

Filter Bubble effect in the multistate voter model

Giulio Iannelli,^{1,2} Giordano De Marzo*,^{1,3,4} and Claudio Castellano^{5,1}

¹*Centro Ricerche Enrico Fermi, Piazza del Viminale, 1, I-00184 Rome, Italy.*

²*Dipartimento di Fisica, Università di Roma “Tor Vergata”, 00133 Roma, Italy.*

³*Dipartimento di Fisica Università “Sapienza”, P.le A. Moro, 2, I-00185 Rome, Italy.*

⁴*Sapienza School for Advanced Studies, “Sapienza”, P.le A. Moro, 2, I-00185 Rome, Italy.*

⁵*Istituto dei Sistemi Complessi (ISC-CNR), Via dei Taurini, 19, I-00185 Rome, Italy*

(*Electronic mail: giordano.demarzo@cref.it)

(Dated: February 3, 2022)

Social media influence online activity by recommending to users content strongly correlated with what they have preferred in the past. In this way they constrain users within filter bubbles that strongly limit their exposure to new or alternative content. We investigate this type of dynamics by considering a multistate voter model where, with a given probability λ , a user interacts with a “personalized information” suggesting the opinion most frequently held in the past. By means of theoretical arguments and numerical simulations, we show the existence of a nontrivial transition between a region (for small λ) where consensus is reached and a region (above a threshold λ_c) where the system gets polarized and clusters of users with different opinions persist indefinitely. The threshold always vanishes for large system size N , showing that consensus becomes impossible for a large number of users. This finding opens new questions about the side effects of the widespread use of personalized recommendation algorithms.

Information is nowadays mainly diffused via digital channels as people increasingly use online platforms instead of newspapers or television to access news. If, on the one hand, this makes information easily available to all, on the other hand it allows extremely detailed personalization. Most web sites use recommendation algorithms to provide users with content in line with their taste and way of thinking. Examples are the personalized page rank of Google, “suggested for you” posts by Facebook or the recommendations of Amazon and Netflix. This leads to the formation of the so-called “filter bubbles”, in which users are exposed almost exclusively only to content they have already shown to be interested to. In this manuscript we model this phenomenon by a modification of a very simple model for opinion dynamics (the multi-state voter model) where individuals are influenced not only by their peers but also by an external “field” which encodes information about the opinions the individual held in the past. This field, if large enough, leads to a polarized steady state in which users opinions are crystallized and consensus is no more possible.

I. INTRODUCTION

The concept of “filter bubble” has crossed the boundaries of the academic world reaching public discourse and mainstream media. This reflects the realization that online social media (OSM) have a tremendous impact on how people share information and form their opinions. For this reason they may constitute not only a great opportunity for the diffusion of knowledge, but also a great threat for the stability of social fabric and the function-

ing of democracy. A filter bubble occurs when a user is selectively exposed predominantly to content that tends to reinforce his/her current opinion/belief/state, while suppressing other alternatives¹⁻³. In OSM this typically happens because of personalized recommender systems, which leverage information on past user activity to provide suggestions which, aiming at maximizing user satisfaction, tend to be very similar to what the user has already shown to prefer⁴⁻⁶. Together with the “echo chamber” effect⁷⁻⁹, filter bubbles are thought to be at the heart of the overall increase of polarization and radicalization that is observed in many social contexts¹⁰⁻¹².

A great deal of activity has been devoted in the last years to the goal of understanding what are the basic microscopic mechanisms underlying the rise of polarization and how phenomena observed at population scale are linked to them¹³⁻²¹. These efforts follow the line of research aimed at understanding how basic mechanisms underlying the interaction of individuals give rise to collective consensus phenomena^{22,23}. The voter model²⁴⁻²⁶ played an important role in this activity, because of its extremely simple nature amenable to exact analytical treatment. Its dynamics describes the evolution of a population of agents which have to choose between two perfectly equivalent alternatives and do it by selecting at random a peer and copying their selection. Starting from a disordered state, clusters of individuals sharing the same opinion form and grow over time. For any structure of the interaction pattern among individuals, consensus (i.e. all agents having the same opinion) is invariably reached. A very natural generalization is the Multistate Voter Model (MVM), where the number of available equivalent options is a fixed value $M > 2$ ²⁷⁻²⁹. In mean-field, MVM reaches consensus for any value of M , and the average time required for it depends on M only weakly. In this paper we investigate the behavior of MVM in the presence of an additional interaction mecha-

nism which biases the opinion of an agent toward the state the agent has chosen most frequently in the past. This interaction mimics the filter bubble effect of personalized recommenders in OSM, which present to users suggestions based on their previous behavior.

The effect of personalized information on binary voter model dynamics has been investigated in a previous publication³⁰. In a homogeneous mean-field framework it was shown that for sufficiently strong coupling with the agents' past, consensus is no longer reached and the system remains stuck in a polarized state with coexistence of both opinions. Here we generalize the work in Ref.³⁰ by studying both analytically and numerically the effect of personalized information in the context of the MVM. The goal is to understand whether, depending on the number of agents N , the number of possible opinions M and the strength of the personalized information (λ , to be defined below) consensus is reached or not. Depending on the scaling of M with respect to N we identify three different regimes and in each of them we compute the threshold λ_c separating consensus from polarized states. Numerical simulations are in good agreement with theory in all cases except for $M = N$ and they also show that the transition from consensus to polarization is continuous and characterized by a power law distribution of opinions numerosity at the threshold. Remarkably, for $N \rightarrow \infty$ the threshold goes to zero in all three regimes and this implies that in large systems even a very small form of personalized information always breaks consensus, leading to opinion polarization.

II. MULTISTATE VOTER MODEL WITH PERSONALIZED INFORMATION

A. Definition of the model

Many real life situations are characterized by the presence of more than two possible opinions or factions, as in the case of political elections or football clubs, and so the voter model with binary opinions is not suitable for schematizing the dynamics of such systems. However this difficulty can be easily bypassed by considering the usual voter dynamics and allowing the states of the N agents to vary among M distinct opinions instead of only two, so to obtain the so called multistate voter model. Denoting by $\sigma_i(t)$ the opinion of agent i at time t we have $\sigma_i(t) = 1 \dots M$ and the update rule reads

$$\sigma_i(t + \delta t) = \sigma_j(t) \text{ with prob. } \frac{A_{ij}}{\sum_j A_{ij}}. \quad (1)$$

Here A_{ij} is the binary adjacency matrix of the undirected network over which the dynamics take place, while $\delta t = 1/N$. In the following we focus on the mean field case, meaning that the underlying network is a complete graph and it holds $A_{ij} = 1$ for $i \neq j$; in this case the update rule Eq. (1) can be written as

$$\sigma_i(t + \delta t) = k \text{ with prob. } \frac{N_k(t)}{N}, \quad (2)$$

where $N_k(t)$ is the number of agents with opinion k at time t . In order to endow this system with personalized information (PI) we consider also N PI fields e_i , each coupled with the corresponding standard voter agent σ_i . The state of PI fields is a random variable ranging from 1 to M and which assumes the value k with probability $P[e_i(t) = k]$. This probability varies from agent to agent and over time, depending on the history of the corresponding voter: the more a voter has chosen a given opinion in the past, the higher the probability that its corresponding PI field suggests that opinion. In order to quantify this reinforcement process we generalize the expression for $P[e_i(t) = k]$ proposed in³⁰ to the case of M distinct opinions, more precisely

$$P[e_i(t) = k] = P[n_i^{(k)}(t)] = \frac{c^{n_i^{(k)}(t)}}{\sum_{j=1}^M c^{n_i^{(j)}(t)}}, \quad (3)$$

where $n_i^{(k)}(t)$ is the number of times agent i has chosen (or also confirmed) opinion k up to time t ; in the following we will often write $n_i^{(k)}$, keeping the time dependence implicit. The state of the system is thus described by $N(M+1)$ variables, namely $\{(\sigma_i, n_i^{(k)})\}$, and the dynamics takes place as in the voter model with personalized information³⁰. Initially each opinion σ_i is set equal to a random value; at each time step a given agent i is selected uniformly at random and with probability $1 - \lambda$ it follows the usual voter dynamics, while with probability λ the agent copies the state $e_i(t)$ of the corresponding PI field. More explicitly

$$\sigma_i(t + \delta t) = \begin{cases} e_i(t) & \text{with prob. } \lambda \\ k & \text{with prob. } (1 - \lambda) \frac{N_k}{N} \end{cases} \quad (4)$$

As in³⁰, the parameter λ sets the strength of the personalized information with respect to the interaction with other individuals, while c determines how fast personalized information adapts to the preferences of agents.

