

The role of voting intention in public opinion polarization

Federico Vazquez*

Instituto de Cálculo, FCEN, Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina[†]

Nicolas Saintier[‡] and Juan Pablo Pinasco[§]

Departamento de Matemática and IMAS, UBA-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (1428) Pabellón I - Ciudad Universitaria - Buenos Aires - Argentina

(Dated: December 17, 2019)

We introduce and study a simple model for the dynamics of voting intention in a population of agents that have to choose between two candidates. The level of indecision of a given agent is modeled by its propensity to vote for one of the two alternatives, represented by a variable $p \in [0, 1]$. When an agent i interacts with another agent j with propensity p_j , then i either increases its propensity p_i by h with probability $P_{ij} = \omega p_i + (1 - \omega)p_j$, or decreases p_i by h with probability $1 - P_{ij}$, where h is a fixed step. We assume that the interactions form a complete graph, where each agent can interact with any other agent. We analyze the system by a rate equation approach and contrast the results with Monte Carlo simulations. We found that the dynamics of propensities depends on the weight ω that an agent assigns to its own propensity. When all the weight is assigned to the interacting partner ($\omega = 0$), agents' propensities are quickly driven to one of the extreme values $p = 0$ or $p = 1$, until an extremist absorbing consensus is achieved. However, for $\omega > 0$ the system first reaches a quasistationary state of symmetric polarization where the distribution of propensities has the shape of an inverted Gaussian with a minimum at the center $p = 1/2$ and two maxima at the extreme values $p = 0, 1$, until the symmetry is broken and the system is driven to an extremist consensus. A linear stability analysis shows that the lifetime of the polarized state, estimated by the mean consensus time τ , diverges as $\tau \sim (1 - \omega)^{-2} \ln N$ when ω approaches 1, where N is the system size. Finally, a continuous approximation allows to derive a transport equation whose convection term is compatible with a drift of particles from the center towards the extremes.

I. INTRODUCTION

Political bi-polarization is a widespread phenomenon that generates divisions in a society, and even clashes and revolts. Several works try to explain and model the emergence of polarization using different mechanisms, like negative influence between individuals of antagonistic opinion groups or between members of the same group [1–4], or a confirmation bias, by which individuals tend to search for information that affirms their prior beliefs and discard arguments that confront their opinions [5–7]. More recently, it has been proposed a new alternative mechanism that gives rise to bi-polarization, which combines homophily with the theory of persuasive arguments [8–10]. The idea is that homophily increases interactions between individuals with the same opinion orientation who then persuade each other with arguments that support their opinion tendency, reinforcing their initial positions and becoming more extreme in their beliefs (see [11–14]). The model studied in [11] assumes that each agent has a list with a number of pro and con arguments in favor and against a given issue (e.g. marijuana legalization), respectively. Agents interact by pairs and incorporate in their list of arguments one of its partner's

argument chosen at random, while old arguments are dismissed. Within the context of voting intention, it is natural to think that the number of arguments in favor of a given candidate is proportional to the inclination or propensity to vote for that candidate prior to the elections.

In this article we try to provide a contribution to the understanding of the phenomenon of opinion bi-polarization from a theoretical viewpoint. For that, we study a simple model that mimics the dynamics of propensities in a population of voters that have to decide between two candidates A and B . When two agents meet, the first agent asks its partner about its voting intention, whose answer (A or B) depends on its propensity for that candidate. For simplicity we assume that the second agent answers A with a probability equal to the fraction of its arguments in favor of candidate A , that is, its propensity for A (and equivalently for B). If the answer is A , then the first agent increases its propensity and thus becomes more prone to vote for A . Otherwise, if the answer is B the propensity of the first agent is decreased and becomes more prone to B . In a more general version of the model we assume that a given agent increases its propensity for candidate A with a probability that is a weighted average of its own and its partner's propensity. This dynamics makes each agent have a tendency to change its propensity towards its already favored candidate, which increases with the weight assigned to the own propensity. However, we stress the fact that the propensity update of an agent in an interaction step is probabilistic and not deterministic, even in the extreme

*Electronic address: fede.vazmin@gmail.com

[†]URL: <https://fedevazmin.wordpress.com>

[‡]Electronic address: nsaintie@dm.uba.ar

[§]Electronic address: jpinasco@dm.uba.ar

case of maximum weight.

This model is simple enough to be analytically tractable and is able to induce bi-polarization without relying on the previous mentioned mechanisms, and by implementing a pairwise interaction rule that is independent on the opinion orientation, unlike the models studied in [13, 14]. The model has some analogies with continuous or discrete opinion models whose dynamics is driven by a compromise process [15], in which the opinions of two agents become closer when they interact. However, in the propensity model studied here the opinion (or propensity) difference between the two interacting agents may either decrease or increase due to the probabilistic nature of the update.

The article is organized as follows. In Section II we introduce the model and its dynamics. In Section III we derive the rate equations and obtain their steady state solutions. We present simulation results in section IV. A stability analysis of the rate equations is performed in sections V and VI. In section VII we develop a continuous approximation that allows to derive a transport equation. We conclude in section VIII with a short summary and a discussion of the results.

II. DESCRIPTION OF THE MODEL

We consider a population of N interacting voters that have to choose between two candidates, A or B . During the period previous to the election day, agents have some degree of indecision about the two possible alternatives, which is modeled by assuming that each agent has an inclination or propensity p to vote for candidate A that varies between 0 and 1, so that when $p \simeq 0$ ($p \simeq 1$) the agent is prone to choose B (A). We assume that the propensity of each agent evolves under the influence of the other agents, in such a way that agents update their propensities after pairwise interactions. Initially, the propensities of all agents are uniformly distributed in $[0, 1]$. Then, at each time step $\Delta t = 1/N$ of the dynamics, two agents i and j are chosen at random to interact. Agent j tells its partner i that it is going to vote for A with a probability equal to its propensity p_j . Then, agent i can either increase or decrease its propensity with probabilities that depend on its own and its partner's propensity:

$$p_i(t + 1/N) = \begin{cases} p_i(t) + h & \text{with probability } P_{i,j}, \\ p_i(t) - h & \text{with probability } 1 - P_{i,j}, \end{cases} \quad (1)$$

where

$$P_{i,j} = \omega p_i(t) + (1 - \omega) p_j(t). \quad (2)$$

The step length h ($0 < h \leq 1$) is fixed, while the parameter ω ($0 \leq \omega \leq 1$) is the weight that gives each agent to its own propensity in an interaction. The value of p_i is set to 1 (0) when it becomes larger (smaller) than 1 (0), so that propensities are constrained to the interval $[0, 1]$. This time step is repeated ad infinitum.

III. RATE EQUATIONS AND STATIONARY STATES

At time $t = 0$ the distribution of propensities $f(p, 0)$ is uniform in $[0, 1]$, but after a time of order 1 agents' propensities adopt discrete values $p = 0, h, 2h, \dots, 1$. For the sake of simplicity we can take h such that $S \equiv 1/h$ is an integer number, and thus the propensity adopts discrete values $p = kh$, with $k = 0, \dots, S$. Then, the propensity distribution can be written as $f(p, t) = \sum_{k=0}^S n_k(t) \delta(p - kh)$, where we define $n_k(t)$ as the fraction of agents in state k (with propensity kh) at time t , where $\sum_{k=0}^S n_k(t) = 1$ for all times $t \geq 0$. Then, the evolution of the system is described by the following rate equations:

$$\frac{dn_0}{dt} = [1 - \omega h - (1 - \omega)m] n_1 - (1 - \omega)m n_0, \quad (3a)$$

$$\begin{aligned} \frac{dn_k}{dt} &= [\omega h(k - 1) + (1 - \omega)m] n_{k-1} - n_k \\ &\quad + \{1 - [\omega h(k + 1) + (1 - \omega)m]\} n_{k+1} \end{aligned} \quad (3b)$$

for $1 \leq k \leq S - 1$,

$$\frac{dn_S}{dt} = [\omega(1 - h) + (1 - \omega)m] n_{S-1} - (1 - \omega)(1 - m)n_S, \quad (3c)$$

where $m \equiv \langle p \rangle = h \sum_{k=0}^S k n_k$ is the mean value of the propensity. The first (gain) term in Eq. (3b) corresponds to the transition of particles from state $k - 1$ to state k , which happens with probability $\omega h(k - 1) + (1 - \omega)h k'$ when they interact with another particle in state k' . Adding over all k' values leads to m . The second (loss) term represents $k \rightarrow k - 1$ and $k \rightarrow k + 1$ transitions, which happen with probability 1 after interacting with any other particle. Finally, the third (gain) term is analogous to the first term, where particles switch from state $k + 1$ to state k with probability $1 - [\omega h(k + 1) + (1 - \omega)h k']$ when they interact with a k' -particle.

We are interested in the stationary distributions of propensities in the population, which correspond to the fixed points of the system of equations (3). On the one hand, we can first notice that the two consensus states in the extreme propensity values $p = 0$ ($n_0^* = 1$, $n_k^* = 0$ for $k = 1, \dots, S$) and $p = 1$ ($n_S^* = 1$, $n_k^* = 0$ for $k = 0, \dots, S - 1$) are fixed points of Eqs. (3) that correspond to the two absorbing states of the particle system, where the mean propensities are $m = 0$ and $m = 1$, respectively. On the other hand, we show in Appendix A that for a given mean propensity $m < 1$ the system of Eqs. (3) has non-trivial fixed points given by

$$\begin{aligned} n_0^* &= \frac{1}{1 + \sum_{k=1}^S G_k(h, \omega, m)}, \\ n_k^* &= n_0^* G_k(h, \omega, m) \quad \text{for } 1 \leq k \leq S, \end{aligned} \quad (4)$$

with

$$G_k(h, \omega, m) = \frac{\prod_{j=0}^{k-1} [(1 - \omega)m + j\omega h]}{\prod_{j=1}^k [1 - (1 - \omega)m - j\omega h]}, \quad (5)$$

or

$$G_k(h, \omega, m) = \frac{\Gamma\left(\frac{(1-\omega)m}{\omega h} + k\right) \Gamma\left(\frac{1-(1-\omega)m}{\omega h} - k\right)}{\Gamma\left(\frac{(1-\omega)m}{\omega h}\right) \Gamma\left(\frac{1-(1-\omega)m}{\omega h}\right)} \quad (6)$$

using the gamma functions for $m > 0$ (see Appendix A), and where the mean propensity must satisfy the relation

$$m + \sum_{k=1}^S (m - kh) G_k(h, \omega, m) = 0. \quad (7)$$

Notice that for $m = 0$ we have from Eq. (5) that $G_k(h, \omega, 0) = 0$ for $1 \leq k \leq S$ due to the $j = 0$ term in the numerator, and thus we recover the consensus solution $n_0^* = 1$. Note also that Eq. (5) is not valid when $m = 1$ because the denominator equals 0 due to the term $j = S = 1/h$.

