



# A Survey on Nonstrategic Models of Opinion Dynamics

Michel Grabisch, Agnieszka Rusinowska

## ► To cite this version:

Michel Grabisch, Agnieszka Rusinowska. A Survey on Nonstrategic Models of Opinion Dynamics. 2021. halshs-03167886

**HAL Id: halshs-03167886**

**<https://shs.hal.science/halshs-03167886>**

Submitted on 12 Mar 2021

**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.



Distributed under a Creative Commons Attribution 4.0 International License

CES

Centre d'Économie de la Sorbonne  
UMR 8174

**A Survey on Nonstrategic Models of Opinion Dynamics**

Michel GRABISCH, Agnieszka RUSINOWSKA

**2021.05**



# A Survey on Nonstrategic Models of Opinion Dynamics

Michel GRABISCH <sup>1,2\*</sup> and Agnieszka RUSINOWSKA <sup>1,3</sup>

<sup>1</sup> Centre d'Economie de la Sorbonne, 106-112 Bd de l'Hôpital, 75647 Paris, France;  
{michel.grabisch,agnieszka.rusinowska}@univ-paris1.fr

<sup>2</sup> Paris School of Economics, Université Paris 1 Panthéon-Sorbonne

<sup>3</sup> Paris School of Economics, CNRS

\* Correspondence: michel.grabisch@univ-paris1.fr

**Abstract:** The paper presents a survey on selected models of opinion dynamics. Both discrete (more precisely, binary) opinion models as well as continuous opinion models are discussed. We focus on frameworks that assume non-Bayesian updating of opinions. In the survey, a special attention is paid to modeling nonconformity (in particular, anticonformity) behavior. For the case of opinions represented by a binary variable, we recall the threshold model, the voter and  $q$ -voter models, the majority rule model, and the aggregation framework. For the case of continuous opinions, we present the DeGroot model and some of its variations, time-varying models, and bounded confidence models.

**Keywords:** opinion dynamics; binary opinion; continuous opinion; non-Bayesian updating; nonconformity

---

## 1. Introduction

Although for some issues and specific topics, opinions can be formed immediately and do not change over time despite possible interactions between individuals, in most real-life situations, opinion formation is a rich, complex, and dynamic process. For example, in any society individuals' views on numerous economic, societal, and political issues are usually formed via a process of opinion dynamics, caused by interactions among the individuals. The central question is: *How can the opinion of a society evolve, due to various influence phenomena?*

Opinion dynamics has gained a lot of attention for over 60 years and in various scientific fields: sociology, psychology, economics, mathematics, physics, computer science, statistics, control theory, etc. Sociologists and psychologists have already addressed this problem in the '60s and '70s. One of the best known seminal work comes from DeGroot [1], a statistician who proposed a consensus reaching model of opinion formation where opinions are continuous and individuals take weighted averages of their neighbors' opinions when updating their own views. However, several related models introduced even earlier are due to sociologists: one of them is French Jr [2], whose model coincides with the DeGroot model (and for this reason is also called the French-DeGroot model), while Abelson [3] and Taylor [4] proposed continuous-time versions. We mention also a discrete version of Taylor's model proposed by two other sociologists, Friedkin and Johnsen [5].

One of the ideas behind variations of the French-DeGroot model lies in the assumption that the opinion updating can vary with time and circumstances. Important examples of time-varying frameworks are bounded confidence models, assuming that an individual pays attention only to other individuals whose beliefs do not differ much from his own belief. Important models in this stream of literature are those of Hegselmann-Krause [6,7] and of Deffuant *et al.* [8]. The Bayesian literature offers

also many models where opinion updating is context-dependent (Bayesian persuasion (Bergemann and Morris [9], Kamenica and Gentzkow [10]) and motivated belief models (Bénabou and Tirole [11])).

Assuming that an opinion is represented by a binary variable, another seminal framework of opinion formation is the threshold model coming from another sociologist, namely Granovetter [12]; see also Schelling [13].

Economists, who consider individuals as rational economic agents, have proposed various models, often incorporating utilities and strategic considerations; see e.g., Jackson [14], Acemoglu and Ozdaglar [15], Bramoullé *et al.* [16] for surveys. Apart from the standard questions of convergence and consensus reaching, additional related questions in the context of the DeGroot model are, e.g., those on the speed of convergence (How quickly beliefs reach their limit?) or on the “correctness” of consensus beliefs (Do beliefs converge to the right answer?). Some examples of extensions of the DeGroot model and time-varying models are presented by DeMarzo *et al.* [17], Golub and Jackson [18] and Büchel *et al.* [19,20], the latter focusing on the inheritance of cultural traits along generations. Another related literature is the one on social learning in the context of social networks, where individuals observe choices over time and update their beliefs accordingly (Banerjee [21], Ellison [22], Ellison and Fudenberg [23,24], Bala and Goyal [25,26], Banerjee and Fudenberg [27]). Note that this type of modeling is different from the framework in which the choices depend on the influence of others. The seminal works are those on the “herd behavior” (Banerjee [21], Scharfstein and Stein [28]) and the “informational cascades” (Bikhchandani *et al.* [29], Anderson and Holt [30]), where individuals are assumed to get information by observing others’ actions and are inclined to imitate those who are supposed to be better informed; see, e.g., Celen and Kariv [31]. Opinion dynamics is also related to diffusion of rumours and innovation (see, e.g., Banerjee [32], Chatterjee and Dutta [33], Bloch *et al.* [34]), and to contagion phenomena (see, e.g., Morris [35], and also Grabisch *et al.* [36] for an overview of the related works).

Physicists, who consider agents as particles, have used techniques from statistical physics (in particular, mean-field approximation) and coined the terms “sociophysics” (Galam [37]) and “econophysics”. The related literature in physics is very vast; see, e.g., Deffuant *et al.* [8], Weisbuch *et al.* [38], Weisbuch [39], Galam [40,41], Jędrzejewski and Sznajd-Weron [42], Nowak and Sznajd-Weron [43], Nyczka and Sznajd-Weron [44], Nyczka *et al.* [45]. An impressive survey is given by Castellano *et al.* [46], containing a huge bibliography.

On the other hand, computer scientists, mathematicians and statisticians have studied (probabilistic) automata models and processes that can model opinion dynamics on (infinite) lattices (Gravner and Griffeth [47], Holley and Liggett [48], Mossel and Tamuz [49]). Also control theorists have brought their contribution by introducing techniques from system theory; see, e.g., Altafini [50,51], Shi *et al.* [52], as well as Proskurnikov and Tempo [53,54] for very comprehensive surveys, and Breer *et al.* [55], Fagnani and Frasca [56], Bullo [57] for monographs.

In what follows, we try to classify the vast literature on opinion dynamics and present in more details some selected contributions. We focus in this brief survey on nonstrategic models, i.e., models which do not use game-theoretic considerations nor utility functions that agents try to maximize, nor are we considering models where opinion is about some unknown state of nature which has to be determined (Bayesian models). In short, we do not consider models usually studied by economists and rather concentrate on those studied by physicists, computer scientists and control theorists. We hope that this survey will bring, in particular to economists, a convenient overview of models they are not necessarily familiar with. We have tried to gather for each model the essential theoretical results on convergence, usually scattered in the literature, presented here in a unified way, and showing from which domain these results come from, at the risk of being dense and a bit concise, but providing a quick and handy reference. We do not pretend in these few pages to be exhaustive nor to give all technical details. We have given as far as possible all useful references for full details. Our original touch compared to other surveys is to put some emphasis on less developed aspects which

are nonconformist models and time-varying models. The interested reader can further consult other surveys on the topic, among which we recommend Castellano *et al.* [46], Proskurnikov and Tempo [53,54], Anderson and Ye [58], Flache *et al.* [59].

## 2. A tentative classification

The great variety of models studied in the literature obliges to start this survey by an attempt to classify the models, and at the same time to introduce relevant features of these models and the corresponding terminology.

We start with describing the type of society of *agents* (which are also called, depending on the context, individuals, players, voters, etc.) which underlies the model. Denoting by  $N$  the set of agents, it can be

- (i) finite, with  $|N| =: n$ . Most studies consider  $N$  to be finite, having in mind a society of relatively small size, a committee, etc.
- (ii) countably infinite, and in more rare cases, uncountable. The typical example of a countably infinite society is  $\mathbb{Z}^d$  with  $d \geq 1$ . The case  $d = 1$  corresponds to the case where the society is arranged on a line, i.e., each agent has a left and a right neighbor. The case  $d = 2$  is called the *infinite grid*, i.e., a surface where each agent has neighbors at least on south, north, west and east directions. It is useful for modeling a large society of individuals having many connections. Some studies take  $N = [0, 1]$  for mathematical convenience.

Supposing  $N$  to be fixed, it can be with or without structure. In the latter case, it is understood that every individual is in contact with every other individual. Otherwise, one can define a set of links (directed or undirected) between agents, which in turn defines a *graph* or *network*. A directed link (also called an *arc*) is an ordered pair  $(i, j)$ , with  $i, j \in N$ , and represents the fact that agent  $i$  has a connection with  $j$  (but not necessarily the converse) in some sense (e.g.,  $i$  listens to  $j$ 's opinion). An undirected link (also called an *edge*) is denoted by  $\{i, j\}$  or simply  $ij$ , and represents a connection without direction (i.e., it is symmetric). A *directed graph* or *digraph* is a graph with directed links. Otherwise, the graph is said to be *undirected*. The *neighborhood* of an agent  $i$  is the set of agents  $j$  such that there is a link from  $i$  to  $j$  (directed or undirected). The *degree* of  $i$  is the size of its neighborhood (number of neighbors). A graph is *complete* if any agent  $i$  has a link with any other agent  $j$ . Clearly, the unstructured case corresponds to the case of a complete graph.

Taking an agent  $i \in N$ , its opinion is a variable  $x_i$ , which can be

- (i) discrete, taking only finitely many values. By far the most common case is that the variable is *binary*, taking values 0 and 1, or  $-1$  and  $+1$ , representing for instance the opinions 'yes' and 'no' on some issue (political, ethical, fashion, adoption of a new technology, etc.). When the variable is not necessarily binary, it could be about the choice of a candidate at some election, choice of a school, restaurant, etc.
- (ii) continuous, taking value in a closed real interval like  $[0, 1]$ ,  $[-1, +1]$ , or in  $\mathbb{R}$ . Here the variable indicates either a degree of approval on some issue, the guess of an unknown value (see the famous example of guessing the weight of an ox cited by Surowiecki [60] in his book *The Wisdom of Crowds*) a percentage (e.g., of a total amount of money for allotting a budget), etc.

The type of opinion being defined, it remains to define its mechanism of updating. Here several further distinctions have to be made. The first one is to distinguish between *Bayesian updating* and *non-Bayesian updating*. In the former case, the opinion (which is more appropriately called a *belief*) is about an unknown state of nature  $\theta$ , called the *ground truth* (think for example of a group of people trying to guess the weight of an ox, or the result of a horse race, or the coming of a financial crisis, etc.). Then, updating of belief is done on the basis of the knowledge of the prior probabilities on the different states of nature and the Bayes rule with conditional probabilities (we refer the reader to, e.g., Acemoglu and Ozdaglar [15] for a survey of Bayesian updating models). By contrast, in what we

mean by non-Bayesian models<sup>1</sup>, there is no ground truth (no unknown state of nature) to guess, and therefore no probabilistic mechanism transforming prior probabilities into a posteriori probabilities. This is the case, for example, when opinion is on a new technology, on some social/ethical issue, on the amount of budget that should be allotted to some activity, etc. Rather, the agent updates his/her opinion on the basis of the opinions held by the other agents. We should mention here at the frontier between these two types of model the interesting work of Jadbabaie *et al.* [64], where the updating is done by combining the Bayesian rule and averaging of the opinion of the neighbors à la DeGroot (see a generalization of this framework and axiomatization in (Molavi *et al.* [65])).

Another important distinction is whether the updating mechanism is in continuous time or discrete time. A model in *continuous time* has the form  $\dot{x}(t) = F(x(t))$ , while in *discrete time* it reads  $x(t+1) = A(x(t))$ . Furthermore, the transition functions  $F$  and  $A$  can be fixed, i.e., invariant with time (this is often called *homogeneous*), or time-varying (also called *inhomogeneous*).

The updating rule can also be

- (i) *synchronous*, which means that at a given time, all agents simultaneously update their opinion;
- (ii) *asynchronous*, in which case at a given time step, only one randomly chosen agent (or a small subset of agents) updates its opinion.

Lastly, the updating rule can represent different types of behavior of the agent w.r.t. the opinion of the other agents. Mainly, one distinguishes between *conformity behavior*, where the agent has the tendency to follow the trend, to conform to the opinion of the others, and *nonconformity behavior*, which is the opposite. In nonconformity we further distinguish between *anticonformity*, which means that the agent, knowing the trend of the society, tends to go in the opposite direction, and *independence* or *stubbornness*, which means that the agent simply ignores the opinion of the rest of the society and has a fixed opinion.

