



Perturbed Anisotropic Opinion Dynamics with Delayed Information

Juan Neirotti¹

Received: 16 January 2023 / Accepted: 7 July 2023
© The Author(s) 2023

Abstract

We revisited the problem of modeling a publicity campaign in a society of intelligent agents that form their opinions by interchanging information with each other and with the society as a whole. We use a Markov approximation to consider the effects of an interaction delay τ in the system of perturbed differential equations that model the social dynamics. The stable points of the dynamical system are the manifestation of the agent's attitudes, either in favor or against the social rule, as it was previously found, but the approach to the stable points is greatly modified by the presence of the delay.

Keywords Opinion Dynamics · Sociophysics

1 Introduction

Opinions are mental representations of the individual's beliefs, constructed by inference processes mostly done with incomplete information. We are interested on the process of opinion formation in a population of individuals, represented by a set of agents $\{a : 1 \leq a \leq M\}$, interacting within a society where there are some rules (also represented by an agent, B) that determine whether a social issue (represented by \mathbf{S}) is socially acceptable or not. We say that the label B assigns to an issue \mathbf{S} is $\sigma_B(\mathbf{S})$ that is either 1 (acceptable, legal, fashionable) or -1 (not acceptable, illegal, unfashionable). In similar way we state that the opinion of individual a on issue \mathbf{S} is $\sigma_a(\mathbf{S})$, either 1 or -1 . Modeling opinions with binary variables [1] is consistent with the observation that most people opt for one out of two opposite positions while answering questions with high social impact [2].

In the presented scenario we understand the opinion formation process in agent a as the process where a learns how to classify issues \mathbf{S} like B (otherwise a wouldn't know how to behave properly in the society ruled by B). If the agents are rational (i.e., Bayesian [3]) two noninteracting agents should, upon receiving the same complete information, reach the same opinion, disregarding their priors (see Ref. [4], Theorem 1). Therefore, a model of

Communicated by Pierpaolo Vivo.

✉ Juan Neirotti
j.p.neirotti@aston.ac.uk

¹ Department of Mathematics, Aston University, The Aston Triangle, B4 7ET Birmingham, UK

the opinion formation process of rational agents in a society with rules must consider an interaction mechanism between agents as a means for the emergence of different social positions [5–14]. There is a body of evidence supporting the effect of social influence on opinion formation process [15–21]; as a consequence, to model the agents' interactions, we follow social impact theory [22, 23] assuming that the interaction between an agent and its neighbors has a strength proportional to the credibility, number and proximity of the neighbors to the agent [24–26] (see Sect. 2 for details).

There are several studies on the effects of external influences on the opinion formation process in voters models [27, 28] and the influence of adverts on the formation of opinions [29]; in particular in Refs. [30, 31] it is concluded that, given a particular measure, periodic advert campaigns are the most effective. Therefore, we increase the complexity of our model by adding a (small) periodic term to the interaction of the agents with the society as a whole, to emulate the effects of an advert campaign in favor of the position hold by B . Keeping the magnitude of this term smaller than the magnitude of the typical inter-agent interaction allows us to characterize the advert campaign as a perturbation.

In recent works [32, 33] we introduced the above-mentioned concepts to explain the opinion formation process in a society of interacting agents where there is a rule B that indicates the social position. The social position, originally supposed to be fixed, was allowed to evolve [34], and then reinforced by a perturbation (i.e. an added term, small in magnitude when compared with the strength of other interactions) representing an official publicity campaign [35]. In the present article we complete this effort by studying the effects of delay of information in the interaction between agents.

Models for opinion formation process consider inter-agent interactions as means for the emergence of different social positions. The model we present in this article differs from the classical Deffuant model of opinion dynamics [36, 37], where the interaction between voters takes place if the difference between the variables that indicate the opinions of the interacting agents is below a given threshold (bounded confidence model). In our model the interaction is between connected neighbors that learn to have an opinion; the stronger the connection the larger the influence peers have in the local neighborhood, opening the possibility for the emergence of local consensuses opposed to the social *status quo*. Our model differs also from models that consider opinions as continuous variables [38, 39]; the evolution of the connections between agents [40–42] has not been considered either.

Several opinion formation models consider a time delay in the macroscopic description of the system [43–46] for instance as a means to describe a mechanism for observed instabilities in the dynamics of democratic political systems [47–49]. In the present model, interactions are based on the perception agents have on past positions of their neighbors, which motivates the introduction of a time delay τ . We expect this approach to be also analytically tractable, thus facilitating a more direct and thorough analysis.

The article is organized as follows. In Sect. 2 we discuss our model and we present the differential equations that rule the time evolution of the parameters of the system. The derivation of the set of differential equations can be found in Appendix A. By building on the results presented in Appendix B, we solve the system of differential equations for constant (zero frequency) and periodic (non-zero frequency) perturbations in Sect. 3. This enables us to plot the phase diagram of the system, based on the amplitude of the perturbation and the initial condition for the most conservative agent. We present our final discussions in Sect. 4.

2 The Model

Let us suppose we have a population of M agents $\{a\}_{a=1}^M$, living in a society with a set of rules B , also represented as an agent. We represent the topology of the society by a directed graph $G = \{\{a\}, \{\eta_{a,b}\}\}$ where $\{a\}$ is a set of vertexes associated with the agents and $\{\eta_{a,b}\}$ is a set of strengths $\eta_{a,b}$ that represent the influence of agent b on agent a . Self-influence effects are neglected, i.e. $\eta_{a,a} = 0$. The neighborhood of a is defined as $\mathbb{N}_a = \{c \in [M] : \eta_{a,c} > 0\}$ [2].

B determines whether a social issue $\mathbf{S} \in \{-1, +1\}^N$ is acceptable $\sigma_B(\mathbf{S}) = 1$ or not $\sigma_B(\mathbf{S}) = -1$.¹ With the end to give an internal structure to the opinion variables and to make the mathematical description of the system tractable we provide the agents (a and B alike) with a *perceptron* [50]. In this form we introduce an adaptive, cognitive mechanism for the formation of opinions [51–54]. Each perceptron is characterized by an internal representation vector ($\mathbf{B} \in \mathbb{R}^N$ for the social rule $\mathbf{J}_a \in \mathbb{R}^N$ for the agents) such that the labels become $\sigma_B(\mathbf{S}) = \text{sgn}(\mathbf{B} \cdot \mathbf{S})$ and the opinions become $\sigma_a(\mathbf{S}) = \text{sgn}(\mathbf{J}_a \cdot \mathbf{S})$, where $\mathbf{V} \cdot \mathbf{S} \equiv \sum_{i=1}^N V_i S_i$ for all $\mathbf{V} \in \mathbb{R}^N$, and $\text{sgn}(x) = 1$ if $x > 0$, -1 if $x < 0$, and 0 if $x = 0$.²

In the current scenario we consider all internal representations, \mathbf{B} and $\{\mathbf{J}_a\}$, to evolve over time. To make the evolution happen we consider the sampling $\mathbb{S} = \{(\sigma_{B,n}, \sigma_{\mathbb{N}_a,n}, \mathbf{S}_n)\}_{n=1}^T$, where $\sigma_{B,n} = \text{sgn}(\mathbf{B} \cdot \mathbf{S}_n)$, $\sigma_{\mathbb{N}_a,n} = \{\sigma_c(\mathbf{S}_{n-m}) : c \in \mathbb{N}_a\}$ is the information agent a receives from its neighbors, with a delay $m < n$, \mathbf{S}_n belongs to $\{-1, +1\}^N$ and it has been drawn according to the probability $\mathcal{P}(\mathbf{S}) = \prod_{j=1}^N \mathcal{P}(S_j) = \prod_{j=1}^N \{\frac{1}{2} \delta_{S_j,+1} + \frac{1}{2} \delta_{S_j,-1}\}$. The symbol $\delta_{A,B} = 1$ if $A = B$ and 0 otherwise, is the Kronecker delta. We are assuming that the entries of the vectors \mathbf{S} are independent and identically distributed (iid) variables with the same probability to have the value $+1$ or -1 . The length of the sampling \mathbb{S} is typically assumed to be $T = \alpha N$ where $\alpha \in \mathbb{R}$ is independent of N .

In an online scenario, elements of the sampling \mathbb{S} are taken one at a time, applied to the internal-representation update algorithms, and then discarded. The update algorithms for the internal representations of $\{a\}$ and B are:

$$\mathbf{J}_{a,n+1} = \mathbf{J}_{a,n} + \psi_{a,n} \frac{\sigma_{B,n} \mathbf{S}_n}{\sqrt{N}} + f_n \frac{|\mathbf{J}_{a,n}|}{\sqrt{N}} \sqrt{\frac{1 + R_{a,n}}{N(1 - R_{a,n})}} P_n \mathbf{b}_n \tag{1}$$

$$\mathbf{B}_{n+1} = \mathbf{B}_n + \frac{\lambda_o}{\sqrt{N}} f_n \mathbf{L}_n, \tag{2}$$

where $\psi_{a,n}$ is the learning amplitude, f_n is the annealing factor, $R_{a,n}$ is the parameter we will use to describe the evolution of the system, and it is defined as:

$$R_{a,n} = \frac{\mathbf{J}_{a,n} \cdot \mathbf{B}_n}{|\mathbf{J}_{a,n}| |\mathbf{B}_n|}, \tag{3}$$

P_n is a suitable periodic function, \mathbf{b}_n is a vector of length one in the direction of \mathbf{B}_n , λ_o is the rate of change of the rule B , and \mathbf{L}_n is the average projection of the vectors $\{\mathbf{J}_{a,n}\}$ onto the plane perpendicular to \mathbf{B}_n . In order to facilitate the reading of this manuscript, we have left the full details for the development of the equations of motion from these update algorithms

¹ Any social issue can be written in a binary code. We are supposing the letters of such a code are 1 and -1 and that all issues codified in this form have a length N that is sufficiently large.

² It appears that there are three possible labels in the classification of issues, $+1$, -1 and 0 . The 0 label is assigned to the issues \mathbf{S}_0 such that $\mathbf{B} \cdot \mathbf{S}_0 = 0$, which is the equation of the hyperplane perpendicular to the vector \mathbf{B} . Given that the fraction of vectors \mathbf{S}_0 living in this hyperplane is negligible small compared to the total number of possible issues \mathbf{S} , we will simply consider the events with label 0 to have a null probability.

for Appendix A. Here, we will only make a mention about the learning amplitude $\psi_{a,n}$ and the periodic function P_n . The last term to the right-hand side of Eq. (1) corresponds to the modeling of the publicity campaign, that favors the increment of the $\mathbf{J}_{a,n}$ in the direction of \mathbf{B}_n . The magnitude of this term is small when compared to the second term, and given that P_n is periodic, this term is treated as a periodic perturbation (the perturbation character of this term is explicitly used in Sect. 3). The construction of the learning amplitude $\psi_{a,n}$ is inspired from social impact theory [22, 23], and involves a corroboration mechanism [55, 56] done with delayed information from the neighbors.

The relevant parameter of the system is $R_{a,n}$, defined in Eq. (3). This quantity is the projection of the normalized vector $\mathbf{J}_{a,n}$ onto the direction of \mathbf{B}_n , i.e. is the cosine of the angle $\theta_{a,n}$ between $\mathbf{J}_{a,n}$ and \mathbf{B}_n , $R_{a,n} = \cos(\theta_{a,n})$. It is also known as the overlap between the $\mathbf{J}_{a,n}$ and \mathbf{B}_n . $R_{a,n}$ represents the level of agreement of agent a with the social rule B at time n . Following the techniques presented in Appendix A and considering the large N limit we finally obtain that, for a pair of interacting agents, the evolution of the agreement in continuous time is:

$$\begin{aligned} \dot{R}_a = & -\frac{\sum_{b \in \mathbb{N}_a} \eta_{a,b} - 2}{2} (1 - R_a^2) + \lambda_o \sqrt{1 - R_a^2} \\ & + \left[\frac{1}{2} \sum_{b \in \mathbb{N}_a} \eta_{a,b} \Theta(\bar{R}_b - R_a) \left(\sqrt{1 - R_a^2 \bar{R}_b} - \sqrt{1 - \bar{R}_b^2 R_a} \right) + P(t) \right] \sqrt{1 - R_a^2}. \end{aligned} \tag{4}$$

In the case of only two interacting agents we have that

$$\begin{aligned} \dot{R}_a = & -\frac{\eta_{a,b} - 2}{2} (1 - R_a^2) + \lambda_o \sqrt{1 - R_a^2} + \\ & + \left[\frac{\eta_{a,b}}{2} \Theta(\bar{R}_b - R_a) \left(\sqrt{1 - R_a^2 \bar{R}_b} - \sqrt{1 - \bar{R}_b^2 R_a} \right) + P(t) \right] \sqrt{1 - R_a^2}. \end{aligned} \tag{5}$$

with

$$\frac{dV(R_a)}{dR_a} = \frac{\eta_{a,b} - 2}{2} (1 - R_a^2) - \lambda_o \sqrt{1 - R_a^2}, \tag{6}$$

and $P(t)$ is the continuous time version of the periodic perturbation representing the publicity campaign in favor of B , and $\bar{R}_b = R_b(t - \tau)$ where τ is the continuous-time delay. The equation for R_b is obtained by switching indexes a and b on Eq. (5).