B. Phenomenology of the multistate voter model with personalized information

Let us illustrate the overall qualitative phenomenology of the model, common for generic values of M , N and c . A crucial observable is the number of surviving opinions $M_s(t)$ as function of time, which plays the role of the order parameter of the system. In the absence of personalized information (that is for $\lambda = 0$), for a homogeneous initial condition with $M_s(0) = M$ states, the average value of $M_s(t)$ satisfies for $t \ll N^{27}$

$$M_s(t) = \frac{M}{1 + \frac{M}{N}t}. \quad (5)$$

For times of the order of N , M_s reaches the value $M_s = 1$, implying an ordered configuration (consensus) in which all agents share the same opinion, as also shown in Fig. 1a. Notice also that there is an intrinsic timescale $t_0 = N/M$ in Eq. (5) such that M_s remains constant and equal to $M_s(0)$ for times up to t_0 . As λ increases deviations from

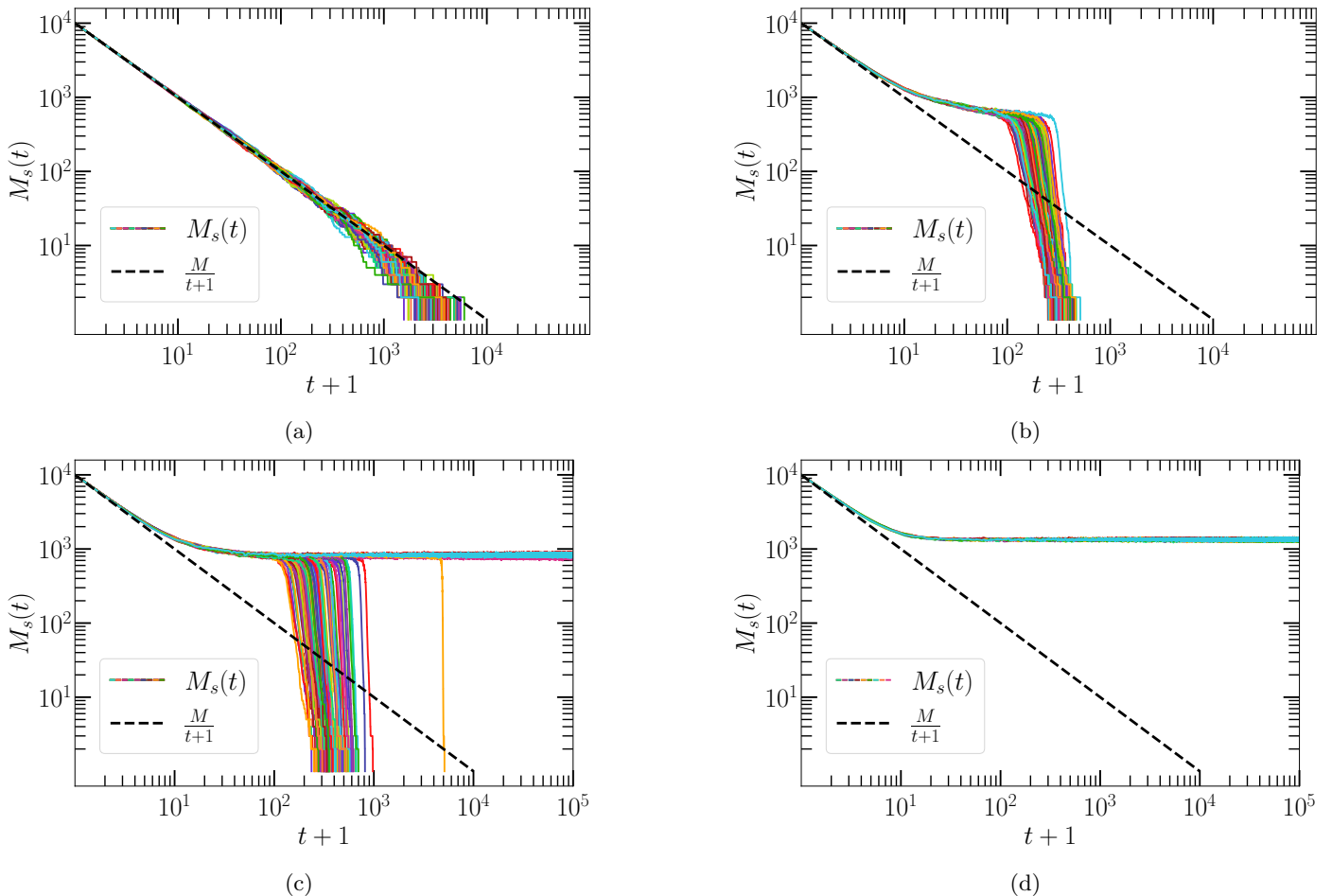


Figure 1: Evolution of the number of surviving opinions for 100 realizations of the dynamical process with various values of λ . (a) Very small $\lambda = 0.001$; (b) Slightly subcritical $\lambda = 0.04$; (c) Critical $\lambda = 0.05$; (d) Supercritical $\lambda = 0.08$. In all panels $N = M = 10000$, $c = \infty$. For convenience we plot $t + 1$ along the horizontal axis.

Eq. (5) start to appear, but the system keeps reaching the ordered state after a sufficiently large amount of time, see Fig. 1b. The behavior changes as λ approaches the threshold λ_c ; in this case consensus can still be reached, but in some realizations the system remains trapped in a stable disordered state (also denoted as polarized state) characterized by the presence of more than one opinion (Fig. 1c). Finally, for λ substantially larger than λ_c , the system never reaches consensus and the asymptotic state is always the disordered one (Fig. 1d). This qualitative phenomenology is observed independently of the value of c (provided that $c > 1$). This parameter only determines the possible presence of an initial transient dominated by the randomness of personalized information in the first time steps, as shown Fig. 2. Indeed for $c = 1 + \delta$ with $\delta \ll 1$ the probability of PI fields, Eq. (3), can be approximated as

$$P[n_i^{(k)}] = \frac{(1 + \delta)^{n_i^{(k)}}}{\sum_{j=1}^M (1 + \delta)^{n_i^{(j)}}} \approx \frac{1 + n_i^{(k)} \delta}{M + \delta \sum_{j=1}^M n_i^{(j)}}.$$

Exploiting the fact that $\sum_{j=1}^M n_i^{(j)}$ is nothing but the number of times agent i has been updated and that going from t to $t + 1$ each agent is on average updated once we have

$$\sum_{j=1}^M n_i^{(j)} \approx t \text{ so that}$$

$$P[n_i^{(k)}] \approx \frac{1 + n_i^{(k)} \delta}{M + t \delta} \approx \frac{1}{M} + \frac{\delta}{M} \left(n_i^{(k)} - \frac{t}{M} \right).$$

Since $0 < n_i^{(k)} < t$ this implies

$$\frac{1}{M} - \delta \frac{t}{M^2} < P[n_i^{(k)}] < \frac{1}{M} + \delta \frac{t}{M} \left(1 - \frac{1}{M} \right).$$

As a consequence $P[n_i^{(k)}] = \frac{1}{M} + O(\delta t)$ and so for δ small the dynamics is initially equal to that of a multistate voter model with external random field. This produces the observed transient. A detailed analysis of the multistate voter model in presence of random external information is beyond the scope of this work, the interested reader can find further details in³¹. In order to determine an upper bound of the c_{th} above which no transient is observed, let us consider the first update of agent i and let us suppose that opinion m is selected. As a consequence the probability of

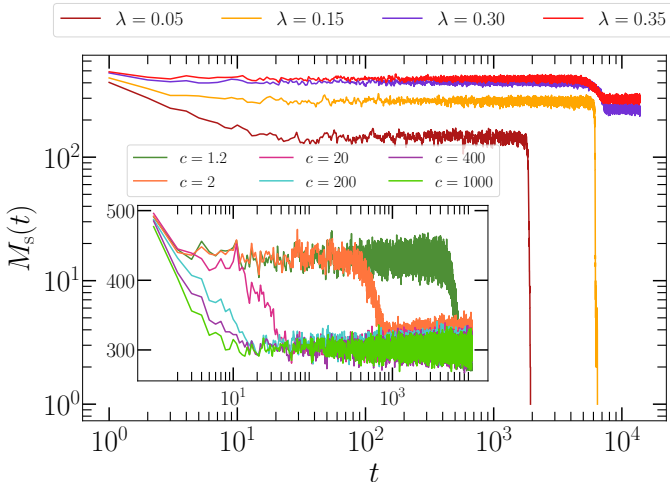


Figure 2: Temporal evolution of the number of surviving opinions for $N = 1000$, $M = 1000$, $c = 1.2$ and various values of λ . The inset shows the same plot for $\lambda = 0.35$ and various values of c .

the corresponding PI field for the successive update is

$$P[e_i = k] = \begin{cases} \frac{c}{c + M - 1} & \text{for } k = m \\ \frac{1}{c + M - 1} & \text{for } k \neq m. \end{cases}$$

If the probability of the PI to be in state m is larger than the probability of being in any other state, a reinforcing loop gets established and the PI gets more and more polarized along opinion m . This implies that a sufficient condition for observing a polarized PI already after the first step is

$$P[e_i(\delta t) = m] > \sum_{k \neq m} P[e_i(\delta t) = k],$$

which implies

$$\frac{c}{c + M - 1} > \frac{M - 1}{c + M - 1}$$

yielding a threshold value

$$c_{th} = M - 1. \quad (6)$$

Fig. 2 shows that this estimate is an upper bound of the real c_{th} . The duration of the transient governed by a random external field gets shorter as c is increased and it is completely absent for $c \approx c_{th} = 999$. Since c plays only a marginal role, in the rest of the paper we focus on the case $c > c_{th}$ so as to remove the initial transient.

Finally, we illustrate the nature of the transition observed as λ is varied. As evident from Fig. 1, around the transition different realizations of the process lead to different outcomes: either consensus or a stationary state. The transition is characterized by the variation, as a function of λ , of the fraction P_s of runs reaching a stationary state (Fig. 3a,inset). As it is possible to see, the larger is N , the sharper the transition becomes. Moreover, since for large

N the critical threshold λ_c goes to zero (see Subsec. III C and figures therein) and the transition gets very sharp, this implies that in large systems even an infinitesimal amount of personalized information is sufficient to make the reaching of consensus impossible. The main panel of Fig. 3a displays how the fraction of surviving opinions in the stationary state $M_s(t \rightarrow \infty)/M = M_s^\infty/M$, averaged only over surviving runs, varies as a function of λ . The quantity M_s^∞/M grows in a continuous fashion, starting from a finite value decreasing with N . This suggests that in the large N limit the transition is continuous, as also confirmed by inspecting the distribution of $S_k(t \rightarrow \infty) = S_k^\infty$, the number of agents polarized along opinion k in the stationary state. When $c > c_{th}$ the number of polarized agents S_k along opinion k is given by the number of agents such that $n_i^{(k)} > n_i^{(j)}$ for any $j \neq k$, since, in this case, the personalized information suggest the most chosen opinion. Indeed, by looking at Fig. 3b it is clear that such a distribution decays as a power-law for $\lambda = \lambda_c$, with a nontrivial exponent approximately equal to 2.8 (for $N = M$).