It is important to note that the distinction conformity/nonconformity underlies a large domain of research in economics, as well as in sociology. In the latter domain, there is a distinction between models of assimilative social influence (conformist behavior) and models with repulsive influence (anticonformist behavior), and we refer the reader to the survey of Flache *et al.* [59] on social influence, built on this distinction. In the economic and financial domains, conformity is referred to as *trend-following* or *herd behavior*, while nonconformity is known as *contrarian* behavior. Among the first works mentioning herding behavior, we should cite Banerjee [21] and Bikhchandani *et al.* [29], where people acquire information in sequence by observing the actions of other individuals who precede them in the sequence, which necessarily leads to *informational cascades* (examples of such situations are: choice of stores, restaurants, schools, medical practices, etc.). Another famous work is Asch's experiment [66], showing that individuals may express judgments with which they disagree in order to conform to the majority judgment (see a discussion of these works in [67]). In financial markets, the herding behavior, informational cascades and contrarian behaviors have been widely studied (see, e.g., [68–71]).

In the behavioral aspects of the updating mechanism, we could have also introduced strategic considerations, where agents have utility functions depending on the reaction of the other agents, and their aim is to maximize their utility<sup>2</sup>. A simple canonical example is the guessing game of the

---

<sup>1</sup> It is important to note that there exists a whole bunch of literature on belief updating in a non-Bayesian way, in the sense that a mechanism different from the Bayes rule is used, e.g., in the Dempster-Shafer framework or with imprecise probabilities, multiple prior, etc. See, e.g., [61–63].

<sup>2</sup> It is interesting to note that in strategic games, one distinguishes between strategies which are complements and strategies which are substitutes. In the former case, agents have an incentive to match the strategies of other player (this is herding behavior), while in the latter case, agents have an incentive to do the opposite of the other agents, i.e., they have a contrarian behavior (see, e.g., Camerer and Fehr [72], Fehr and Tyran [73]). Hence, the behavior of an agent (conformist/anticonformist, herding/contrarian) may be dependent of the situation and is not a feature of the individual. Most models we present below, as strategic considerations are not taken into account in this survey, exhibit a fixed, exogenous type for each agent, except in Section 3.2, where the type is randomly chosen.

$p$ -beauty contest due to Moulin [74] (see Nagel [75]). Although this gives rise to a quite interesting and well developed branch of opinion dynamics (let us mention here, among many others, [19,20,76,77]), we limit the scope of this survey to the (already vast) case on nonstrategic models and non-Bayesian models, in the above sense.

As we said in the introduction, in order to distinguish this paper from existing surveys in the literature, we would like to make an emphasis on time-varying models and models of anticonformity, for which the literature is less generous.

We divide our survey in two sections: binary opinion models, and continuous opinion models, for which the techniques are completely different.

### 3. Binary opinion models

#### 3.1. The threshold model

The threshold model is one of the simplest and most natural ones. It has been rediscovered many times and studied in depth. The seminal papers are due to Granovetter [12] and Schelling [13]. The main idea is that agents display inertia in switching states, but once their personal threshold has been reached, the action of even a single neighbor can tip them from one state to another.

Consider  $N$  to be finite, with an underlying network. Each agent  $i \in N$  has a neighborhood  $\Gamma_i$ , with  $\Gamma_i \ni i$ , and the opinion of agent  $i$  at time  $t$  is denoted by  $x_i(t) \in \{0, 1\}$ . In addition, each agent  $i$  has a threshold  $\mu_i \in [0, 1]$ . We define the *state of the society* at time  $t$  to be the set of agents with opinion equal to 1 at time  $t$ . The opinion of agent  $i$  at next time step is given by

$$x_i(t+1) = \begin{cases} 1, & \text{if } \frac{1}{|\Gamma_i|} \sum_{j \in \Gamma_i} x_j(t) \geq \mu_i \\ 0, & \text{otherwise.} \end{cases}$$

The updating is synchronous since every agent updates at the same time. Agent  $i$  adopts opinion 1 if the proportion of its neighbors with opinion 1 is at least equal to the threshold  $\mu_i$ , otherwise its opinion is 0. Note that unless  $\mu_i = 0.5$ , the updating process is not symmetric w.r.t. 0 and 1. Indeed, if  $\mu_i > 0.5$  it is “easier” for agent  $i$  to adopt opinion 0.

A folklore result<sup>3</sup> says that the threshold model either converges to a constant vector  $x(\infty)$  or enters a cycle of period 2. The most general form of this result is:

**Theorem 1** (Goles and Olivos [80]). *Let  $W = [w_{ij}]$  be a  $n \times n$  symmetric matrix with  $w_{ij} = w_{ji} \in \mathbb{R}$  for all  $i, j$ , and threshold values  $\theta_i \in \mathbb{R}$ ,  $i \in N$ . Consider that*

$$x_i(t+1) = \begin{cases} 1, & \text{if } \sum_{j=1}^n w_{ij} x_j(t) \geq \theta_i \\ 0, & \text{otherwise.} \end{cases}$$

*Then, for every  $x(0) \in \{0, 1\}^n$ , there exists  $s \in \mathbb{N}$  such that*

$$x(s+2) = x(s).$$

In his seminal paper, Granovetter considers a complete graph so that  $\Gamma_i = N$  for every agent  $i$ , and takes as example a riot (1 = participate to the riot, 0 = do not participate). A threshold equal to 0 corresponds to an agent called the *instigator* as he starts rioting alone. Then a domino effect happens if there is one agent with threshold  $1/n$ , one with  $2/n$ , etc., till one with  $(n-1)/n$ . Indeed, once the instigator has started to riot, the agent with threshold  $1/n$  is triggered and starts to riot also. Then the

---

<sup>3</sup> See, e.g., [78,79]. More references can be found in [80].



threshold  $2/n$  is reached and another agent starts to riot, etc. Finally, the whole society participates to the riot.

The classical threshold model is clearly conformist. There exist recent studies on an anticonformist version of the threshold model, see, e.g., Nowak and Sznaid-Weron [43], Vanelli *et al.* [81], Grabisch and Li [82]. We give some excerpts of the latter reference.

The society  $N = N^a \cup N^c$  is now divided into the sets of conformist agents ( $N^c$ ) and anticonformist agents ( $N^a$ ). Conformist agents update their opinion exactly as in the classical threshold model. An anticonformist agent  $i$  updates its opinion as follows:

$$x_i(t+1) = \begin{cases} 0, & \text{if } \frac{1}{|\Gamma_i|} \sum_{j \in \Gamma_i} x_j(t) \geq \mu_i \\ 1, & \text{otherwise,} \end{cases}$$

that is, when too many agents have opinion 1, agent  $i$  takes the opposite opinion 0. Here also, the updating is synchronous and asymmetric in 0 and 1. The key function in the study of this model is the *transition function*  $G$ :

$$c(t+1) = G(c(t)) = \frac{1}{n} \left( \sum_{i \in N^c} \mathbb{1}_{\mu_i \leq c(t)} + \sum_{i \in N^a} \mathbb{1}_{\mu_i > c(t)} \right),$$

with  $c(t) \in [0, 1]$  the proportion of agents with opinion 1 at time  $t$ . An important fact, generalizing the result found by Granovetter, is that fixed points of  $G$  correspond to the absorbing states of the process. The study of convergence becomes fairly complicated, as cycles of various length may occur, as well as various absorbing states. Let us study more in detail the situation in the case of a complete network and a Gaussian distribution of the thresholds (i.i.d.) with mean  $m$  and variance  $\sigma^2$ . Then the transition function becomes

$$G(x) = \left( q - \frac{1}{2} \right) \operatorname{erf} \left( \frac{x - m}{\sigma \sqrt{2}} \right) + \frac{1}{2},$$

with  $q$  the proportion of conformists. When  $q > 1/2$ , there always exists a stable fixed point, no cycle, but possibly two unstable fixed points which appear for low values of  $\sigma$ , in which case the stable fixed point tends to  $q$ , i.e., all conformists have opinion 1, and all anticonformists have opinion 0 (see Figure 1). If  $q < 1/2$ , i.e., anticonformists are predominant, there exists a unique fixed point  $x_*$ . It is stable if  $|G'(x_*)| \leq 1$  (this happens if  $\sigma$  is not too small), otherwise there exists a limit cycle of length 2, where the two points of the cycle tend to  $q$  and  $1 - q$  respectively when  $\sigma$  tends to 0 (see Figure 2).

We suppose now that at each time, agents meet randomly other agents, i.e., their neighborhood is a random set of size either fixed or drawn from a distribution. The consequence is that the mechanism of updating is no more deterministic but obeys a Markov chain. A complete analysis has been done when the size of the neighborhood is fixed and under the assumption that all conformist agents have the same known threshold, and similarly for the anticonformist agents. The exhaustive list of all possible absorbing classes has been obtained, among which one finds absorbing states, but also cycles and more complex absorbing classes. Such a complete analysis does not seem possible in more general situations, however, a general simple result can be obtained in the fully general case, with arbitrary distributions of the size of the neighborhood and thresholds.

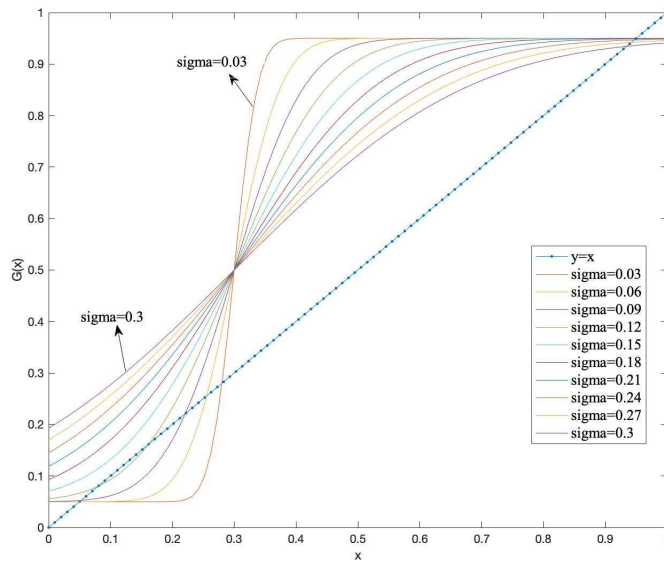
**Theorem 2.** [82] Assume an arbitrary distribution of the thresholds with lowest value  $\underline{\mu} > 0$  and highest value  $\bar{\mu}$ , and an arbitrary degree (size of the neighborhood) distribution, with lowest value  $\underline{d}$ .

Under the condition that  $n^c, n^a$ , the numbers of conformist and anti-conformist agents, satisfy

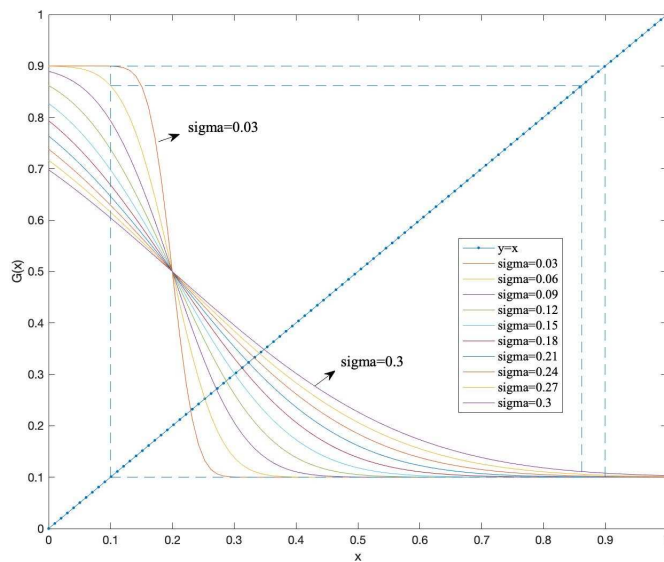
$$\bar{\mu} \underline{d} \leq n^a, n^c < n - \underline{d}(1 - \underline{\mu})$$

$2^N$  is the only absorbing class, i.e., the transition matrix of the Markov chain is irreducible.





**Figure 1.** Gaussian distribution with a small proportion of anti-conformist agents ( $q = 0.95$ ),  $m = 0.3$



**Figure 2.** Gaussian distribution with a large proportion of anti-conformist agents ( $q = 0.1$ ),  $m = 0.2$

To interpret this result, one must recall that the state of the society is defined as the set of agents with opinion 1. When  $2^N$  is an absorbing class, it means that at each time step, there is a positive probability to be in any state  $S \subseteq N$ , and that this state changes every time step. Therefore, the behavior in the long run is totally erratic. The theorem says that this happens under very mild conditions on the distributions. Indeed, as  $\bar{\mu} \in [0, 1]$ , it is enough that the numbers of conformists and anticonformists exceed the smallest value of the size of the neighborhood.

We end this section by mentioning a related model proposed by Gardini *et al.* [83] where conformist and anticonformist agents are considered, both in continuous time and in discrete time versions. In the discrete time version, a fraction  $\delta$  of agents updates their opinion at each time step, based on a threshold mechanism, and conformist agents copy the opinion of the agent they match with, while anticonformist agents take the opposite opinion. The threshold for conformist (resp., anticonformist) agents depends on the relative frequency of across-type matching for conformist (resp., anticonformist) agents. It is found that under some conditions, limit cycles of length 2 arise, which according to the authors, explain fashion cycles. A thorough analysis in terms of chaos theory is undertaken.