According to [34] and given the constant $\kappa_o \approx 1.12282$, if the interaction $\eta_{a,b}$ is chosen to be:

$$\eta_{a,b} = 2(\kappa_o \lambda_o + 1), \tag{7}$$

the system described by Eq. (5) becomes bi-stable, with stable points at $R_a = 1$ and $R_a = -\sqrt{1 - \kappa_o^{-2}}$, i.e.

$$\left. \frac{dV}{dR_a} \right|_{R_a=1} = \left. \frac{dV}{dR_a} \right|_{R_a=-\sqrt{1-\kappa_o^{-2}}} = 0 \tag{8}$$

$$V(1) = V\left(-\sqrt{1 - \kappa_o^{-2}}\right). \tag{9}$$

$R_a = 1$ is the stable point at which the system converges if the agents, at the end of the opinion formation process, classify strings in agreement with B , and $R_a = -R_o = -\sqrt{1 - \kappa_o^{-2}}$ is

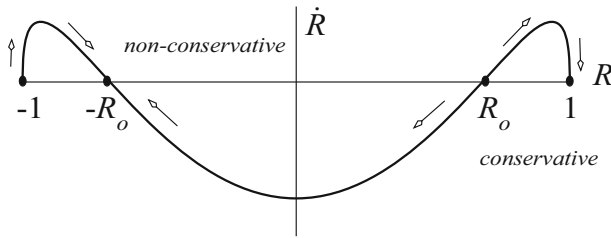


Fig. 1 Phase space representation for the system with potential (6), in the bi-stable regime determined by the conditions (8) and (9)

the stable point at which the system converges if the agents classify strings mostly differently to B . Due to these properties, we dubbed these points as the *conservative stable point* and the *non-conservative stable point* respectively. In consequence the conservative basin is defined as the interval $(R_o, 1)$, and the non-conservative basin is the interval $(-1, R_o)$. A sketch of the phase space of this system is presented in Fig. 1.

Let us define $|\mathbb{N}_a|$ as the cardinality of the set \mathbb{N}_a . By defining the population averages:

$$v \equiv \frac{1}{M} \sum_{a=1}^M |\mathbb{N}_a| \tag{10}$$

$$\eta \equiv \frac{1}{M(M-1)} \sum_{a \neq b=1}^M \eta_{a,b}, \tag{11}$$

and assuming that the fluctuations

$$\Delta v = \sqrt{\frac{1}{M} \sum_{a=1}^M |\mathbb{N}_a|^2 - v^2} \tag{12}$$

$$\Delta \eta = \sqrt{\frac{1}{M(M-1)} \sum_{a \neq b=1}^M \eta_{a,b}^2 - \eta^2}, \tag{13}$$

are sufficiently small, the evolution of the agreement R_a of agent a , in the neighborhood \mathbb{N}_a , and effective interaction $v\eta = (\kappa_o \lambda_o + 1)$ is given by:

$$\begin{aligned} \dot{R}_a = & -\frac{v\eta - 2}{2}(1 - R_a^2) + \lambda_o \sqrt{1 - R_a^2} \\ & + \left[\frac{\eta}{2} \sum_{b \in \mathbb{N}_a} \Theta(\bar{R}_b - R_a) \left(\sqrt{1 - R_a^2} \bar{R}_b - \sqrt{1 - \bar{R}_b^2} R_a \right) + P(t) \right] \sqrt{1 - R_a^2}. \end{aligned} \tag{14}$$

The objective of our investigation is to study the effects of a periodic perturbation on the opinion formation process in a community of interacting agents, when the system is in a bi-stable regime and the initial condition locates the agents in the basin of the stable non-conservative point $-R_o$, and there is a time delay τ in the communication between agents. For such a scenario we can translate the equation in overlaps (14) into an equation in phases: where $\theta_a(0) \in (\phi_o, \pi)$, $\phi_o \equiv \arccos(R_o) = 1.0987(1)$, and $P(t) \equiv \lambda_o A p(\omega t) \geq 0$ is a periodic perturbation with amplitude $\lambda_o A$ and frequency ω . By re-scaling the time $(1 + \lambda_o \kappa_o)t \rightarrow t$

and the frequency $(1 + \lambda_o \kappa_o)^{-1} \omega \rightarrow \omega$ we obtain:

$$\dot{\theta}_a = -\frac{1}{\nu} \sum_{c \in \mathbb{N}_a} \Theta(\theta_a - \bar{\theta}_c) \sin(\theta_a - \bar{\theta}_c) + \Lambda [\kappa_o \sin \theta_a - 1 - Ap(\omega t)], \tag{15}$$

where $\Lambda \equiv (1 + \lambda_o \kappa_o)^{-1} \lambda_o$. The first term of the right-hand-side of (15) is the average interaction over the neighborhood of the agent, the second term is a perturbation mainly proportional to the rate of change of B .

In the phase description given by Eq. (15) the conservative basin is defined by $(0, \phi_o)$ and the non-conservative basin by (ϕ_o, π) .

3 Solution of the System (15)

3.1 Solution for the Perturbed System with Two Agents

The two-agent problem is the simplest case of interacting agents learning from the rule B . It has been observed in [35] that the agent with the smallest phase dictates the behavior of the system, in particular it is remarked that, when all the agents are initially in the non-conservative basin, it is sufficient to convince the most conservative (the one with the smallest initial phase) to direct them all to the conservative basin. We analyze here the two-agent problem in order to use the behavior of the agent with the smallest phase to explain the expected behavior of larger societies.

Let us consider the problem of two agents perturbed with a constant perturbation (or zero-frequency periodic perturbation) with initial conditions such that $\theta_< = \min\{\theta_a(0), \theta_b(0)\} \in (\phi_o, \pi - \phi_o)$ and $\theta_> = \max\{\theta_a(0), \theta_b(0)\} \in (\theta_<, \pi)$. Without lose of generality we assume that $\theta_> = \theta_a(0)$ and $\theta_< = \theta_b(0)$. Then, for all $0 < t < \tau$ we have that the evolution of the system is ruled by the equations:

$$\dot{\theta}_a = \Lambda (\kappa_o \sin \theta_a - 1) - \Theta(\theta_a - \bar{\theta}_b) \sin[\theta_a(t) - \bar{\theta}_b] - \Lambda A \tag{16}$$

$$\dot{\theta}_b = \Lambda (\kappa_o \sin \theta_b - 1) - \Theta(\theta_b - \bar{\theta}_a) \sin[\theta_b(t) - \bar{\theta}_a] - \Lambda A, \tag{17}$$

which can be simplified into

$$\dot{\theta}_a = \Lambda (\kappa_o \sin \theta_a - 1) - \sin[\theta_a(t) - \theta_<] - \Lambda A \tag{18}$$

$$\dot{\theta}_b = \Lambda (\kappa_o \sin \theta_b - 1) - \Lambda A. \tag{19}$$

Let us define $\kappa_A \equiv (1 + A)^{-1} \kappa_o$, and $\phi_A \equiv \arcsin(\kappa_A^{-1})$. The numerical integration of (16) and (17) for values of the amplitude $A > \kappa_o - 1$ and initial condition $\theta_< \in (\phi_o, \pi)$ or for values of the amplitude $A < \kappa_o - 1$ and initial condition $\theta_< \in (\phi_o, \phi_A)$, produce curves with a pattern similar to the one presented in Fig. 2. Observe that the integration of Eqs. (16) and (17) is valid until the correspondent phase crosses over ϕ_o , entering the basin correspondent to the stable point 0, where the correct equations to be used are obtained from (14).

For values of the amplitude $A < \kappa_o - 1$ and $\theta_< \in (\pi - \phi_A, \pi)$ the general behavior of the numerical solution to (16) and (17) is represented in Fig. 3. In all cases the curves approach asymptotically $\pi - \phi_A$.

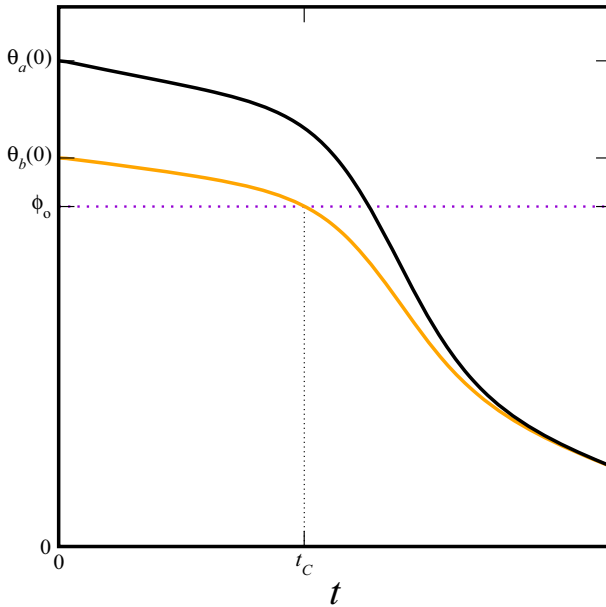


Fig. 2 Typical behavior of the solution of the numerical integration of the Eqs. (16) and (17) either for values of the amplitude $A > \kappa_o - 1$ and initial condition $\theta_{<} \in (\phi_o, \pi)$ or for values of the amplitude $A < \kappa_o - 1$ and initial condition $\theta_{<} \in (\phi_o, \phi_A)$ Both curves approach 0 asymptotically after crossing over ϕ_o . For the definition of the time t_c see Eqs. (23) and (25)

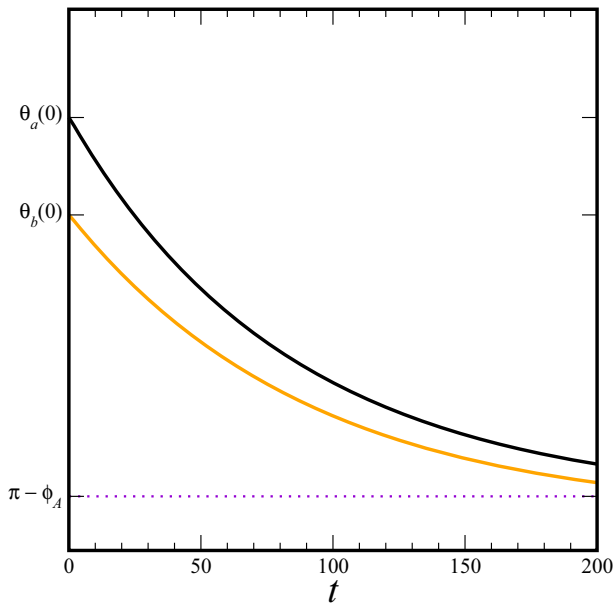


Fig. 3 Typical behavior of the solution of the numerical integration of the Eqs. (16) and (17) for values of the amplitude $A < \kappa_o - 1$ and initial condition $\theta_{<} \in (\pi - \phi_A, \pi)$ Both curves approach $\pi - \phi_A$ asymptotically