We now look for non-consensus stationary states $m \neq 0, 1$. It is instructive to start analyzing the simplest cases $S = 1$ and $S = 2$. For $S = 1$ ($h = 1$) is $G_1(1, \omega, m) = m/(1 - m)$ from Eq. (5), and thus Eq. (7)

$$m + (m - 1) G_1(1, \omega, m) = 0 \quad (8)$$

is satisfied for all values of m , that is, each m is a stationary value. This is because for $S = 1$ the propensity model turns out to be equivalent to the voter model [16, 17], where it is known that the fractions of voters in each state are conserved, i.e., $n_0(t) = n_0(0)$ and $n_1(t) = n_1(0) = 1 - n_0(0)$ for all $t \geq 0$. Then, the stationary value of m is equal to its initial value $m(0) = n_1(0)$ given by the initial propensity distribution. For $S = 2$ ($h = 1/2$), Eq. (7) for m is

$$m + (m - 1/2) G_1(h, \omega, m) + (m - 1) G_2(h, \omega, m) = 0. \quad (9)$$

Replacing in Eq. (9) the expressions for G_1 and G_2

$$G_1(h, \omega, m) = \frac{(1 - \omega)m}{1 - \omega/2 - (1 - \omega)m} \quad \text{and} \quad (10a)$$

$$G_2(h, \omega, m) = \frac{m[(1 - \omega)m + \omega/2]}{(1 - m)[1 - \omega/2 - (1 - \omega)m]}, \quad (10b)$$

we arrive, after doing some algebra, to the simple relation

$$(1 - \omega)m(1 - m)(1 - 2m) = 0.$$

Therefore, besides the extreme consensus states $m^* = 0$ and $m^* = 1$, we obtain the new stationary solution $m^* = 1/2$. For general S , the possible stationary values of m are given by the solutions of Eq. (7), which are the roots of a polynomial of order $S + 1$. We were unable to find the roots of that polynomial analytically for any $S \geq 3$. However, we have verified numerically for many different values of ω and h that the only real roots are $m^* = 0$, $m^* = 1$ and $m^* = 1/2$, as in the case $S = 2$.

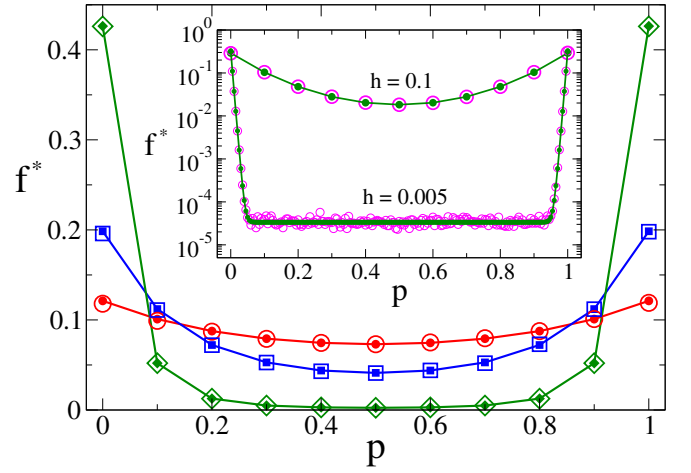


FIG. 1: Main: distribution of propensities at the stationary polarized state for step $h = 0.1$ and weights $\omega = 0.1$ (circles), $\omega = 0.3$ (squares) and $\omega = 0.8$ (diamonds). Filled symbols joined with lines are the stationary solution Eqs. (11), while open symbols correspond to MC simulation results. Inset: upper (lower) curve corresponds to $\omega = 0.5$ and $h = 0.1$ ($h = 0.005$), respectively.

We now analyze the non-trivial solution $m^* = 1/2$ that corresponds, as we shall see below, to a stationary distribution of propensities $f^*(p)$ that is symmetric and polarized around $p = 1/2$. For $m = 1/2$, Eqs. (4) and (6) become

$$n_{P,0}^* = \frac{1}{1 + \sum_{k=1}^S G_k(h, \omega, 1/2)}, \quad (11a)$$

$$n_{P,k}^* = n_{P,0}^* G_k(h, \omega, 1/2) \quad 1 \leq k \leq S, \quad (11b)$$

$$G_k(h, \omega, 1/2) = \frac{\Gamma\left(\frac{1-\omega}{2\omega h} + k\right) \Gamma\left(\frac{1+\omega}{2\omega h} - k\right)}{\Gamma\left(\frac{1-\omega}{2\omega h}\right) \Gamma\left(\frac{1+\omega}{2\omega h}\right)}, \quad (11c)$$

where the subindex P in Eqs. (11) stands for polarization. This series solution Eqs. (11) is plotted in Fig. 1 for $S = 10$ ($h = 0.1$) and different values of ω (filled symbols), while the inset shows the behavior for two values of h (filled circles). We observe that, for each ω , the shape of $f^*(p)$ is symmetric around $p = 1/2$ and peaked at the opposite extreme values $p = 0$ and $p = 1$. This describes a situation in which propensities in the population are polarized, where most individuals adopt opposite and extreme propensity values. In the main plot we see that the system becomes more polarized as ω increases, while we observe in the inset that the polarization is more pronounced as h decreases. To quantify the level of polarization we computed the ratio $R(h, \omega) \simeq \sigma^*(h, \omega)/\sigma_u$ between the standard deviation $\sigma^*(h, \omega) = \sqrt{\langle p^2 \rangle - \langle p \rangle^2} = \sqrt{\left(h^2 \sum_{k=0}^S k^2 n_k^*\right) - 1/4}$ of the propensity distribution $f^*(p)$ for given values of h and ω , and the corresponding value $\sigma_u = \sqrt{1/12 + h/6}$ of the uniform distribution $f_u(p) = \frac{1}{S+1} \sum_{k=0}^S \delta(p - kh)$.

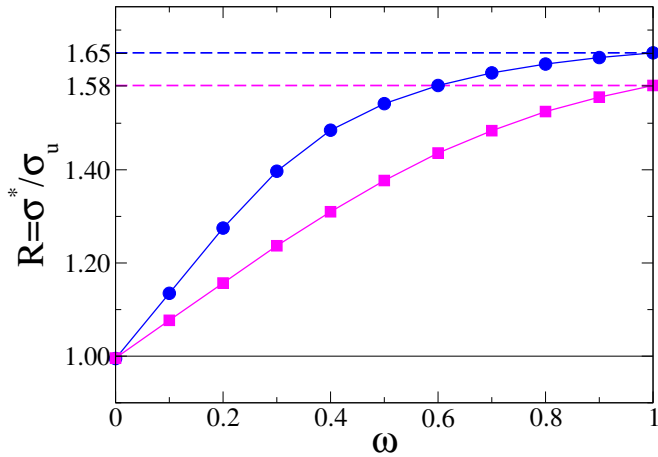


FIG. 2: Polarization level R versus weight ω for $h = 0.05$ (circles) and $h = 0.1$ (squares). The horizontal dashed lines denote the saturation values at $\omega = 1$.

As the width of $f^*(p)$ increases with respect to the uniform distribution when the system is polarized, we expect $R > 1$ and proportional to the magnitude of the polarization. Results are shown in Fig. 2. We see that, for a given h , R increases with ω from the value 1 for $\omega = 0$ to the value $1/(2\sigma_u)$ corresponding to the double-peak distribution $f(p) = [\delta(p) + \delta(p - 1)]/2$ obtained when $\omega = 1$. It is quite remarkable that the system becomes polarized ($R > 1$) for any $\omega > 0$, while for $\omega = 0$ there is no polarization ($R = 1$) and the symmetric stationary solution is the uniform distribution $n_k = 1/(S+1)$ for all $k = 0, \dots, S$, as we can check from Eqs. (4) and (5) using $\omega = 0$ and $m = 1/2$, or alternatively from Eqs. (11). In other words, there is no polarization when agents assign zero weight to their own propensity, but a tiny amount of weight is enough to polarize the population. In summary, for any h and ω , the only stationary states predicted by the rate equations (3) are the consensus absorbing states $m^* = 0$ and $m^* = 1$, and the symmetric polarized state $m^* = 1/2$ in which most agents hold extreme propensities. This polarization phenomenon appears when $\omega > 0$ and is magnified as ω increases.