### 3.2. The voter and $q$ -voter models

The voter model is another classical model with a very simple mechanism, proposed around the same time as the threshold model by Clifford and Sudbury [84], who studies spatial conflict with animals fighting for territory. It was further analyzed by Holley and Liggett [48]. Among various generalizations proposed in the literature, Castellano *et al.* [85] introduces the (nonlinear)  $q$ -voter model, which we detail below. A good survey of the  $q$ -voter model can be found in Jędrzejewski and Sznajd-Weron [42].

We consider that the opinion (state) of an agent  $i$  is either  $s_i = -1$  or  $s_i = +1$ . We denote the update of the state by  $s_i \rightarrow a$ , where  $a \in \{-1, +1\}$ . We give below a brief presentation of the voter model and its main variants.

- **The voter model.** At each time step, an agent  $i$  is selected together with one of its neighbors, say  $j$ . Then  $s_i \rightarrow s_j$ .
- **The Sznajd model** [86]. The agents are on a line. Two consecutive agents  $i, i+1$  influence their neighbors  $i-1, i+2$  as follows:

$$\text{if } s_i = s_{i+1}, \text{ then } s_{i-1} \rightarrow s_i \text{ and } s_{i+2} \rightarrow s_{i+1}.$$

If they disagree, nothing happens. In the original version of the model, when  $s_i \neq s_{i+1}$ , then  $s_{i-1} \rightarrow s_{i+1}$  and  $s_{i+2} \rightarrow s_i$ .

- **The  $q$ -voter model.** At each time step, an agent  $i$  is selected as well as  $q$  of its neighbors. If all  $q$  neighbors agree, then  $s_i$  becomes the opinion of the  $q$  neighbors. Otherwise,  $s_i \rightarrow -s_i$  with probability  $\epsilon$ , and there is no change with probability  $1 - \epsilon$ .

The probability  $f(x)$  that an agent changes its opinion given that  $x$  is the proportion of disagreeing neighbors is

$$f(x) = x^q + \epsilon[1 - x^q - (1 - x)^q].$$

Note that if  $q = 1$  (voter model), then  $f(x) = x$ . The following variant exists: instead of unanimity of the  $q$  neighbors, only  $1 \leq r \leq q$  neighbors in the same state suffice for updating opinion (see, e.g., [44]).

- **The  $q$ -voter model with nonconformity** (Nyczka *et al.* [45]). Two types of nonconformity are considered: independence, where the decision about opinion change is made independently of the group norm (Note: such agents are commonly called “stubborn”); anticonformity, which means disagreement with the group norm.

There exist various versions of this model (see Jędrzejewski and Sznajd-Weron [42]). The general frame is as follows. Agents are either conformist (in which case they follow the original  $q$ -voter model) or nonconformist, with proportion  $1 - p$  and  $p$ , respectively. In the latter case for agent  $i$ , with a probability  $z_i$ , agent  $i$  acts as an independent player, that is, it takes the opposite opinion with probability  $f_i$ , otherwise keeps its opinion, regardless of the neighborhood. Otherwise, agent  $i$  acts as an anticonformist: picking at random  $q_i$  neighbors, if all<sup>4</sup> agents agree, the opposite opinion is adopted, otherwise the agent is not influenced. Figure 3 represents the general mechanism.

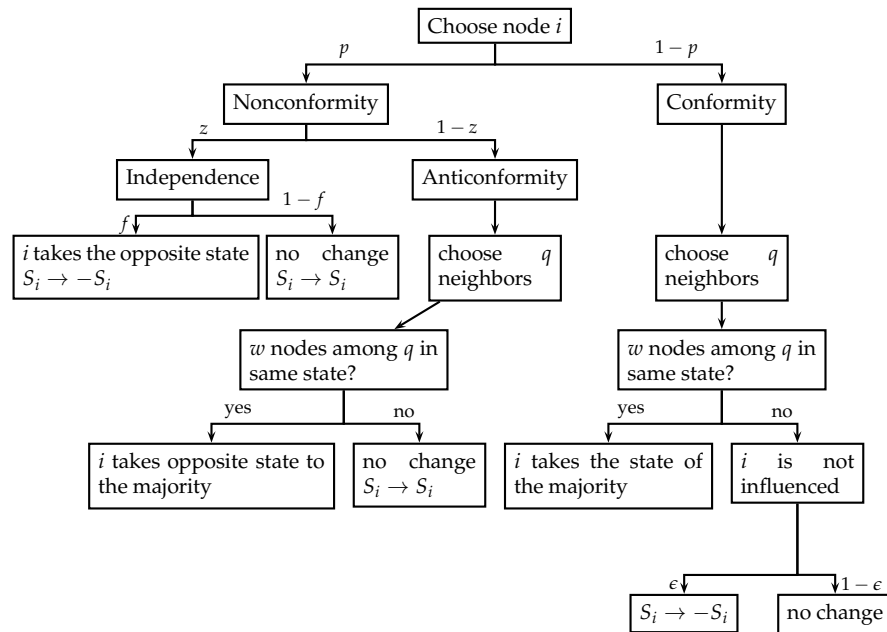


Figure 3. The  $q$ -voter model with nonconformity (adapted from [42])

We provide now an analysis of convergence of the  $q$ -voter model with non-conformity (see Jędrzejewski and Sznajd-Weron [42]).

### Mean-field approximation

In physics, when describing the evolution of states of particles, the mean-field approximation consists in replacing fluctuating values of variables describing the state of neighboring particles by their expectation. In the context of opinion dynamics, for every agent, the conditional probability that a neighbor has opinion  $\alpha$ , given the opinion of the agent, is approximated by the proportion of agents in the network having opinion  $\alpha$  (in some sense, the population is “well-mixed”; this is also equivalent to have a complete network, as the neighborhood of an agent is the whole network).

Let us explain how this approach can be used to derive  $\phi$ , the average opinion in the society of agents (or, equivalently, the concentration  $c$  of agents with opinion  $+1$ ), and how to obtain the steady state of  $\phi$  and to study its stability. Let  $P(x, \tau)$  be the probability to have a proportion  $x$  of agents with opinion  $+1$  at time  $\tau$ . The concentration  $c(\tau)$  is the expected value of  $x$ , i.e.,  $c(\tau) = \sum_x xP(x, \tau)$ . The evolution of  $P(x, \tau)$  is supposed to have the following (Markovian) form:

$$P(x, \tau + 1) = \sum_{x'} \gamma(x | x') P(x', \tau) \quad (1)$$

<sup>4</sup> As for conformist agents, “all” can be replaced by “ $w_i$  among  $q_i$ ”.

where  $\gamma(x | x')$  is the probability of transition from  $x'$  to  $x$ . In the  $q$ -voter models, as only one agent may change its opinion at each time step, the variation in  $x$  is either  $1/n$ ,  $0$  or  $-1/n$ , for which we use respectively the notation  $\gamma^+(x')$ ,  $\gamma^0(x')$ ,  $\gamma^-(x')$ . This yields

$$c(\tau + 1) = c(\tau) + \frac{1}{n} \sum_{x'} P(x', \tau) (\gamma^+(x') - \gamma^-(x')).$$

Then, the mean-field approximation implies that  $x$  is assumed to be equal to  $c(\tau)$ , yielding

$$c(\tau + 1) = c(\tau) + \frac{1}{n} (\gamma^+(c(\tau)) - \gamma^-(c(\tau))).$$

Taking time step equal to  $1/n$  and  $n \rightarrow \infty$  gives the continuous time version

$$\frac{dc(t)}{dt} = \gamma^+(c(t)) - \gamma^-(c(t)) \quad (2)$$

(rate equation). Doing the same with (1) gives the evolution of  $P(x, t)$  with time, which has the form of a Fokker-Planck equation<sup>5</sup>:

$$\frac{\partial P(x, t)}{\partial t} = -\frac{\partial}{\partial x} (F(x)P(x, t)) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (D(x)P(x, t)), \quad (3)$$

where  $F(x) = \gamma^+(x) - \gamma^-(x)$ ,  $D(x) = \frac{1}{n}(\gamma^+(x) + \gamma^-(x))$  are the drift and diffusion coefficients, respectively. The stationary solution of (3), giving  $P(x)$  when  $t$  tends to infinity, is:

$$P(x) = \frac{Z}{D(x)} \exp \left( \int \frac{F(x)}{D(x)} dx \right),$$

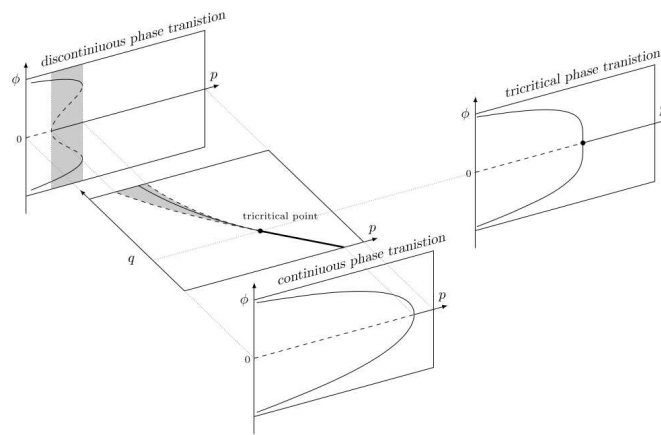
where  $Z$  is a normalizing constant.

A current practice in physics is to use a *phase transition diagram*, i.e., a diagram which depicts the evolution of the state variable of interest (here the concentration  $c$ , or the average opinion  $\phi = 2c - 1$ ) w.r.t. control parameters, which in the case of the nonconformist  $q$ -voter model are the values  $q$  (assumed to be continuous) and  $p$ , the proportion of nonconformists.

By (2), the system reaches a steady state  $c_s$  if  $F(c_s) = 0$ . As it is typical in dynamical and chaotic systems (see, e.g., Strogatz [88]), the steady state  $c_s$  is stable (i.e., after a small perturbation the system reaches again the steady state) if the derivative  $F'(c_s) < 0$ , and unstable if it is positive. Equivalently, the potential  $V(c) := -\int F(c)dc$  has its minima corresponding to the stable values, and its maxima to the unstable values. When the stable values evolve in a discontinuous way with the control parameters, we speak of a discontinuous phase transition, otherwise the phase transition is continuous. On Figure 4, one can see the phase diagram of  $\phi$  w.r.t  $p, q$ . Solid lines indicate stable values of  $\phi$  while dashed lines indicate unstable values. For high values of  $q$ , the phase transition becomes discontinuous. The critical point corresponds to the frontier between continuous and discontinuous phase transitions. In the discontinuous case, there is a zone (in grey) where three stable values coexist:  $\phi = 0$  (called “disordered phase”), and  $\phi$  close to 1 or to  $-1$  (ordered phase). Actually, the global minimum of  $V(\phi)$  is for  $\phi = 0$  in this zone, and the two other minima are *metastable* (i.e., a perturbation may make the system to jump in another (more) stable state).

---

<sup>5</sup> The Fokker-Planck equation is used in statistical mechanics for describing the time evolution of the probability density function of the velocity of a particle, under the influence of random forces (diffusion term in  $\frac{\partial^2}{\partial x^2}$ , like in the Brownian motion) and drag forces (drift term in  $\frac{\partial}{\partial x}$ ). It has been applied in economics with stochastic games (see, e.g., Anderson *et al.* [87]).



**Figure 4.** Phase transition diagram of  $\phi$  (borrowed from [42])

### Pair approximation

The pair approximation is a commonly used method in physics, also used for the study of random graphs (see, e.g., Newman [89]). It provides a more precise modeling than the mean-field approximation, which works moderately well on sparse networks, structures with a small average node degree, etc. We consider that the network is directed and introduce  $E_{++}, E_{+-}, E_{-+}, E_{--}$ , where  $E_{+-}$  indicates the number of links from an agent with opinion  $+1$  to an agent with opinion  $-1$ , etc. Links with different opinions at extremities are said to be active. We have  $E_{++} + E_{+-} + E_{-+} + E_{--} = \langle k \rangle n$ , where  $\langle k \rangle$  is the average degree of an agent. We define  $\theta_+$  to be the probability of selecting an active link with opinion  $+1$  at the head of the link (and similarly for  $\theta_-$ ):

$$\theta_+ = \frac{E_{+-}}{E_{++} + E_{+-}}, \quad \theta_- = \frac{E_{-+}}{E_{--} + E_{-+}}.$$

Supposing the network to be undirected (and therefore  $E_{+-} = E_{-+}$ ), we can write

$$\theta_+ = \frac{b}{1+m}, \quad \theta_- = \frac{b}{1-m}$$

with  $b = 2E_{+-}/\langle k \rangle n$  and  $m = (E_{++} - E_{--})/\langle k \rangle n$ . Then it is possible to write differential equations in  $c, b, m$  (see Jędrzejewski and Sznajd-Weron [42] for details).

### 3.3. The majority rule model

This model has been proposed by Galam [40]. We suppose that the network is complete and that the opinion is  $+1$  or  $-1$  (e.g.,  $+1$  means the adoption of a new reform). At each iteration, a group of  $r$  agents is selected at random. Usually,  $r$  is not fixed but drawn from a distribution. Then all agents inside that group take the majority opinion. In case of a tie, one of the opinions (usually  $-1$ ) arbitrarily prevails (principle of social inertia, preference of the status quo, etc.).