III. ANALYTICAL APPROACH

As shown above, the phenomenology of the model is only marginally influenced by the parameter c . Therefore consider the case $c \gg c_{th}$ which can be more easily handled analytically. Indeed in such a situation the probability of the PI field e_i is strongly peaked on the opinion more frequently held by agent i and Eq. (3) reduces to

$$P[e_i = k] = \delta_{k,m}, \quad (7)$$

where m is the opinion which satisfies

$$m = \operatorname{argmax}_k [n_i^{(k)}].$$

In other words the PI suggests only the favorite opinion in the past. Our goal is to derive, under the assumption of large c , analytical estimates of the critical value λ_c for generic large values of N and M .

A. Stability of polarized states

Let us consider an agent i in state σ_i and let us assume that its PI is polarized along opinion m . Combining Eqs. (4) and (7) we can write the transition probability of this agent³² as

$$W(\sigma_i \rightarrow k) = \begin{cases} \frac{1}{N} \left[\lambda + (1 - \lambda) \frac{N_k}{N} \right] & \text{for } k = m \\ \frac{1}{N} (1 - \lambda) \frac{N_k}{N} & \text{for } k \neq m. \end{cases} \quad (8)$$

For the overall stability of the polarized state, the PI field e_i should remain polarized on opinion m so as to keep the agent we are considering fixed (on average) on this same opinion. This requires the transition probability to opinion m to be larger than the transition probability to any other

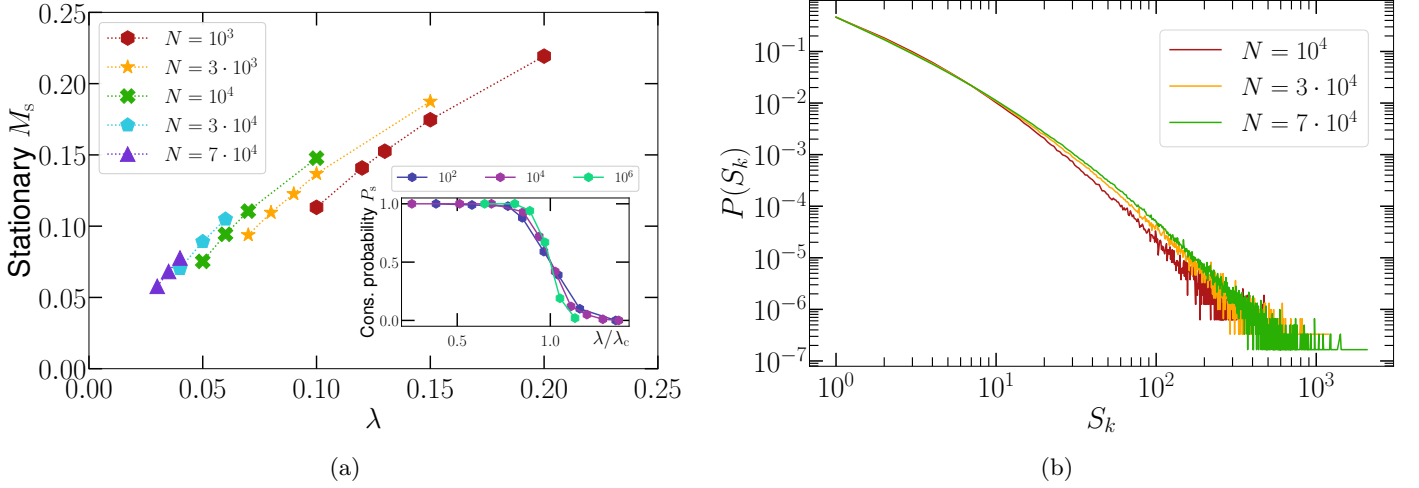


Figure 3: a) Main: Stationary value of the fraction M_s of opinions, for $M = N$, averaged only over surviving runs. Inset: Probability that consensus is reached as a function of λ for $N = 10^2, 10^4, 10^6$ and $M = 20$. As N increases the transition between consensus and polarization becomes sharper and sharper. b) Probability distribution of the number of agents S_k polarized along opinion k at criticality for $N = M$. This has been obtained by letting the system evolve toward the stationary state and then performing a binning over the S_k so to obtain their histogram. Note that such an histogram corresponds to a single realization of the system.

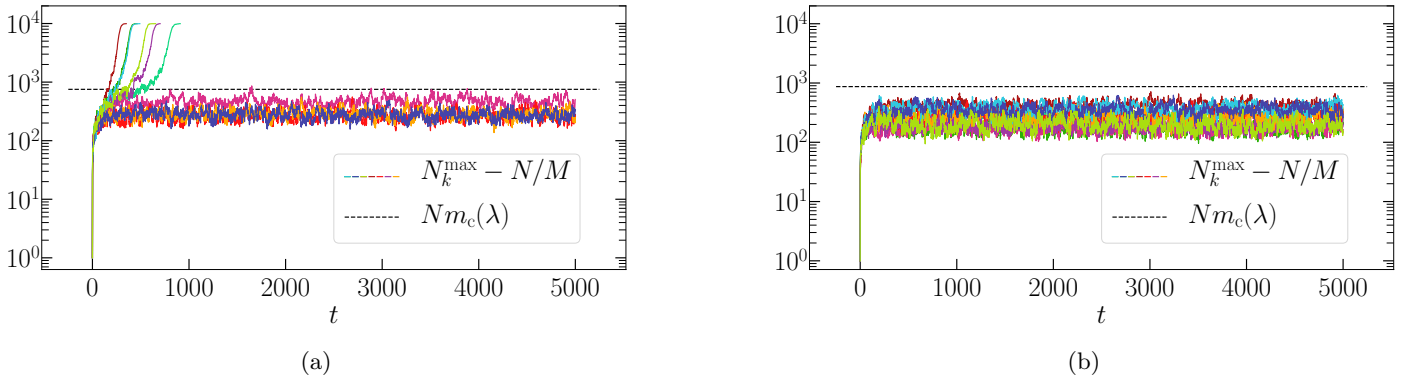


Figure 4: Temporal evolution of $N_{\max} - N/M$, compared with the critical value $Nm_c(\lambda)$ for $N = 10^4$, $M = 200$: a) $\lambda = 0.07$; b) $\lambda = 0.08$. Different colors correspond to different simulations.

opinion l , otherwise $n_i^{(l)}$ would grow faster than $n_i^{(m)}$ and the PI would eventually depolarize and then polarize along opinion l . As a consequence, in order for the polarized state to be stable it must be

$$W(\sigma_i \rightarrow m) > W(\sigma_i \rightarrow l) \quad \forall l \neq m,$$

which yields, using Eq. (8)

$$\frac{N_m - N_l}{N} > -\frac{\lambda}{1 - \lambda} = -m_c(\lambda). \quad (9)$$

Here, in analogy with³⁰, we defined the critical magnetization $m_c(\lambda)$ as

$$m_c(\lambda) = \frac{\lambda}{1 - \lambda}. \quad (10)$$

We can then repeat this same reasoning but considering an agent j whose PI field e_j is polarized along opinion l . This

leads to

$$\frac{N_m - N_l}{N} < m_c(\lambda).$$

Combining this expression with Eq. (9) we obtain the following necessary condition for a polarized state to be stable

$$\left| \frac{N_m - N_l}{N} \right| < m_c(\lambda) \quad \forall (l, m). \quad (11)$$

By introducing the partial magnetization m_k along opinion k , we can also rewrite this constraint as

$$m_k = \frac{N_k - \sum_{l \neq k} N_l}{N} = \frac{2N_k - N}{N}.$$

In this way Eq. (11) becomes

$$\left| \frac{m_m - m_l}{2} \right| < m_c. \quad (12)$$

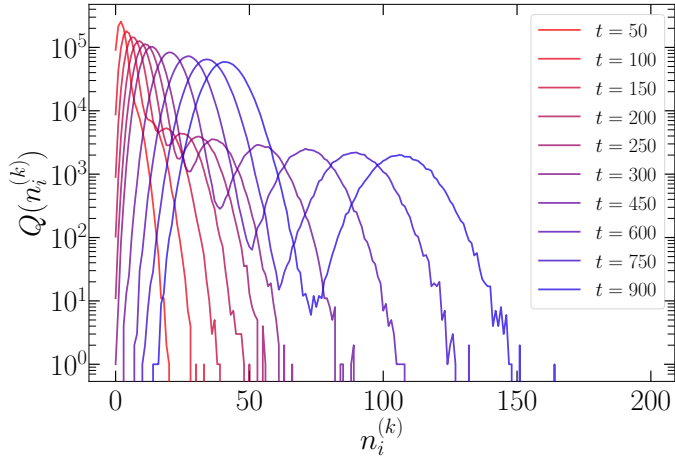


Figure 5: Temporal evolution of the $Q(n_i^{(k)})$ distribution as a function of time. Data are for $c = 400$, $N = 10^6$, $M = 20$, $\lambda = 0.07$. The component with high mean value represents agents polarized along opinion k ; the component with low mean value represents agents unpolarized or polarized along other opinions.

The conclusion if the difference between any two magnetizations is smaller than the threshold m_c then the polarized state is stable. Note, however, that this condition is very strict and a polarized state can be stable on average even if there are some opinions held by a small number of agents not fulfilling it. Indeed in such a situation these opinions will be absorbed by the others, but still consensus will not be reached. As a consequence what really matters is that (11) is fulfilled when considering the most common opinion whose size is N_{max} and the average opinion size given by N/M . The condition ensuring the system not to reach consensus is thus

$$N_{max} - \frac{N}{M} < Nm_c(\lambda). \quad (13)$$

Fig. 4 checks and confirms the validity of this argument in numerical simulations. Note that m_c is larger than 1 for $\lambda > 1/2$ and so, as in the voter model with personalized information, above $\lambda = 1/2$ an opinion can survive even with only a single agent supporting it.