IV. MONTE CARLO SIMULATIONS

To compare the previous analytical results with that obtained from Monte Carlo (MC) simulations of the model we studied the time evolution of the fractions n_k in single realizations of the dynamics starting from a uniform distribution, as we show in Fig. 3 for a population of $N = 10^6$ agents, $S = 10$ ($h = 0.1$) and $\omega = 0.5$. We can see that after a short initial transient the fractions n_k reach a nearly constant value (plateau) that depends on k . However, this state is not stable and eventually all n_k finally decay to zero except for n_{10} ($p = 1$) that increases and reaches 1, corresponding to a consensus in $p = 1$. We observe that the duration of this quasista-

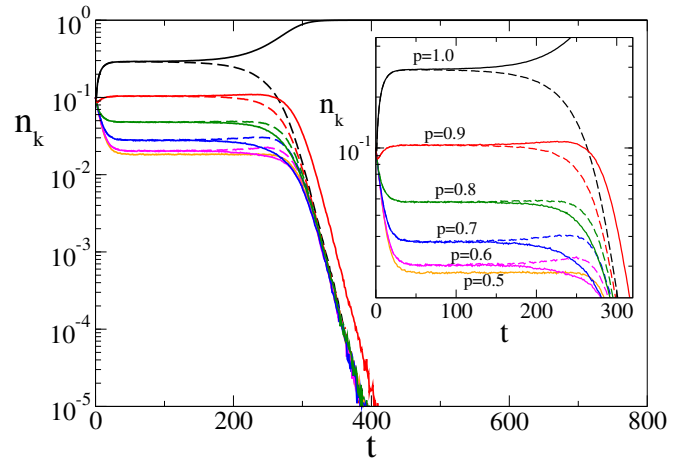


FIG. 3: Monte Carlo results for the time evolution of the fraction of agents n_k with propensities $p = kh$ ($k = 1, \dots, 10$) in a single realization of the dynamics starting from a uniform distribution of propensities in $[0, 1]$. The system consists of a population of $N = 10^6$ agents with $h = 0.1$ and $\omega = 0.5$. Solid (dashed) curves correspond to $p = 1.0$ (0.0), $p = 0.9$ (0.1), $p = 0.8$ (0.2), $p = 0.7$ (0.3), $p = 0.6$ (0.4) and $p = 0.5$ (from top to bottom). The main panel shows the entire evolution up to the consensus time $t \simeq 420$, while the inset shows a detailed view of the plateaus at the quasistationary state.

tionary state described by the plateaus is approximately $\Delta t \simeq 150$, which is similar to that of the final exponential approach to consensus, a phenomenon observed in related models with intermediate states [13, 18, 19]. The height of the plateaus (see inset of Fig. 3 for a detailed view) are plotted by open symbols in the main panel of Fig. 1 for $h = 0.1$ and different values of ω , and in the inset of the same figure for $\omega = 0.5$ and two values of h . We observe a very good agreement with the analytic stationary values Eqs. (11) obtained from the rate Eqs. (3) (filled symbols), which describe an infinite large system where finite-size fluctuations are neglected. We have checked that, indeed, the length of the plateaus increase with N and thus they become infinitely large in the thermodynamic limit, corresponding to the stationary states predicted by the theory. Therefore, the polarized state seems to be unstable, and an extremist consensus is eventually achieved in a finite system.

We have also performed simulations starting from an asymmetric initial condition that consists of a uniform distribution of propensities in the interval $[0, 0.5]$, with $N = 10^6$ agents, $S = 10$ and $\omega = 0.8$. Results are shown in Fig. 4. The top and bottom curves correspond to the total fraction of agents with propensities $p \leq 0.5$ ($\sum_{k=0}^5 n_k$) and $p \geq 0.6$ ($\sum_{k=6}^{10} n_k$), respectively, while the insets show the individual fractions. We see that the system does not reach a quasistationary state as in the previous studied case where the initial propensity distribution was symmetric but, instead, agents' propensities are driven to $p = 0$ given that low propensity values are favored at the initial state. As we shall see in sec-

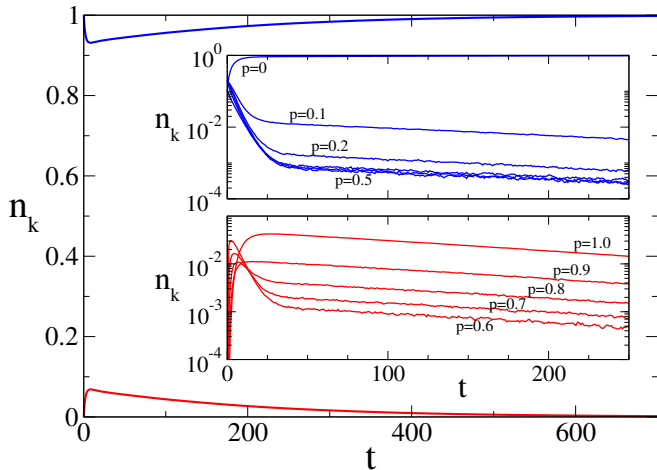


FIG. 4: Main: time evolution of the total fraction of agents with propensities $p = 0, 0.1, 0.2, 0.3, 0.4$ or 0.5 (top curve) and propensities $p = 0.6, 0.7, 0.8, 0.9$ or 1.0 (bottom curve), for a system with $N = 10^6$ agents, $S = 10$ and $\omega = 0.8$, and with an asymmetric initial state where propensities are uniformly distributed in the interval $[0, 0.5]$. Top and bottom insets show the evolution of propensity fractions n_k for $p \leq 0.5$ and $p \geq 0.6$, respectively.

tion V, this is a consequence of the fact that the two consensus states are stable fixed points while the symmetric polarized state is a saddle point and, therefore, for most asymmetric initial conditions the system is quickly driven away from the symmetric polarized state towards the consensus state. Even though a stationary state is not reached, the dynamics exhibits an interesting behavior. We can see in Fig. 4 that propensities above $p = 0.5$, which are initially empty, become populated by agents, and thus the fractions n_k for $k = 6, \dots, 10$ exhibit a non-monotonic time evolution (bottom inset). As a consequence, the system reaches a small level of transient polarization at time $t \simeq 10$ (maximum of bottom curve of the main panel) before decaying to the $p = 0$ consensus.

To study the lifetime of the quasi-stationary polarized state in finite systems we computed the time to reach consensus starting from a uniform distribution in $[0, 1]$. In Fig. 5 we show MC results of the mean consensus time τ vs ω for various system sizes N . We see that τ increases with ω and seems to diverge when ω approaches 1 as a power law $\tau \sim (1 - \omega)^{-\alpha}$, with $\alpha \simeq 2$ (see inset). This means that polarization not only gets stronger as ω increases, but also lasts for longer times. The collapse of the data in the inset also shows that τ increases very slowly with N , as $\ln N$. In the next section we perform a linear stability analysis for the $S = 2$ case that allows to obtain the exponent α and the logarithmic scaling with N .

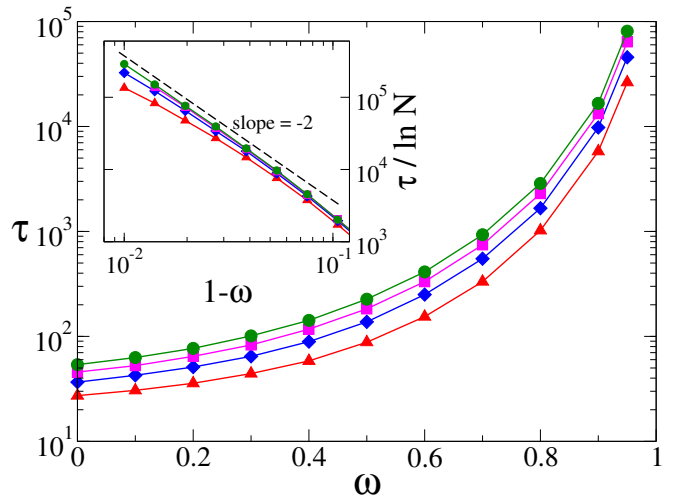


FIG. 5: Mean consensus time τ vs ω for $h = 0.1$ and system sizes $N = 20$ (triangles), $N = 80$ (diamonds), $N = 320$ (squares) and $N = 1280$ (circles). Inset: the data collapse shows the approximate scaling $\tau \sim (1 - \omega)^{-2} \ln N$ for $\omega \lesssim 1$. The dashed line has slope -2 .

V. STABILITY ANALYSIS FOR THE $S = 2$ CASE

An insight into the results shown in the last section can be obtained by studying the simplest non-trivial case $S = 2$, for which the rate equations (3) are

$$\frac{dn_0}{dt} = [1 - \omega/2 - (1 - \omega)m]n_1 - (1 - \omega)m n_0, \quad (12a)$$

$$\frac{dn_1}{dt} = (1 - \omega)m n_0 - n_1 + (1 - \omega)(1 - m)n_2, \quad (12b)$$

$$\frac{dn_2}{dt} = [\omega/2 + (1 - \omega)m]n_1 - (1 - \omega)(1 - m)n_2. \quad (12c)$$

It proves convenient to work with the closed system of equations for n_0 and n_2

$$\frac{dn_0}{dt'} = [1 + \epsilon(n_0 - n_2)](1 - n_0 - n_2) - \epsilon(1 - n_0 + n_2)n_0, \quad (13a)$$

$$\frac{dn_2}{dt'} = [1 - \epsilon(n_0 - n_2)](1 - n_0 - n_2) - \epsilon(1 + n_0 - n_2)n_2, \quad (13b)$$

obtained from Eqs. (12) by using the identities $n_1 = 1 - n_0 - n_2$ and $2m = n_1 + 2n_2 = 1 - n_0 + n_2$ to express n_1 and m in terms of n_0 and n_2 , and defining the parameter $\epsilon \equiv 1 - \omega$ and the rescaled time $t' = t/2$. As we proved in section III for $S = 2$, the rate equations have three fixed points. The two trivial fixed points that represent consensus states in $p = 0$ and $p = 1$ are $\vec{n}_0^* = (1, 0, 0)$ and $\vec{n}_2^* = (0, 0, 1)$, respectively. The non-trivial fixed point $\vec{n}_P^* = (n_{P,0}^*, n_{P,1}^*, n_{P,2}^*)$ corresponding to polarization is

$$\vec{n}_P^* = \left(\frac{1}{3 - \omega}, \frac{1 - \omega}{3 - \omega}, \frac{1}{3 - \omega} \right), \quad (14)$$

which is calculated from Eqs. (4) using the expressions $n_{P,0}^* = 1/(1 + G_1 + G_2)$, $n_{P,1}^* = G_1 n_{P,0}^*$ and $n_{P,2}^* = G_2 n_{P,0}^*$, with $G_1(1/2, \omega, 1/2) = 1 - \omega$ and $G_2(1/2, \omega, 1/2) = 1$ obtained from Eqs. (10) for $S = 2$ ($h = 1/2$) by plugging $m = 1/2$.