Let us fix the largest size of a group to be  $L$ , and consider a probability distribution of the size of a group  $(a_k)_{k=1,\dots,L}$ . Let  $P_+(t)$  be the probability that an individual at time  $t$  has opinion  $+1$ . Then, applying the above rule, assuming that ties are broken with  $-1$ , we find

$$P_+(t+1) = \sum_{k=1}^L a_k \sum_{j=\lfloor \frac{k}{2} + 1 \rfloor}^k \binom{k}{j} P_+(t)^j (1 - P_+(t))^{k-j}. \quad (4)$$

Interpreting  $P_+(t)$  as the proportion of agents with opinion +1 at time  $t$ , the analysis of the behavior of the system amounts to analyzing the transition function  $G$  (exactly as it is done in the threshold model, see Section 3.1):

$$P_+(t+1) = G(P_+(t))$$

We have  $G(0) = 0$  and  $G(1) = 1$ . Observe that the slope of  $G$  at 0 is smaller than 1. Indeed, suppose  $P_+(t) = \frac{1}{n}$  (one individual with opinion +1). Then, if ties are broken with +1, there is at most one additional individual with opinion +1 (obtained in groups of size 2), so on average, strictly less than one, and if ties are broken with -1, there cannot be an increase. In the latter case, taking  $P_+(t) = \frac{2}{n}$ , there is at most one additional individual with opinion +1 (obtained in groups of size 3), so on average, strictly less. Therefore, there exists an intermediate fixed point  $P_F \in ]0, 1[$ , which is unstable: if  $P_+(0) < P_F$ , then the society tends to the consensus -1, and if  $P_+(0) > P_F$ , the society tends to the consensus +1.

Let us take for the size of the group  $r = 3$  to be fixed. Then (4) becomes

$$P_+(t+1) = 3P_+^2(t)(1 - P_+(t)) + P_+^3(t) = -2P_+^3(t) + 3P_+^2(t).$$

Hence, the transition function is  $G(x) = -2x^3 + 3x^2$ . Its inflexion point is given by  $G''(x) = 0$ , which yields  $P_F = \frac{1}{2}$ .

Galam [41] introduces *contrarians*, i.e., anticonformist agents who update in the opposite way of conformist agents. Again, as with the threshold model, this yields radically different dynamics. Indeed, supposing that  $a$  is the proportion of anticonformist agents in the society, and for simplicity that  $r$  is fixed and odd, (4) becomes

$$\begin{aligned} P_+(t+1) &= (1-a) \sum_{j=\frac{r+1}{2}}^r \binom{r}{j} P_+^j(t) (1 - P_+(t))^{r-j} + a \sum_{j=0}^{\frac{r-1}{2}} \binom{r}{j} P_+^j(t) (1 - P_+(t))^{r-j} \\ &= (1-a) \sum_{j=\frac{r+1}{2}}^r \binom{r}{j} P_+^j(t) (1 - P_+(t))^{r-j} + a \left[ 1 - \sum_{j=\frac{r+1}{2}}^r \binom{r}{j} P_+^j(t) (1 - P_+(t))^{r-j} \right] \\ &= (1-2a) \sum_{j=\frac{r+1}{2}}^r \binom{r}{j} P_+^j(t) (1 - P_+(t))^{r-j} + a. \end{aligned}$$

One obtains  $G(0) = a$  and  $G(1) = 1 - a$ . Then  $G$  has still an intermediate fixed point  $P_F$ , which is stable, and 0,1 are no more stable fixed points.

See a similar study in the presence of stubborn (independent) agents by Galam and Jacobs [90].

### 3.4. The aggregation model

This model has been initiated by Grabisch and Rusinowska [91], and further developed and studied in [36,92,93]. It can be seen as a generalization of the threshold model.

Let us rewrite the threshold model as follows, supposing a complete network:

$$x_i(t+1) = \mathbb{1}_{\frac{1}{n} \sum_{j \in N} x_j(t) \geq \mu_i'}$$

where  $\mathbb{1}_p = 1$  if the logical proposition  $p$  is true, and 0 otherwise. The aggregation model is a generalization of the threshold model in 2 directions:

- The arithmetic mean for computing the average opinion in the society is replaced by any *aggregation function*  $A_i$ , i.e., a real-valued function on  $[0, 1]^n$  satisfying  $A_i(0, \dots, 0) = 0$  and  $A_i(1, \dots, 1) = 1$ , and being nondecreasing in each argument;

- The deterministic threshold mechanism is replaced by a *probabilistic mechanism*: the probability that  $x_i(t+1) = 1$  is the output of the aggregation function:

$$\Pr(x_i(t+1) = 1 | x(t)) = A_i(x_1(t), \dots, x_n(t)).$$

Typical aggregation functions are the weighted arithmetic mean  $\sum_i w_i x_i$ , particular case of the family of generalized means  $f^{-1}\left(\sum_i w_i f(x_i)\right)$ , with  $f$  increasing or decreasing and  $\sum_i w_i = 1$  with  $w_i \geq 0$  for all  $i$ , the minimum, maximum, median and more generally the ordered weighted average

$$\text{OWA}(x_1, \dots, x_n) = \sum_{i=1}^n w_i x_{(i)}$$

where the input vector  $x$  has been reordered in increasing values  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ . Note that the weight  $w_i$  is not associated to the opinion of agent  $i$  but to the  $i$ th smallest opinion. This makes the aggregation *anonymous*, i.e., the result is invariant to a permutation of the agents. It is very useful in contexts where the identity of agents does not matter or cannot be known, and only the number of agents with a given opinion matters, like in voting contexts and applications on the internet (see Förster *et al.* [92]). Note that if  $A_i$  is 0-1-valued, the classical threshold model is recovered.

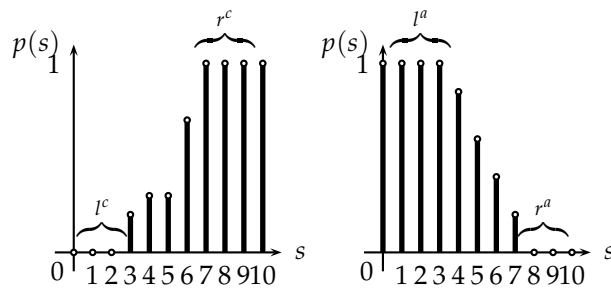
It is straightforward to apply the model on arbitrary networks. It suffices to replace the aggregation of  $x$  on the whole society by the aggregation limited to the neighborhood of the agent:  $A_i((x_j)_{j \in \Gamma_i})$ . The case of countably infinite networks has been investigated by Grabisch *et al.* [36].

As defined above, the updating process is synchronous and obeys a Markov chain, where the state of the society is defined as the set of agents with opinion 1 ( $2^n$  possible states). The study of convergence has been addressed for complete networks in [91]. Generally speaking, there are three types of absorbing classes (i.e., in which the process must end), and conditions for their existence have been identified:

- (i) absorbing states  $\{S\}$ . This means that the society is polarized: agents in  $S$  stick to opinion 1 and the other agents stick to opinion 0 for ever. Observe that by definition of an aggregation function,  $N$  and  $\emptyset$  are always absorbing states, but others may exist.
- (ii) cycles  $S_1 \rightarrow S_2 \rightarrow \dots \rightarrow S_k$ , with the condition that all sets are pairwise incomparable in the sense of inclusion. Here, the state of the society endlessly goes through this sequence of sets, hence there is no convergence.
- (iii) intervals and union of intervals of sets, e.g.,  $[S_1, T_1]$ ,  $[S_1, T_1] \cup [S_2, T_2]$ , etc., with  $\emptyset \neq S_i \subseteq T_i \neq N$ , where by interval of sets we mean  $[S, T] = \{K : S \subseteq K \subseteq T\}$ . This case can be interpreted as follows: when the absorbing class is  $[S, T]$ , agents in  $S$  will have opinion 1 for ever, while those outside  $T$  will have opinion 0 for ever. The rest of the agents in  $T \setminus S$  oscillates from 0 to 1 and vice versa in an erratic way. This can be seen as a “fuzzy polarization”.

The aggregation model in the presence of anticonformists has been studied by Grabisch *et al.* [93]. The society is supposed to be divided into the set of conformist agents  $N^c$  and the set of anticonformist agents  $N^a$ . It is assumed here that the aggregation is anonymous: only the number of agents with a given opinion matters, not their identity. This amounts to saying that the aggregation function reduces to a function  $p(s)$  of the number  $s$  of agents with opinion 1. If the agent is conformist, then  $p(s)$  is a nondecreasing function, while for an anticonformist agent,  $p(s)$  is nonincreasing. Figure 5 below gives an illustration of both types, and also shows some key parameters. Note that  $p(0) = 0$  and  $p(1) = 1$  for a conformist agent, while  $p(0) = 1$  and  $p(1) = 0$  for an anticonformist agent. The numbers of additional values where  $p$  is 0 or 1 are denoted by  $l^c, r^c$  for conformists, and  $l^a, r^a$  for anticonformists (see Figure 5). These are key parameters which can be interpreted as follows (let us replace for the rest of this section opinion 1 by ‘yes’ and opinion 0 by ‘no’).  $l^c$  is the maximum number of ‘yes’ for which no effect on the probability of saying ‘yes’ arises (this is called the *firing threshold*).



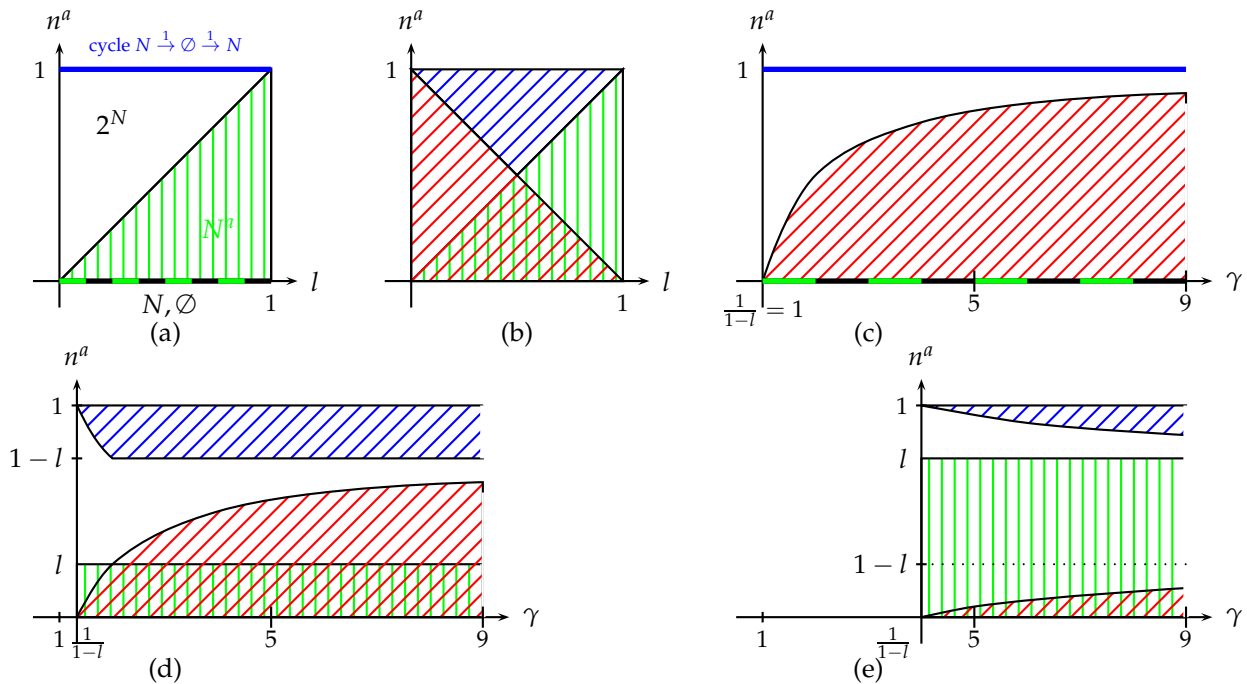


**Figure 5.** Aggregation functions when  $n = 10$  for conformist (left) and anticonformist agent (right) (borrowed from [93])

$n - r^c$  is the *saturation threshold* beyond which there is no more change in the opinion of the agent (and similarly for  $l^a, r^a$ ). A simplifying assumption is that all conformist agents have the same values  $l^c, r^c$ , although their aggregation function may differ (and similarly for anticonformists).

The analysis of convergence reveals to be very complex, and 20 types of absorbing classes have been identified, which can be classified into several categories: polarization, cycles, fuzzy cycles (precisely, cycles of intervals of states, like  $[S_1, T_1] \rightarrow [S_2, T_2] \rightarrow [S_1, T_1]$ ), fuzzy polarization (as defined above), and disorder (the class  $2^N$ ). In order to get more insight into convergence, we consider  $n \rightarrow \infty$  and use phase diagrams like in physics to see the transition between the different types of absorbing classes, with the control parameters  $l^c, r^c, l^a, r^a$  and  $n^a$ , the number of anticonformists. Now, these parameters are the relative version of the previous ones (i.e.,  $l^c/n, r^c/n$ , etc.) and are therefore numbers in  $[0, 1]$ .

We present in Figure 6 phase diagrams for the situation  $l^c = l^a =: l, r^c = r^a =: r$ , introducing also the parameter  $\gamma = \frac{1}{1-r-l}$ , interpreted as the reactiveness of agents once they have started to change their opinion ( $\gamma \in [\frac{1}{1-l}, \infty[$ ).