B. Polarization time

In deriving Eq. (12) we assumed all PI fields to be polarized on a certain opinion, meaning that for any i there is only one k such that $n_i^{(k)}$ is maximum. However initially $n_i^{(k)} = 0$ for all pairs (i, k) and the randomness of voter updates implies the occurrence of several ties between the various $n_i^{(k)}$. The assumption that PI fields are polarized starts to be valid only after a polarization time $t^* > 0$. In order to compute this time we write down the master equation for the distribution of the $n_i^{(k)}$ values, $Q(n_i^{(k)})$,

that reads

$$\frac{dQ(n_i^{(k)})}{dt} = NR(n_i^{(k)} - 1)Q(n_i^{(k)} - 1) - NR(n_i^{(k)})Q(n_i^{(k)}), \quad (14)$$

where we introduced the single transition probability

$$\begin{aligned} NR(n_i^{(k)}) &= (1 - \lambda) \frac{N_k}{N} + \lambda P(e_i = k) \\ &\approx (1 - \lambda) \frac{1}{M_s(t)} + \lambda P(e_i = k). \end{aligned} \quad (15)$$

Here we assumed that all surviving opinions share approximately the same number of agents and so we can make the approximation $N_k/N \approx 1/M_s(t)$. Now, considering opinion k , only a fraction S_k/N of the agents is polarized along this opinion and thus we assume that the distribution $Q(n_i^{(k)})$ is formed by two normalized components Q_1 and Q_2 corresponding to the two types of agents: those polarized on opinion k (component Q_2) and those polarized on any other opinion or not polarized at all (component Q_1). We can then write Q as a bimodal distribution

$$Q(n_i^{(k)}) = \left(1 - \frac{S_k}{N}\right) Q_1(n_i^{(k)}) + \frac{S_k}{N} Q_2(n_i^{(k)}).$$

Both components evolve according to the same general master equation Eq. (14), but the transition rates are different. For what concerns the polarized component Q_2 we have

$$NR(n_i^{(k)}) = (1 - \lambda) \frac{1}{M_s(t)} + \lambda$$

and using this expression and Eq. (14) we can compute both the mean value n_2 and the variance σ_2^2 of the distribution Q_2 for polarized agents, finding

$$n_2 = \sigma_2^2 = \int_0^t dt' \left(\frac{1 - \lambda}{M_s(t')} + \lambda \right), \quad (16)$$

Analogously, the transition rate of the component Q_1 , corresponding to agents not polarized on opinion k , is

$$NR(n_i^{(k)}) = (1 - \lambda) \frac{1}{M_s(t)}$$

and so the mean value n_1 and the variance σ_1^2 are

$$n_1 = \sigma_1^2 = \int_0^t dt' \left(\frac{1 - \lambda}{M_s(t')} \right). \quad (17)$$

Detailed computations are reported in Appendix A.

Looking at Eqs. (16) and (17) it is clear that initially the two components are indistinguishable since they both have null mean and variance. However, having different velocities they tend to separate over time. This behavior is shown in Fig. 5. Until Q_1 and Q_2 are superposed there is no real distinction between polarized and unpolarized agents, as all agents can change their preferred opinion in one or few voter updates. Consequently we identify the polarization time t^* introduced above as the time when the two components split for the first time. This can be

readily determined by imposing the distance between the two peaks to be equal to their widths, that is

$$n_2(t^*) - n_1(t^*) = 2 \left[\sqrt{\sigma_1^2(t^*)} + \sqrt{\sigma_2^2(t^*)} \right],$$

where the factor 2 is somewhat arbitrary and does not influence the scaling of t^* . Substituting Eqs. (16) and (17) into this last expression we come up with an implicit integral expression for the polarization time t^*

$$t^* = \frac{2}{\lambda} \left\{ \sqrt{\int_0^{t^*} \left[\frac{1-\lambda}{M_s(t)} + \lambda \right] dt'} + \sqrt{\int_0^{t^*} \frac{1-\lambda}{M_s(t)} dt'} \right\} \quad (18)$$

C. The transition point

Once the polarization time has been obtained we can turn to the determination of the critical parameter λ_c . The idea is the following: In the initial stage of the dynamics, the numbers of agents holding the different opinions, N_i , tend to fluctuate and their differences tend to grow over time. If this growth is slow enough, at the polarization time t^* the condition for the stability of polarized states, Eq. (13) is satisfied. In such a case the disordered state with multiple coexisting opinions is stable and persists forever. Conversely, if at t^* Eq. (13) is violated, the system will eventually evolve toward the consensus state. The condition for λ_c thus reads

$$x_{\max}(t^*) = N_{\max}(t^*) - \frac{N}{M} = N \frac{\lambda_c}{1-\lambda_c}. \quad (19)$$

By means of numerical simulations and a simple scaling argument (see Appendix B) we determine that for $M \ll N$ $x_{\max}(t)$ scales as

$$x_{\max}(t) \sim \gamma \frac{N^{1/2}}{M} t^{3/2}, \quad (20)$$

where $\gamma \approx 1/4$ is a numerical constant. We also derived an equation for the time growth of $N_k - N/M$ and we checked that the temporal scaling of $x_{\max}(t)$ is consistent with this expression, see Appendix C. Replacing this last expression into Eq. (19) we obtain an equation for λ_c

$$\gamma \frac{N^{1/2}}{M} t^{*3/2} = N \frac{\lambda_c}{1-\lambda_c},$$

implying

$$\lambda_c = \frac{\gamma \frac{N^{1/2}}{M} [t^*(\lambda_c)]^{3/2}}{N + \gamma \frac{N^{1/2}}{M} [t^*(\lambda_c)]^{3/2}}. \quad (21)$$

Eq. (21), together with Eq. (18) for the polarization time, provide a closed system of equations in λ_c and t^*

$$\begin{cases} t^* = \frac{2}{\lambda_c} \left\{ \sqrt{\int_0^{t^*} \left[\frac{1-\lambda_c}{M_s(t')} + \lambda_c \right] dt'} + \sqrt{\int_0^{t^*} \frac{1-\lambda_c}{M_s(t')} dt'} \right\} \\ \lambda_c = \frac{\gamma \frac{N^{1/2}}{M} (t^*)^{3/2}}{N + \gamma \frac{N^{1/2}}{M} (t^*)^{3/2}}. \end{cases} \quad (22)$$

In order to solve this system we need an explicit expression for the number $M_s(t)$ of surviving opinions. We consider two possible assumptions, which are expected to be fairly accurate for different values of M and N . Here we limit ourselves to report the main results, the interested reader can find detailed calculations in Appendix D.

D. $M \ll N$

If the number of opinions is much smaller than the number of agents, each opinion is initially shared by a large number of individuals. As a consequence during a first time interval no opinion disappears and it is reasonable to make the approximation

$$M_s(t) = M$$

In this way Eq. (22) can be rewritten as

$$\begin{cases} t^* = \frac{4}{\lambda_c^2} \left[\sqrt{\frac{1-\lambda_c}{M} + \lambda_c} + \sqrt{\frac{1-\lambda_c}{M}} \right]^2 \\ \lambda_c = \frac{\gamma \frac{N^{1/2}}{M} (t^*)^{3/2}}{N + \gamma \frac{N^{1/2}}{M} (t^*)^{3/2}}. \end{cases} \quad (23)$$

Depending on how large M is with respect to N , the solutions of this system scale in different ways.

• $M \ll N^{1/3}$

$$\begin{cases} t^* \approx 4(MN)^{1/4} \\ \lambda_c \approx 2(M^5 N)^{-1/8}. \end{cases} \quad (24)$$

• $N^{1/3} \ll M \ll N$

$$\begin{cases} t^* \approx 2^{8/5} (NM^2)^{1/5} \\ \lambda_c \approx 2^{2/5} (NM^2)^{-1/5}. \end{cases} \quad (25)$$

E. $M = N$

If $M = N$ each agent has initially a different opinion. Hence, even during the first time steps, some opinions disappear by chance. In this case we can approximate the surviving opinions $M_s(t)$ with the expression valid for the simple multistate voter model, Eq. (5), that is for $M = N$,

$$M_s(t) = \frac{N}{1+t}. \quad (26)$$

Substituting this expression in Eq. (22) and imposing $M = N \gg 1$, $\lambda_c \ll 1$ we obtain the following system

$$\begin{cases} t^* = \frac{32N^{1/2} + 8\lambda_c N}{2(\lambda_c^2 N - 16)} \\ \lambda_c = \frac{\gamma N^{-1/2} [t^*(\lambda_c)]^{3/2}}{N + \gamma N^{-1/2} [t^*(\lambda_c)]^{3/2}}, \end{cases} \quad (27)$$

whose solution scales as

$$\begin{cases} t^* \approx 2^{4/3} N^{2/3} \\ \lambda_c \approx 4N^{-1/2} \end{cases} \quad (28)$$

IV. NUMERICAL SIMULATIONS

We now compare theoretical predictions with numerical simulations. We iterate MVM dynamics for various values of M and N up to $t = 2000$ and we determine the fraction of times $P_s(\lambda)$ the system has reached consensus. We take as numerical estimate of λ_c the values such that $P_s(\lambda_c) = 1/2$.

For fixed value of M and increasing size N , the analytical results predict initially a scaling regime given by Eq. (25) as long as N is much larger than M but much smaller than M^3 , followed by a scaling given by Eq. (24). For $M = 20$, the first regime spans a short interval of N values. In Fig. 6a we see only a hint of the associated scaling, while for larger values of N the agreement between theory and numerics is very good. For $M = 200$ instead (see Fig. 6b), the first regime extends to much larger values of N , so that feasible values of N lie only in this regime. In this case λ_c nicely scales as $N^{-1/5}$, as predicted by Eq. (25), although the prefactor is not predicted exactly.

For the case $M = N$ instead, Figure 6c show that the theory is not able to catch the correct scaling. This mismatch has various potential origins. Indeed using Eq. (26) for the temporal dependence of $M_s(t)$ is a quite rough approximation, as the number of different opinions actually decays much more slowly over time. But also the use of Eq. (19) for $x_{\max}(t)$ is not warranted for $M = N$. A deeper understanding of the phenomenology of the case $M = N$ remains an interesting open question.