The integration of θ_b , Eq. (19), is straightforward:

$$\tan \frac{\theta_b}{2} = \frac{\tan \frac{\pi - \phi_A}{2} \left| \tan \frac{\phi_A}{2} - \tan \frac{\theta_{<}}{2} \right| e^{\Lambda_A t}}{\left| \tan \frac{\phi_A}{2} - \tan \frac{\theta_{<}}{2} \right| e^{\Lambda_A t} - \left| \tan \frac{\pi - \phi_A}{2} - \tan \frac{\theta_{<}}{2} \right|} \frac{\tan \frac{\phi_A}{2} \left| \tan \frac{\pi - \phi_A}{2} - \tan \frac{\theta_{<}}{2} \right|}{\left| \tan \frac{\phi_A}{2} - \tan \frac{\theta_{<}}{2} \right| e^{\Lambda_A t} - \left| \tan \frac{\pi - \phi_A}{2} - \tan \frac{\theta_{<}}{2} \right|} \quad (20)$$

if $\kappa_A > 1, \theta_{<} \notin (\phi_A, \pi - \phi_A)$,

$$\tan \frac{\theta_b}{2} = 1 + \frac{\tan \frac{\theta_{<}}{2} - 1}{1 + \frac{\Lambda \kappa_o t}{2} \left(\tan \frac{\theta_{<}}{2} - 1 \right)} \quad (21)$$

if $\kappa_A = 1$, and

$$\tan \frac{\theta_b}{2} = \kappa_A + \sqrt{1 - \kappa_A^2} \tan \left[\arctan \left(\frac{\tan \frac{\theta_{<}}{2} - \kappa_A}{\sqrt{1 - \kappa_A^2}} \right) - \frac{\Lambda_A t}{2} \right] \quad (22)$$

if $\kappa_A < 1$, where $\Lambda_A \equiv \Lambda \sqrt{|\kappa_o^2 - (1 + A)^2|}$. Observe that in (20), we have explicitly avoided the discussion of the initial condition $\theta_{<} \in (\phi_A, \pi - \phi_A)$. The results presented correspond to an asymptotic decaying behavior. In particular, the phase θ_b would cross over the conservative basin at time t_C , i.e. $\theta_b(t_C) = \phi_o$, in the following cases:

$$\Lambda_A t_C = 2 \ln \left[\frac{\left(\tan \frac{\phi_A}{2} - \tan \frac{\phi_o}{2} \right) \left(\tan \frac{\pi - \phi_A}{2} - \tan \frac{\theta_{<}}{2} \right)}{\left(\tan \frac{\phi_A}{2} - \tan \frac{\theta_{<}}{2} \right) \left(\tan \frac{\pi - \phi_A}{2} - \tan \frac{\phi_o}{2} \right)} \right], \quad (23)$$

if $\kappa_A > 1$ and $\phi_o < \theta_{<} < \phi_A$,

$$\Lambda \kappa_o t_C = 2 \left(\frac{1}{1 - \tan \frac{\theta_{<}}{2}} - \frac{1}{1 - \tan \frac{\phi_o}{2}} \right), \quad (24)$$

if $\kappa_A = 1$ and $\phi_o < \theta_{<} < \frac{\pi}{2}$, and

$$\Lambda_A t_C = 2 \left\{ \arctan \left(\frac{\tan \frac{\theta_{<}}{2} - \kappa_A}{\sqrt{1 - \kappa_A^2}} \right) - \arctan \left(\frac{\tan \frac{\phi_o}{2} - \kappa_A}{\sqrt{1 - \kappa_A^2}} \right) \right\}, \quad (25)$$

if $\kappa_A < 1$ and $\phi_o < \theta_{<} < \pi$.

The cross-over time t_C is finite and only present if the amplitude A and the initial condition $\theta_{<}$ are such that the agents, although initially in the non-conservative basin (ϕ_o, π) would cross over the conservative basin $(0, \phi_o)$ at times $t \geq t_C$. If the amplitude $A < \kappa_o - 1$

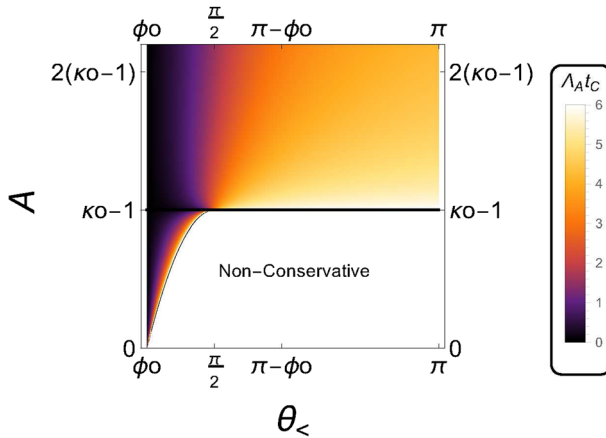


Fig. 4 The cross-over time t_C , given by Eqs. (23,25), as a function of the initial condition $\theta_<$ and the amplitude A . The darker the color the shorter it takes for the agent with the lower phase to reach the cross-over point ϕ_o . In the non-conservative region, defined for $A < \kappa_o - 1$ and $\theta_< > \phi_A$, the cross-over time t_C is not defined (Color figure online)

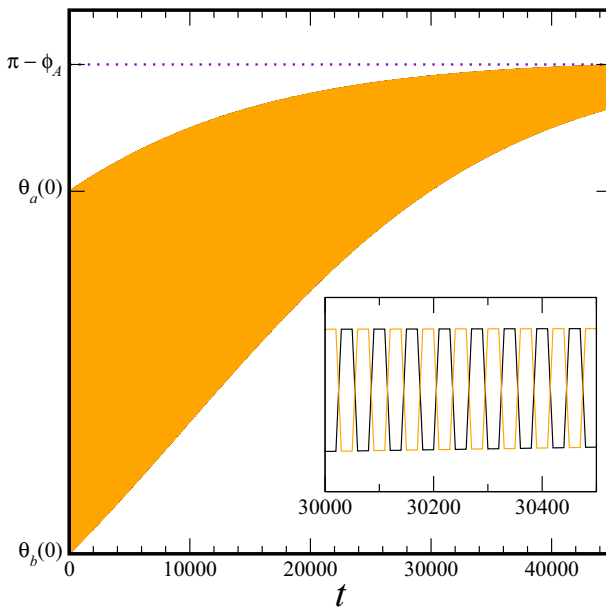


Fig. 5 Oscillatory behavior of the solutions of Eqs. (16) and (17) for values of $A < \kappa_o - 1$ and $\phi_A < \theta_< < \pi - \phi_A$. Both trajectories converge to $\pi - \phi_A$. In the inset we have a detail for just a few oscillation cycles

and initial condition $\theta_< > \phi_A$ the agent remain in the non-conservative basin. We present a density plot of t_C as a function of A and $\theta_<$ in Fig. 4.

The behavior of the phases that remain in the non-conservative basin is as follows: if $A < \kappa_o - 1$ and $\pi - \phi_A < \theta_<$ the phases converge to $\pi - \phi_A$ asymptotically. If $A < \kappa_o - 1$ and $\phi_A < \theta_< < \pi - \phi_A$ the phases converge to $\pi - \phi_A$ oscillatorily (see Fig. 5).

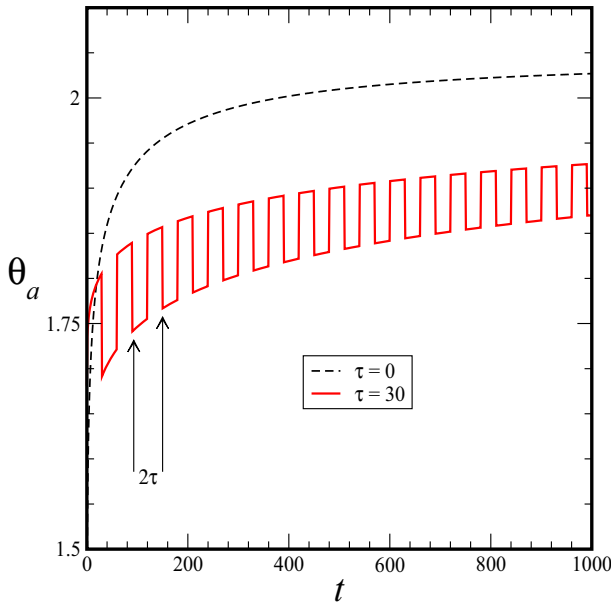


Fig. 6 Integration of Eqs. (16) and (17) for a system without delay (dashed curve) and for a system with delay $\tau = 30$. Both systems are unperturbed ($A = 0$) and with $\Lambda = 0.01$. Only the curves for θ_a are presented

The behavior in the area determined by $\kappa_A > 1$ and $\phi_A < \theta_< < \pi - \phi_A$ can be approximated, following an argument similar to the one presented in Appendix B, by:

$$\theta_a(\tau t) \approx x_A([t]; \theta_{a,0}) [1 - \text{mod}([t], 2)] + x_A([t]; \theta_{b,0}) \text{mod}([t], 2) \tag{26}$$

$$\theta_b(\tau t) \approx x_A([t]; \theta_{a,0}) \text{mod}([t], 2) + x_A([t]; \theta_{b,0}) [1 - \text{mod}([t], 2)], \tag{27}$$

where $[t] \in \mathbb{Z}$ and $[t] \leq t < [t] + 1$, and

$$\begin{aligned} \tan \frac{x_A(t; x_0)}{2} = & \frac{\tan \frac{\pi - \phi_A}{2} \left(\tan \frac{x_0}{2} - \tan \frac{\phi_A}{2} \right) e^{\Lambda_A t}}{\left(\tan \frac{x_0}{2} - \tan \frac{\phi_A}{2} \right) e^{\Lambda_A t} + \tan \frac{\pi - \phi_A}{2} - \tan \frac{x_0}{2}} \\ & + \frac{\tan \frac{\phi_A}{2} \left(\tan \frac{\pi - \phi_A}{2} - \tan \frac{x_0}{2} \right)}{\left(\tan \frac{x_0}{2} - \tan \frac{\phi_A}{2} \right) e^{\Lambda_A t} + \tan \frac{\pi - \phi_A}{2} - \tan \frac{x_0}{2}}. \end{aligned} \tag{28}$$

These solutions are similar to the ones presented in Fig. 9. Observe that for sufficiently long times:

$$x_A(t; x_0) \approx \pi - \phi_A - 2 \frac{1 - \tan \frac{\phi_A}{2}}{1 + \left(\tan \frac{\pi - \phi_A}{2} \right)^2} \frac{\tan \frac{\pi - \phi_A}{2} - \tan \frac{x_0}{2}}{\tan \frac{x_0}{2} - \tan \frac{\phi_A}{2}} e^{-\Lambda_A t}, \tag{29}$$

which indicates that the envelope (28) converges towards $\pi - \phi_A$.