To investigate how the system approaches consensus we start by performing a linear stability analysis of the trivial fixed point $\vec{n}_0^* = (1, 0, 0)$ which, by symmetry, is analogous to the analysis of \vec{n}_2^* . We consider small independent perturbations $0 < x_0, x_2 \ll 1$ of the fixed point components and write $n_0 = 1 - x_0$ and $n_2 = x_2$. Plugging these expressions for n_0 and n_2 into Eqs. (13) we obtain, to first order in x_0 and x_2 (neglecting terms of order 2), the following system of linear equations written in matrix representation:

$$\frac{d\mathbf{x}}{dt'} = \mathbf{A} \mathbf{x},$$

where

$$\mathbf{A} \equiv \begin{pmatrix} -1 & 1 + 2\epsilon \\ 1 - \epsilon & -1 - \epsilon \end{pmatrix},$$

and $\mathbf{x} \equiv (x_0, x_2)$. The eigenvalues of the matrix \mathbf{A}

$$\lambda_{\pm} = \frac{-2 - \epsilon \pm \sqrt{(2 + \epsilon)^2 - 8\epsilon^2}}{2} \quad (15)$$

are both negative, and thus the fixed point \vec{n}_0^* is stable under a small perturbation in any direction. To study the behavior of the system for $\omega \lesssim 1$ we expand Eqs. (15) to leading order in $0 < \epsilon \ll 1$. This gives

$$\lambda_+ = -\epsilon^2 + \mathcal{O}(\epsilon^3), \quad \lambda_- = -2 + \mathcal{O}(\epsilon) \quad (16)$$

and, therefore, we have

$$\begin{aligned} x_0(t') &\simeq a e^{-\epsilon^2 t'} + b e^{-2t'} \quad \text{and} \\ x_2(t') &\simeq c e^{-\epsilon^2 t'} + d e^{-2t'}, \end{aligned}$$

where a, b, c and d are constants given by the initial condition. At long times, only the term corresponding to the lowest eigenvalue λ_+ survives, and thus the time evolution of n_0 and n_2 after a perturbation from \vec{n}_0^* are approximately given by

$$n_0(t) \simeq 1 - a e^{-\epsilon^2 t/2} \quad \text{and} \quad n_2(t) \simeq c e^{-\epsilon t/2}. \quad (17)$$

The mean time to reach consensus in a population of N agents can be estimated as the time for which the fraction of agents with propensity $p = 0$ becomes larger than $1 - 1/N$ (less than one agent with propensity $p > 0$). Then, from Eq. (17) we obtain that at consensus time τ is $n_0(\tau) = 1 - a e^{-\epsilon^2 \tau/2} = 1 - 1/N$, from where we arrive to the approximate expression for the mean consensus time

$$\tau \simeq \frac{2 \ln(aN)}{(1 - \omega)^2}, \quad (18)$$

after replacing back ϵ by $1 - \omega$. The $(1 - \omega)^{-2}$ divergence of τ as $\omega \rightarrow 1$ predicted by Eq. (18) is in good agreement with the exponent $\alpha \simeq 2$ found from MC simulations (see inset of Fig. 5). Equation (18) also agrees with the logarithmic increase of τ with N observed in the inset of Fig. 5.

We now study the stability of the polarized state. For that, we linearize the system of Eqs. (13) around the fixed point \vec{n}_P^* by rewriting n_0 and n_2 as $n_0 = n_{P,0}^* + x_0$ and $n_2 = n_{P,2}^* + x_2$, with $0 < |x_0|, |x_2| \ll 1$ and $n_{P,0}^* = n_{P,2}^* = 1/(2 + \epsilon)$ from Eq. (14). Expanding the resulting equations to first order in x_0 and x_2 we arrive to the following linear system:

$$\frac{d\mathbf{x}}{dt''} = \mathbf{B} \mathbf{x},$$

with

$$\mathbf{B} \equiv \begin{pmatrix} -2(1 + \epsilon) & -2(1 + \epsilon) - \epsilon^2 \\ -2(1 + \epsilon) - \epsilon^2 & -2(1 + \epsilon) \end{pmatrix},$$

and $t'' = t/(4 + 2\epsilon)$. The eigenvalues of \mathbf{B} are

$$\lambda_+ = \epsilon^2 \quad \text{and} \quad \lambda_- = -4(1 + \epsilon) - \epsilon^2, \quad (19)$$

while the associated eigenvectors are

$$v_+ = (1, -1) \quad \text{and} \quad v_- = (1, 1).$$

In Fig. 6 we show the flow diagram that summarizes the stability analysis of the $S = 2$ case. The fixed points \vec{n}_0^* and \vec{n}_2^* (circles) are stable in any direction, while \vec{n}_P^* (diamond) is a saddle point that is stable only along the v_- direction ($\lambda_- < 0$) and unstable along any other direction. This means that starting from a state that is symmetric around $p = 1/2$ [$n_0(0) = n_2(0)$] the system evolves along the line $n_0(t) = n_2(t)$ towards the fixed point \vec{n}_P^* . However, any perturbation from the fixed point \vec{n}_P^* that is not symmetric around the center propensity $p = 1/2$ leads the system to one of the absorbing consensus configurations $p = 0$ or $p = 1$ for all agents. This explains the MC simulation results shown in section IV, where the fractions n_k show an initial fast approach from a uniform (symmetric) distribution $n_k(0) \simeq 1/(S + 1)$ (with $k = 0, \dots, S$) to the polarized stationary state \vec{n}_P^* (see Fig. 3) but, eventually, finite-size fluctuations allow the system to escape from this unstable state and reach consensus.

VI. STABILITY ANALYSIS FOR THE $S \geq 3$ CASE

In this section we analyze the stability of the fixed points of the rate equations (3) for general S , namely, the two consensus points $\vec{n}_0^* = (1, 0, \dots, 0)$ and $\vec{n}_S^* = (0, 0, \dots, 1)$, and the symmetric point \vec{n}_P^* given by Eqs. (11). As we showed in the last section for the 3-propensity system ($S = 2$), any symmetric distribution

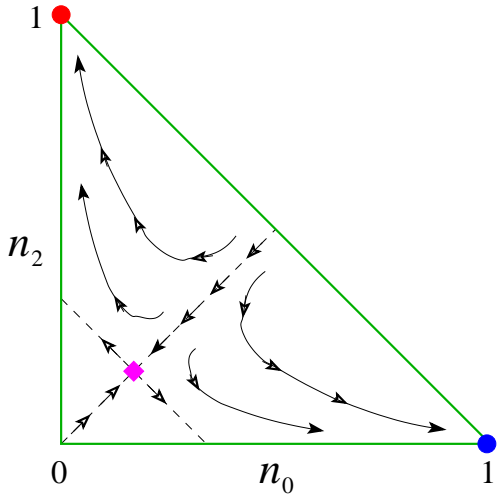


FIG. 6: Schematic flow diagram in the $n_0 - n_2$ plane for the 3-propensity system ($S = 2$). The two stable fixed points denoted by circles correspond to the absorbing consensus states in an extreme propensity, while the saddle point denoted by a diamond represents the steady-state of polarization. The lines with arrows show the flow direction of the system inside the composition triangle $0 \leq n_0 + n_2 \leq 1$.

evolves towards the fixed point \vec{n}_P^* , and one can guess that this behavior also holds for any $S \geq 2$. More interestingly, we found that for $S \geq 3$ there are non-trivial distributions that are not symmetric around $p = 1/2$ which also evolve towards \vec{n}_P^* , as we shall see below. We start by reducing the number of independent variables to S using the relation $n_S = 1 - \sum_{k=0}^{S-1} n_k$ and expressing the mean propensity as

$$m = \sum_{k=1}^{S-1} kh n_k + n_S = 1 - \sum_{k=0}^{S-1} (kh - 1)n_k,$$

and so we can rewrite Eqs. (3) as $\frac{d}{dt}n_k = F_k(n_0, \dots, n_{S-1})$, where the functions F_k correspond to the right-hand-side of the rate equations. We can then differentiate F_k around the fixed points to obtain a linear system of equations defined by a linearized matrix \mathbf{M} . For practical reasons, we used *Mathematica* to calculate the matrix \mathbf{M} , its eigenvalues and eigenvectors.