V. CONCLUSIONS

In this paper we have introduced and analyzed a multistate voter model where the coupling with an external history-dependent individual field mimics the effect of personalized recommendation algorithms in online social media. A population of agents, initially having different opinions, reaches consensus on a single opinion or remains polarized on multiple different opinions, depending on the strength of personalized information. The phenomenology is governed by the competition between the fluctuations induced by voter dynamics and the tendency, due to personalized information, to bind agents to the opinion they adopted most frequently in the past. By means of arguments based on this physical picture, we estimated analytically the critical threshold between the two regimes, obtaining a reasonable agreement with simulations for $M \ll N$. Conversely, the dynamics when the initial number of opinions is comparable with the number of agents seems to elude our approach.

From a more general point of view, our study indicates how difficult reaching consensus is, in the presence of personalized recommendations. For any number M of initial

opinions the threshold λ_c tends to vanish when the number of agents N diverge. This means that no matter how weak is the coupling with the personalized information, if the system is large enough polarization unavoidably arises. In this respect, note that voter dynamics is extremely favorable to the establishment of consensus: for any interaction pattern, consensus is necessarily reached for any finite number of interacting individuals. The addition of a personalized recommendation completely changes this picture, at least for large systems. For a finite number of interacting agents instead consensus is still reached if the strength of personalized information is small enough. The transition between the two regimes exhibits nontrivial features that may be the focus of future activity along with generalizations to include nontrivial interaction patterns or a different functional dependence between the probability distribution of the personalized information and the number of times an opinion has been selected in the past.

A natural extension of this work would be to analyze more realistic recommendation algorithms. With respect to this, a promising possibility is to consider collaborative filtering algorithms³³, which are based on similarities among the history of different users or opinions instead of considering only the history of the user itself. This typology of algorithm is used in real life applications^{34,35}, but still it is sufficiently simple to try an analytical study of its effects on opinion dynamics models.

REFERENCES

- ¹E. Pariser, *The filter bubble: What the Internet is hiding from you* (Penguin UK, 2011).
- ²T. R. Dillahunt, C. A. Brooks, and S. Gulati, "Detecting and visualizing filter bubbles in google and bing," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15 (Association for Computing Machinery, New York, NY, USA, 2015) p. 1851–1856.
- ³S. Nagulendra and J. Vassileva, "Understanding and controlling the filter bubble through interactive visualization: A user study," in *Proceedings of the 25th ACM Conference on Hypertext and Social Media*, HT '14 (Association for Computing Machinery, New York, NY, USA, 2014) p. 107–115.
- ⁴T. T. Nguyen, P.-M. Hui, F. M. Harper, L. Terveen, and J. A. Konstan, "Exploring the filter bubble: The effect of using recommender systems on content diversity," in *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14 (Association for Computing Machinery, New York, NY, USA, 2014) p. 677–686.
- ⁵L. V. Bryant, "The youtube algorithm and the alt-right filter bubble," *Open Information Science* **4**, 85–90 (2020).
- ⁶D. O'Callaghan, D. Greene, M. Conway, J. Carthy, and P. Cunningham, "The extreme right filter bubble," arXiv preprint arXiv:1308.6149 (2013).
- ⁷M. Cinelli, G. De Francisci Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, "The echo chamber effect on social media," *Proceedings of the National Academy of Sciences* **118** (2021), 10.1073/pnas.2023301118, <https://www.pnas.org/content/118/9/e2023301118.full.pdf>.
- ⁸W. Cota, S. C. Ferreira, R. Pastor-Satorras, and M. Starnini, "Quantifying echo chamber effects in information spreading over political communication networks," *EPJ Data Science* **8**, 35 (2019).
- ⁹P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau, "Tweeting from left to right: Is online political communication more than an echo chamber?" *Psychological Science* **26**, 1531–1542 (2015), pMID: 26297377, <https://doi.org/10.1177/0956797615594620>.

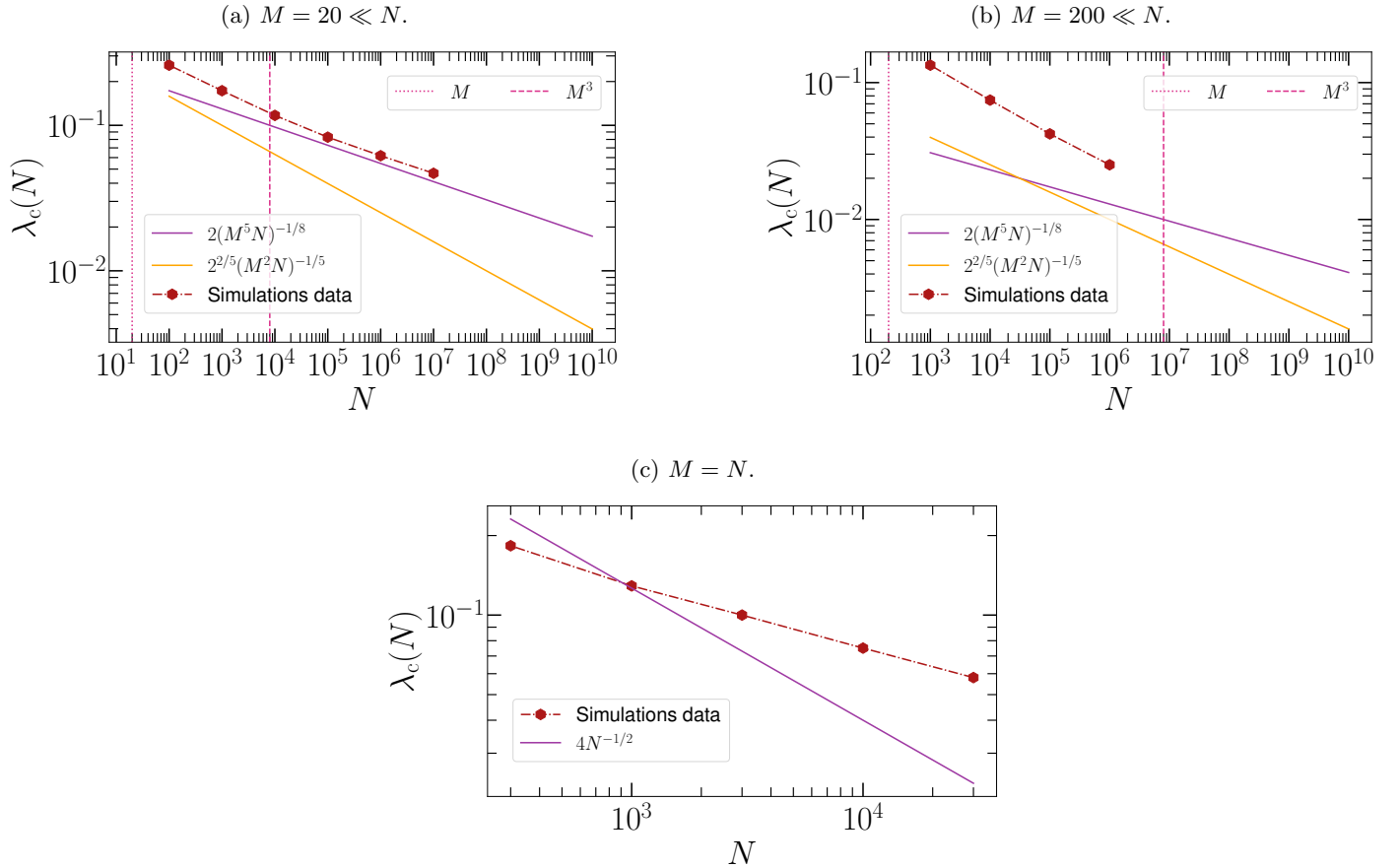


Figure 6: Comparison of analytical predictions for the threshold λ_c and numerical estimates of it

¹⁰“The partisan divide on political values grows even wider,” <https://www.pewresearch.org/politics/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider/> (2017), accessed: 2021-10-04.

¹¹U. Chitra and C. Musco, “Analyzing the impact of filter bubbles on social network polarization,” in *Proceedings of the 13th International Conference on Web Search and Data Mining*, WSDM ’20 (Association for Computing Machinery, New York, NY, USA, 2020) p. 115–123.

¹²M. Maes and L. Bischofberger, “Will the personalization of online social networks foster opinion polarization?” Available at SSRN 2553436 (2015).

¹³S. Galam, “Rational group decision making: A random field ising model at $t=0$,” *Physica A: Statistical Mechanics and its Applications* **238**, 66–80 (1997).

¹⁴N. Crokidakis, “Role of noise and agents’ convictions on opinion spreading in a three-state voter-like model,” *Journal of Statistical Mechanics: Theory and Experiment* **2013**, P07008 (2013).

¹⁵G. L. Ciampaglia, A. Nematzadeh, F. Menczer, and A. Flammini, “How algorithmic popularity bias hinders or promotes quality,” *Scientific Reports* **8**, 15951 (2018).

¹⁶N. Perra and L. E. Rocha, “Modelling opinion dynamics in the age of algorithmic personalisation,” *Scientific reports* **9**, 7261 (2019).

¹⁷A. Sirbu, D. Pedreschi, F. Giannotti, and J. Kertesz, “Algorithmic bias amplifies opinion fragmentation and polarization: A bounded confidence model,” *PLOS ONE* **14**, 1–20 (2019).

¹⁸F. Freitas, A. R. Vieira, and C. Anteneodo, “Imperfect bifurcations in opinion dynamics under external fields,” *Journal of Statistical Mechanics: Theory and Experiment* **2020**, 024002 (2020).

¹⁹F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, “Modeling echo chambers and polarization dynamics in social networks,” *Phys. Rev. Lett.* **124**, 048301 (2020).

²⁰A. F. Peralta, M. Neri, J. Kertész, and G. Iñiguez, “Effect of algorithmic bias and network structure on coexistence, consensus, and polarization of opinions,” *Phys. Rev. E* **104**, 044312 (2021).

²¹J. W. Baron, “Consensus, polarization, and coexistence in a continuous opinion dynamics model with quenched disorder,” *Phys. Rev. E* **104**, 044309 (2021).

²²C. Castellano, S. Fortunato, and V. Loreto, “Statistical physics of social dynamics,” *Rev. Mod. Phys.* **81**, 591–646 (2009).

²³P. Sen and B. K. Chakrabarti, *Sociophysics: an introduction* (Oxford University Press, 2014).