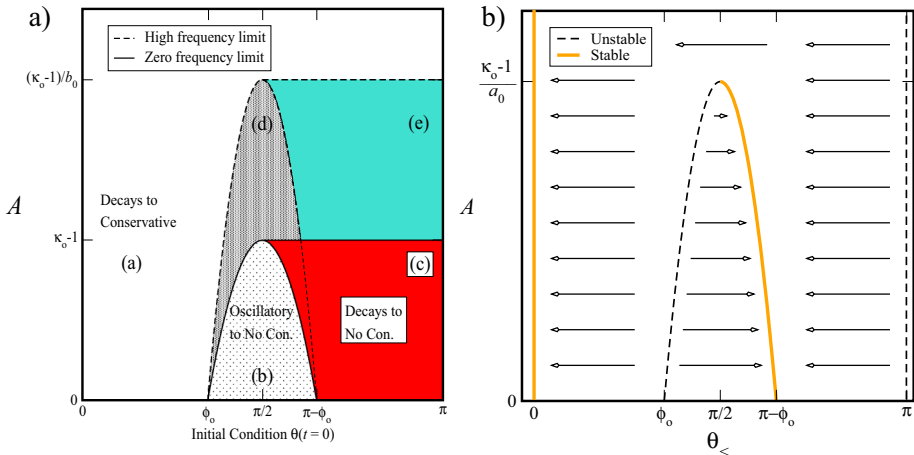


Fig. 7 (a) Phase diagram for the system described by Eqs. (35) and (36) in the plane $(\theta_<, A)$. There are three different behaviors in the system, depending on the values of the parameters $\theta_<$ and A . The system may decay towards the conservative fixed point (a), decay inside the non-conservative basin (e) and (c) or present an oscillatory decay towards the non-conservative fixed point (b) and (d). Region (b) corresponds to a perturbation $p(\omega t) = 1$ (or zero frequency limit). Region (d) corresponds to a perturbation with a very high frequency ($\omega \rightarrow \infty$). Observe that this region is obtain by scaling the vertical axis $A \rightarrow A/b_0$, where b_0 is the high-frequency limit of the average (33). Perturbations with finite values of ω would produce a phase diagram with an oscillatory behavior region in between regions (b) and (d). The line separating regions (a) and (b) is given by $\kappa_o \sin \theta_< - 1 - A = 0$, the one separating regions (b) and (c) is given by $\kappa_o \sin(\pi - \theta_<) - 1 - A = 0$. Equivalently, the line separating regions (a) and (d) is given by $\kappa_o \sin \theta_< - 1 - Ab_0 = 0$, the one separating regions (d) and (d) is given by $\kappa_o \sin(\pi - \theta_<) - 1 - Ab_0 = 0$. b) The fixed points of the system are represented by the dark dashed lines (unstable points) and by the grey (orange on line) thick lines (stable points). The arrows indicate the direction of flow, at constant A , for a given value of the initial condition $\theta_<$. The number $a_0 \in [b_0, 1]$ depends on the frequency of the perturbation (Color figure online)

It is important to note that, in the absence of delay and by applying the observation that the order of the phases is preserved, i.e. if $\theta_< = \theta_b(0) < \theta_a(0) = \theta_>$, then $\theta_b(t) < \theta_a(t)$ for all t , the phases of the undelayed system satisfy the following expressions $\theta_{a(b)}(t) = x_A(t; \theta_>(<))$. The fundamental difference between delayed and non-delayed systems for $0 \leq A \leq \kappa_o - 1$ and $\phi_A < \theta_< < \pi - \phi_A$ is that the relaxation time of the exponential decay in the delayed system is equal to the relaxation time of the undelayed system times τ .

To illustrate this change of behavior we present the result of the integration of Eqs. (16) and (17) in Fig. 6 for unperturbed ($A = 0$) systems. The dashed curve corresponds to a system without delay, the full curve corresponds to a system with a delay $\tau = 30$, with $\Lambda = 0.01$ in both cases.

3.2 General Periodic Perturbation

Consider the system of equations at $t < \tau$:

$$\dot{\theta}_a = \Lambda (\kappa_o \sin \theta_a - 1) - \sin[\theta_a(t) - \theta_<] - \Lambda A p(\omega t) \tag{30}$$

$$\dot{\theta}_b = \Lambda (\kappa_o \sin \theta_b - 1) - \Lambda A p(\omega t), \tag{31}$$

with $\theta_a(0) = \theta_> \in (\phi_o, \pi)$, $\theta_b(0) = \theta_< \in (\phi_o, \pi - \phi_o)$ and where $p(\omega t)$ is a periodic function with period $2\pi/\omega$, satisfying that $0 \leq p(\omega t) \leq 1$ for all t . Let us suppose that we

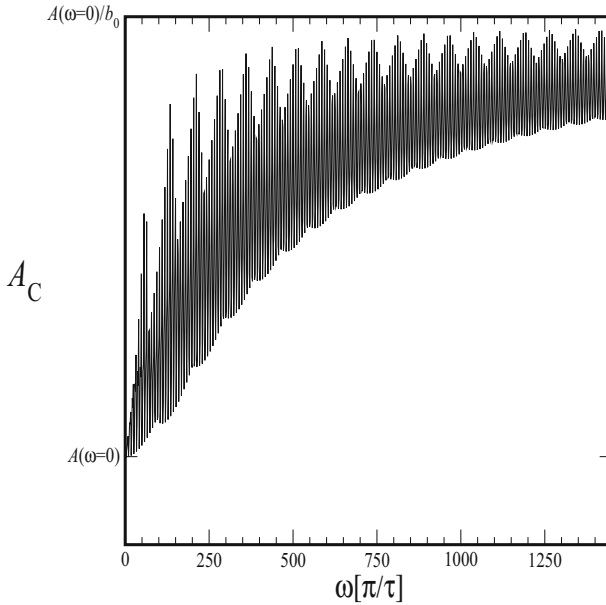


Fig. 8 Critical amplitude A_C as a function of the frequency ω (in units of π/τ) for the perturbation $p(\omega t) = \frac{1}{2} - \frac{1}{2} \cos(\omega t)$

can expand the perturbation in the following way:

$$p(\omega t) = \sum_{\ell=0}^{\infty} b_{\ell} \cos(\ell \omega t), \tag{32}$$

therefore we can compute the average perturbation as:

$$\bar{p}(\omega) \equiv \frac{1}{\tau} \int_0^{\tau} ds p(s) = b_0 + \sum_{\ell=1}^{\infty} b_{\ell} \text{sinc}(\ell \omega \tau), \tag{33}$$

where

$$\text{sinc}(x) \equiv \frac{\sin(x)}{x}. \tag{34}$$

$\bar{p}(\omega)$ approaches b_0 for sufficiently large values of the frequency ω . Therefore, we can obtain a crude estimate of the agent’s behavior by approximating the system formed by Eqs. (30) and (31) by:

$$\dot{\theta}_a = \Lambda (\kappa_o \sin \theta_a - 1) - \sin[\theta_a(t) - \theta_{<}] - \Lambda A \bar{p}(\omega) \tag{35}$$

$$\dot{\theta}_b = \Lambda (\kappa_o \sin \theta_b - 1) - \Lambda A \bar{p}(\omega). \tag{36}$$

We observe that the phase diagram of the system in the plane $(\theta_{<}, A)$, where $\theta_{<}$ is the initial condition of the initially smallest phase and A is the amplitude of the perturbation, is presented in Fig. 7a). There are three different behaviors in the system, depending on the values of the parameters $\theta_{<}$ and A . The system may decay towards the conservative fixed point $\theta = 0$ (a), or decay to the non-conservative point $\theta = \pi - \phi_A$ in a monotonous way, (e) and (c), or oscillatory way, (b) and (d). Region (b) corresponds to a perturbation $p(\omega t) = 1$

(or zero frequency limit). Region (d) corresponds to a perturbation with a very high frequency ($\omega \rightarrow \infty$). Observe that this region is obtain by scaling the vertical axis $A \rightarrow A/b_0$, where b_0 is the high-frequency limit of the average (33). Perturbations with finite values of ω would produce a phase diagram with an oscillatory behavior region in between regions (b) and (d).

The line separating regions (a) and (b) is given by $\kappa_o \sin \theta_0 - 1 - A = 0$, the one separating regions (b) and (c) is given by $\kappa_o \sin(\pi - \theta_<) - 1 - A = 0$. Equivalently, the line separating regions (a) and (d) is given by $\kappa_o \sin \theta_< - 1 - Ab_0 = 0$, the one separating regions (d) and (e) is given by $\kappa_o \sin(\pi - \theta_0) - 1 - Ab_0 = 0$. In Fig. 7b) we present the distribution of unstable fixed points (dashed lines) and stable points (full line) in a system with a perturbation of frequency ω , that fixes the value of the parameter $a_0 \in (0, b_0)$. The arrows indicate the direction of evolution of the system at constant A , from $\theta(0) = \theta_<$ to $\theta(\infty) = 0$ or $\pi - \phi_A$.

Observe that the previous analysis is based on the values of the effective amplitude of the perturbation ($A\bar{p}(\omega)$) and the initial position of the smallest phase $\theta_<$. Following [35] we hypothesize that the phase space presented in Fig. 7 is also valid for a system with $M > 2$ interacting agents.

It is also important to note that, in the analysis presented in [35], the parameter consider is the *expected* value of the smallest initial phase, which is a parameter drawn from a distribution that depends on M . In that scenario, the larger the number of agents M the closer the initial smallest phase to the instable point ϕ_o , and therefore the smallest the minimal amplitude of the perturbation A to convince the most conservative agent to cross over the boundary between basins. In the present scenario, the parameter considered is the smallest initial phase $\theta_<$ itself. In this manner, the analysis we present here corresponds to all possible values of the parameters $\theta_<$ and A , not the most likely ones. Finally, the approach presented in [35] considers a smooth, periodic perturbation $\nu(t)$, with at least two continuous derivatives and satisfying the constraints $\nu(0) = \dot{\nu}(0) = 0$. In the present approach we cannot consider such constraints because would render meaningless the analysis of the constant perturbation ($\omega = 0$).

3.3 Study of the Oscillatory Phase Under a General Periodic Perturbation

In this section we focus on the solution to the system of differential Eqs. (30) and (31) with an initial condition such that $\phi_A < \theta_< < \pi - \phi_A$. In this case we have an oscillatory decay towards an asymptotic non-conservative behavior. In particular we want to explore the dependency of the critical amplitude (A_C) on the frequency of the perturbation ω . The critical amplitude is the value of A over the line separating the regions (a) and (d) in Fig. 7.

With this end we propose a perturbative solution to (30) and (31) of the form $\theta_{a(b)}(t) = \theta_{a(b)}(t; A = 0) + A\delta_{a,b}(t)$, where the $\theta_{a(b)}(t; A = 0)$ are given by (26) and (27). The equations for the $\delta_{a(b)}$ are

$$\dot{\delta}_{a(b)} = - [1 - \Lambda\kappa_0 \cos \theta_{a(b)}] \delta_{a(b)} + \bar{\delta}_{b(a)} - \Lambda p(\omega t), \tag{37}$$

and their solution can be approximated by the expression:

$$\delta_{a(b)}(t) \approx -\Lambda(n + 1) \sum_{\ell=0}^{\infty} b_{\ell} \left| \frac{\text{sinc} \left[\frac{n+1}{2} (\ell\omega\tau + \phi_{\ell}) \right]}{\text{sinc} \left[\frac{1}{2} (\ell\omega\tau + \phi_{\ell}) \right]} \right| \cos(\ell\omega t - \Psi_{\ell,n}), \tag{38}$$

$$\Psi_{\ell,n} \equiv \phi_{\ell} + \frac{n}{2} (\ell\omega\tau + \phi_{\ell}) \tag{39}$$

where $\phi_{\ell} \equiv \arctan(\ell\omega)$ and $n \equiv \lfloor t/\tau \rfloor$. Observe that $\delta_{a(b)}$ is identical for all agents (independent of the index $a(b)$) and it presents a strong correlation with modes associated with

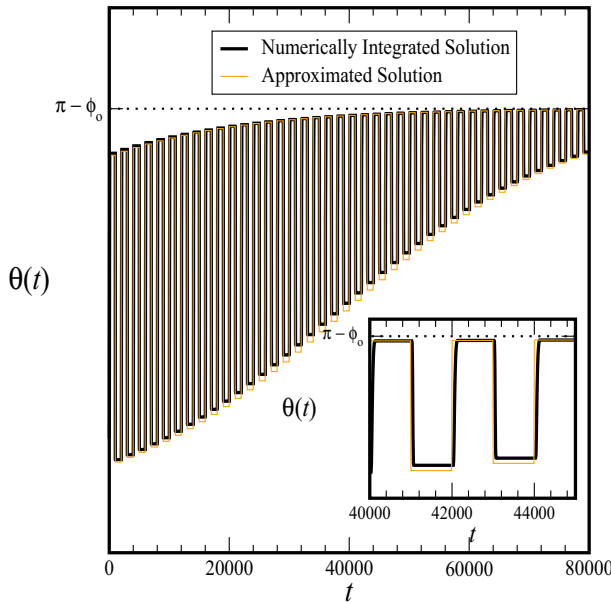


Fig. 9 Numerical solution to the Eq. (79) and approximation (91) against time, for initial conditions $\theta_a(0) = \pi - \phi_o - 0.1$ and $\theta_b(0) = \phi_o + 0.1$, for $t \in [0, 80000]$. The parameters of the system were set to $\Lambda = 0.1$ and $\tau = 1000$. In the inset we present a detail of the same curves for times $t \in [40000, 40000 + 5\tau]$

the natural frequency of the system ($\Omega = \pi/\tau$). To illustrate this point we present a plot of $A_C(\omega)$ for the perturbation $p(\omega t) = \frac{1}{2} - \frac{1}{2} \cos(\omega t)$ in Fig. 8.