Let us first analyze the stability of \vec{n}_0^* (the same results hold from the analysis of \vec{n}_S^*). In Fig. 7 we plot the maximum of the real part of the eigenvalues λ_0^{\max} of \mathbf{M}_0 , calculated numerically, as a function of ω for two values of h . We can see that $\text{Re}(\lambda_0^{\max}) < 0$ for all $0 \leq \omega \leq 1$. Therefore, all eigenvalues of \mathbf{M}_0 have a negative real part and thus the consensus fixed point \vec{n}_0^* is locally asymptotically stable. This result generalizes the stability of the two consensus states found for $S = 2$ (section V) to all values $S \geq 2$. We also see in the inset of Fig. 7 that $\text{Re}(\lambda_0^{\max}) \sim (1 - \omega)^2$ as $\omega \rightarrow 1$. Thus, assuming that the mean consensus time scales as $\tau \sim 1/\text{Re}(\lambda_0^{\max})$, we found that the scaling $\tau \sim (1 - \omega)^{-2}$ when ω approaches 1.0

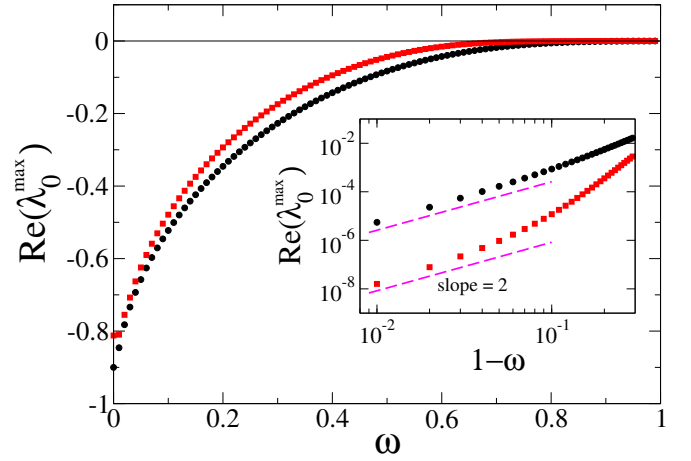


FIG. 7: Maximum real part of the eigenvalues of matrix \mathbf{M}_P vs ω for $h = 0.1$ (circles) and $h = 0.05$ (squares). The inset shows how $\text{Re}(\lambda)$ approaches 0 from below as ω goes to 1. Dashed lines have slope 2.

seems to hold for any $S \geq 2$ (see Fig. 5 for $S = 10$).

We now repeat the analysis above for the symmetric fixed point \vec{n}_P^* . We observed numerically that, for various values of S and ω , the matrix \mathbf{M}_P has only one positive eigenvalue and $S - 1$ negative eigenvalues. As we know from standard dynamical system theory, the eigenvalues of \mathbf{M}_P with negative real part generate the tangent plane \mathcal{T} to the stable manifold of \vec{n}_P^* , which therefore has dimension $S - 1$. Besides, the space of propensity distribution (n_0, \dots, n_S) with mean $m = 1/2$ is an affine plane of codimension 1 which contains \vec{n}_P^* , whose intersection with \mathcal{T} is a manifold of positive dimension $S - 1$. When the system starts from a point of this manifold it follows a trajectory that converges to \vec{n}_P^* , that is, the points on \mathcal{T} represent propensity distributions with mean $m = 1/2$ that evolve towards the polarized state. To illustrate with an example, one of the eigenvectors of \mathbf{M}_P that we found numerically for the case $S = 9$ and $\omega = 1/2$ is

$$\vec{V} \simeq (0.31551, -0.74694, 0.90799, -0.97402, 1.00000, -1.00000, 0.97402, -0.90799, 0.74694, -0.31551),$$

whose associated eigenvalue is $\lambda = -1.878148$. We can now consider the point $\vec{n}(0) = \vec{n}_P^* + 0.02\vec{V}$ on the plane \mathcal{T} as an initial state of the system, which is obtained by slightly perturbing the fixed point \vec{n}_P^* in the direction of \vec{V} . The time evolution of the components n_k of \vec{n} are plotted in the main panel of Fig. 8, while the inset shows the initial perturbed state (empty diamonds) as compared to \vec{n}_P^* (filled circles). We can see that the $\vec{n}(0)$ is not symmetric with respect to $p = 0.5$. This asymmetry is the result of the components of \vec{V} , which exhibit an anti-symmetry that is necessary to preserve the normalization condition $\sum_{k=0}^S n_k(t) = 1$ for all $t \geq 1$. We observe that the fractions n_k (solid and dotted lines) quickly converge to the corresponding values of the components of \vec{n}_P^* denoted by horizontal dashed lines. However, if we

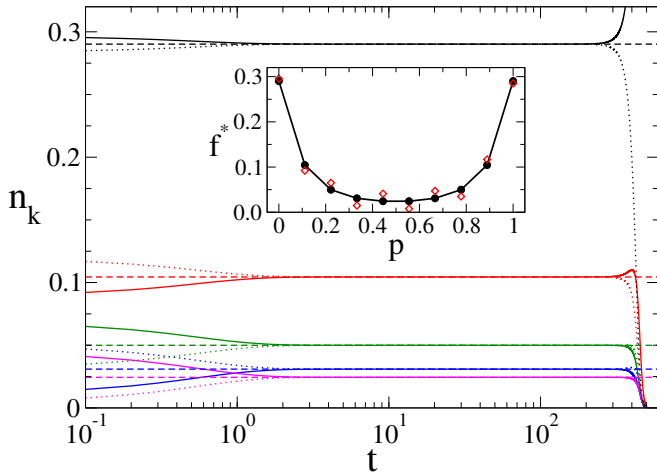


FIG. 8: Time evolution of the propensity fractions n_k for a system with $h = 1/9$ ($S = 9$) and $\omega = 0.5$. Solid curves correspond to $p = 0, 1/9, 2/9, 3/9$ and $4/9$, while dotted curves are for $p = 1, 8/9, 7/9, 6/9$ and $5/9$ (from top to bottom). Inset: the initial values $n_k(0)$ (empty diamonds) correspond to a small perturbation of the polarized stationary distribution given by the components of the fixed point \vec{n}_P^* (filled circles).

zoom in we can see that \vec{n} gets extremely close to \vec{n}_P but not exactly to \vec{n}_P . This very tiny difference is a consequence of the fact that the initial state $\vec{n}(0)$ belongs to the tangent plane \mathcal{T} to the stable manifold of \vec{n}_P , but a priori not to the stable manifold itself. Therefore, the system spends some time very near to \vec{n}_P before eventually going away and converging to the consensus state $p = 0$.

VII. CONTINUOUS APPROXIMATION

In order to analyze the polarized state in more detail it proves useful to consider the system of rate Eqs. (3) in the limiting case of a very small step $h \ll 1$. This allows to derive continuous in p partial differential equations that describe the long-time behavior of the system, as we shall see in this section.

As we explained in section III, after a short initial transient all agents take discrete propensities in the set $p = kh$, with $k = 0, \dots, S$, and thus the propensity distribution can be written as

$$f(p, t) = \sum_{k=0}^S n_k(t) \delta(p - kh), \quad (20)$$

where $\delta(p - kh)$ is the Dirac delta function at kh . Notice that the consensus states correspond to $f(p) = \delta(p)$ and $f(p) = \delta(p - 1)$. We consider a generic function $\phi(p)$ of the propensity, whose mean value over the population is defined as

$$\langle \phi \rangle_f(t) \equiv \int_0^1 \phi(p) f(p, t) dp = \sum_{k=0}^S n_k(t) \phi(kh). \quad (21)$$

This is a macroscopic scalar variable of the particle system, like the mean propensity $m(t) = \langle p \rangle(t)$ and its variance when we take $\phi(p) = p$ and $\phi(p) = (p - \langle p \rangle)^2$, respectively. In Appendix B we show that the time evolution of $\langle \phi \rangle_f$ is described by the following equation:

$$\begin{aligned} \frac{1}{h} \frac{d}{dt} \langle \phi \rangle_f &= \left\langle v(p, t) \phi'(p) + \frac{h}{2} \phi''(p) \right\rangle_f \\ &+ [1 - B_0(t)] f(0, t) \left[\phi'(0) - \frac{h}{2} \phi''(0) \right] \\ &- [1 - B_1(t)] f(1, t) \left[\phi'(1) + \frac{h}{2} \phi''(1) \right] + \mathcal{O}(h^2), \end{aligned} \quad (22)$$

where

$$\begin{aligned} v(p, t) &\equiv 2m(t) - 1 + 2\omega [p - m(t)], \\ B_0(t) &\equiv (1 - \omega)m(t) \text{ and} \\ B_1(t) &\equiv (1 - \omega)[1 - m(t)]. \end{aligned}$$

There are no $\mathcal{O}(h^2)$ terms when ϕ is linear in p . As we can see in the derivation of Appendix B, the first term in the rhs of Eq. (22) comes from the rate equations for $n_k(t)$ ($0 < k < S$) that describe the evolution of the propensity distribution in $(0, 1)$, while the second and third terms come from the dynamics near the boundary points at $p = 0$ and $p = 1$, respectively, and describe the balance between the particles entering and leaving the boundary. The coefficient $v(p, t)$ is related to the drift of particles from the center value $p = 1/2$ towards the ends of the interval $[0, 1]$, while $B_0(t)$ and $B_1(t)$ are boundary coefficients.

Taking $\phi(p) = 1$ in Eq. (22) leads to the conservation of the total mass $\int_0^1 f(p, t) dp = 1$, as expected. Besides, for $\phi(p) = p$ we obtain the following equation for the evolution of the mean propensity:

$$\begin{aligned} \frac{1}{h} \frac{d}{dt} m(t) &= 2m(t) - 1 + f(0, t) [1 - (1 - \omega)m(t)] \\ &- f(1, t) [\omega + (1 - \omega)m(t)]. \end{aligned} \quad (23)$$

We can check from Eq. (23) that if the population is initially in a consensus state, i.e. (i) $m(0) = 0$ or (ii) $m(0) = 1$, then (i) $m(t) = 0$ or (ii) $m(t) = 1$ for any $t \geq 0$, meaning that the population remains in the consensus state as expected from the fixed point solutions $m^* = 0, 1$. We can also see in Eq. (23) that term $2m - 1$ describes a drift towards $m = 0$ ($m = 1$) when $m < 1/2$ ($m > 1/2$) caused by the instability of the fixed point $m^* = 1/2$. Therefore, starting from a nearly uniform distribution with $m(0)$ slightly larger than $1/2$ as in the MC simulations, agents' propensities are slowly dragged to $p = 1$. Besides, we see that $m(0) = m^* = 1/2$ is a stationary value if $f(0, 0) = f(1, 0)$, in agreement with the fact that any symmetric distribution $f(1/2 - p) = f(1/2 + p)$ evolves towards the polarized fixed point \vec{n}_P^* , as shown in sections V and VI.