**Figure 6.** Phase diagrams: (a)  $\gamma$  has minimum value  $\frac{1}{1-l}$  ( $r = 0$ ); (b)  $\gamma \rightarrow \infty$ ; (c)  $l = 0$ ; (d)  $l \in [0, 1/2]$ ; (e)  $l \in [1/2, 1]$ . Color code: white= $2^N$ , blue=cycle, red= $N^c$ , green= $N^a$  (borrowed from [93])

The general observation is that when  $n^a$  increases from 0 to 1, we start from consensus, next go to polarization, next to disorder ( $2^N$ ), and finally to cycles. It shows that the “degree of non-convergence”

of the opinion of the society is increasing with the proportion of anticonformists, if we consider cycles as the ultimate state of non-convergence. A notable phenomenon is visible on Figures (c) and (e). When the firing threshold  $l$  is very low (Figure (c)), there is a “cascade” effect leading to polarization where all conformist agents say ‘yes’, which becomes more salient with reactivity. “Cascade” means that whatever the initial situation of the opinion in the society is, the final situation is inescapably driven to this polarization situation. Similarly, when the firing threshold is high (Figure (e)), a cascade effect leads to a polarization where all conformist say ‘no’, at the condition that the proportion of anticonformist is not too small but below  $l$ . This shows how important is the effect of the presence of anticonformist agents in a society, as it could lead in some sense to manipulation. We mention also Figure (b), where the point at the intersection of the two diagonals is very special: it is a *triple point* in the sense of physics, as all “states” coexist (i.e., cycles,  $N^c$  and  $N^a$ ).

A more general framework for binary opinion models, based on the notion of groups and their interrelationship (like in [50,51]), and enabling the representation of various behaviors (conformist, anticonformist, communitarian, with leaders, etc.) has been recently proposed by Poindron [94].

### 3.5. A final remark on binary opinion models

The discrete nature of the range of the opinion makes the mechanism of updating to be discontinuous by essence: this can be seen in the definition of the threshold model, the various  $q$ -voter models, and the majority model. In each of them, the shift of opinion of one agent may possibly reverse the opinion of the whole society, while at the same time, important variations in the opinion of agents may have no effect at all on the society, if the threshold is not reached. This mechanism, if it makes the process apparently simpler to analyze, is too crude to be realistic. Let us observe that the aggregation model, a direct generalization of the threshold model obtained by introducing a probabilistic device, does not exhibit this drawback, as there is no discontinuity any more in the updating: the deterministic change of opinion is replaced by a random draw.

## 4. Continuous opinion models

### 4.1. The DeGroot model and its variants

As for the classical models in binary opinion dynamics, the most classical models in continuous opinions are from the '60s and '70s. The well-known DeGroot [1] model was already proposed by French in 1956 [2], with continuous-time versions proposed as soon as 1964 by Abelson [3] and then by Taylor [4]. Its success is due to its simplicity, and the fact that everything is known about its convergence.

The DeGroot model is a discrete time model, in which every agent  $i \in N$  puts a confidence weight  $w_{ij}$  on the opinion of agent  $j$ . This yields a  $n \times n$  row-stochastic matrix  $W = [w_{ij}]$ . The updating is synchronous and consists in averaging the opinions of all other agents:

$$x(t) = Wx(t-1) = W^t x(0).$$

The analysis of convergence amounts therefore to the study of  $W^t$  when  $t$  tends to infinity. This is a classical topic in the theory of nonnegative matrices, and fully treated, e.g., in (Seneta [95]). We restrict here the exposition to the minimum.

Let us associate with  $W$  the corresponding directed graph  $G$  with set of nodes  $N$ , where an arc  $(i, j)$  exists if and only if  $w_{ij} > 0$ . There is a direct correspondence between the properties of connectedness of  $G$  and the power matrices  $W^k =: [w_{ij}^{(k)}]$ . There is a *walk* of length  $k$  from agent  $i$  to agent  $j$  if there is a sequence  $i = i_0, i_1, \dots, i_k = j$  of nodes such that  $(i_{\ell-1}, i_\ell)$  is an arc in  $G$  for  $\ell = 1, \dots, k$ . Note that the existence of such a walk is equivalent to  $w_{ij}^{(k)} > 0$ . We write  $i \rightarrow j$  if there is a walk from  $i$  to  $j$ . We use the notation  $i \leftrightarrow j$  if there is also a walk from  $j$  to  $i$ .

A *class* or *strongly connected component* of  $G$  is a set of nodes  $C$  such that either  $C$  is a singleton or  $i \leftrightarrow j$  for every distinct  $i, j \in C$ , and any  $C' \supset C$  does not fulfill the latter property. A class can be *essential* or *absorbing* if no arc is going out of it, otherwise it is *inessential*. The *period*  $\tau(i)$  of  $i \in N$  is the greatest common divisor of the length of all walks from  $i$  to  $i$ . A class is *aperiodic* if  $\tau(i) = 1$  for some  $i$  in the class (note that it suffices that some  $i$  in the class has a self-loop  $(i, i)$ ).

A matrix  $W$  is *primitive* if  $W^k > \mathbf{0}$  for some integer  $k$ , i.e., each entry is positive. It is *irreducible* if for every  $i, j \in N$  there exists an integer  $m(i, j)$  such that  $w_{ij}^{(m(i, j))} > 0$ . An irreducible matrix corresponds to a single class.  $W$  is primitive iff  $W$  is irreducible and aperiodic. In the latter case, 1 is the greatest eigenvalue, while if  $G$  is periodic, there exist other eigenvalues of modulus 1 and  $W^t$  does not converge.

If  $W$  is primitive, then  $W^t$  converges to a matrix with identical rows:

$$\lim_{t \rightarrow \infty} W^t = \mathbf{1} \cdot v^T$$

where  $v$  is the left eigenvector associated to eigenvalue 1, such that  $\mathbf{1}^T \cdot v = 1$ . As a consequence

$$\lim_{t \rightarrow \infty} x_i(t) = x(0)^T \cdot v \text{ for all } i \in N. \quad (5)$$

As  $x_i(\infty)$  is the same for every  $i$ , this shows that the society has reached a consensus. The form of  $x_i(\infty)$  suggests to interpret the left eigenvector  $v$  as the vector of *social influence weights* of the agents.

The speed of convergence is governed by the second largest eigenvalue  $\lambda_2$ :

$$|x_i(t) - \bar{x}_i| \leq C |\lambda_2|^t$$

When  $W$  is not irreducible, there exist several strongly connected components in  $G$ , then consensus is reached only within each strongly connected component. The computation of the limit  $\lim_{t \rightarrow \infty} x_i(t)$  for  $i$  in an absorbing class is similar to (5), taking instead of  $v$  the left eigenvector associated to 1 for the submatrix corresponding to the absorbing class.

We now review some variants of the DeGroot model. Our exposition follows Proskurnikov and Tempo [53]. Other variants can be found in [58].

### The Abelson model

It is a continuous version of the DeGroot model, proposed by Abelson [3]. Using the fact that  $1 - w_{ii} = \sum_{j \neq i} w_{ij}$ , one can rewrite the DeGroot model as

$$x_i(k+1) - x_i(k) = \sum_{j \neq i} w_{ij} (x_j(k) - x_i(k)) \quad (i \in N).$$

Letting the time interval tends to 0, one obtains the continuous time version, known as Abelson's model:

$$\dot{x}_i(t) = \sum_{j \neq i} a_{ij} (x_j(t) - x_i(t)) \quad (i = 1, \dots, n)$$

where  $A$  is a nonnegative matrix of weights. We can write it in a more compact form:

$$\dot{x}(t) = -L[A]x(t), \quad (6)$$

where  $L[A]$  is called the *Laplacian matrix*, defined by

$$L[A] = (l_{ij})_{i,j \in N}, \text{ where } l_{ij} = \begin{cases} -a_{ij}, & \text{if } i \neq j \\ \sum_{j \neq i} a_{ij}, & \text{if } i = j. \end{cases}$$

We give some facts about convergence.

- (i) For any matrix  $A$ , the limit  $P^\infty = \lim_{t \rightarrow \infty} e^{-L[A]t}$  exists, and the vector of opinion converges to  $x^\infty = P^\infty x(0)$ , where  $e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k$ .
- (ii) Consensus is reached in the Abelson's model if and only if the corresponding graph  $G$  is quasi-strongly connected (i.e., it has a directed spanning tree), and in this case the consensus vector is

$$x_1^\infty = \dots = x_n^\infty = p_\infty^T x(0)$$

with  $p_\infty$  uniquely given by  $p_\infty^T L[A] = 0$  and  $p_\infty^T \mathbf{1} = 1$ . Similarly to the French-DeGroot model,  $p_\infty$  can be interpreted as the vector of social influence weights.

### The Taylor model

It is an extension of the Abelson's model proposed by Taylor [4]. There are  $n$  agents with opinions  $x_1, \dots, x_n$  and  $m \geq 1$  communication sources providing static opinions  $s_1, \dots, s_m$ . The model reads:

$$\dot{x}_i(t) = \sum_{j=1}^n a_{ij}(x_j(t) - x_i(t)) + \sum_{k=1}^m b_{ik}(s_k - x_i(t)),$$

with  $B$  the matrix of persuability constants. The model can be seen as an Abelson's model with  $k$  additional stubborn agents, i.e., who do not change their opinion (independent agents). Also, it can be put under this form:

$$\dot{x}_i(t) = \sum_{j=1}^n a_{ij}(x_j(t) - x_i(t)) + \gamma_i(u_i - x_i(t)),$$

with  $\gamma_i = \sum_k b_{ik}$  and  $u_i = \frac{1}{\gamma_i} \sum_k b_{ik} s_k$ . The quantity  $u_i$  is called the *prejudice* of agent  $i$ ; it represents the internal agent's opinion (i.e., without influence), and is often taken as the initial opinion  $x_i(0)$ , and agent  $i$  is *prejudiced* if  $\gamma_i > 0$ . In matrix form:

$$\dot{x} = -(L[A] + \Gamma)x(t) + \Gamma u \quad (7)$$

with  $\Gamma = \text{diag}(\gamma_1, \dots, \gamma_n)$ , the matrix whose diagonal is formed by  $\gamma_1, \dots, \gamma_n$ , other elements being 0. Agent  $i$  is *P-dependent* (prejudice-dependent) if it is prejudiced or there exists a path from a prejudiced agent to  $i$ . Writing  $x = \begin{bmatrix} x^1 & x^2 \end{bmatrix}$  with  $x^1, x^2$  the vectors pertaining to the P-dependent and P-independent agents respectively, (7) can be rewritten as (note that  $L^{21} = \Gamma^{22} = 0$ )

$$\dot{x}^1(t) = -(L^{11} + \Gamma^{11})x^1(t) - L^{12}x^2(t) + \Gamma^{11}u^1 \quad (8)$$

$$\dot{x}^2(t) = -L^{22}x^2(t) \quad (9)$$

Note that (9) is the Abelson's model.

It can be proved that  $x^1(t)$  converges to

$$x^1(\infty) = M \begin{bmatrix} u^1 \\ x^2(\infty) \end{bmatrix}, \text{ with } M = (L^{11} + \Gamma^{11})^{-1} \begin{bmatrix} \Gamma^{11} & -L^{12} \end{bmatrix}.$$

$M$  is a row-stochastic matrix, and  $x^2(\infty)$  is obtained as for the Abelson's model.

### The Friedkin-Johnsen model

This is a discrete version of Taylor's model proposed by Friedkin and Johnsen [5]. The model reads:

$$x(k+1) = \Lambda W x(k) + (I - \Lambda)u \quad (10)$$

where  $W$  is as in the French-DeGroot model,  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  with  $\lambda_i \in [0, 1]$  the susceptibility of agent  $i$  to social influence, and  $u$  is a constant vector of agents' prejudices (e.g.:  $u_i = x_i(0)$ ). If  $\Lambda = I$ , the French-DeGroot model is recovered.

As for the Taylor's model, we distinguish between P-dependent and P-independent agents. With similar notation, the model is rewritten as

$$x^1(k+1) = \Lambda^{11}[W^{11}x^1(k) + W^{12}x^2(k)] + (I - \Lambda^{11})u^1 \quad (11)$$

$$x^2(k+1) = W^{22}x^2(k). \quad (12)$$

Note that (12) is a French-DeGroot model.

It can be shown that the model (10) is convergent if and only if all agents are P-dependent or (12) is convergent (i.e., all essential classes of  $W^{22}$  are aperiodic). Then,

$$x^1(\infty) = V \begin{bmatrix} u^1 \\ x^2(\infty) \end{bmatrix} \quad (13)$$

with  $V = (I - \Lambda^{11}W^{11})^{-1} \begin{bmatrix} I - \Lambda^{11} & \Lambda^{11}W^{12} \end{bmatrix}$ , which is row-stochastic. Equivalently, (10) is convergent if and only if  $\Lambda W$  is *regular* (i.e., there is a single essential class which is aperiodic, but possibly other inessential classes).