4 Discussion

We analyzed the effects of a perturbation on the opinion formation process in a population of adaptive agents. The scenario considered is anisotropic since the agents interact with a set of rules B that determine the socially acceptable position. It was also assumed that the perception an agent had on a peer’s position about a social issue may not be up to date, hence a time delay τ was considered. A periodic perturbation was added to the model to mimic the action of a publicity campaign in pro of the official position represented by B .

In the microscopic model we chose to represent the internal state of the agents by a perceptron that is trained with a Hebbian algorithm (40). The rule B was also represented by a perceptron that learns from the average public opinion (44). By applying statistical mechanics techniques and the Markov approximation we obtained the large system-size limit of these equations. The dynamical system ruled by the differential equations so obtained presents two stable points, one that represents the conservative position agents express when mostly follow B , and other, dubbed non-conservative, that represents the attitude of agents following their peers.

By setting the system parameters in such a way that the conservative and non-conservative fixed points are equally likely (bi-stable regime), we studied the time evolution of the system of delayed differential equations (15) representing the process of opinion formation, under the action of a small periodic perturbation representing an official publicity campaign. The

initial conditions set for most of this study locate the agents in the basin of attraction of the non-conservative point.

We found out that the perturbed system presents three different behaviors, depending on the values of the initial condition of the smallest phase $\theta_<$, and the value of the perturbation's amplitude A . For small values of $\theta_<$ and sufficiently large A the system decays towards the conservative fixed point, for sufficiently large $\theta_<$ and sufficiently small A the system decays towards the non-conservative fixed point, and for intermediate values of $\theta_<$ and sufficiently low values of A the system presents damped oscillations that decay towards the non-conservative point. The time the system needs to cross-over to the conservative basin, Eqs. (23–25) is presented in Fig. 4. The emergence of this oscillatory behavior for $A < \kappa_o - 1$ and $\phi_A < \theta_< < \pi - \phi_A$ is due to the presence of the delay in the exchange of information between peers. The presence of this delay increases the relaxation time towards the fixed point $\pi - \phi_A$ by a factor of τ .

The frequency of the perturbation produces quantitative changes in the phase diagram of the system, without altering the qualitative distribution of the regions (see Fig. 7). We also observe that the perturbations consider in this approach are different from the periodic perturbation $v(t)$ considered in [35], which satisfy the constraints $v(0) = \dot{v}(0) = 0$. Such constraints were imposed to permit the relaxation at short times in the low frequencies regime. In our present approach we make use of the constant perturbation as the limit of a very low frequency. Is the perturbation is initially zero, it should remain so in the limit of the frequency going to zero, rendering our analysis meaningless.

The asymptotic behavior observed on the non-conservative phases can be represented by a monotonous decay towards $\pi - \phi_A$ for $\kappa_A > 1$ and $\theta_< > \pi - \phi_A$ (region (c)-(e)) and by an oscillatory decay towards $\pi - \phi_A$ for $\kappa_A > 1$ and $\phi_A < \theta_< < \pi - \phi_A$ (region (b)-(d)), thus, the transition from phase (b)-(d) to (c)-(e) can be characterized by the value of the imaginary part of the exponent, responsible for the oscillations of period 2τ observed in region (b)-(d). This is not a phase transition in the sense that there is a non-analyticity in the free energy of the system (which has not been defined) but a change in the asymptotic behavior of the system.

The role played by the parameter Λ , which has absorbed the parameter λ_o that controls the update of the rule B through the wisdom-of-the-crowd effect, is to provide a suitable scale to the relaxation and oscillation processes studied [Eqs. (23–25), (26) and (27)], without changing the asymptotic values reached by the average agreement R_a . If no update rule on B is considered ($\lambda_o = 0$), the four fixed points of the potential (6) get reduced to two (see Ref. [32]) and the complexity of the dynamical processes analyzed is lost.

Interactions among spatially distributed agents are subjected to a time delay. Any physical interaction propagates at a finite speed. The most salient effect of the time delay in the agent-agent interaction is observed in the phase-space region where $\phi_A < \theta_< < \pi - \phi_A$. In this region the system oscillatory decays towards the non-conservative point. As it is shown in Fig. 6, the presence of the delay retards the decay towards the correspondent fixed point. Similar effects have been observed in [46].

Our model, in its current form, is related to the scenario presented in [57], where an analysis on an anti-drug campaign focussed on adolescents is presented. In our model B represents the government, the agents are the school attending teenagers, and the issues opinions are formed over is the use of drugs. It was found that exposure to social interaction about campaign messages can affect behavior, and such interactions are favored if the intensity of the campaign is not perceived as oppressive (low A in our model).

Finally, we presented a short analysis of the perturbative solution to the system with a perturbation $p(\omega t) = \frac{1}{2}[1 - \cos(\omega t)]$. The value of the critical amplitude A_C has been

numerically computed as a function of ω an plot in Fig. 8. We observe that $A_C(\omega) \leq \lim_{\omega \rightarrow \infty} A_C(\omega) = A_C(0)/b_0$, thus partially validating the phase diagram presented in Fig. 7.

Acknowledgements The author would like to acknowledge the constructive discussions with Dr. O D’Huys. The advise of Dr. C. M Juarez is kindly appreciated.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A: Update Equations for a dimer

Assuming that the population of interacting agents receives information taken from the set $\mathbb{S} \equiv \{(\sigma_{B,n}, \sigma_{\mathbb{N}_a,n-m}, \mathbf{S}_n)\}_{n=1}^T$, where the string \mathbf{S}_n is presented at time n and then discarded, $\sigma_{B,n} = \text{sgn}(\mathbf{B}_n \cdot \mathbf{S}_n)$ and $\sigma_{\mathbb{N}_a,n-m} = \{\sigma_{c,n-m} : c \in \mathbb{N}_a\}$ is the set form by the classifications given by agent a ’s neighbors at time $n - m$, the update equation for the internal representation of a is:

$$\mathbf{J}_{a,n+1} = \mathbf{J}_{a,n} + \psi_{a,n} \frac{\sigma_{B,n} \mathbf{S}_n}{\sqrt{N}}, \tag{40}$$

where $\sigma_B \mathbf{S} / \sqrt{N}$ is the (unit length) Hebb vector [58], that indicates the direction of the socially acceptable position on \mathbf{S}_n and $\psi_{a,n}$ is the learning amplitude, that regulates how the information is incorporated in the internal representation of a . The length of the opinion formation process T is considered to be proportional to the number of issues presented to the agents. We propose

$$\psi_{a,n} \equiv f_n \frac{|\mathbf{J}_{a,n}|}{\sqrt{N}} \Psi_{a,n} \tag{41}$$

where f_n is a decaying function of n that implements the annealing of the learning process [59], $|\mathbf{J}_a| / \sqrt{N} = \sqrt{\sum_{j=1}^N J_{a,j}^2} / N$ is a factor that has no impact on the learning efficiency of the algorithm [60] and it has been only considered for technical purposes, and:

$$\Psi_{a,n,\ell} \equiv 1 - \Theta(-\sigma_{B,n} \sigma_{a,n}) \sum_{c \in \mathbb{N}_a} \eta_{a,c} \Theta(\sigma_{a,n} \sigma_{c,n-m}), \tag{42}$$

where $\Theta(x) = 1$ if $x > 0$ and 0 otherwise is the Heaviside step function, and $m = n - \ell$ is the delay in the information received by a from the neighbors $c \in \mathbb{N}_a$. This time delay accounts for the lack of update on the knowledge a has on the opinion of its neighbor c about current affairs.

The learning algorithm (42) works in the following way: In the absence of any inter-agent interaction the agents $\{a\}$ learn from the social rule B only, and their internal representations $\{\mathbf{J}_a\}$ grow in the direction of \mathbf{B} . This accounts for the first term of the right-hand-side of Eq. (42). But disagreement between agent a and social rule B may arise, i.e. $\Theta(-\sigma_a(\mathbf{S})\sigma_B(\mathbf{S})) > 0$. There is a cost for disagreement, as it is documented in Ref. [61], thus the agent a checks

with their *closest* peers in \mathbb{N}_a (where *close* refers to a criterion extracted from social impact theory [22, 23]), whether $\sigma_a(\mathbf{S}) = \sigma_c(\mathbf{S})$ or not. The corroboration [55, 56] is done with delayed information, corresponding to the state of the neighbors in \mathbb{N}_a , m time units in the past. If the integrated contribution of agreeing neighbors is sufficiently large (second contribution at the right-hand side of (42)), Ψ_a becomes negative and the internal representation vector \mathbf{J}_a grows opposite to \mathbf{B} .

Let us define the unit vectors $\mathbf{b} \equiv |\mathbf{B}|^{-1}\mathbf{B}$ in the direction of the internal representation of B , $\mathbf{j}_a \equiv |\mathbf{J}_a|^{-1}\mathbf{J}_a$ in the direction of the internal representation of agent a and $\mathbf{j}_{a,\perp} \equiv [1 - (\mathbf{j}_a \cdot \mathbf{b})^2]^{-1/2}[\mathbf{j}_a - (\mathbf{j}_a \cdot \mathbf{b})\mathbf{b}]$ in the direction of the component of \mathbf{J}_a perpendicular to \mathbf{B} . Given that an agent's classification is obtained through information processing using the internal representation vector \mathbf{J}_a , and that any modification to the vector \mathbf{B} in the direction of \mathbf{B} does not produce any change on B 's classifications, we will construct the update algorithm for B by considering the vector:

$$\mathbf{L} \equiv \frac{1}{M} \sum_{c=1}^M \mathbf{j}_{c,\perp}. \tag{43}$$

which is the arithmetic average over all the components of the internal representations \mathbf{J}_c perpendicular to \mathbf{B} . Observe that $\mathbf{B} \cdot \mathbf{L} = 0$ and $|\mathbf{L} \cdot \mathbf{L}| \sim O(M^{-1})$. Then:

$$\mathbf{B}_{n+1} = \mathbf{B}_n + \frac{\lambda_o}{\sqrt{N}} f_n \mathbf{L}_n, \tag{44}$$

where λ_o/\sqrt{N} is a suitable scale factor. Observe that if $\lambda_o \sim O(1)$ the updates of \mathbf{B} at each time step are very small, thus λ_o/\sqrt{N} is a measure of the inverse inertia (if the *mass* of B is infinite we wouldn't expect any change at all). Observe also that $|\mathbf{B}_{n+1}|^2 = |\mathbf{B}_n|^2 + O(f_n^2 N^{-1})$, which implies that the length of the vector \mathbf{B} does not change with the update.