²⁴P. Clifford and A. Sudbury, “A model for spatial conflict,” *Biometrika* **60**, 581–588 (1973).

²⁵L. Frachebourg and P. L. Krapivsky, “Exact results for kinetics of catalytic reactions,” *Phys. Rev. E* **53**, R3009–R3012 (1996).

²⁶P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A kinetic view of statistical physics* (Cambridge University Press, 2010).

²⁷M. Starnini, A. Baronchelli, and R. Pastor-Satorras, “Ordering dynamics of the multi-state voter model,” *Journal of Statistical Mechanics: Theory and Experiment* **2012**, P10027 (2012).

²⁸W. Pickering and C. Lim, “Solution of the multistate voter model and application to strong neutrals in the naming game,” *Phys. Rev. E* **93**, 032318 (2016).

²⁹A. F. Peralta, N. Khalil, and R. Toral, “Ordering dynamics in the voter model with aging,” *Physica A: Statistical Mechanics and its Applications* **552**, 122475 (2020), tributes of Non-equilibrium Statistical Physics.

³⁰G. De Marzo, A. Zaccaria, and C. Castellano, “Emergence of polarization in a voter model with personalized information,” *Phys. Research* **2**, 043117 (2020).

³¹F. Herreras-Azcué and T. Galla, “Consensus and diversity in multistate noisy voter models,” *Phys. Rev. E* **100**, 022304 (2019).

³²Note that these transition probabilities do not depend on the state

σ_i of the agent.

³³J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, "Collaborative filtering recommender systems," in *The adaptive web* (Springer, 2007) pp. 291–324.

³⁴G. Linden, B. Smith, and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," *IEEE Internet Computing* **7**, 76–80 (2003).

³⁵B. Smith and G. Linden, "Two decades of recommender systems at amazon.com," *IEEE Internet Computing* **21**, 12–18 (2017).

³⁶Only two because the scaling $\lambda < N^{-1/2}$ corresponds to negative times.

Appendix A: Moments of the distribution of $n_i^{(k)}$

Let us consider the distribution $Q(n_i^{(k)}, t)$, where $n_i^{(k)}$ is defined as the number of times agent i has chosen or confirmed opinion k in the past. The evolution of this distribution is given by

$$Q(n_i^{(k)}, t + \delta t) = R(n_i^{(k)} - 1)Q(n_i^{(k)} - 1, t) + [1 - R(n_i^{(k)})]Q(n_i^{(k)}, t).$$

Expanding the left hand side for small $\delta t = 1/N$ we obtain which yields Eq. (14)

$$\frac{dQ(n_i^{(k)})}{dt} = NR(n_i^{(k)} - 1)Q(n_i^{(k)} - 1) - NR(n_i^{(k)})Q(n_i^{(k)}). \quad (\text{A1})$$

Now we consider the two components $Q_1(n_i^{(k)}, t)$ and $Q_2(n_i^{(k)}, t)$ introduced in III B and we compute their mean value and variance. We recall that Q_2 corresponds to agents whose PI are polarized along k , while Q_1 to the remaining agents, and that both components evolve according to the same master equation Eq. (14), but with different transition rates. The general equation for the drift of the distribution (also called average drift in the following) is

$$\nu = \frac{d\langle n_i^{(k)} \rangle}{dt} = \sum_{n_i^k} n_i^{(k)} \frac{dQ(n_i^{(k)})}{dt}$$

and using Eq. (A1) we obtain

$$\begin{aligned} \nu &= \sum_{n_i^k} n_i^{(k)} [NR(n_i^{(k)} - 1)Q(n_i^{(k)} - 1) - NR(n_i^{(k)})Q(n_i^{(k)})] = \\ &= \sum_{n_i^k} NR(n_i^{(k)})Q(n_i^{(k)}) \end{aligned} \quad (\text{A2})$$

Analogously the evolution of the second moment is

$$\frac{d\langle (n_i^{(k)})^2 \rangle}{dt} = \sum_{n_i^k} (n_i^{(k)})^2 \frac{dQ(n_i^{(k)})}{dt},$$

that using Eq. (A1) becomes

$$\begin{aligned} \frac{d\langle (n_i^{(k)})^2 \rangle}{dt} &= \sum_{n_i^k} (n_i^{(k)})^2 [NR(n_i^{(k)} - 1)Q(n_i^{(k)} - 1) - NR(n_i^{(k)})Q(n_i^{(k)})] = \\ &= \sum_{n_i^k} (n_i^{(k)} + 1)^2 NR(n_i^{(k)})Q(n_i^{(k)}) - (n_i^{(k)})^2 NR(n_i^{(k)})Q(n_i^{(k)}) = \\ &= \nu + 2 \sum_{n_i^k} n_i^{(k)} NR(n_i^{(k)})Q(n_i^{(k)}), \end{aligned} \quad (\text{A3})$$

where we used Eq. (A2). Let us firstly consider agents polarized along k and so Q_2 , the transition rate is

$$NR_2(n_i^{(k)}) = (1 - \lambda) \frac{1}{M_s(t)} + \lambda.$$

and putting this expression into Eqs. (A2) and (A3) we get

$$\begin{cases} \nu_2 = (1 - \lambda) \frac{1}{M_s(t)} + \lambda \\ \frac{d\langle (n_i^{(k)})^2 \rangle}{dt} = \left[(1 - \lambda) \frac{1}{M_s(t)} + \lambda \right] (1 + 2n_2), \end{cases} \quad (\text{A4})$$

where we defined $n_2 = \langle n_i^{(k)} \rangle_2$. Analogously for the component Q_1 it holds

$$NR_1(n_i^{(k)}) = (1 - \lambda) \frac{1}{M_s(t)}$$

and so

$$\begin{cases} \nu_1 = (1 - \lambda) \frac{1}{M_s(t)} \\ \frac{d\langle (n_i^{(k)})^2 \rangle}{dt} = (1 - \lambda) \frac{1}{M_s(t)} (1 + 2n_2). \end{cases} \quad (\text{A5})$$

Finally, considering that

$$n = \int_0^t \nu \cdot dt'$$

$$\sigma_n^2 = \langle (n_i^{(k)})^2 \rangle - \langle n_i^{(k)} \rangle^2$$

$$\frac{d\sigma^2}{dt} = \frac{d\langle (n_i^{(k)})^2 \rangle}{dt} - 2\langle n_i^{(k)} \rangle \frac{d\langle n_i^{(k)} \rangle}{dt}$$

we can write the expressions for the mean value and variance of both components

$$\begin{cases} n_2(t) = \sigma_2^2(t) = \int_0^t dt' \left(\frac{1 - \lambda}{M_s(t')} + \lambda \right) \\ n_1(t) = \sigma_1^2(t) = \int_0^t dt' \frac{1 - \lambda}{M_s(t')}. \end{cases} \quad (\text{A6})$$

These are Eqs. (16) and (17).

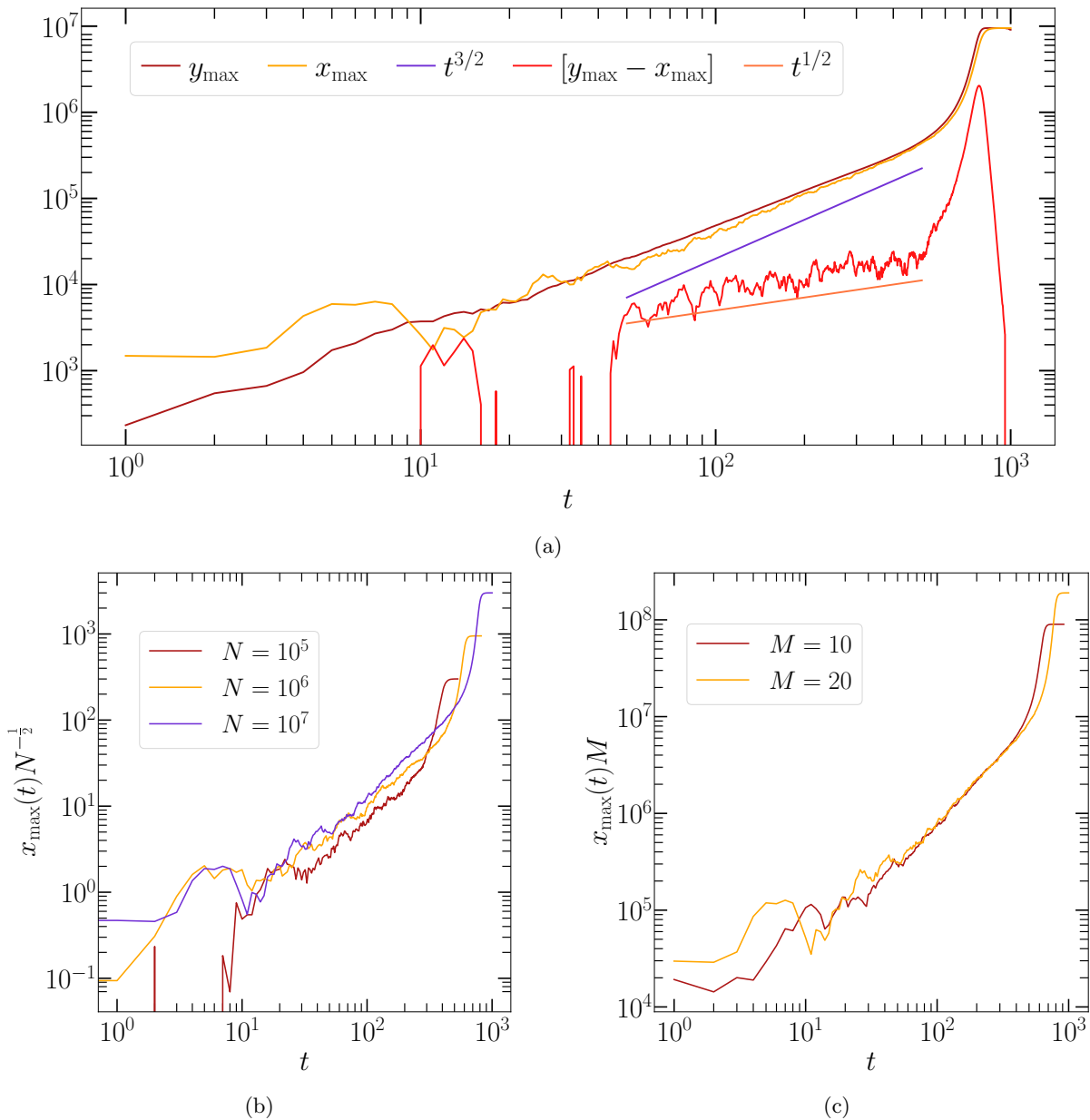


Figure 7: (a) Temporal evolution of $x_{\max}(t)$, $y_{\max}(t)$ and their difference for $N = 10^7$, $M = 20$, $\lambda \approx \lambda_c$. (b) Collapse plot to check the scaling (20) with respect to N ; $M = 20$ and $\lambda \approx \lambda_c$. (c) Collapse plot to check the scaling (20) with respect to M ; $N = 10^7$ and $\lambda \approx \lambda_c$.