Thanks to the above result, it is possible to define a social influence power index as in the DeGroot model. We suppose here  $u = x(0)$  and that all agents are P-dependent, hence by the above result,  $x(\infty) = Vx(0)$  with  $V = (I - \Lambda W)^{-1}(I - \Lambda)$ . Then

$$\bar{x} := \frac{1}{n} \sum_{i=1}^n x_i(\infty) = \frac{1}{n} \underbrace{\mathbf{1}^T V}_{c^T} x(0).$$

Then the vector  $c$  is defined as *Friedkin's influence centrality* vector, and satisfies  $c^T \mathbf{1} = 1$ .

This notion can be applied as a centrality measure in graphs. Let  $W$  be a stochastic matrix representing a graph. Putting  $\Lambda = \alpha I$  with  $0 < \alpha < 1$ , one can use  $c_\alpha$  as a centrality measure, with the property that  $c_\alpha$  tends to the French-DeGroot social power when  $\alpha \rightarrow 1$ .

### The Altafini model

Altafini [50,51] proposes a generalization of the DeGroot model where agents may put negative weights on other's opinions, thus exhibiting anticonformity or anti-trend behavior. The fundamental difference with other models of anticonformity presented so far is that a given agent may weigh positively some agents while weighing negatively some other ones (recall that in all models we have seen so far, an anticonformist agent reacts negatively to *all* agents). Hence, Altafini's model is the Abelson model where the matrix  $A$  has real coefficients. The corresponding graph is therefore a *signed graph*, where a link has a weight which is either positive or negative. The model reads:

$$\begin{aligned} \dot{x}_i(t) &= \sum_{j=1}^n |a_{ij}| (x_j(t) \text{sgn}(a_{ij}) - x_i(t)) \\ &= \sum_{j=1}^n (a_{ij} x_j(t) - |a_{ij}| x_i(t)) \quad (i \in N) \end{aligned} \quad (14)$$

where  $\text{sgn}$  is the sign function, taking respectively values  $+1, -1$  or  $0$  when the entry is positive, negative or null. In matrix form we obtain exactly (6), but where the Laplacian is defined differently:

$$L[A] = (l_{ij})_{i,j \in N}, \text{ where } l_{ij} = \begin{cases} -a_{ij}, & \text{if } i \neq j \\ \sum_{j=1}^n |a_{ij}|, & \text{if } i = j. \end{cases}$$

An important notion here is the notion of structural balancedness, originally introduced by Cartwright and Harary [78]. An undirected signed graph on  $N$  is *structurally balanced* if one can make a partition  $N = N_1 \cup N_2$  in two groups such that every edge  $ij$  with  $i, j$  in the same group has a positive weight, while if  $i, j$  belong to different groups its weight is negative. Under the assumption that the graph is structurally balanced, it is possible to transform it by a *gauge transform* into a graph where all weights are positive, transforming then the model into the Abelson's model and thus obtaining immediately convergence results. The gauge transform maps the opinion vector  $x$  to the opinion vector  $y$  defined by

$$y_i = \delta_i x_i, \quad \delta_i = \begin{cases} +1, & i \in N_1 \\ -1, & i \in N_2 \end{cases} \quad (i \in N)$$

Noticing that  $\delta_i \delta_j a_{ij} = |a_{ij}|$ , (14) becomes

$$\dot{y}_i(t) = \sum_{j=1}^n |a_{ij}| (y_j(t) - y_i(t)) \quad (i \in N)$$

which is the Abelson's model with matrix  $A' := [|a_{ij}|]$ . A direct application of previous results shows that the model converges, and if the graph is quasi-strongly connected, then

$$\lim_{t \rightarrow \infty} x_i(t) = \begin{cases} p_{\infty}^T x(0), & i \in N_1 \\ -p_{\infty}^T x(0), & i \in N_2 \end{cases}$$

with  $p_{\infty}$  the left eigenvector of  $L[A']$  associated to the eigenvalue 0 and such that  $p_{\infty}^T \mathbf{1} = 1$ . The system converges anyway, even if the graph is not quasi-strongly connected. When the graph is not structurally balanced, it can be shown that all opinions converge to 0 if the graph is strongly connected.

### The EPO model

There exists a vast literature in social psychology and sociology showing that individuals may express public opinions which do not reflect their privately held opinion (we refer the reader to [58] for references in this domain), because of social norm pressure, group pressure, etc. In this respect, the Asch experiment is famous [66]<sup>6</sup>, and it has inspired a number of models trying to reproduce this kind of behavior (see, again, [58] for references), but in a static manner. The EPO (Expressed-Private-Opinion) model, proposed by Anderson and Ye [58], is a dynamic model distinguishing private and public opinions, in the vein of the DeGroot model.

---

<sup>6</sup> The Asch experiment consists in the following: a naive participant is in a room with 7 other participants (called confederates), who have agreed in advance about their answers. Each person has to state aloud which of 3 lines A, B, C the length of a target line is most similar. In each trial, the answer is obvious, the naive participant speaks the last and the confederates obviously gives a wrong answer. It is observed that about one third of people placed in the situation of the naive participant always conformed to the wrong opinion of the group, while 75% conformed at least once over 12 trials.

Let  $x(k)$  and  $\hat{x}(k)$  represent the private and the expressed vectors of opinion, respectively. The updating mechanism is as follows:

$$\begin{aligned} x_i(k+1) &= \lambda_i \left( w_{ii}x_i(k) + \sum_{j \neq i} w_{ij}\hat{x}_j(k) \right) + (1 - \lambda_i)x_i(0) \\ \hat{x}_i(k) &= \phi_i x_i(k) + (1 - \phi_i)\hat{x}_{ave}(k-1) \end{aligned}$$

where  $\lambda_i, W$  are as in the Friedkin-Johnsen model,  $\phi_i \in [0, 1]$  is the *resilience* of agent  $i$ , i.e., its ability to withstand group pressure, and  $\hat{x}_{ave}(k-1)$  is the average of the publicly expressed opinions at time  $k-1$ . Note the similarity with the Friedkin-Johnsen model, with the difference that an individual can only know the publicly expressed opinions of its neighbors.

Supposing that the matrix  $W$  is primitive, and all  $\lambda_i, \phi_i \neq 0, 1$ , it is found in [58] that a steady state for  $x$  and  $\hat{x}$  is reached exponentially fast, and that  $x_i(\infty) \neq \hat{x}_i(\infty)$  in general for any  $i$ . Moreover, there is no consensus for  $x(\infty), \hat{x}(\infty)$ , and the *diameters*  $V(k) := \max_i x_i(k) - \min_i x_i(k)$  and  $\hat{V}(k) := \max_i \hat{x}_i(k) - \min_i \hat{x}_i(k)$  are such that  $V(\infty) > \hat{V}(\infty)$ . This shows that agents, under the pressure of social norm, exhibit more consensus than they really feel. In addition, the steady state for  $x$  and  $\hat{x}$  does not depend on  $\hat{x}(0)$ , the initial public opinion, but only on  $x(0)$ , the initial private opinion.

#### 4.2. Time-varying models

The general form of a time-varying model (in discrete time form) is

$$x(k+1) = A(k)x(k) = A(k)A(k-1) \cdots A(0)x(0),$$

where  $A(k)$  is (usually) a row-stochastic matrix. We denote by  $G(k)$  the graph corresponding to  $A(k)$ .

Generally speaking, the study of such systems relies on results on backwards inhomogeneous matrices products. Most of the results assume the following property, called *Condition (C)* in [95]:

$$\text{For all } k \geq 0, \min_{i,j}^+ a_{ij}^{(k)} \geq \gamma > 0$$

$\gamma$  being independent of  $k$ , and  $\min^+$  denoting the minimum taken over *positive* elements. As a consequence, nothing can be said in general when some elements of  $A(k)$  tend to 0 when  $k \rightarrow \infty$ .

The backwards product is said to be *ergodic* if the product tends to a matrix with identical rows (like the  $k$ th power of a primitive matrix). In this case, the analysis amounts to the classical DeGroot model with a primitive matrix. We give two essential results for ergodicity (see Seneta [95]):

- (i) If the product  $A_{p,r} := A(p+r) \cdots A(p+1)A(p)$  is regular (single aperiodic essential class) for  $p, r \geq 0$  and condition C is satisfied, then ergodicity obtains.
- (ii) If  $\lim_{k \rightarrow \infty} A(k) = A$  and  $A$  is regular, the ergodicity obtains, and the limit vector  $v$  is the unique stationary distribution of  $P$ .

The following results are given in (Bullo [57]).

**Theorem 3.** Suppose that the  $A(k)$ 's are symmetric doubly-stochastic satisfying condition (C) and being primitive. Then the process is ergodic (i.e.,  $x(k)$  converges to the consensus  $\frac{1}{n} \sum_i x_i(0)$ ).

**Theorem 4.** (Moreau [96]) Suppose that the  $A(k)$ 's are row-stochastic matrices satisfying condition (C), with each diagonal entry being nonzero, and there exists an integer  $\delta$  s.t. for all  $k \geq 0$  the graph  $G(k) \cup \cdots \cup G(k+\delta-1)$  has a single essential class. Then the process is ergodic.

**Theorem 5.** Suppose that the  $A(k)$ 's are symmetric doubly-stochastic satisfying condition (C), with each diagonal entry being nonzero, and for all  $k \geq 0$  the graph  $\bigcup_{\tau \geq k} G(\tau)$  is connected. Then the process is ergodic.



The next results are given in (Proskurnikov and Tempo [54]).

**Lemma 1.** (Blondel et al. [97], Lorenz [98]) Suppose the  $A(k)$ 's are row-stochastic, satisfy condition (C) with the constant  $\gamma$ ,  $a_{ii}(k) \geq \gamma$  for all  $i, k$ , and  $a_{ij}(k) > 0$  iff  $a_{ji}(k) > 0$ . Then  $\bar{x} = \lim_{k \rightarrow \infty} x(k)$  exists for any  $x(0)$ , and satisfy  $A(k)\bar{x} = \bar{x}$  for sufficiently large  $k$ . If agents  $i, j$  interact persistently (i.e.,  $\sum_{k=0}^{\infty} a_{ij}(k) = \infty$ ), then  $\bar{x}_i = \bar{x}_j$ .

As a consequence:

**Theorem 6.** (Lorenz [98]) Under the assumptions of the above Lemma, the product  $A(k) \cdots A(0)$  tends to a limit  $\bar{A}$ , which is block-diagonal

$$\bar{A} = \begin{bmatrix} \bar{A}_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \bar{A}_{rr} \end{bmatrix}$$

with each block being a "consensus matrix", i.e.,  $\bar{A}_{ii} = \mathbf{1} \cdot p_i^T$  with  $p_i \in \mathbb{R}^{n_i}$  satisfying  $p_i^T \mathbf{1} = 1$ .

### The DeMarzo-Vayanos-Zwiebel model

This model is proposed by DeMarzo *et al.* [17], and is an example of time-varying model, where typically the above results do not apply. The model reads

$$x(t+1) = [(1 - \lambda_t)I + \lambda_t T]x(t)$$

with  $\lambda_t \in ]0, 1]$  and  $T$  being row-stochastic. Supposing that  $T$  is irreducible and  $\sum_{t=0}^{\infty} \lambda_t = \infty$ , there is convergence to consensus:

$$\lim_{t \rightarrow \infty} x(t) = w^T x(0) \mathbf{1}$$

where  $w$  is the left eigenvector of  $T$  associated to 1. Observe that if  $\lambda_t$  tends to 0 when  $t$  tends to infinity, then condition (C) does not hold.

### 4.3. Bounded confidence models

These models are a nice example of time-varying models and have initiated an important literature. The basic idea is simple: agents listen preferably to agents whose opinion is not too far from their opinion.

Let  $d > 0$  be the *range of confidence*. For a given opinion profile  $x \in \mathbb{R}^N$ , we define for each agent  $i$  the set of "trusted" individuals  $I_i(x) = \{j : |x_j - x_i| \leq d\} \ni i$ , i.e., whose opinions lie in its confidence interval  $[x_i - d, x_i + d]$ . Agent  $i$  updates its opinion by taking the average of the opinions of the trusted individuals:

$$x_i(k+1) = \frac{1}{|I_i(x(k))|} \sum_{j \in I_i(x(k))} x_j(k) \quad (i \in N)$$

This is known in the literature as the Hegselmann-Krause (HK) model [6,7,99] (a closely related model has been proposed by Deffuant *et al.* [8]). It can be seen as a French-DeGroot model with time-varying matrix  $W$ . The HK model is a simple model of *homophily*: *people tend to locomote into groups that share their attitudes and out of groups that do not agree with them (Abelson)*. Note that the HK model supposes that every agent is aware of the opinion of every other agent, implying the existence of an underlying complete network.

An important property is the preservation of order: suppose  $x_{j_1}(k) \leq x_{j_2}(k) \leq \cdots \leq x_{j_n}(k)$ , then it holds  $x_{j_1}(k') \leq x_{j_2}(k') \leq \cdots \leq x_{j_n}(k')$  for every  $k' > k$ . Therefore, we may assume from now on that  $x_1 \leq x_2 \leq \cdots \leq x_n$  at every time step. We say that  $(x_i, \dots, x_m)$  is a *d-chain* if all the distances

between two consecutive opinions are  $\leq d$ . Then the opinion vector  $x$  is formed by a set of maximal  $d$ -chains.