To help describe the state of the system we define the variables:

$$\phi_{a,n,n-\Delta n} \equiv \sigma_{B,n} \frac{\mathbf{J}_{a,n-\Delta n} \cdot \mathbf{S}_n}{|\mathbf{J}_{a,n-\Delta n}|} \tag{45}$$

$$\beta_n \equiv \sigma_{B_n} \frac{\mathbf{B}_n \cdot \mathbf{S}_n}{|\mathbf{B}_n|} \tag{46}$$

and parameters:

$$R_{a,n} \equiv \frac{\mathbf{J}_{a,n} \cdot \mathbf{B}_n}{|\mathbf{J}_{a,n}| |\mathbf{B}_n|} \tag{47}$$

$$W_{a,n;b,n-\Delta n} \equiv \frac{\mathbf{J}_{a,n} \cdot \mathbf{J}_{b,n-\Delta n}}{|\mathbf{J}_{a,n}| |\mathbf{J}_{b,n-\Delta n}|} \tag{48}$$

$$Y_{a,n;b,n-\Delta n} \equiv \frac{\mathbf{J}_{a,n,\perp} \cdot \mathbf{J}_{b,n-\Delta n,\perp}}{|\mathbf{J}_{a,n,\perp}| |\mathbf{J}_{b,n-\Delta n,\perp}|}. \tag{49}$$

The variables depend explicitly on the information $\{\sigma_{B,n}, \mathbf{S}_n\}$ whereas the parameters depend on the internal representations $\{\{\mathbf{J}_{a,n}\}, \mathbf{B}_n\}$ only. The variable β_n is non-negative and the smaller the $\beta_n(\mathbf{S})$ the higher the likelihood of \mathbf{S} to be in the classification boundary (given by $\mathbf{B} \cdot \mathbf{S} = 0$). The variable $\phi_{a,n,n}(\mathbf{S})$ indicates how much the vector $\mathbf{J}_{a,n}$ has to be modified to agree with \mathbf{B}_n . The parameter R_a represents the level of agreement of agent a with the social rule B , $W_{a,n;b,\ell}$ represents the level of agreement between agents a (in the current state) and b ($n - \ell$ time-steps in the past) and the parameter $Y_{a,n;b,\ell}$ represents the level of agreement

between the current agent a and the past agent b on strings \mathbf{S}_n laying on the current classification boundary, $\mathbf{B}_n \cdot \mathbf{S}_n = 0$. Given that $W_{a,n;b,\ell} = R_{a,n}R_{b,\ell} + Y_{a,n;b,\ell}\sqrt{(1 - R_{a,n}^2)(1 - R_{b,\ell}^2)}$ we only need to know $\{R_{a,n}\}$ and $\{Y_{a,n;b,\ell}\}$ to know the state of the system.

The data accessible to the agent a is $(\sigma_{B,n}, \phi_{a,n}, \phi_{b,\ell}, \mathbf{S}_n)$. The length of the training set is $T = \alpha_{\max}N$, which implies that $\alpha_{\max} = T/N$. For a given number $1 \leq n < N$ of examples presented to the perceptrons there is a number $0 < \alpha < \alpha_{\max}$ such that $n = \alpha(n)N$. Observe that, given that the minimum increment in the number of examples presented is 1, $\Delta\alpha(n) \equiv \alpha(n + 1) - \alpha(n) = 1/N$. By defining $\Delta t \equiv f_n\Delta\alpha = f_n/N$ and by using the update rules (40) and (44), we have that the equation for the evolution of the parameters are:

$$\frac{\Delta R_{a,n}}{\Delta t} = \psi_{a,n,\ell} (\beta_n - \phi_{a,n,n}R_{a,n}) + \lambda_o\sqrt{1 - R_{a,n}^2}\frac{1 + Y_{a,n;b,\ell}}{2} + O(\Delta t) \tag{50}$$

and

$$\begin{aligned} \frac{\Delta Y_{a,n;b,\ell}}{\Delta t} = & \psi_{a,n,\ell} \left[\frac{\phi_{b,n,\ell} - R_{b,\ell}\beta_n}{\sqrt{(1 - R_{a,n}^2)(1 - R_{b,\ell}^2)}} - Y_{a,n;b,\ell} \frac{\phi_{a,n,n} - R_{a,n}\beta_n}{1 - R_{a,n}^2} \right] \\ & - \frac{\lambda_o R_{a,n}}{\sqrt{1 - R_{a,n}^2}} \left(\frac{1 - Y_{a,n;b,\ell}^2}{2} \right) \\ & + \psi_{b,\ell,\ell-m} \left[\frac{\phi_{a,\ell,n} - R_{a,n}\beta_\ell}{\sqrt{(1 - R_{a,n}^2)(1 - R_{b,\ell}^2)}} - Y_{a,n;b,\ell} \frac{\phi_{b,\ell,\ell} - R_{b,\ell}\beta_\ell}{1 - R_{b,\ell}^2} \right] \\ & - \frac{\lambda_o R_{b,\ell}}{\sqrt{1 - R_{b,\ell}^2}} \left(\frac{1 - Y_{a,n;b,\ell}^2}{2} \right) + O(\Delta t). \end{aligned} \tag{51}$$

Following [62], these parameters can be prove to be self-averaging in the large N limit. Averaging over the variables $\phi_{a,n,\ell}$, $\phi_{b,\ell,\ell-m}$ and β_n , we have the differential equations

$$\begin{aligned} \dot{R}_a = & \left\langle \psi_a(t, t - \tau) (\beta(t) - \phi_a(t, t)R_a(t)) + \lambda_o\sqrt{1 - R_a^2(t)}\frac{1 + Y_{a,b}(t, t - \tau)}{2} \right\rangle \tag{52} \\ \dot{Y}_{a,b} = & \left\langle \psi_a(t, t - \tau) \left[\frac{\phi_b(t, t - \tau) - R_b(t - \tau)\beta(t)}{\sqrt{(1 - R_a^2(t))(1 - R_b^2(t - \tau))}} \right. \right. \\ & \left. \left. - Y_{a,b}(t, t - \tau) \frac{\phi_a(t, t) - R_a(t)\beta(t)}{1 - R_a^2(t)} \right] \right\rangle \\ & + \left\langle \psi_b(t - \tau, t - 2\tau) \left[\frac{\phi_a(t - \tau, t) - R_a(t)\beta(t - \tau)}{\sqrt{(1 - R_a^2(t))(1 - R_b^2(t - \tau))}} \right. \right. \\ & \left. \left. - Y_{a,b}(t, t - \tau) \frac{\phi_b(t - \tau, t - \tau) - R_b(t - \tau)\beta(t - \tau)}{1 - R_b^2(t - \tau)} \right] \right\rangle \\ & - \lambda_o \left\langle \left[\frac{R_a(t)}{\sqrt{1 - R_a^2(t)}} + \frac{R_b(t - \tau)}{\sqrt{1 - R_b^2(t - \tau)}} \right] \left(\frac{1 - Y_{a,b}^2(t, t - \tau)}{2} \right) \right\rangle. \end{aligned} \tag{53}$$

The averages in (52) and the first term of (53) are over $\beta(t)$, $\phi_a(t, t)$ and $\phi_b(t, t - \tau)$; the averages on the remaining terms of (53) are over $\phi_a(t - \tau, t)$, $\beta(t - \tau)$, $\phi_b(t - \tau, t - \tau)$ and $\phi_a(t - \tau, t - 2\tau)$. For the variables $\phi(t_1, t_2)$, the first argument t_1 corresponds to the time at which information is presented, and t_2 is the time of the last update of the internal representation of the perceptron. By imposing a Markov approximation [63] in the distribution of probabilities, i.e.:

$$\mathcal{P}(\phi_a(t_1, t_2) | \{\beta, \phi_a, \phi_b\}(t)) \approx \mathcal{P}(\phi_a(t_1, t_2) | \phi_b(t_1, t_2 - \tau), \beta(t_1 - \tau)) \tag{54}$$

$$\mathcal{P}(\beta(t_1) | \{\beta, \phi_a, \phi_b\}(t)) \approx \mathcal{P}(\beta(t_1) | \phi_b(t_1, t_2), \phi_a(t_1, t_2 - \tau)), \tag{55}$$

where $\{\beta, \phi_a, \phi_b\}(t)$ represents the dependency on all the variables (β , ϕ_a , and ϕ_b evaluated at times $t' < \max\{t_1, t_2\}$), and by defining the quantities:

$$\mathcal{L}(\phi_c, \bar{\phi}_d) \equiv \frac{[R_c - \bar{W}_{c,d}\bar{R}_d]\phi_c + [\bar{R}_d - \bar{W}_{c,d}R_c]\bar{\phi}_d}{\bar{\Omega}_{c,d}\sqrt{1 - \bar{W}_{c,d}^2}} \tag{56}$$

$$\bar{\Omega}_{c,d}^2 \equiv [1 - R_c^2][1 - \bar{R}_d^2][1 - \bar{Y}_{c,d}^2] \tag{57}$$

$$\bar{\Sigma}_{c,d}^2 \equiv \frac{\bar{\Omega}_{c,d}^2}{1 - \bar{W}_{c,d}^2}, \tag{58}$$

where the over-line indicates a quantity evaluated at $t - \tau$, we propose (following the results obtained for $\tau = 0$ [32]) the following approximation for the probabilities $\mathcal{P}(\phi_a | \bar{\beta}, \bar{\phi}_b, \bar{\phi}_a) \approx \mathcal{P}(\phi_a | \bar{\beta}, \bar{\phi}_b) \approx \mathcal{N}(\phi_a | \bar{\mu}, \bar{\sigma}^2)$, with

$$\bar{\mu} \equiv \frac{[R_a - \bar{R}_b\bar{W}_{a,b}]\bar{\beta} + [\bar{W}_{a,b} - R_a\bar{R}_b]\bar{\phi}_b}{1 - \bar{R}_b^2} \tag{59}$$

$$\bar{\sigma}^2 \equiv \frac{\bar{\Omega}_{a,b}^2}{1 - \bar{R}_b^2} \tag{60}$$

and the rest of the probabilities are derived from the distribution of \mathbf{S} , $\mathcal{P}(\mathbf{S}) = 2^{-N} \prod_{j=1}^N (\delta_{S_j,1} + \delta_{S_j,-1})$, i.e.

$$\mathcal{P}(\beta | \phi_b, \bar{\phi}_a) = \frac{\mathcal{N}\left(\beta \mid \bar{\Sigma}_{b,a}\mathcal{L}(\phi_b, \bar{\phi}_a); \bar{\Sigma}_{b,a}^2\right)}{2\mathcal{H}(-\mathcal{L}(\phi_b, \bar{\phi}_a))} \tag{61}$$

$$\mathcal{P}(\phi_b | \bar{\phi}_a) = \frac{\mathcal{H}(-\mathcal{L}(\phi_b, \bar{\phi}_a))\mathcal{N}\left(\phi_b \mid \bar{W}_{b,a}\bar{\phi}_a; 1 - \bar{W}_{b,a}^2\right)}{\mathcal{H}\left(-\frac{\bar{R}_a\bar{\phi}_a}{\sqrt{1 - \bar{R}_a^2}}\right)}, \tag{62}$$

$$\mathcal{P}(\bar{\phi}_a) = 2\mathcal{H}\left(-\frac{\bar{R}_a\bar{\phi}_a}{\sqrt{1 - \bar{R}_a^2}}\right)\mathcal{N}(\bar{\phi}_a), \tag{63}$$

where (63) is the marginal probability once all other variables evaluated at times shorter than $t - \tau$ have been integrated, and $\mathcal{N}(x | \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp[-(x - \mu)^2/2\sigma^2]$,

$\mathcal{H}(x) = \int_x^\infty dy \mathcal{N}(y)$ and $\mathcal{N}(y) = \mathcal{N}(y|, 0, 1)$. Thus, the conditional averages are:

$$\langle \phi_a | \bar{\beta}, \bar{\phi}_b \rangle = \frac{[R_a - \bar{R}_b \bar{W}_{a,b}] \bar{\beta} + [\bar{W}_{a,b} - R_a \bar{R}_b] \bar{\phi}_b}{1 - \bar{R}_b^2} \tag{64}$$

$$\langle \beta | \phi_b, \bar{\phi}_a \rangle = \bar{\Sigma}_{b,a} [\mathcal{F}(\mathcal{L}(\phi_b, \bar{\phi}_a)) + \mathcal{L}(\phi_b, \bar{\phi}_a)], \tag{65}$$

where $\mathcal{F}(x) \equiv \mathcal{N}(x)/\mathcal{H}(-x)$, and the complete averages are:

$$\langle \phi_a \rangle = \sqrt{\frac{2}{\pi}} R_a \quad \forall t \tag{66}$$

$$\langle \phi_b \rangle = \sqrt{\frac{2}{\pi}} R_b \quad \forall t \tag{67}$$

$$\langle \beta \rangle = \sqrt{\frac{2}{\pi}} \quad \forall t \tag{68}$$

$$\langle \Theta(-\phi_b) \Theta(-\bar{\phi}_a) \bar{\phi}_a \rangle = \sqrt{\frac{2}{\pi}} \left(\bar{R}_a \frac{\arccos(-\bar{Y}_{b,a})}{2\pi} - \frac{1 + \bar{W}_{b,a}}{2} \bar{\rho}_{b,a} \right) \tag{69}$$

$$\langle \Theta(-\phi_b) \Theta(-\bar{\phi}_a) \phi_b \rangle = \sqrt{\frac{2}{\pi}} \left(R_b \frac{\arccos(-\bar{Y}_{b,a})}{2\pi} - \frac{1 + \bar{W}_{b,a}}{2} \bar{\rho}_{b,a} \right) \tag{70}$$

$$\langle \Theta(-\phi_b) \Theta(-\bar{\phi}_a) \beta \rangle = \sqrt{\frac{2}{\pi}} \left(\frac{\arccos(-\bar{Y}_{b,a})}{2\pi} - \frac{\bar{R}_a + R_b}{2} \bar{\rho}_{b,a} \right) \tag{71}$$

with

$$\bar{\rho}_{b,a} \equiv \frac{1}{2} - \frac{1}{\pi} \arctan \left(\frac{R_b - \bar{W}_{b,a} \bar{R}_a}{\bar{\Omega}_{b,a}} \right). \tag{72}$$