Appendix B: Scaling of $x_{\max}(t)$

As discussed in the main text, near the transition point the quantity $x_{\max}(t) = N_{\max}(t) - N/M$ satisfies Eq. (20), that is

$$x_{\max}(t) = \gamma \frac{N^{1/2}}{M} t^{3/2}.$$

This is shown in Fig. 7, where we test both the evolution in time and the scaling with respect to M and N . The latter can be also explained by considering the dynamics during the first steps. Initially all $n_i^{(k)}$ are null and all opinions are equally common, so each agent, when updated, chooses an opinion with uniform probability. This regime lasts up to

$t < \bar{t}$, where \bar{t} is defined as the time when the $n_i^{(k)}$ becomes statistically different from zero and so the PI stops to be completely random. This implies that for fixed i and up to \bar{t} the M variables $n_i^{(k)}$ are distributed according to a multinomial distribution with uniform probability $1/M$. The multinomial regime ends when the maximum $\max_k [n_i^{(k)}]$ exceeds the average $\langle n_i^{(k)} \rangle_k$ by one, since when this occurs the external information stops to suggest random opinions. The mean value after t extractions satisfies $\langle n_i^{(k)} \rangle_k = t/M$, while by drawing from a multinomial we numerically determined that $\max_k [n_i^{(k)}] - \langle n_i^{(k)} \rangle_k \approx t^{1/2} M^{-1/4}$, see Fig. 8.

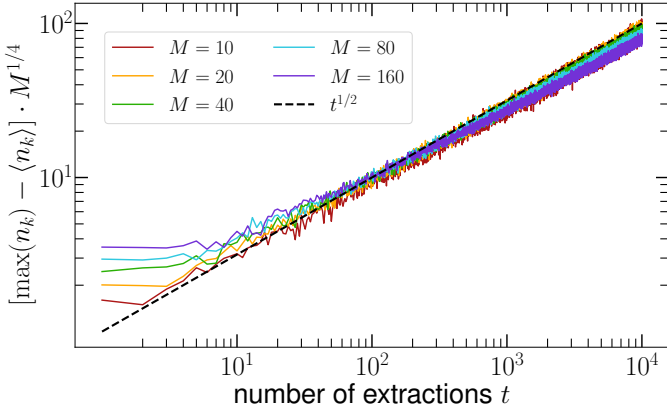


Figure 8: Growth of the maximum of a multinomial distribution with M equally probable classes as a function of the number of extractions t for different values of M . We denote by n_k the number of counts in the k th class. We plotted the quantity $\{\max[n_k] - \langle n_k \rangle\} M^{1/4}$ to highlight the scaling with respect to M and we averaged over 25 different realizations of the process.

We can then obtain the time \bar{t} as

$$\bar{t}^{1/2} M^{-1/4} = 1 \rightarrow \bar{t} = M^{1/2}. \quad (\text{B1})$$

Now we consider the evolution of the number of agents N_k in opinion k in the simple multistate voter model. The transition rates are

$$\begin{cases} W(N_k \rightarrow N_k \pm 1) = \frac{1}{\delta t} \frac{N - N_k}{N} \frac{N_k}{N} \\ W(N_k \pm 1 \rightarrow N_k) = \frac{1}{\delta t} \frac{N - (N_k \pm 1)}{N} \frac{N_k \pm 1}{N}, \end{cases} \quad (\text{B2})$$

where $\delta t = 1/N$. The master equation for N_k is thus

$$\begin{aligned} P(N_k, t + \delta t) = & P(N_k + 1, t)W(N_k + 1 \rightarrow N_k) + \\ & + P(N_k - 1, t)W(N_k - 1 \rightarrow N_k) + \\ & + [1 - W(N_k \rightarrow N_k - 1) - W(N_k \rightarrow N_k + 1)]P(N_k, t). \end{aligned} \quad N_{\max}(t) \approx \langle N_k(t) \rangle + \sigma_k(t) \rightarrow x_{\max}(t) \approx \sigma_k(t).$$

Expanding for δt small we get

$$\begin{aligned} \frac{\partial P(N_k, t)}{\partial t} \cdot \delta t = & W(N_k + 1 \rightarrow N_k)P(N_k + 1, t) \\ & + W(N_k - 1 \rightarrow N_k)P(N_k - 1, t) + \\ & - [W(N_k + 1 \rightarrow N_k) + W(N_k - 1 \rightarrow N_k)]P(N_k, t). \end{aligned} \quad (\text{B3})$$

Starting from this master equation we can compute the average drift and the variance of N_k . The average drift satisfies

$$\begin{aligned} \nu_k = \frac{d\langle N_k \rangle}{dt} &= \sum_{N_k} N_k \frac{\partial P(N_k, t)}{\partial t} \\ &= \frac{1}{\delta t} \sum_{N_k} [W(N_k \rightarrow N_k + 1) - W(N_k \rightarrow N_k - 1)]P(N_k, t) = \\ &= \sum_{N_k} \frac{dN_k}{dt} P(N_k, t) = \sum_{N_k} \nu_k(N_k) P(N_k, t), \end{aligned} \quad (\text{B4})$$

where we also introduced the drift $\nu_k(N_k) = [W(N_k \rightarrow N_k + 1) - W(N_k \rightarrow N_k - 1)]/\delta t$. Note that while the average drift ν_k determines how the mean value of the distribution $P(N_k)$ moves in time, the drift $\nu_k(N_k)$ allows to compute (neglecting diffusion) how a specific value of N_k evolves. Using Eqs. (B2) we get

$$\nu_k = 0 \rightarrow \langle N_k(t) \rangle = N_k(0) = \frac{N}{M}. \quad (\text{B5})$$

Analogously the evolution of the variance is

$$\frac{d\sigma_k^2}{dt} = \frac{d\langle N_k^2 \rangle}{dt} - 2\langle N_k \rangle \frac{d\langle N_k \rangle}{dt} = \frac{d\langle N_k^2 \rangle}{dt},$$

where the last equality follows from Eq. (B5). Exploiting Eq. (B3) we then obtain

$$\frac{d\langle N_k^2 \rangle}{dt} = \sum_{N_k} N_k^2 \frac{\partial P(N_k, t)}{\partial t} = 2 \left(\langle N_k \rangle - \frac{1}{N} \langle N_k^2 \rangle \right),$$

whose solution is

$$\langle N_k^2 \rangle(t) = \left(\frac{N^2}{M^2} - \frac{N^2}{M} \right) e^{-2\frac{t}{N}} + \frac{N^2}{M}.$$

The variance is thus

$$\sigma_k^2(t) = \langle N_k^2 \rangle(t) - \langle N_k \rangle^2(t) = \left(\frac{N^2}{M} - \frac{N^2}{M^2} \right) \left(1 - e^{-2\frac{t}{N}} \right). \quad (\text{B6})$$

As also shown in Fig. 1, for small times and small values of λ our model with personalized information behaves as the usual voter model. This is due to the fact that for $t < \bar{t}$ the external information is completely random and so, being all opinions equally numerous at the beginning, a voter model-like update or a personalized information update are equivalent. This implies that we can use Eqs. (B5) and (B6) to determine how N_{\max} and thus x_{\max} evolve for small times even in the presence of personalized information

By expanding Eq. (B6) for small times we thus obtain

$$x_{\max}(t) \approx \sqrt{2\frac{t}{N} \left(\frac{N^2}{M} - \frac{N^2}{M^2} \right)} \approx 2\sqrt{\frac{N}{M}} t \quad \text{for } t \leq \bar{t},$$

where again we assumed $M \gg 1$. Finally, we know from empirical evidence that for large times it holds $x_{\max}(t) \sim t^{3/2}$; this implies that the functional form of $x_{\max}(t)$ must be

$$x_{\max}(t) = \alpha t^{3/2} + 2\sqrt{\frac{N}{M}} t^{1/2}.$$

The prefactor α can now be determined imposing the $t^{3/2}$ scaling to become dominant when the binomial scaling $t^{1/2}$ ends, so for $t = \bar{t} = \sqrt{M}$. This gives

$$\alpha \bar{t}^{3/2} = 2\sqrt{\frac{N}{M}} \bar{t}^{1/2}$$

and so

$$\alpha = \frac{N^{1/2}}{M}.$$

This shows that for sufficiently large t it holds

$$x_{\max}(t) \sim \frac{N^{1/2}}{M} t^{3/2}.$$

Appendix C: Evolution of $N_k - N/M$

The transition rate for N_k in the presence of personalized information is

$$W(N_k \rightarrow N_k + 1) = (1 - \lambda) \frac{N - N_k}{N} \frac{N_k}{N} + \lambda P_{PI, \rightarrow k},$$

where the first term is the usual voter contribution, while the second one is due to PI. Denoting as $S_k = N s_k$ the number of agents whose PI is polarized along opinion k , we can write the latter as

$$P_{PI, \rightarrow k} = \frac{N - N_k - \sum_{j \neq k} S_k}{N} \frac{1}{M_s(t)}.$$