**Proposition 1.** Suppose  $x_1(0) \leq x_2(0) \leq \dots \leq x_n(0)$  and  $x_{i+1}(0) - x_i(0) > d$ . Then  $x_i(k) \leq x_i(0)$  and  $x_{i+1}(k) \geq x_{i+1}(0)$  for all  $k > 0$ .

We list the main properties of maximal  $d$ -chains:

- (i) Two maximal  $d$ -chains can never merge (consequence of Proposition 1).
- (ii) The diameter  $x_m(k) - x_i(k)$  of a maximal  $d$ -chain  $(x_i, \dots, x_m)$  is nonincreasing (this is also consequence of Proposition 1).
- (iii) Maximal  $d$ -chains of length at least 5 can split. Others necessarily collapse into a consensus (in one step for length 2, in 2 steps for length 3, and in 5 steps for length 4). Example with 5: let  $n = 7$ ,  $d = 1$  and  $x(0) = (0, 0, 1, 2, 3, 4, 4)$ , forming one maximal  $d$ -chain. Then  $x(1) = (1/3, 1/3, 3/4, 2, 13/4, 11/3, 11/3)$ , which forms 3 maximal  $d$ -chains.
- (iv) A maximal  $d$ -chain with diameter not greater than  $d$  converges in one shot to a consensus, as all agents in the chain are connected.

More precisely, we have:

**Proposition 2.** During two consecutive steps  $k$  and  $k + 1$ , any maximal  $d$ -chain either collapses into a singleton, splits into several maximal  $d$ -chains or reduces in diameter by at least  $d/n^2$ .

As a consequence:

**Theorem 7.** For any  $x(0)$ , the HK model converges to a vector  $\bar{x}$  in a finite number of steps (not more than  $O(n^3)$  steps), with the property that for each distinct  $i, j$ , either  $\bar{x}_i = \bar{x}_j$  (consensus) or  $|\bar{x}_i - \bar{x}_j| > d$  (distrust).

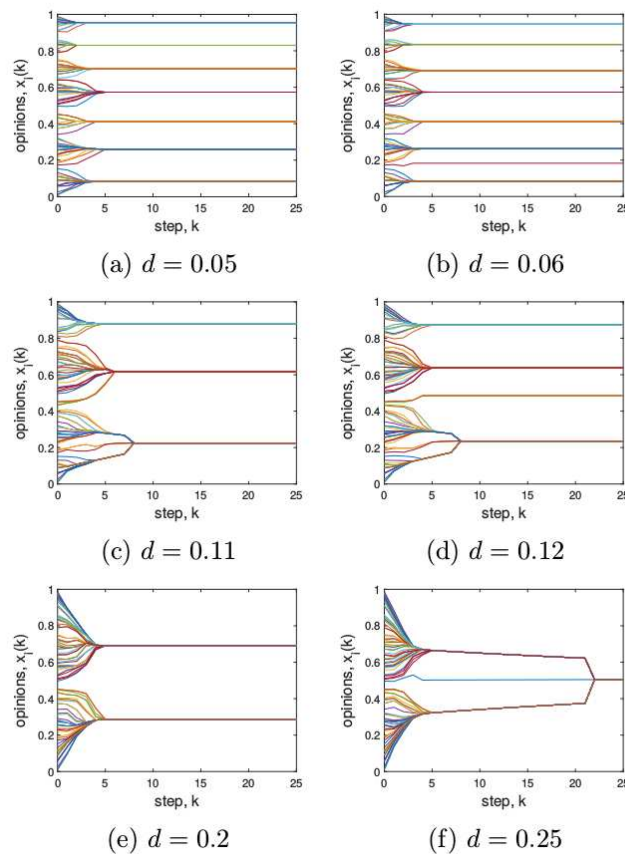
Lastly, we borrow from Proskurnikov and Tempo [54] Figure 7, on which one can see the evolution in time of the opinion of a society of 100 agents, with  $x_i(0)$  uniformly distributed in  $[0,1]$ . One may intuitively think that the number of clusters at convergence is monotonically decreasing with the value of the confidence range  $d$ . But this is not true, as it can be seen from the figure: there is an increase of the number of clusters when  $d$  goes from 0.05 to 0.06, and also from 0.11 to 0.12. Also, the convergence time does not seem to depend monotonically on  $d$ .

#### 4.4. A final remark on continuous opinion models

Most of the presented models show a convergence of the opinion vector in the long run to some steady state, and non-convergence (like cycles, oscillations) arises only in rare cases (we have mentioned this for the Altafini model, but it can even happen with the simple DeGroot model, e.g., if the matrix  $W$  is irreducible but not primitive, in which case the class is periodic).

It is useful to distinguish several types of convergence. The simplest case is the *consensus*, where all agents converge to the same opinion. This happens with the DeGroot model when  $W$  is primitive, and for its continuous-time version, the Abelson model. The second case of interest is *polarization*, where the society is divided into two parts, each part reaching consensus, but different and often well apart. This is typically the situation obtained with Altafini's model under the condition that the graph is structurally balanced. Going a step further, one speaks of *clustering* or *segmentation*. It means that the society breaks into disjoint components (segments), each reaching a consensus. This is exactly the result obtained in the Hegselmann-Krause model, as it can be seen from Figure 7.

According to Flache *et al.* [59], this is typical of the three main classes of (continuous) opinion dynamics models. In models of *assimilative social influence* (like the DeGroot model), individuals influence each other towards reducing opinion difference, and provided that the network is strongly connected, consensus is reached in the long run. In models with *similarity biased influence* (like



**Figure 7.** The HK model with  $n = 100$  and various  $d$  (borrowed from Proskurnikov and Tempo (2018))

bounded confidence models), only individuals with sufficiently close opinions can influence each other, and as a result, there is a fragmentation of the society into clusters, each of them reaching consensus. The third type consists in models with *repulsive influence*, where individuals influence each other towards increasing mutual opinion differences, where typically polarization with maximal distance between opinions obtains.

Reaching a consensus value within each group is called *weak diversity* [58,100], which means that there is no difference between opinions in the same cluster. There is a growing interest to study models which are able to capture *strong diversity*, frequently observed in practice, i.e., where there is a range of diverse opinion values within each cluster. The Friedkin-Johnsen model is able to represent this, due to the presence of the  $(I - \Lambda)u$  term in (10), making some of the agents (precisely, those who are prejudiced) stubborn to a certain degree, and therefore sticking more or less to their initial opinion. Indeed, the matrix  $V$  in (13) does not have identical rows, making the components of  $x^1(\infty)$  to differ. The same conclusion applies to the EPO model.

## 5. Conclusion

In the present paper we aimed at providing a short survey of selected works on opinion dynamics. As the literature is extremely vast and comes from different scientific fields, we could focus only on specific frameworks, and chose some contributions that assume non-Bayesian updating of opinions and no strategic consideration. Another survey, focusing on Bayesian updating and recent works that combine both approaches, would be useful in order to complete the present overview. Despite the extremely broad literature on opinion dynamics, the phenomenon continues to attract a lot of attention and interest.

In the paper we did not provide any survey of the literature that explores experimental and behavioral approaches to opinion dynamics. While we focused in our presentation on theoretical works, in the network literature there exist important experimental contributions to this issue. In particular, lab and field experiments are used in order to test either Bayesian models of social learning in networks or non-Bayesian (DeGroot) updating, or to compare both types of learning behavior. For instance, the experimental data on networks of three nodes presented in Choi *et al.* [101] appear to be consistent with Bayesian behavior. Choi *et al.* [102] conduct experimental study of learning in three-person networks and use the so-called quantal response equilibrium for interpreting experimental data. Corazzini *et al.* [103] show the empirical evidence of the DeGroot-like behavior in networks of four nodes. Mueller-Frank and Neri [105] investigate a model of non-Bayesian learning in networks in an environment with a finite set of alternatives and run experiments whose outcomes appear to be consistent with their model. Mobius *et al.* [108] use a field experiment on social learning to analyze both information diffusion and aggregation. They compare two mechanisms of information aggregation, a naïve learning model and a sophisticated model. They can distinguish between the two models after incorporating imperfect information diffusion, and find more evidence for the sophisticated one. Mengel and Grimm [106] conduct lab experiments and compare the DeGroot-like learning with Bayesian learning. Also, Grimm and Mengel [107] study belief formation in social networks using lab experiments, and examine both naïve (DeGroot) and Bayesian models. Similarly, Chandrasekhar *et al.* [104] investigate a model of social learning with agents being potentially Bayesian or DeGroot-like. The authors conduct lab experiments in Indian villages and with university students in Mexico City, to study if learning behavior is consistent with DeGroot, Bayesian, or a mixed population.

We conclude this survey by giving a summary of the main properties of all models we studied in Table 1.

Model	Opinion	Time	Updating	Behavior	Convergence	Other
Threshold	disc.	disc.	sync.	conf/anticonf	stable/unstable steady state/cycle	
$q$ -voter	disc.	disc.	async.	conf/anticonf	stable/unstable steady state	
majority rule	disc.	disc.	async.	conf/anticonf	stable/unstable steady state	
aggregation	dis.	disc.	sync.	conf/anticonf	polarization, cycles, fuzzy polar., fuzzy cycles	
DeGroot	cont.	disc.	sync.	conf.	consensus	
Abelson	cont.	cont.	sync.	conf.	consensus	
Friedkin-Johnsen	cont.	disc.	sync.	conf.	strong diversity	
Taylor	cont.	cont.	sync.	conf.	strong diversity	
Altafini	cont.	disc.	sync.	conf/anticonf links	polarization	
EPO	cont.	disc.	sybc.	conf.	strong diversity	private/public opinion
DeMarzo et al.	cont.	disc.	sync.	conf.	consensus	time-varying
Hegselmann-Krause	cont.	disc.	sync.	conf.	clustering	time-varying; bounded confidence

**Table 1.** Summary of the main properties of the various models. disc.=discrete, cont.=continuous, sync.=synchronous, async.=asynchronous, conf.=conformist, anticonf.=anticonformist. Only most typical convergence results are indicated.

**Author Contributions:** The authors have equally contributed to the writing of the paper.

**Funding:** The authors acknowledge the financial support from the European Union’s Horizon 2020 research and innovation programme under the grant agreement No 721846, “Expectations and Social Influence Dynamics in Economics (ExSIDE)”.

**Acknowledgments:** The authors would like to thank the three anonymous referees for their careful reading and useful suggestions, which permitted to considerably improve the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. DeGroot, M.H. Reaching a consensus. *J. of the American Statistical Association* **1974**, *69*, 118–121.
2. French Jr, J. A formal theory of social power. *Psychol. Rev.* **1956**, *63*, 181–194.
3. Abelson, R. Mathematical models of the distribution of attitudes under controversy. In *Contributions to Mathematical Psychology*; Frederiksen, N.; Guliksen, H., Eds.; Holt, Rinehart & Winston, Inc: New York, 1964; pp. 142–160.
4. Taylor, M. Towards a mathematical theory of influence and attitude change. *Human Relations* **1968**, *21*, 121–139.
5. Friedkin, N.; Johnsen, E. Social influences and opinion. *J. Math. Sociol.* **1990**, *15*, 193–205.
6. Krause, U. A discrete nonlinear and non-autonomous model of consensus formation. *Commun. in Difference Equations* **2000**, pp. 227–236.
7. Hegselmann, R.; Krause, U. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation* **2002**, *5*.
8. Deffuant, G.; Neau, D.; Amblard, F.; Weisbuch, G. Mixing beliefs among interacting agents, analysis, and simulation. *Advances in Complex Systems* **2000**, *3*, 87–98.
9. Bergemann, D.; Morris, S. Information design, Bayesian persuasion, and Bayes correlated equilibrium. *American Economic Review* **2016**, *106*, 586–591.
10. Kamenica, E.; Gentzkow, M. Bayesian persuasion. *American Economic Review* **2011**, *101*, 2590–2615.
11. Bénabou, R.; Tirole, J. Mindful economics: The production, consumption, and the value of beliefs. *Journal of Economic Perspectives* **2016**, *30*, 141–164.
12. Granovetter, M. Threshold models of collective behavior. *American Journal of Sociology* **1978**, *83*, 1420–1443.
13. Schelling, T. *Micromotives and macrobehaviour*; Norton: New York, 1978.
14. Jackson, M. *Social and Economic Networks*; Princeton University Press, 2008.
15. Acemoglu, D.; Ozdaglar, A. Opinion dynamics and learning in social networks. *Dynamic Games and Applications* **2011**, *1*, 3–49.
16. Bramoullé, Y.; Galeotti, A.; Rogers, B.W., Eds. *The Oxford Handbook of the Economics of Networks*; Oxford University Press, 2016.
17. DeMarzo, P.; Vayanos, D.; Zwiebel, J. Persuasion bias, social influence, and unidimensional opinions. *Quarterly Journal of Economics* **2003**, *118*, 909–968.
18. Golub, B.; Jackson, M.O. Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics* **2010**, *2*, 112–149.
19. Büchel, B.; Hellmann, T.; Pichler, M. The dynamic of continuous cultural traits in social networks. *J. of Economic Theory* **2014**, *154*, 274–309.
20. Büchel, B.; Hellmann, T.; Klößner, S. Opinion dynamics and wisdom under conformity. *Journal of Economic Dynamics and Control* **2015**, *52*, 240–257.
21. Banerjee, A.V. A simple model of herd behavior. *Quarterly Journal of Economics* **1992**, *107*, 797–817.
22. Ellison, G. Learning, local interaction, and coordination. *Econometrica* **1993**, *61*, 1047–1072.
23. Ellison, G.; Fudenberg, D. Rules of thumb for social learning. *Journal of Political Economy* **1993**, *101*(4), 612–643.
24. Ellison, G.; Fudenberg, D. Word-of-mouth communication and social learning. *Journal of Political Economy* **1995**, *111*(1), 93–125.
25. Bala, V.; Goyal, S. Learning from neighbours. *Review of Economic Studies* **1998**, *65*(3), 595–621.
26. Bala, V.; Goyal, S. Conformism and diversity under social learning. *Economic Theory* **2001**, *17*, 101–120.
27. Banerjee, A.V.; Fudenberg, D. Word-of-mouth learning. *Games and Economic Behavior* **2004**, *46*, 1–22.