If $\sqrt{\pi/2}$ is absorbed into dt we have that the differential equations become:

$$\begin{aligned} \dot{R}_a &= \frac{2 - \eta_{a,b}}{2} (1 - R_a^2) \\ &+ \frac{\eta_{a,b}}{2} \left[[1 - R_a^2] \frac{\arccos(\bar{Y}_{a,b})}{\pi} + \bar{\rho}_{a,b} [\bar{R}_b - \bar{W}_{a,b} R_a] \right] \\ &+ \lambda_o \sqrt{1 - R_a^2} \frac{1 + \bar{Y}_{a,b}}{2}, \end{aligned} \tag{73}$$

$$\dot{\bar{Y}}_{a,b} = [1 - \bar{Y}_{a,b}^2] \left\{ \eta_{a,b} \sqrt{\frac{1 - \bar{R}_b^2}{1 - R_a^2}} \frac{\bar{\rho}_{a,b}}{2} - \frac{\lambda_o}{2} \left[\frac{R_a}{\sqrt{1 - R_a^2}} + \frac{\bar{R}_b}{\sqrt{1 - \bar{R}_b^2}} \right] \right\}, \tag{74}$$

where in (73) and (74) we have canceled all dependencies with parameters evaluated before $t - \tau$. By performing a stability analysis similar to the one presented in [34] we conclude that the only stable solution has $\bar{Y}_{a,b} = \bar{Y}_{b,a} = 1$. This implies also that $\bar{\rho}_{a,b} \rightarrow \Theta(\bar{R}_b - R_a)$ after short initial transient. By defining the phases $\theta_d = \arccos(R_d)$ we can re-express Eq. (73) in the form of:

$$\dot{R}_a = \left(1 - \frac{\eta_{a,b}}{2} \right) \sin^2 \theta_a + \left[\frac{\eta_{a,b}}{2} \Theta(\theta_a - \bar{\theta}_b) \sin(\theta_a - \bar{\theta}_b) + \lambda_o \right] \sin \theta_a. \tag{75}$$

According to Definition (47) and Eq. (50) we can expect that, if the agent a learns from the rule B , the overlap $R_a = 1 - r_a/N$, where the constant $0 < r_a \sim O(1)$. Using this hypothesis to model the action of a publicity campaign in favor of the rule B and following the results reported in [31] we propose the following modification to the learning algorithm (40):

$$\mathbf{J}_{a,n+1} = \mathbf{J}_{a,n} + \psi_{a,n} \frac{\sigma_{B,n} \mathbf{S}_n}{\sqrt{N}} + f_n \frac{|\mathbf{J}_{a,n}|}{\sqrt{N}} \sqrt{\frac{1 + R_{a,n}}{N(1 - R_{a,n})}} P_n \mathbf{b}_n, \tag{76}$$

where the first two factors of the last term in the right-hand side of (76) are the annealing factor and the scaling factor (equivalent to (41)). The third factor is an increasing function of the overlap R_a , that produces a positive feedback in favor of the growth of \mathbf{J}_a towards B . Observed that this factor has an upper bound given by:

$$\sqrt{\frac{1 + R_a}{N(1 - R_a)}} \leq \sqrt{\frac{2}{r_a}}. \tag{77}$$

$0 \leq P_n$ is a periodic function of the iteration index n . By considering this perturbation, Eq. (50) becomes, disregarding terms of order Δt :

$$\frac{\Delta R_{a,n}}{\Delta t} = \psi_{a,n,\ell} (\beta_n - \phi_{a,n,n} R_{a,n}) + \sqrt{1 - R_{a,n}^2} \left(\lambda_o \frac{1 + Y_{a,n;b,\ell}}{2} + P_n \right). \tag{78}$$

The correspondent differential equation obtained in the limit $\Delta t \rightarrow 0$ is (5).

Appendix B: Two Agents and No Perturbation, Non-Conservative Initial Conditions

Let us study first the system formed by only two agents and with initial conditions such that $\theta_< = \min\{\theta_a(0), \theta_b(0)\} \in (\phi_o, \pi - \phi_o)$ and $\theta_> = \max\{\theta_a(0), \theta_b(0)\} \in (\phi_o, \pi)$. Without lose of generality we assume that $\theta_> = \theta_a(0)$ and $\theta_< = \theta_b(0)$. Then, for all $0 < t < \tau$ we have that the evolution of the system is ruled by the equations:

$$\dot{\theta}_a = \Lambda (\kappa_o \sin \theta_a - 1) - \sin[\theta_a(t) - \theta_<] \tag{79}$$

$$\dot{\theta}_b = \Lambda (\kappa_o \sin \theta_b - 1). \tag{80}$$

Observe that, by hypothesis, $\kappa_o \sin \theta_b(0) - 1 > 0$, thus θ_b is an increasing function of time

$$\begin{aligned} \tan \frac{\theta_b(t)}{2} = & \frac{\tan \frac{\pi - \phi_o}{2} \left(\tan \frac{\theta_<}{2} - \tan \frac{\phi_o}{2} \right) e^{\Lambda_o t}}{\left(\tan \frac{\theta_<}{2} - \tan \frac{\phi_o}{2} \right) e^{\Lambda_o t} + \tan \frac{\pi - \phi_o}{2} - \tan \frac{\theta_<}{2}} \\ & + \frac{\tan \frac{\phi_o}{2} \left(\tan \frac{\pi - \phi_o}{2} - \tan \frac{\theta_<}{2} \right)}{\left(\tan \frac{\theta_<}{2} - \tan \frac{\phi_o}{2} \right) e^{\Lambda_o t} + \tan \frac{\pi - \phi_o}{2} - \tan \frac{\theta_<}{2}} \end{aligned} \tag{81}$$

which clearly indicates that $\theta_b(t) \in (\phi_o, \pi - \phi_o)$ for all t . We have also use that $\Lambda_o = \Lambda \sqrt{\kappa_o^2 - 1}$. Observe that if the initial condition of a $\theta_a(0) = \theta_> < \pi - \phi_o$ then there is a

finite time t_{b1} such that $\theta_b(t_{b1}) = \theta_>$:

$$t_{b1} = \frac{\tan \phi_o}{\Lambda} \left\{ \ln \frac{\tan \frac{\pi - \phi_o}{2} - \tan \frac{\theta_<}{2}}{\tan \frac{\theta_<}{2} - \tan \frac{\phi_o}{2}} + \ln \frac{\tan \frac{\theta_>}{2} - \tan \frac{\phi_o}{2}}{\tan \frac{\pi - \phi_o}{2} - \tan \frac{\theta_>}{2}} \right\}, \tag{82}$$

and we choose the parameters of the system in such a way that $t_{b1} \ll \tau$. If $\theta_> > \pi - \phi_o$ then $\theta_b \rightarrow \pi - \phi_o$ asymptotically. For $\tau > t > t_{b1}$ the equation for θ_b changes and it must be treated as Eq. (79).

Before analyzing the equation for $\theta_b(t > t_{b1})$ we will solve Eq. (79) through a perturbative approach: Let us assume that the solution to (79), $\theta_a(t; \Lambda)$ admits an expansion in powers of Λ , i.e. up to terms of $O(\Lambda^2)$,

$$\begin{aligned} \theta_a(t) = & \theta_< + 2 \arctan \left[e^{-(t-\tilde{t}_0)} \right] \\ & + \Lambda \left\{ \kappa_o \left[\sin \theta_< \left(1 - \frac{\cosh(\tilde{t}_0)}{\cosh(t-\tilde{t}_0)} \right) + \cos \theta_< \frac{t}{\cosh(t-\tilde{t}_0)} \right] \right. \\ & \left. - \frac{\sinh(t-\tilde{t}_0) + \sinh(\tilde{t}_0)}{\cosh(t-\tilde{t}_0)} \right\} + O(\Lambda^2) \end{aligned} \tag{83}$$

where

$$\tilde{t}_0 \equiv \ln \left[\tan \left(\frac{\theta_> - \theta_<}{2} \right) \right]. \tag{84}$$

If $\theta_> < \pi - \phi_o$ (and $t_{b1} < \infty$) the equation for $\theta_b(t > t_{b1})$ becomes:

$$\dot{\theta}_b = \Lambda [\kappa_o \sin \theta_b - 1] - \sin[\theta_b - \theta_>], \tag{85}$$

with a solution given by:

$$\theta_b(t > t_{b1}) = \theta_> + \Lambda(\kappa_o \sin \theta_> - 1)(1 - e^{-t}) + O(\Lambda^2). \tag{86}$$

At a $0 \ll t \leq \tau$ we expect saturation levels to be almost achieved:

$$\theta_a(t) \rightarrow \theta_< + \Lambda(\kappa_o \sin \theta_< - 1) + O(\Lambda^2) \tag{87}$$

$$\begin{aligned} \theta_b(t) & \rightarrow \pi - \phi_o & t_{b1} & \rightarrow \infty \\ & \rightarrow \theta_> + \Lambda(\kappa_o \sin \theta_> - 1) + O(\Lambda^2) & t_{b1} & < \infty. \end{aligned} \tag{88}$$

For future cycles of integration we assume that $\tau \gg 1$, thus the activation time is shorter than τ . Thus, subsequent saturation values (minimum and maximum), are controlled by the same map:

$$x_{n+1} = x_n + \Lambda(\kappa_o \sin x_n - 1), \tag{89}$$

with $x_0 = \theta_>(\theta_<)$ for the maximum (minimum) value of the wave. If $x_o \in (\phi_o, \pi)$ the only stable fixed point of the map is $\pi - \phi_o$.

It is straightforward that the envelope curve will satisfy the equation:

$$\tan \frac{x(t; x_0)}{2} = \frac{\tan \frac{\pi - \phi_o}{2} \left(\tan \frac{x_0}{2} - \tan \frac{\phi_o}{2} \right) e^{\Lambda_o t}}{\left(\tan \frac{x_0}{2} - \tan \frac{\phi_o}{2} \right) e^{\Lambda_o t} + \tan \frac{\pi - \phi_o}{2} - \tan \frac{x_0}{2}}$$

$$\begin{aligned}
 & \tan \frac{\phi_o}{2} \left(\tan \frac{\pi - \phi_o}{2} - \tan \frac{x_0}{2} \right) \\
 & + \frac{\phantom{\tan \frac{\phi_o}{2} \left(\tan \frac{\pi - \phi_o}{2} - \tan \frac{x_0}{2} \right)}}{\left(\tan \frac{x_0}{2} - \tan \frac{\phi_o}{2} \right) e^{\Lambda_o t} + \tan \frac{\pi - \phi_o}{2} - \tan \frac{x_0}{2}} \tag{90}
 \end{aligned}$$

The approximated solution then is a square wave with period 2τ with maximum and minimum values given by $x(t, \theta_>)$ and $x(t, \theta_<)$ respectively:

$$\theta_a(\tau t) \approx x(\lfloor t \rfloor; \theta_{a,0}) [1 - \text{mod}(\lfloor t \rfloor, 2)] + x(\lfloor t \rfloor; \theta_{b,0}) \text{mod}(\lfloor t \rfloor, 2) \tag{91}$$

$$\theta_b(\tau t) \approx x(\lfloor t \rfloor; \theta_{a,0}) \text{mod}(\lfloor t \rfloor, 2) + x(\lfloor t \rfloor; \theta_{b,0}) [1 - \text{mod}(\lfloor t \rfloor, 2)], \tag{92}$$

where $\lfloor x \rfloor = n \in \mathbb{Z}$ such that $n \leq x < n + 1$.