Let us explain this expression. Among all the N agents, those already with opinion k do not contribute to the transition rate to opinion k , while those whose PI is polarized along $j \neq k$ can make a transition toward k only by a voter update. As a consequence only $N - N_k - \sum_{j \neq k} S_k$ should be considered in computing the transition probability. Moreover the PI of such agents will be unpolarized and so we can assume that it suggests a random opinion, giving the factor $1/M_s(t)$. Analogously

$$W(N_k \rightarrow N_k - 1) = (1 - \lambda) \frac{N - N_k}{N} \frac{N_k}{N} + \lambda P_{PI, k \rightarrow}$$

with

$$P_{PI, k \rightarrow} = \frac{N_k - S_k}{N} \frac{M_s(t) - 1}{M_s(t)}.$$

Inserting these transition rates in Eq. (B4) we can write the drift of N_k as

$$\nu_k(N_k) = \frac{\lambda}{M_s} [N - S - M_s(N_k - S_k)],$$

where we introduced the total number of polarized agents as $S = \sum_k S_k$. The time evolution of N_k , neglecting diffusive fluctuations, is thus

$$\frac{dN_k}{dt} = \nu_k(N_k) = \lambda \left[\left(S_k - \frac{S}{M_s} \right) - \left(N_k - \frac{N}{M_s} \right) \right].$$

For $N \gg M$ and short times we can make the approximation $M_s \approx M$ and so defining $y_k = S_k - S/M$ we arrive at an expression for the time evolution of $x_k = N_k - N/M$

$$\frac{dx_k}{dt} = \lambda [y_k - x_k] + \text{diffusive terms.} \quad (\text{C1})$$

This expression provides additional support to our analytical approach. Indeed, as shown in Fig. 7, it holds $y_{\max}(t) - x_{\max}(t) \sim t^{1/2}$ and so Eq. (C1) predicts $x_{\max}(t) \sim t^{3/2}$, as actually observed in Fig. 7. Note that by y_{\max} we denote $\max_k [y_k]$.

Appendix D: Scaling regimes of the critical threshold

1. $M \ll N$

For $M \ll N$ during the first steps no opinion disappears and so we can make the approximation $M_s(t) = M$. Inserting this into Eq. (22) yields Eq. (23), that is

$$\begin{cases} t^* = \frac{4}{\lambda_c^2} \left[\sqrt{\frac{1 - \lambda_c}{M}} + \lambda_c + \sqrt{\frac{1 - \lambda_c}{M}} \right]^2 \\ \lambda_c = \frac{\gamma \frac{N^{1/2}}{M} (t^*)^{3/2}}{N + \gamma \frac{N^{1/2}}{M} (t^*)^{3/2}}. \end{cases} \quad (\text{D1})$$

From the expression for the polarization time t^* we see that there are two possible regimes.

A) $\frac{1 - \lambda_c}{M} \ll \lambda_c$

The first equation in Eq. (D1) becomes

$$t^* = \frac{4}{\lambda_c}$$

and replacing this expression into the second equation we obtain

$$\lambda_c = \frac{\gamma N^{1/2} M^{-1} \left(\frac{4}{\lambda_c} \right)^{3/2}}{N + \gamma N^{1/2} M^{-1} \left(\frac{4}{\lambda_c} \right)^{3/2}} \approx 2N^{-1/2} M^{-1} \lambda_c^{-3/2}.$$

This gives

$$\lambda_c = 2^{2/5} (NM^2)^{-1/5},$$

so that

$$\lambda_c M = 2^{2/5} N^{-1/5} M^{3/5}.$$

Our initial assumption $\lambda_c M \gg 1$ is then verified provided that

$$M^{3/5} \gg N^{1/5} \rightarrow M \gg N^{1/3}.$$

Recalling that we are assuming $M \ll N$ we then have

$$\lambda_c = 2^{2/5} (NM^2)^{-1/5} \text{ for } N^{1/3} \ll M \ll N.$$

Putting together the expressions for t^* and λ_c we finally obtain Eq. (25)

$$\begin{cases} t^* = 2^{8/5} (NM^2)^{1/5} \\ \lambda_c = 2^{2/5} (NM^2)^{-1/5} \end{cases}$$

B) $\frac{1 - \lambda_c}{M} \gg \lambda_c$

In this case the polarization time becomes

$$t^* = \frac{16}{M \lambda_c^2}$$

and again, by putting this expression into the expression for λ_c we obtain

$$\lambda_c = \frac{\gamma N^{1/2} M^{-1} 2^6 M^{-3/2} \lambda_c^{-3}}{N + \gamma N^{1/2} M^{-1} 2^6 M^{-3/2} \lambda_c^{-3}} \approx 2^4 M^{-5/2} N^{-1/2} \lambda_c^{-3}.$$

This gives

$$\lambda_c = 2(M^5 N)^{-1/8}$$

and so

$$M\lambda_c \approx 2N^{-1/8} M^{3/8}.$$

Consequently the two hypotheses $M \ll N$ and $M\lambda \ll 1$ are both satisfied if

$$M \ll N^{1/3}$$

and the expressions for t^* and λ_c are

$$\begin{cases} t^* \approx 4(MN)^{1/4} \\ \lambda_c \approx 2(M^5 N)^{-1/8} \end{cases}$$

i.e., Eq. (24).

2. $M = N$

For $M = N$ one can no longer make the assumption $M_s(t) = M_s(0) = M$, since some of the opinions disappear even during the first steps. In this case we can approximate the decrease of $M_s(t)$ exploiting Eq. (5), which is exact for a simple multistate voter model with $M = N$. The general expression is

$$M_s(t) = \frac{M}{1 + \frac{M}{N}t}$$

and putting it into Eq. (22) we get

$$t^* = \frac{\pm \sqrt{(16\frac{\lambda N}{M} - 8\lambda N - 16\frac{N}{M})^2 - 64N(\lambda^2 N + 16\lambda - 16) - \frac{16\lambda N}{M} + 8\lambda N + 16\frac{N}{M}}}{2(\lambda^2 N + 16\lambda - 16)},$$

which, for $M = N$, becomes

$$t^* = \frac{\pm \sqrt{(16\lambda - 8\lambda N - 16)^2 - 64N(\lambda^2 N + 16\lambda - 16) - 16\lambda + 8\lambda N + 16}}{2(\lambda^2 N + 16\lambda - 16)},$$

In the limit of large N and small λ this expression can be approximated as

$$t^* = \frac{\pm 32N^{1/2} + 8\lambda N}{2(\lambda^2 N - 16)} \quad (\text{D2})$$

and, as it is possible to see, there are two distinct solutions. Requiring the denominator to vanish gives

$$\bar{\lambda}^2 N - 16 = 0 \rightarrow \bar{\lambda} = \frac{4}{N^{1/2}}$$

Note that while the solution with the plus, t_+^* , diverges in this limit, the one with the minus, t_-^* is finite and positive, indeed

$$\begin{aligned} & \lim_{\lambda \rightarrow 4N^{-1/2}} \frac{-32N^{1/2} + 8\lambda N}{2(\lambda^2 N - 16)} = \\ & = 8N^{1/2} \lim_{\lambda \rightarrow 4N^{-1/2}} \frac{\lambda N^{1/2} - 4}{2(\lambda N^{1/2} + 4)(\lambda N^{1/2} - 4)} = \\ & = \frac{N^{1/2}}{2}. \end{aligned}$$

Moreover, while for $\lambda < \bar{\lambda}$ t_-^* is positive, the other one is negative, meaning that t_+^* is meaningful only in the region $\lambda > \bar{\lambda}$ since a time must be a positive quantity. For the solution with the minus we can then take the limit $\lambda \rightarrow 0$, which should give back the behavior of the multistate voter model

$$\lim_{\lambda \rightarrow 0} t_-^* = \lim_{\lambda \rightarrow 0} \frac{-32N^{1/2} + 8\lambda N}{2(\lambda^2 N - 16)} = N^{1/2}.$$

This result suggest that the solution t_-^* is non physical, since if $\lambda = 0$ it holds $n_1 = n_2$ (see Eqs. (17) and (16)) and so the two peaks should never split meaning that $t^* = \infty$. In conclusion the expression for the splitting time is the one with the plus and it is not defined for any value of λ , more precisely

$$t^* = \frac{32N^{1/2} + 8\lambda N}{2(\lambda^2 N - 16)} \text{ with } \lambda > \bar{\lambda} = \frac{4}{N^{1/2}}. \quad (\text{D3})$$

In the limit $\lambda \rightarrow 0$ and $N \rightarrow \infty$ we have two possible scaling regimes³⁶

A) $\lambda N > N^{1/2} \rightarrow \lambda > N^{-1/2}$

In this case we can approximate t^* as

$$t^* \approx \frac{8\lambda N}{2(\lambda^2 N - 16)} \approx \frac{4}{\lambda}$$

and substituting this expression into Eq. (21) we get

$$\lambda_c \approx \left(\frac{4}{N^3}\right)^{1/5} \sim N^{-3/5},$$

where we also exploited the fact that $M = N$. Note, however, that this result is in contrast with the initial assumption $\lambda > N^{-1/2}$ and so this scaling regime is impossible.

B) $\lambda \sim N^{-1/2}$

In this case we can set $\lambda^2 N = 16 + \epsilon$ and Eq. (D3)

becomes

$$t^* \approx \frac{32N^{1/2}}{\epsilon}.$$

Moreover, using Eq. (21), we obtain

$$\lambda_c = \gamma N^{-3/2} (t^*)^{3/2} = 2^{13/2} N^{-3/4} \epsilon^{-3/2}$$

and setting $\lambda_c = \sqrt{\frac{16+\epsilon}{N}}$ we get

$$\epsilon = 2^{11/3} N^{-1/6}.$$

This is consistent with ϵ being a small correction (that is our initial assumption) and moreover we see that it is the smaller the larger is N .

In conclusion we have

$$\begin{cases} t^* \approx 2^{4/3} N^{2/3} \\ \lambda_c \approx \frac{4}{N^{1/2}}, \end{cases}$$

that is Eq. (28).