, then every odd position is surrounded by only inactive agents and therefore is inactive with probability 1 at the next stage. Hence, there is almost surely no active node on odd positions in  $\omega'$ .

Secondly, let  $x$  be an odd position. Since the aggregation function is strict, the probability that  $\omega'(x)$  is equal to 1 is strictly positive, if and only if, there exists a neighbor  $x'$  of  $x$  such that  $\omega(x') = 1$ . The set of neighbors of odd positions is the set of even positions and each agent has at most  $\gamma$  neighbors, therefore almost surely

$$|\{x \in \mathcal{X}_o \mid \omega'(x) = 1\}| \leq \gamma |\{x \in \mathcal{X}_e \mid \omega(x) = 1\}|.$$

Hence, if the right-hand side is finite, so is the left-hand side.

Finally, we assume that  $\{x \in \mathcal{X}_e \mid \omega(x) = 1\}$  is infinite. We compute the probability that  $\omega'$  has a finite number of active nodes on odd states. If this is the case, it implies that there exists an infinite number of elements  $x \in \Theta$  such that

$$\forall x' \in \Gamma(x), \omega'(x') = 0.$$

But by assumption,  $x'$  has at least one neighbor which is active, it follows that the probability that  $\omega'(x') = 0$  is strictly smaller than 1. Since the probabilities are independent, it follows that the total probability of the event is 0.  $\square$

Theorem 3.(ii) is then an immediate consequence of the above proposition.

## E.2 Transient set

We now prove that if there is a strictly positive but finite number of active agents at odd positions, then there is a positive probability to reach a configuration where there are no active agents anymore on odd positions. Similar results hold for inactive and/or for odd positions. This result can be expressed in terms of reachability.

**Lemma 10.** Let  $X \subset \mathcal{X}_o$ .

- If  $X$  is a finite set, then  $\text{int}(\text{int}(X)) \subset X$ .
- If  $X$  is a co-finite set, then  $X \subset \text{clo}(\text{clo}(X))$ .

*Proof.* We prove the first result. The second one can be deduced immediately by complementation.

Let us prove the result by contradiction. First, we know that  $\text{int}(\text{int}(X)) \subset \mathcal{X}_o$ . Take  $x \in \text{int}(\text{int}(X))$  and assume that  $x \notin X$ . By definition of  $\mathcal{X}_o$  and  $\mathcal{X}_e$ , we know that  $\text{int}(X) \cap \mathcal{X}_o = \emptyset$ , hence  $\Gamma(x) \not\subseteq X$ , i.e., there exists  $y \in \Gamma(x)$  s.t.  $y \notin X$ . Since  $x \in \text{int}(\text{int}(X))$ ,  $\Gamma(x) \subseteq \text{int}(X)$  and therefore  $y \in \text{int}(X)$ . It follows that  $\Gamma(y) \subseteq X$ . However, since the graph is undirected, we have  $x \in \Gamma(y)$ , so  $x \in X$ , a contradiction. This proves that  $\text{int}(\overset{\circ}{X}) \subseteq X$ .

Suppose that for all  $x \in X$ ,  $x \in \text{int}(\text{int}(X))$ . It follows that for all  $x \in X$ ,  $\Gamma(x) \subseteq \text{int}(X)$ , hence

$$\forall x \in X, \forall y \in \Gamma(x), \forall z \in \Gamma(y), z \in X. \quad (12)$$

Since  $X$  is finite, pick some  $t \in \mathcal{X}_o \cap X^c$ . By connectedness of the graph, there should exist a finite path  $x_0 = x, x_2, \dots, x_{2k} = t$  from  $x \in X$  to  $t \notin X$ . However, no such path can exist by Equation (12).

Therefore, there exists  $x \in X$  such that  $x \notin \text{int}(\text{int}(X))$ . This proves that the inclusion is strict.  $\square$

We can immediately deduce from Lemma 10 the following proposition.