To better explore the dynamics, we can derive an approximate equation for the time evolution of the propensity distribution $f(p, t)$. For that, we can rewrite Eq. (22)

neglecting order h terms as

$$\frac{1}{h} \frac{d}{dt} \langle \phi \rangle_f = \int_0^1 [v(p, t) + u(p, t)] f(p, t) \phi'(p) dp, \quad (24)$$

where we have introduced the field

$$\begin{aligned} u(p, t) &= [1 - B_0(t)] \delta(p) - [1 - B_1(t)] \delta(p - 1) \\ &= [1 - (1 - \omega)m(t)] \delta(p) - [\omega + (1 - \omega)m(t)] \delta(p - 1). \end{aligned}$$

Integrating by parts the r.h.s. of Eq. (24) and regrouping terms leads to

$$\begin{aligned} &\int_0^1 \phi(p) \left\{ \frac{1}{h} \frac{\partial}{\partial t} f(p, t) + \frac{\partial}{\partial p} \left\{ [v(p, t) + u(p, t)] f(p, t) \right\} \right\} dp \\ &= [v(1, t) + u(1, t)] f(1, t) \phi(1) \\ &\quad - [v(0, t) + u(0, t)] f(0, t) \phi(0). \end{aligned}$$

Since this relation holds for any function ϕ we see that f satisfies formally the transport equation

$$\begin{aligned} \frac{\partial}{\partial t} f(p, t) &= - \frac{\partial}{\partial p} \left\{ h [v(p, t) + u(p, t)] f(p, t) \right\} \\ &\quad + h [v(1, t) + u(1, t)] f(1, t) \delta(p - 1) \\ &\quad - h [v(0, t) + u(0, t)] f(0, t) \delta(p). \end{aligned} \quad (25)$$

Equation (25) expresses the conservation of total number of particles under the transport induced by the effective drift $v + u$ and with source terms $h [v(1, t) + u(1, t)] f(1, t) \delta(p - 1)$ and $-h [v(0, t) + u(0, t)] f(0, t) \delta(p)$ at the boundary points $p = 1$ and $p = 0$, respectively. An intuitive interpretation of this equation is that the mass density $f(p, t)$ is transported by the field v in $[0, 1]$ and suffers an additional impulse at the borders $p = 0, 1$ given by the field u , which is associated to the rate Eqs. (3a) and (3c) for n_0 and n_S , respectively. This is reminiscent of the bouncing effect of particles at the boundaries, by which a particle that hits $p = 0$ ($p = 1$) can later jump back to the interval $(0, 1)$ with probability $(1 - \omega)m(t) = v(0, t) + u(0, t)$ [$(1 - \omega)(1 - m) = -v(1, t) - u(1, t)$].

A. Approximate stationary state solution

It is useful to decompose $f(p, t)$ into a sum of a boundary term $f(0, t) \delta(p) + f(1, t) \delta(p - 1)$ taking into account the dynamics near $p = 0$ and $p = 1$, and an inside term $\tilde{f}(p, t)$ that describes the dynamics in $(0, 1)$:

$$f(p, t) = \tilde{f}(p, t) + f(0, t) \delta(p) + f(1, t) \delta(p - 1),$$

with $\tilde{f}(0, t) = \tilde{f}(1, t) = 0$. Then, Eq. (22) becomes

$$\begin{aligned} \frac{1}{h} \frac{d}{dt} \langle \phi \rangle_f &= \left\langle v \phi' + \frac{h}{2} \phi'' \right\rangle_{\tilde{f}} + B_0 f(0, t) \left[\phi'(0) - \frac{h}{2} \phi''(0) \right] \\ &\quad - B_1 f(1, t) \left[\phi'(1) - \frac{h}{2} \phi''(1) \right], \end{aligned} \quad (26)$$

where we have neglected terms of order 2 and higher, and simplified the notation by writing $v = v(p, t)$, $B_0 = B_0(t)$, $B_1 = B_1(t)$ and $\phi = \phi(p)$. We are interested in the stationary solutions to Eq. (26). As expected from previous results, the consensus states $f^*(p) = \delta(p)$ and $f^*(p) = \delta(p - 1)$ are stationary solutions. One can check that by noticing that for $p = 0$ ($p = 1$) consensus is $\tilde{f}^*(p) = 0$, $f^*(0) = 1$ ($f^*(0) = 0$), $f^*(1) = 0$ ($f^*(1) = 1$) and $m = 0$ ($m = 1$). In view of our findings in section III we also expect a symmetric polarized state with mean $m^* = 1/2$ to be a stationary solution. In that perspective it makes sense to drop the terms involving $\phi''(0)$ and $\phi''(1)$ for symmetry reasons. We then look for a stationary solution $\tilde{f}^*(p) = f^*(p) - f^*(0) \delta(p) - f^*(1) \delta(p - 1)$ satisfying

$$\left\langle v \phi' + \frac{h}{2} \phi'' \right\rangle_{\tilde{f}^*} + B_0 f^*(0) \phi'(0) - B_1 f^*(1) \phi'(1) = 0, \quad (27)$$

for any $\phi(p)$. In Appendix C we show that the solution to Eq. (27), different from $\delta(p)$ and $\delta(p - 1)$, is given by

$$\begin{aligned} f^*(p) &= A \left\{ \exp \left[\frac{2\omega}{h} \left(p - \frac{\alpha}{2\omega} \right)^2 \right] \right. \\ &\quad \left. + \frac{h}{2(1-\omega)} \exp \left(\frac{\alpha^2}{2\omega h} \right) \left[\frac{\delta(p)}{m} + \frac{\delta(p-1)}{1-m} \right] \right\}, \end{aligned} \quad (28)$$

where $\alpha \equiv 1 - 2(1 - \omega)m$ and $A > 0$ is a normalization constant that satisfies the condition $\int_0^1 f^*(p) dp = 1$. Notice that the magnitude m in Eq. (29) is the mean propensity that must satisfy the relation $\int_0^1 (p - m) f^*(p) dp = 0$, which is equivalent to

$$\int_0^1 (p - m) \exp \left[\frac{2\omega}{h} \left(p - \frac{\alpha}{2\omega} \right)^2 \right] dp = 0.$$

This is a nonlinear equation in m that we studied numerically for various values of ω and h . We found that $m = 1/2$ is the only solution in all cases, which is in agreement with the symmetric solution of the rate Eqs. (3) in section III. Therefore, the symmetric stationary distribution of Eq. (26) is given by

$$\begin{aligned} f^*(p) &= A \left\{ \exp \left[\frac{2\omega}{h} \left(p - \frac{1}{2} \right)^2 \right] \right. \\ &\quad \left. + \frac{h}{(1-\omega)} \exp \left(\frac{\omega}{2h} \right) [\delta(p) + \delta(p - 1)] \right\}, \end{aligned} \quad (29)$$

with

$$A = \left\{ \int_0^1 \exp \left[\frac{2\omega}{h} \left(p - \frac{1}{2} \right)^2 \right] dp + \frac{2h \exp \left(\frac{\omega}{2h} \right)}{(1-\omega)} \right\}^{-1}.$$

As we can see, $f^*(p)$ is symmetric around $m = 1/2$ and is the sum of the continuous function $A \exp \left[\frac{2\omega}{h} \left(p - \frac{1}{2} \right)^2 \right]$ in the interval $(0, 1)$ that has the shape of an inverted

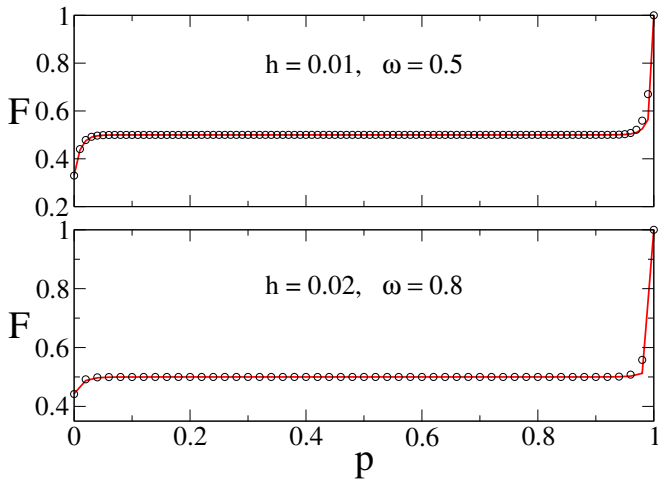


FIG. 9: Propensity cumulative distribution vs p for the parameter values indicated in the legends. Solid lines correspond to the approximate continuous solution for $h \ll 1$ from Eq. (30), while circles represent the exact discrete solution $F_k^* = \sum_{k'=0}^k n_{k'}$.

Gaussian, and the two Dirac masses located at the boundaries, what makes $f^*(p)$ a discontinuous function at $p = 0$ and $p = 1$.

We can alternatively describe the stationary solution by the cumulative distribution function of $f^*(p)$

$$F^*(p) = \begin{cases} 0 & \text{if } p < 0, \\ \frac{Ah}{(1-\omega)} \exp\left(\frac{\omega}{2h}\right) & \text{if } p = 0, \\ A \int_0^p e^{\frac{2\omega}{h}(p-\frac{1}{2})^2} dp + \frac{Ah}{(1-\omega)} \exp\left(\frac{\omega}{2h}\right) & \text{if } 0 \leq p < 1, \\ 1 & \text{if } p \geq 1. \end{cases} \quad (30)$$

In Fig. 9 we plot the approximate stationary cumulative distribution $F^*(p)$ for continuous p (solid curves) and the exact discrete cumulative distribution $F_k^* = \sum_{k'=0}^k n_{k'}$ (circles) for two different small values of h . We see that the data for $F^*(p)$ agrees very well with that of F_k^* , showing that $f^*(p)$ given by Eq. (29) is indeed the limit of $\sum_{k=0}^S n_k^* \delta(p - kh)$ when $h \rightarrow 0$.