28. Scharfstein, D.S.; Stein, J.C. Herd behavior and investment. *The American Economic Review* **1990**, *80*(3), 465–479.
29. Bikhchandani, S.; Hirshleifer, D.; Welch, I. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy* **1992**, *100*, 992–1026.
30. Anderson, L.; Holt, C. Information cascades in the laboratory. *The American Economic Review* **1997**, *87*, 847–862.
31. Celen, B.; Kariv, S. Distinguishing informational cascades from herd behavior in the laboratory. *American Economic Review* **2004**, *94*(3), 484–497.
32. Banerjee, A.V. The economics of rumours. *Review of Economic Studies* **1993**, *60*, 309–327.
33. Chatterjee, K.; Dutta, B. Credibility and strategic learning in networks. *International Economic Review* **2016**, *57*(3), 759–786.
34. Bloch, F.; Demange, G.; Kranton, R. Rumors and social networks. *International Economic Review* **2018**, *59*, 421–448.
35. Morris, S. Contagion. *The Review of Economic Studies* **2000**, *67*, 57–78.
36. Grabisch, M.; Rusinowska, A.; Venel, X. Diffusion in large networks. arXiv:2010.09256, 2020.
37. Galam, S. *Sociophysics — A Physicist's Modeling of Psycho-political Phenomena*; Springer, 2012.
38. Weisbuch, G.; Kirman, A.; Herreiner, D. Market organization. *Economica* **2000**, *110*, 411–426.
39. Weisbuch, G. Bounded confidence and social networks. *The European Physical Journal B* **2004**, *38*, 339–343.
40. Galam, S. Minority opinion spreading in random geometry. *Eur. Phys. J. B* **2002**, *25*, 403–406.
41. Galam, S. Contrarian deterministic effects on opinion dynamics: "the hung elections scenario". *Physica A* **2004**, *333*, 453–460.
42. Jędrzejewski, A.; Sznajd-Weron, K. Statistical Physics Of Opinion Formation: Is it a SPOOF? *Comptes Rendus Physique* **2019**.
43. Nowak, N.; Sznajd-Weron, K. Homogeneous symmetrical threshold model with nonconformity: independence versus anticonformity. *Complexity* **2019**, pp. 1–14. ID 5150825.
44. Nyczka, P.; Sznajd-Weron, K. Anticonformity or independence? — Insights from statistical physics. *J. Stat. Physics* **2013**, *151*, 174–202.
45. Nyczka, P.; Sznajd-Weron, K.; Cisko, J. Phase transition in the  $q$ -voter model with two types of stochastic driving. *Phys. Rev. E* **2012**, *86*, 011105.
46. Castellano, C.; Fortunato, S.; Loreto, V. Statistical physics of social dynamics. *Rev. Modern Physics* **2009**, *81*, 591–646.
47. Gravner, J.; Griffeath, D. Cellular automaton growth on  $\mathbb{Z}^2$ : Theorems, examples and problems. *Advances in Applied Mathematics* **1998**, *21*, 241–304.
48. Holley, R.A.; Liggett, T.M. Ergodic theorems for weakly interacting infinite systems and the voter model. *The Annals of Probability* **1975**, pp. 643–663.
49. Mossel, E.; Tamuz, O. Opinion exchange dynamics. *Probability Surveys* **2017**, *14*, 155–204.
50. Altafini, C. Consensus problems on networks with antagonistic interactions. *IEEE Tr. on Automatic Control* **2012**, *58*, 935–946.
51. Altafini, C. Dynamics of opinion forming in structurally balanced social networks. 51st IEEE Conf. on Decision and Control (CDC), 2012, p. 5876–5881.
52. Shi, G.; Altafini, C.; Baras, J.S. Dynamics over signed networks. *SIAM Review* **2019**, *61*, 229–257.
53. Proskurnikov, A.V.; Tempo, R. A tutorial on modeling and analysis of dynamic social networks (I). *Annual Reviews in Control* **2017**, *43*, 65–79. arXiv:1701.06307v2.
54. Proskurnikov, A.V.; Tempo, R. A tutorial on modeling and analysis of dynamic social networks (II). *Annual Reviews in Control* **2018**, *45*, 166–190.
55. Breer, V.; Novikov, D.; Rogatkin, A. *Mob control: Models of threshold collective behavior*; Springer, 2017.
56. Fagnani, F.; Frasca, P. *Introduction to Averaging Dynamics over Networks*; Lecture Notes in Control and Information Sciences 472, Springer, 2018.
57. Bullo, F. *Lectures on Network Systems*; Kindle Direct Publishing, 2019.
58. Anderson, B.D.O.; Ye, M. Recent advances in the modelling and analysis of opinion dynamics on influence networks. *Int. J. of Automation and Computing* **2019**, *16*, 129–149.
59. Flache, A.; Mäs, M.; Feliciani, T.; Chattoe-Brown, E.; Deffuant, G.; Huet, S.; Lorenz, J. Models of social influence: towards the next frontiers. *J. of Artificial Societies and Social Simulation* **2017**, *20*.

60. Surowiecki, J. *The Wisdom of Crowds*; Anchor Books, 2004.
61. Denneberg, D. Conditioning (updating) non-additive measures. *Annals of Operations Research* **1994**, 52, 21–42.
62. Gilboa, I.; Schmeidler, D. Updating ambiguous beliefs. *J. of Economic Theory* **1993**, 59, 33–49.
63. Pires, C.P. A rule for updating ambiguous beliefs. *Theory and Decision* **2002**, 53, 137–152.
64. Jadbabaie, A.; Molavi, P.; Sandroni, A.; Tahbaz-Salehi, A. Non-Bayesian social learning. *Games and Economic Behavior* **2012**, 76, 210–225.
65. Molavi, P.; Tahbaz-Salehi, A.; Jadbabaie, A. A theory of non-Bayesian social learning. *Econometrica* **2018**, 86, 445–490.
66. Asch, S. *Social psychology*; Englewood Cliffs, NJ: Prentice Hall, 1952.
67. Shiller, R.J. Conversation, information, and herd behavior. *The American Economic Review* **1995**, 85, 181–185.
68. Brunnermeier, M.K. *Asset pricing under asymmetric information: Bubbles, crashes, technical analysis, and herding*; Oxford University Press, 2001.
69. Avery, C.; Zemsky, P. Multidimensional uncertainty and herd behavior in financial markets. *American Economic Review* **1998**, pp. 724–728.
70. Park, A.; Sabourian, H. Herding and contrarian behavior in financial markets. *Econometrica* **2011**, 79, 973–1026.
71. Hommes, C. *Behavioral rationality and heterogeneous expectations in complex economic systems*; Cambridge University Press, 2013.
72. Camerer, C.F.; Fehr, E. When does "economic man" dominate social behavior? *Science* **2006**, 311, 47–52.
73. Fehr, E.; Tyran, J.R. Individual irrationality and aggregate outcomes. *Journal of Economic Perspectives* **2005**, 19, 43–66.
74. Moulin, H. *Game theory for social sciences*; Wiley, 1988.
75. Nagel, R. Unraveling in guessing games: An experimental study. *The American Economic Review* **1995**, 85, 1313–1326.
76. Mueller-Frank, M. A general framework for rational learning in social networks. *Theoretical Economics* **2013**, 8, 1–40.
77. Mossel, E.; Mueller-Frank, M.; Sly, A.; Tamuz, O. Social learning equilibria. *Econometrica* **2020**, pp. 1235–1267.
78. Cartwright, D.; Harary, F. Structural balance: a generalization of Heider's theory. *Psychological Review* **1956**, 63, 277–292.
79. McCulloch, W.S.; Pitts, W. Logical calculus of the ideas immanent in nervous activity. *Bulletin of the Mathematical Biophysics* **1943**, 5, 115–133.
80. Goles, E.; Olivos, J. Periodic behavior of generalized threshold functions. *Discrete Applied Mathematics* **1980**, 30, 187–189.
81. Vanelli, M.; Arditti, L.; Como, G.; Fagnani, F. On games with coordinating and anti-coordinating agents. arXiv:1912.02000.
82. Grabisch, M.; Li, F. Anti-conformism in the threshold model of collective behavior. *Dynamic Games and Applications* **2020**, 10, 444–477.
83. Gardini, L.; Sushko, I.; Matsuyama, K. 2D discontinuous piecewise linear map: Emergence of fashion cycles. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2018**, 28.
84. Clifford, P.; Sudbury, A. A model for spatial conflict. *Biometrika* **1973**, 60, 581–588.
85. Castellano, C.; Muñoz, M.A.; Pastor-Satorras, R. Nonlinear  $q$ -voter model. *Phys. rev. E* **2009**, 80, 041129.
86. Sznajd-Weron, K.; Sznajd, J. Opinion evolution in closed community. *Int. J. of Modern Physics C* **2000**, 11, 1157–1165.
87. Anderson, S.P.; Goeree, J.K.; Holt, C.A. Noisy directional learning and the logit equilibrium. *Scandinavian Journal of Economics* **2004**, 106, 581–602.
88. Strogatz, S. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*; Perseus Book Publishing, 1994.
89. Newman, M.E.J. *Networks: An Introduction*; Oxford University Press, 2010.
90. Galam, S.; Jacobs, F. The role of inflexible minorities in the breaking of democratic opinion dynamics. *Physica A* **2007**, 381, 366–376.



91. Grabisch, M.; Rusinowska, A. A model of influence based on aggregation functions. *Mathematical Social Sciences* **2013**, *66*, 316–330. DOI: 0.1016/j.mathsocsci.2013.07.003.
92. Förster, M.; Grabisch, M.; Rusinowska, A. Anonymous Social Influence. *Games and Economic Behavior* **2013**, *82*, 621–635.
93. Grabisch, M.; Poindron, A.; Rusinowska, A. A model of anonymous influence with anti-conformist agents. *J. of Economic Dynamics and Control* **2019**, *103*, 103773.
94. Poindron, A. A general model of binary opinions updating. *Mathematical Social Sciences* **to appear**. <https://doi.org/10.1016/j.mathsocsci.2020.10.004>.
95. Seneta, E. *Non-negative Matrices and Markov Chains*; Springer Series in Statistics, Springer, 2006.
96. Moreau, L. Stability of multiagent systems with time-dependent communication links. *IEEE Tr. on Automatic Control* **2005**, *50*, 169–182.
97. Blondel, V.; Hendrickx, J.; Olshevsky, A.; Tsitsiklis, J. Convergence in multiagent coordination, consensus, and flocking. Proc. IEEE Conf. Decision and Control, 2005, pp. 2996–3000.
98. Lorenz, J. A stabilization theorem for dynamics of continuous opinions. *Physica A* **2005**, *353*, 217–223.
99. Dittmer, J.C. Consensus formation under bounded confidence. *Nonlinear Analysis* **2001**, *47*, 4615–4621.
100. Mäs, M.; Flache, A.; Kitts, J.A. Cultural integration and differentiation in groups and organizations. In *Perspectives on Culture and Agent-based Simulations: Integrating Cultures*; Dignum, V.; Dignum, F., Eds.; Springer, 2014; pp. 71–90.
101. Choi, S.; Gale, D.; Kariv, S. Behavioral aspects of learning in social networks: An experimental study. *Advances in Applied Microeconomics: A Research Annual* **2005**, *13*, 25–61.
102. Choi, S.; Gale, D.; Kariv, S. Social learning in networks: A quantal response equilibrium analysis of experimental data. *Review of Economic Design* **2012**, *16*, 93–118.
103. Corazzini, L.; Pavezi, F.; Petrovich, B.; Stanca, L. Influential listeners: An experiment on persuasion bias in social networks. *European Economic Review* **2012**, *56*, 1276–1288.
104. Chandrasekhar, A.G.; Larreguy, H.; Xandri, J.P. Testing models of social learning on networks: Evidence from a lab experiment in the field. *Econometrica* **2020**, *88*(1), 1–32.
105. Mueller-Frank, M.; Neri, C. Social learning in networks: Theory and experiments. Working Paper.
106. Mengel, F.; Grimm, V. An experiment on learning in a multiple games environment. *Journal of Economic Theory* **2012**, *147*, 2220–2259.
107. Grimm, V.; Mengel, F. Experiments on Belief Formation in Networks. *Journal of the European Economic Association* **2020**, *18*(1), 49–82.
108. Mobius, M.; Phan, T.; Szeidl, A. Treasure hunt: Social learning in the field. Discussion Paper.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Submitted to *Games* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).