**Proposition 12.** For every  $\omega \in \Omega$ , if  $\{x \in \mathcal{X}_o \mid \omega(x) = 1\}$  is finite then the set of configurations with only 0 on odd positions is reachable.

*Proof.* As  $X := \{x \in \mathcal{X}_o \mid \omega(x) = 1\}$  is finite, Lemma 10 implies that the sequence  $\text{int}^2(X), \text{int}^4(X), \dots$  converges to  $\emptyset$  in a finite number of steps. Since  $A$  is a strict aggregation rule, for every  $n \geq 0$ , the probability to reach  $\text{int}^{2n+2}(X)$  from  $\text{int}^{2n}(X)$  is positive, hence we constructed a trajectory between  $X$  and  $\emptyset$  with even stages: hence a trajectory from a configuration in  $(\cdot, \cdot, \cdot, F)$  to a configuration in  $(\cdot, \cdot, \cdot, 0)$ .  $\square$

**Remark 2.** Let us stress out that this lemma does not say anything about the presence of active agents at an even stage on an even position.

We can prove in a similar way the same result for  $X \subset \mathcal{X}_e$ . This yields the proof of Theorem 3.(iii).

## F Proof of Theorem 4

Let us explain the difference with the proof when the graph is non-bipartite. As highlighted when presenting the result, the main difference is the separation of the statuses and nodes in two groups. Let us consider the bipartition  $\mathcal{X}_o$  and  $\mathcal{X}_e$  such that without loss of generality, there exists a complex star centered at  $s_*$  in  $\mathcal{X}_e$ . By definition of a bipartition, if  $x$  and  $x'$  are in  $\mathcal{X}_o$  (resp.  $\mathcal{X}_e$ ) all paths between them are of even length. On the contrary, if  $x$  and  $x'$  are respectively in  $\mathcal{X}_o$  and  $\mathcal{X}_e$  (or the converse), then all paths between them are of odd length.

We obtain as a consequence a separation between two sets of statuses:

- the status at even stages of nodes in  $\mathcal{X}_o$  and the status at odd stages of nodes in  $\mathcal{X}_e$ ,
- the status at even stages of nodes in  $\mathcal{X}_e$  and the status at odd stages of nodes in  $\mathcal{X}_o$ .

Since the graph is bipartite, these statuses are never interacting.

As a consequence of this remark, it is not always possible to reach the cylinder  $\mathcal{I}$  defined by

$$\mathcal{I} = (\{s_1, s'_2\}, \{s_2, s'_1\})^+ \cup (\{s_2, s'_1\}, \{s_1, s'_2\})^+$$

from a configuration with an active agent and an inactive agent. More precisely, this step of the proof in the non-bipartite case was decomposed into two parts: reach an intermediate configuration and then use the existence of an odd cycle to reach the final configuration. The first step is still valid but not the second one.

One obtains the following lemma depending on the position of active and inactive agents in the initial configuration.

**Lemma 11.** Let  $\omega \in \Omega$ . Depending on the existence of an active/inactive status in  $\mathcal{X}_o/\mathcal{X}_e$ , it is possible to reach the following cylinders depicted in Figure 19.

One can then adapt the rest of the proof of Theorem 2 by creating the correct pattern separately on  $\mathcal{X}_o$  and on  $\mathcal{X}_e$ .



Active in $\mathcal{X}_o$	Inactive in $\mathcal{X}_o$	Active in $\mathcal{X}_e$	Inactive in $\mathcal{X}_e$	
Yes	No	Yes	No	
Yes	No	No	Yes	
Yes	No	Yes	Yes	
No	Yes	Yes	No	
No	Yes	No	Yes	
No	Yes	Yes	Yes	
Yes	Yes	Yes	No	
Yes	Yes	No	Yes	
Yes	Yes	Yes	Yes	

Figure 19: Configuration on the star in the bipartite case