VIII. SUMMARY AND DISCUSSION

We studied a system of interacting particles that models the dynamics of voting intentions in a population of individuals that interact by pairs. The propensity of an individual to vote for a given candidate may either increase or decrease after interacting with another partner, depending on the propensity of the partner and the weight ω in $[0, 1]$ assigned to its own propensity. We have investigated the dynamics of the system by means of a rate equation approach and we have checked the results with MC simulations. Starting from a nearly uniform distribution of propensities in $[0, 1]$, we found that for $\omega = 0$ the system is quickly driven towards an extreme

propensity ($p = 0$ or $p = 1$) that corresponds to the initial majority. The dynamics stops evolving when all individuals share the same extreme propensity; an absorbing consensus state. However, for $\omega > 0$ the evolution is quite different: the system initially evolves towards a stationary state characterized by a distribution of propensities that is symmetric around $p = 1/2$ and peaked at the extreme values $p = 0$ and $p = 1$, and it becomes more pronounced when ω gets larger. This distribution describes a state of polarization where most individuals adopt extreme values of p , whose effect is magnified as ω increases. This implies that a tiny weight assigned to our own propensity is enough to polarize the population into two groups with extreme and opposite propensities. However, this state of symmetric polarization is unstable, and thus any perturbation from that state leads the system towards one of the two extremist consensus. Single MC simulations of the dynamics of the model showed that, indeed, the system may initially reach this symmetric quasistationary state but finite-size fluctuations eventually drive the system towards one of the two absorbing configurations. A stability analysis of the rate equations shows that any symmetric distribution evolves towards the polarized state, but there are also non-trivial propensity distributions with mean propensity $m = 1/2$ that are not symmetric around $p = 1/2$ and that evolve and reach the polarized state.

An insight into the polarized state was obtained by analyzing the continuous limit of the system of rate equations. This approximation lead to a transport equation with a convection term that represents a drift of particles from the center propensity $p = 1/2$ towards the extremes, which induces polarization. The stationary solution has the shape of an inverted Gaussian with two Delta functions at $p = 0$ and $p = 1$ that account for the dynamics at the boundaries. In this peculiar dynamics, particles can hit and stay at one of the boundaries for some time but eventually leave, and then hit the boundary again and so on, following an endless loop. We have quantified the lifetime of the polarized state by measuring the mean consensus time τ , and found that it increases with ω and diverges as $\tau \sim (1 - \omega)^{-2}$ when ω approaches 1. This would imply that polarization is quite stable in populations with narrow-minded individuals that only take into account its own opinion when interacting with others, reinforcing their previous believes and adopting more extreme viewpoints. This result is akin to that obtained in related models for opinion formation [11–14] that include a reinforcement mechanism by which pairs of individuals with the same opinion orientation (both in favor or both against a given political issue) are more likely to interact and become more extremists. Even though the propensity model studied in this article does not include this mechanism implicitly, it is able to capture the same phenomenology by implementing a simple interaction rule that consider pairwise interactions between individuals as independent of the opinion group they belong to.

In the studied model, the propensity update proba-

bility is a simple weighted average of the propensities of the two interacting individuals. It would be worthwhile to explore some extensions of the model that consider updating probabilities that are non-linear functions of the propensities and investigate how the behavior of the model is affected, for instance, whether the polarized state becomes more stable or not. It might also be interesting to study versions of the model where pairwise interactions are not simply taken as all-to-all, but rather take place on lattices or complex networks. As the model studied here is akin to multistate voter models with intermediate states and two extreme absorbing states corresponding to consensus [13, 14, 18–21], we ex-

pect that the dynamics of the propensity model on two-dimensional lattices will exhibit some of the phenomenology observed in these models. Some of the phenomena we expect to see include metastable phases in which the system falls into long-lasting striped states where agents organize into single-opinion horizontal, vertical or diagonal bands [18, 19, 22–24] whose lifetime scales with the system size as $\tau \sim N^\nu$, and ν is a non-trivial exponent [14]. Furthermore, by including in the propensity model a control parameter like temperature we also expect to find non-equilibrium phase transitions that belong to the generalized voter model universality class [23–26]. These are all topics for future investigation.

APPENDIX A: STATIONARY SOLUTIONS OF THE RATE EQUATIONS

Setting the time derivatives of Eqs. (3) to zero leads to the system

$$[1 - \omega h - (1 - \omega)m] n_1 - (1 - \omega)m n_0 = 0, \quad (\text{A1a})$$

$$[\omega h(k - 1) + (1 - \omega)m] n_{k-1} - n_k + \{1 - [\omega h(k + 1) + (1 - \omega)m]\} n_{k+1} = 0 \quad 1 \leq k \leq S - 1, \quad (\text{A1b})$$

$$[\omega(1 - h) + (1 - \omega)m] n_{S-1} - (1 - \omega)(1 - m)n_S = 0. \quad (\text{A1c})$$

Notice that the two consensus states corresponding to

- $m = 0, n_0 = 1, n_k = 0$ for $k = 1, \dots, S$,
- $m = 1, n_1 = 1, n_k = 0$ for $k = 0, \dots, S - 1$,

are solutions of the system of Eqs. (A1). To find other possible non-trivial solutions we first note that n_k (for all $k = 1, \dots, S$) can be expressed as a function of n_0 . Starting from Eq. (A1a) we obtain

$$n_1 = \frac{(1 - \omega)m n_0}{1 - \omega h - (1 - \omega)m}. \quad (\text{A2})$$

Then, solving for n_2 from Eq. (A1b) for $k = 1$, and using the previous expression for n_1 we obtain

$$n_2 = \frac{m n_0 (1 - \omega)(\omega h + (1 - \omega)m)}{(1 - \omega h - (1 - \omega)m)(1 - 2\omega h - (1 - \omega)m)}. \quad (\text{A3})$$

The same procedure applied to $k = 2$ leads to

$$n_3 = \frac{m n_0 (1 - \omega)(\omega h + (1 - \omega)m)(2\omega h + (1 - \omega)m)}{(1 - \omega h - (1 - \omega)m)(1 - 2\omega h - (1 - \omega)m)(1 - 3\omega h - (1 - \omega)m)}. \quad (\text{A4})$$

In general we have

$$n_k = n_0 G_k(h, \omega, m) \quad 1 \leq k \leq S, \quad (\text{A5})$$

where

$$G_k(h, \omega, m) = \frac{\prod_{j=0}^{k-1} [(1 - \omega)m + j\omega h]}{\prod_{j=1}^k [1 - (1 - \omega)m - j\omega h]}, \quad (\text{A6})$$

as quoted in Eq. (5) of the main text. The value of n_0 can be obtained by inserting the expression $n_k = n_0 G_k(h, \omega, m)$ in the normalization condition $\sum_{k=0}^S n_k = 1$ and solving for n_0 , which leads to the expression

$$n_0 = \frac{1}{1 + \sum_{k=1}^S G_k(h, \omega, m)}. \quad (\text{A7})$$

quoted in Eq. (4) of the main text. Using relation A5 we can rewrite the mean propensity m as

$$m = \sum_{k=0}^S hkn_k = hn_0 \sum_{k=1}^S k G_k(h, \omega, m),$$

which after replacing the expression A7 for n_0 and rearranging the terms becomes Eq. (7) of the main text.

When $m > 0$ we can rewrite $G_k(h, \omega, m)$ in terms of the gamma functions by using the Pochhammer formula

$$(z)_k \equiv z(z+1)(z+2)\dots(z+k-1) = \frac{\Gamma(z+k)}{\Gamma(z)} \quad \text{for } z \in \mathbb{C} \setminus \mathbb{Z}_- \text{ and } k \geq 0 \text{ integer,} \quad (\text{A8})$$

which follows from the relation $\Gamma(z+1) = z\Gamma(z)$. We first rewrite the numerator of $G_k(h, \omega, m)$ in Eq. (A6) introducing $z \equiv (1-\omega)m/(\omega h)$ as

$$\begin{aligned} (1-\omega)m(\omega h)^{k-1} \left(\frac{(1-\omega)m}{\omega h} + 1 \right) \left(\frac{(1-\omega)m}{\omega h} + 2 \right) \dots \left(\frac{(1-\omega)m}{\omega h} + k - 1 \right) &= (\omega h)^k z(z+1)\dots(z+k-1) \\ &= (\omega h)^k \frac{\Gamma(z+k)}{\Gamma(z)}. \end{aligned}$$

Notice that if $m = 0$ then $z = 0$ and we cannot use Pochhammer formula. Letting $\tilde{z} \equiv [1 - (1-\omega)m - k\omega h]/(\omega h)$, we rewrite in the same way the denominator of $G_k(h, \omega, m)$ in Eq. (A6) as

$$(\omega h)^k \tilde{z}(\tilde{z}+1)\dots(\tilde{z}+k-1) = (\omega h)^k \frac{\Gamma(\tilde{z}+k)}{\Gamma(\tilde{z})}.$$

Inserting these two last expressions for the numerator and denominator of $G_k(h, \omega, m)$ in Eq. (A6) leads to the expression quoted in Eq. (6) of the main text.