In order to assess the quality of these approximated solutions we performed a numerical integration of the Eqs. (79) and (80) (using a Runge–Kutta method of second order) and plot the solution of (79) together with (91) in Fig. 9. The initial conditions were set to $\theta_a(0) = \pi - \phi_o - 0.1$ and $\theta_b(0) = \phi_o + 0.1$. The parameters of the system were set to $\Lambda = 0.1$ and $\tau = 1000$.

References

1. Shao, J., Havlin, S., Stanley, H.E.: Dynamic opinion model and invasion percolation. *Phys. Rev. Lett.* **103**, 018701 (2009)
2. Kacperski, K., Holyst, J.A.: Phase transitions and hysteresis in a cellular automata-based model of opinion formation. *J. Stat. Phys.* **84**, 169 (1996)
3. Fang, A., Yuan, K., Geng, J., Wei, X.: Opinion dynamics with Bayesian learning. *Complexity* **2020**, 8261392 (2020)
4. Acemoglu, D., Ozdaglar, A.: Opinion dynamics and learning in social networks. *Dyn. Games Appl.* **1**, 3 (2011)
5. Gilbert, E., Bergstrom, T., Karahalios, K.: Blogs are echo chambers: blogs are echo chambers. In: *Proc. 42nd Hawaii International Conference on System Sciences* (2009)
6. Centola, D.: The spread of behavior in an online social network experiment. *Science* **329**, 1194 (2010)
7. Dadenkar, P., Goel, A., Lee, D.T.: Biased assimilation, homophily, and the dynamics of polarization. *Proc. Natl. Acad. Sci. USA* **110**, 5791 (2013)
8. Caticha, N., Vicente, R.: Agent-based social psychology: from neurocognitive processes to social data. *Advs. Complex Syst.* **14**, 711 (2011)
9. Quattrociocchi, W., Caldarelli, G., Scala, A.: Opinion dynamics on interacting networks: media competition and social influence. *Sci. Rep.* **4**, 4938 (2014)
10. Lee, J.K., Choi, J., Kim, C., Kim, Y.: Social media, network heterogeneity, and opinion polarization. *J. Commun.* **64**, 702 (2014)
11. Ramos, M., Shao, J., Reis, S.D.S., Anteneodo, C., Andrade, J.S., Jr., Havlin, S., Makse, H.: How does public opinion become extreme? *Sci. Rep.* **5**, 10032 (2015)
12. Gastner, M.T., Takács, K., Gulyás, M., Sztetelszky, Z., Oborny, B.: The impact of hypocrisy on opinion formation: a dynamic model. *PLoS ONE* **14**, e0218729 (2019)
13. Schweitzer, F., Krivachy, T., Garcia, D.: How emotions drive opinion polarization: an agent-based model. *Complexity* **2020**, 5282035 (2020)
14. Adams, J.A., White, G., Araujo, R.P.: Person-to-person opinion dynamics: an empirical study using an online game. *PLoS ONE* **17**, e0275473 (2022)
15. Fernández-Gracia, J., Suchecki, K., Ramasco, J.J., Miguel, M.S., Eguíluz, V.M.: Is the voter model a model for voters? *Phys. Rev. Lett.* **112**, 158701 (2014)
16. Anderson, B.D.O., Ye, M.: Recent advances in the modelling and analysis of opinion dynamics on influence networks. *Int. J. Autom. Comput.* **B16B**, 129 (2019)
17. Wang, S., Hyang, C., Sun, C.: Multiagent diffusion and opinion dynamics model interaction effects on controversial products. *IEEE Access* **10**, 115252 (2022)
18. Tu, S., Neumann, S.: A viral marketing-based model for opinion dynamics in online social network. *WWW '22: Proceedings of the ACM Web Conference 2022*, 1570 (2022)

19. Okada, I., Okano, N., Ishii, A.: Spatial opinion dynamics incorporating both positive and negative influence in small-world networks. *Front. Psychol.* **10**, 953184 (2022)
20. Chen, J., Wang, H., Chao, X.: Cross-platform opinion dynamics in competitive travel advertising: a coupled networks' insight. *Front. Psychol.* **13**, 1003242 (2022)
21. Helfmann L., Djurdjevac, N., Lorenz-Spreen, P., Schütte, C.: Modelling opinion dynamics under the impact of influencer and media strategies. [arXiv:2301.13661v1](https://arxiv.org/abs/2301.13661v1) [physics.soc-ph] (2023)
22. Latané, B.: The psychology of social impact. *Am. Psychol.* **36**, 343 (1981)
23. Lewenstein, M., Nowak, A., Latané, B.: Statistical mechanics of social impact. *Phys. Rev. A* **45**, 763 (1992)
24. Pinheiro, F., Santos, M.D., Santos, F.C., Pacheco, J.: Origin of peer influence in social networks. *Phys. Rev. Lett.* **112**, 098702 (2014)
25. Nicosia, V., Skardal, P.S., Arenas, A., Latora, V.: Collective phenomena emerging from the interactions between dynamical processes in multiplex networks. *Phys. Rev. Lett.* **118**, 138302 (2017)
26. Baumann, F., Lorenzo-Spreen, P., Sokolov, I.M., Starnini, M.: Modeling echo chambers and polarization dynamics in social networks. *Phys. Rev. Lett.* **124**, 048301 (2020)
27. Moore, T., Finley, P., Brodsky, N., Brown, T., Apelberg, B., Ambrose, B., Glass, R.: Modeling education and advertising with opinion dynamics. *J. Artif. Soc. Social Simul.* **18**, 7 (2015)
28. Gupta, A., Moharir, S., Sahasrabudhe, N.: Influencing opinion dynamics in networks with limited interaction. *IFAC-PaperOnLine* **54**, 684 (2021)
29. Luo, G., Liu, Y., Zeng, Q., Diao, S., Xiong, F.: A dynamic evolution model of human opinion as affected by advertising. *Physica A* **414**, 254 (2014)
30. Carletti, T., Fanelli, D., Grolli, S., Guarino, A.: How to make an efficient propaganda. *Europhys. Lett.* **74**, 222 (2006)
31. Eshghi, S., Preciado, V., Sarkar, S., Venkatesh, S.S., Zhao, Q., D'Souza, R., Swami, A.: Spread, then target, and advertise in waves: optimal budget allocation across advertising channels. *IEEE Trans. Net. Sci. Eng.* **1**, 99 (2018)
32. Neirotti, J.: Anisotropic opinion dynamics. *Phys. Rev. E* **94**, 012309 (2016)
33. Neirotti, J.: Consensus formation times in anysotropic societies. *Phys. Rev. E* **95**, 062305 (2017)
34. Neirotti, J.: Anisotropic opinion dynamics with an adaptive social rule. *Phys. Rev. E* **98**, 052306 (2018)
35. Neirotti, J.: Strategies for an efficient official publicity campaign. *J. Stat. Phys.* **183**, 24 (2021)
36. Deffuant, G., Neau, D., Amblard, F., Weisbuch, G.: Mixing beliefs among interacting agents. *Adv. Complex Syst.* **3**, 87 (2000)
37. Kan, U., Feng, M., Porter, M.A.: An adaptive bounded-confidence model of opinion dynamics on networks. *J. Complex Netw.* **1**, cnac055 (2023)
38. Lorenz, J.: Continuous dynamics under bounded confidence: a survey. *Int. J. Mod. Phys. C* **18**, 1819 (2007)
39. Chowdhury, N. R., Morarescu, I.-C., Martin, S., Srikant, S.: Continuous opinions and discrete actions in social networks: a multi-agent system approach. In: 2016 IEEE 55th Conference on Decision and Control (CDC), Las Vegas, pp. 1739–1744 (2016)
40. Durrett, R., Gleeson, J.P., Lloyd, A.L., Mucha, P.J., Shi, F., Sivakoff, D., Socolar, J.E.S., Varghese, C.: Graph fission in an evolving voter model. *Proc. Natl. Acad. Sci. USA* **109**, 3682 (2012)
41. Zhang, S., Chen, L., Hu, D., Sun, X.: Impact of opinions and relationships coevolving on self-organization of opinion clusters. *Interdiscip. Descr. Complex Syst.* **11**, 310 (2013)
42. Horstmeyer, L., Kuehn, C.: An adaptive voter model on simplicial complexes. *Phys. Rev. E* **101**, 022305 (2020)
43. Liu, J., El Chamie, M., Başar, T., Açıkmeşe, B.: The discrete-time Altafani model of opinion dynamics with communication delays and quantization. In: 2016 IEEE 55th Conference on Decision and Control (CDC), ARIA Resort & Casino, Las Vegas (2016)
44. Kim, M., Noh, J.D.: Time-delay induced dimensional crossover in the voter model. *Phys. Rev. Lett.* **118**, 168302 (2017)
45. Jain, E., Singh, A.: Trust and reputation-based opinion dynamics modelling over temporal networks. *J. Complex Netw.* **4**, cnac019 (2022)
46. Chu, W., Porter, M.: Non-Markovian models of opinion dynamics on temporal networks. [arXiv:2208.12787v2](https://arxiv.org/abs/2208.12787v2) [physics.soc-ph] (2023)
47. Gros, C.: Entrenched time delays versus accelerating opinion dynamics: are advanced democracies inherently unstable? *Eur. Phys. J. B* **90**, 223 (2017)
48. Qesmi, R.: Dynamics of an opinion model with threshold-type delay. *Chaos Solitons Fractals* **142**, 110379 (2021)
49. Choi, Y., Paolucci, A., Pignotti, C.: Consensus of the Hegselmann–Krause opinion formation model with time delay. *Math. Methods Appl. Sci.* **44**, 4560 (2020)

50. Engel, A., Van den Broeck, C.: *Statistical Mechanics of Learning*. CUP, Cambridge (2001)
51. Vicente, R., Martins, A.C.R., Caticha, N.: Opinion dynamics of learning agents: does seeking consensus lead to disagreement? *J. Stat. Mech.* **2009**, P03015 (2009)
52. Giardini, F., Vilone, D., Conte, R.: A bottom-up approach to opinion dynamics: a cognitive model. *Front. Phys.* **3**, 64 (2015)
53. Caticha, N., Cesar, J., Vicente, R.: For whom the Bayesian agents vote? *Front. Phys.* **3**, 25 (2015)
54. Sobkowicz, P.: Opinion dynamics model based on cognitive biases of complex agents. *J. Artif. Soc. Soc. Simul.* **21**, 8 (2018)
55. Lord, C.G., Ross, L., Lepper, M.R.: Biased assimilation and attitude polarization: the effects of prior theories and subsequently considered evidence. *J. Pers. Social Psychol.* **37**, 2098 (1979)
56. Baron, R.S., Hoppe, S.I., Kao, C.F., Brunsman, B., Linneweh, B., Rogers, D.: Social corroboration and opinion extremity. *J. Exp. Soc. Psychol.* **32**, 537 (1996)
57. David, C., Capella, J.N., Fishbain, M.: The social diffusion of influence among adolescents: group interaction in a chat room environment about antidrug advertisements. *Commun. Theory* **18**, 118 (2006)
58. Hebb, D.O.: *The Organization of Behavior*. Wiley, New York (1949)
59. Bishop, C.M.: *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford (1995)
60. Caticha, N., Kinouchi, O.: Time ordering in the evolution of information processing and modulation systems. *Philos. Mag.* **77**, 1565 (1998)
61. Eisenberg, N., Lieberman, M.D., Williams, K.D.: Does rejection hurt? An fMRI study of social exclusion. *Science* **302**, 290 (2003)
62. Reents, G., Urbanczik, R.: Self-averaging and on-line learning. *Phys. Rev. Lett.* **80**, 5445 (1998)
63. Gardiner, C.W.: *Handbook of Stochastic Methods For the Natural and Social Sciences*. Springer Series in Synergetics, 4th edn. Springer, New York (2009)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.