APPENDIX B: CONTINUUM EQUATION FOR $\langle \phi \rangle_f$

In this section we derive an equation for the time evolution of the mean of a generic function $\phi(p)$ over the population of agents, expressed as

$$\langle \phi \rangle_f(t) \equiv \int_0^1 \phi(p) f(p, t) dp = \sum_{k=0}^S n_k(t) \phi(kh), \quad (\text{B1})$$

where $f(p, t)$ is the propensity distribution at time t . Taking the time derivative of Eq. (B1) gives

$$\begin{aligned} \frac{d}{dt} \langle \phi \rangle_f &= \sum_{k=0}^S \frac{dn_k}{dt} \phi(kh) \\ &= A\phi(0) + B\phi(1) - \sum_{k=1}^{S-1} n_k \phi(kh) + \sum_{k=1}^{S-1} n_{k-1} \phi(kh) [\omega h(k-1) + (1-\omega)m] \\ &\quad + \sum_{k=1}^{S-1} n_{k+1} \phi(kh) [1 - \omega h(k+1) - (1-\omega)m], \end{aligned} \quad (\text{B2})$$

with

$$A \equiv (1 - \omega h - (1 - \omega)m)n_1 - (1 - \omega)mn_0 \quad \text{and} \quad B \equiv (\omega(1 - h) + (1 - \omega)m)n_{S-1} - (1 - \omega)(1 - m)n_S.$$

We then write

$$\begin{aligned} \phi(kh) &= \phi((k-1)h) + h\phi'((k-1)h) + \frac{h^2}{2}\phi''((k-1)h) + O(h^3) \quad \text{and} \\ \phi(kh) &= \phi((k+1)h) - h\phi'((k+1)h) + \frac{h^2}{2}\phi''((k+1)h) + O(h^3), \end{aligned}$$

and replace these expressions in the summations of Eq. (B2). Then

$$\begin{aligned} \frac{d}{dt}\langle\phi\rangle_f &= (A + n_0)\phi(0) + (B + n_S)\phi(1) - \langle\phi\rangle_f \\ &+ \sum_{k=0}^{S-2} n_k [\omega hk + (1 - \omega)m] \left[\phi(kh) + h\phi'(kh) + \frac{h^2}{2}\phi''(kh) \right] \\ &+ \sum_{k=2}^S n_k [1 - \omega hk - (1 - \omega)m] \left[\phi(kh) - h\phi'(kh) + \frac{h^2}{2}\phi''(kh) \right] + O(h^3). \end{aligned} \quad (\text{B3})$$

The first summation in the rhs of Eq. (B3) is

$$\begin{aligned} &\sum_{k=0}^S n_k [\omega hk + (1 - \omega)m] \left[\phi(kh) + h\phi'(kh) + \frac{h^2}{2}\phi''(kh) \right] \\ &- n_{S-1} [\omega(1 - h) + (1 - \omega)m] \left[\phi(1 - h) + h\phi'(1 - h) + \frac{h^2}{2}\phi''(1 - h) \right] \\ &- n_S [\omega + (1 - \omega)m] \left[\phi(1) + h\phi'(1) + \frac{h^2}{2}\phi''(1) \right] \end{aligned}$$

which is, up to $\mathcal{O}(h^3)$,

$$\begin{aligned} &\left\langle [\omega p + (1 - \omega)m] \left(\phi + h\phi' + \frac{h^2}{2}\phi'' \right) \right\rangle_f - \phi(1) [[\omega + (1 - \omega)m]n_S + [\omega(1 - h) + (1 - \omega)m]n_{S-1}] \\ &- h\phi'(1)n_S[\omega + (1 - \omega)m] - \frac{h^2}{2}\phi''(1)n_S[\omega + (1 - \omega)m]. \end{aligned} \quad (\text{B4})$$

The second summation in the rhs of Eq. (B3) is

$$\begin{aligned} &\sum_{k=0}^S n_k [1 - \omega hk - (1 - \omega)m] \left[\phi(kh) - h\phi'(kh) + \frac{h^2}{2}\phi''(kh) \right] \\ &- n_0 [1 - (1 - \omega)m] \left[\phi(0) - h\phi'(0) + \frac{h^2}{2}\phi''(0) \right] \\ &- n_1 [1 - \omega h - (1 - \omega)m] \left[\phi(h) - h\phi'(h) + \frac{h^2}{2}\phi''(h) \right], \end{aligned}$$

which is, up to $\mathcal{O}(h^3)$,

$$\begin{aligned} &\left\langle [1 - \omega p - (1 - \omega)m] \left(\phi - h\phi' + \frac{h^2}{2}\phi'' \right) \right\rangle_f - \phi(0) [n_0 [1 - (1 - \omega)m] - n_1 [1 - \omega h - (1 - \omega)m]] \\ &+ h\phi'(0)n_0 [1 - (1 - \omega)m] - \frac{h^2}{2}\phi''(0)n_0 [1 - (1 - \omega)m]. \end{aligned} \quad (\text{B5})$$

Replacing the two summations in Eq. (B3) by the expressions (B4) and (B5) leads to Eq. (22) quoted in the main text.

APPENDIX C: STATIONARY SOLUTION OF THE EQUATION FOR $\langle\phi\rangle_f$

We look for a stationary solution f^* of the form

$$f^*(p) = \tilde{f}^*(p) + f^*(0)\delta(p) + f^*(1)\delta(p - 1),$$

where $\tilde{f}^*(p)$ is a continuous function of p . We can then rewrite Eq. (27) as

$$\begin{aligned} 0 &= \int_0^1 (2\omega p + 2(1 - \omega)m - 1)\phi'(p) \tilde{f}^*(p) dp + \frac{h}{2} \int_0^1 \phi''(p) \tilde{f}^*(p) dp \\ &+ \phi'(0)f^*(0)(1 - \omega)m - f^*(1)(1 - \omega)(1 - m)\phi'(1). \end{aligned}$$

Integrating by parts the second integral of the above equation gives

$$0 = \int_0^1 \phi'(p) \left\{ (2\omega p + 2(1-\omega)m - 1) \tilde{f}^*(p) - \frac{h}{2} \tilde{f}^{*'}(p) \right\} dp \\ + \phi'(0) \left\{ f^*(0)(1-\omega)m - \frac{h}{2} \tilde{f}^*(0) \right\} + \phi'(1) \left\{ \frac{h}{2} \tilde{f}^*(1) - f^*(1)(1-\omega)(1-m) \right\}.$$

Since this equality must hold for any function $\phi(p)$ we obtain that

$$[2\omega p + 2(1-\omega)m - 1] \tilde{f}^*(p) - \frac{h}{2} \tilde{f}^{*'}(p) = 0, \\ f^*(0)(1-\omega)m = \frac{h}{2} \tilde{f}^*(0), \\ \frac{h}{2} \tilde{f}^*(1) = f^*(1)(1-\omega)(1-m),$$

from which we arrive to the expression for $f^*(p)$ quoted in Eq. (29).

ACKNOWLEDGMENTS

This work was partially supported by Universidad de Buenos Aires under grants 20020170100445BA, 20020170200256BA, and by ANPCyT PICT2012 0153 and PICT2014-1771.

-
- [1] D. Baldassarri and P. Bearman, *American sociological review* **72**, 784 (2007).
 - [2] P. Balenzuela, J. P. Pinasco, and V. Semeshenko, *PloS one* **10**, e0139572 (2015).
 - [3] A. Flache and M. W. Macy, *The Journal of Mathematical Sociology* **35**, 146 (2011).
 - [4] L. Salzarulo, *Journal of Artificial Societies and Social Simulation* **9** (2006).
 - [5] A. Anagnostopoulos, A. Bessi, G. Caldarelli, M. Del Vicario, F. Petroni, A. Scala, F. Zollo, and W. Quattrociocchi, arXiv preprint arXiv:1411.2893 (2014).
 - [6] C. R. Sunstein, *Journal of political philosophy* **10**, 175 (2002).
 - [7] F. Zollo, P. K. Novak, M. Del Vicario, A. Bessi, I. Mozetič, A. Scala, G. Caldarelli, and W. Quattrociocchi, *PloS one* **10**, e0138740 (2015).
 - [8] D. G. Myers, *Group Decision Making* (Academic Press, New York, London, 1982).
 - [9] N. Mark, *American Sociological Review* **68**, 319 (2003).
 - [10] D. C. Lau and J. K. Murnighan, *Academy of Management Review* **23**, 325 (1998).
 - [11] M. Mäs and A. Flache, *PLoS ONE* **8**, 1 (2013).
 - [12] M. Mäs, A. Flache, K. Takács, and J. K. A., *Organ. Sc* **24**, 716 (2013).
 - [13] C. E. La Rocca, L. A. Braunstein, and F. Vazquez, *Europhys. Lett.* **106**, 40004 (2014).
 - [14] F. Velásquez-Rojas and F. Vazquez, *Journal of Statistical Mechanics: Theory and Experiment* **2018**, 043403 (2018).
 - [15] G. Weisbuch, G. Deffuant, F. Amblard, and J.-P. Nadal, *Complexity* **7**, 55 (2002).
 - [16] P. Clifford and A. Sudbury, *Biometrika* **60**, 581 (1973).
 - [17] R. Holley and T. M. Liggett, *The Annals of Probability* **3**, 643 (1975).
 - [18] X. Castelló, V. M. Eguíluz, and M. S. Miguel, *New Journal of Physics* **8**, 308 (2006).
 - [19] D. Volovik and S. Redner, *Journal of Statistical Mechanics: Theory and Experiment* **2012**, P04003 (2012).
 - [20] M. Starnini, A. Baronchelli, and R. Pastor-Satorras, *Journal of Statistical Mechanics: Theory and Experiment* **2012**, P10027 (2012).
 - [21] F. Vazquez, E. S. Loscar, and G. Baglietto, *Phys. Rev. E* **100**, 042301 (2019).
 - [22] P. Chen and S. Redner, *Phys. Rev. E* **71**, 036101 (2005).
 - [23] S. M. Krause, P. Böttcher, and S. Bornholdt, *Phys. Rev. E* **85**, 031126 (2012).
 - [24] S. M. Krause and S. Bornholdt, *Phys. Rev. E* **86**, 056106 (2012).
 - [25] I. Dornic, H. Chaté, J. Chave, and H. Hinrichsen, *Phys. Rev. Lett.* **87**, 045701 (2001).
 - [26] F. Vazquez and C. López, *Physical Review E* **78**, 061127 (2008).