Opinion dynamics with emergent collective memory: The impact of a long and heterogeneous news history

Gioia Boschi [a],[*], Chiara Cammarota [a],[b], Reimer Kühn [a]

[a] *Mathematics Department, Strand, London WC2R 2LS, UK*
[b] *Dipartimento di Fisica, Sapienza Università di Roma, P.le A. Moro 5, 00185 Rome, Italy*

A B S T R A C T

In modern society people are being exposed to numerous information, with some of them being frequently repeated or more disruptive than others. In this paper we use a model of opinion dynamics to study how this news impact the society. In particular, our study aims to explain how the exposure of the society to certain events deeply change people's perception of the present and future. The evolution of opinions which we consider is influenced both by external information and the pressure of the society. The latter includes imitation, differentiation, homophily and its opposite, xenophobia. The combination of these ingredients gives rise to a collective memory effect, which is triggered by external information. In this paper we focus our attention on how this memory arises when the order of appearance of external news is random. We will show which characteristics a piece of news needs to have in order to be embedded in the society's memory. We will also provide an analytical way to measure how many information a society can remember when an extensive number of news items is presented. Finally we will show that, when a certain piece of news is present in the society's history, even a distorted version of it is sufficient to trigger the memory of the originally stored information.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Nowadays, thanks to the world-wide diffusion of digital technologies, people are bombarded by a myriad of different news. With the increasing popularity of social media, news items incessantly appear on all screens, triggering people's attention and influencing the way in which they think and act. Understanding how opinions and choices are affected by external events has thus become of particular interest for different research fields, such as sociology, politics and marketing. In the last years, physicists have proposed many mathematical models [1–3] to understand how humans behave and form their opinions, giving birth to a new research field called sociophysics. Some of the most famous works in this field concern the use of models borrowed from statistical mechanics, such as the Ising model to study consensus [4–6]. Others introduced a variety of different models such as the Voter model [7,8] or the Majority rule model [9,10]. A part of these works is dedicated to the study of how the society reacts to external information [11,12] coming from a single source [13] or multiple ones [14].

In our recent paper [15] we presented a model of a society where opinions are formed through the combined effect of mutual influence among agents and driving through multiple external events. The key ingredient of the model is the way

---

in which agents interact: when two agents have often been in agreement, they tend to be in agreement also in the future and the same happens with disagreement. This behaviour is the result of the combination of different human tendencies largely studied and applied in different social models. Together with imitation, which is the main ingredient of all of the previously mentioned models, we consider differentiation [12,16–20], homophily (an attractive influence or the tendency to become more similar to people in agreement with us) [21,22] and its opposite, xenophobia (a repulsive influence or the tendency to become different from people in disagreement with us) [23–27]. Our study shows that the history of mutual agreement or disagreement between agents is encoded in the interaction matrix of our model in a way that prompts the emergence of a Hopfield-like [23,28] collective memory by which the society embeds the information coming from external news and recalls them in the future. Note that the memory effect which appears in our model is a memory of past *relations* and must not be confused with the memory of past actions or opinions of single agents that has been more often considered in the literature of social interactions [29–31]. Studies about collective memory appear instead in [32], while a model which consider memory of past relationships is presented in [33].

While the focus of our previous paper was on finite number of news cyclically presented, in this new work we study the behaviour of the same model but focusing our attention on items of information which are presented in random order and with non-uniform intensity. The exploration of this new scenario goes in the direction of making the model more realistic. In fact, in the real world, the news that a person may receive are plenty, with some of them appearing more often than others. We begin by exploring a simple scenario in which only a finite number of news items is presented, before moving on to more complicated ones in which information consists of extensively many items of news. In this second scenario we will use techniques common in statistical mechanics (i.e. replica calculations) to calculate the maximum number of pieces of information a society can retain in its collective memory.

In that context we also investigate the question, how well a society is able to collectively recall a piece of information embedded in its collective memory at some point in the past, if exposed to a news item which is a more or less distorted version of that ancient memory, thus bearing only imperfect similarity with it.

We will present our results by organizing the paper in the following way. In Section 2 we present the model analysed in the paper. In Section 3 we present the behaviour of the society under the effect of a finite number of news items presented in a random order. In Section 3.1 we show how the frequency and the strength of this news are determinant for their storage in the society's memory. In Section 3.2 we perform a similar analysis testing a large spectrum of parameters. In Section 4 we focus on a different setting in which many different news invest the society and we calculate how many among them will be effectively remembered, *i.e.* the storage capacity of our society, comparing analytical results with simulations. Finally in Section 4.2 we show how a society which has received a certain signal, is able to remember it when the noisy version of the same signal is shown.

2. The model

In this paper we will analyse a model, firstly introduced in [15], of a fully connected society of N interacting individuals which are influenced by the pressure of their peers and by an external disruptive signal. In particular, with each agent we associate a continuous variable u_i which evolves following

$$\dot{u}_i = -u_i + I_i + \sum_{j(\neq i)}^{N} J_{ij} v_j + \eta_i \ . \tag{1}$$

In this model, u_i represents the preference field of an agent i on an issue, while $v_i = g(u_i)$ represents the opinion he/she express to the other agents. We chose $g(u_j) = \mathrm{erf}(u_j)$ to be a sigmoid function of the preference field u_j, such that the expressed opinions remain bounded. The dynamics of u_i is driven by the agent's perception of external information I_i and the pressure of the society, defined as the weighted sum of the expressed opinions of all the agents. The weights J_{ij}, which change in time as described later, characterize the type of relation between the pair (i, j) of agents: a positive weight entails an assimilation of opinions while a negative one entails differentiation. Both, assimilation and differentiation are *graded*, as the weights are not just characterized by their sign, but also by their magnitude.

The evolution of u_i depends also on a mean reversion term $-u_i$ which entails that, when external influences are absent, the preference field of each agent fluctuates around zero. The last term η_i is a white noise with Gaussian distribution with zero mean and finite variance $\langle \eta_i(t)\eta_j(t')\rangle = \sigma^2 \delta_{ij}\delta(t-t')$ which accounts for stochastic effects in opinion dynamics which are not controlled by the other terms.

The novelty of our model is the way in which the couplings, and so the interpersonal relations of the agents', evolve in time. The tendency of two agents to agree or disagree is based on their history of agreement and disagreement. People which have a past history of agreement will be more likely to agree in future, and the analogous happens with disagreement. This behaviour is encoded in the couplings J_{ij} which evolve in time following

$$J_{ij}(t) = \frac{J_0 \cdot \gamma}{N} \int_0^t \mathrm{d}s \ v_i(s)v_j(s)e^{-\gamma(t-s)} \ . \tag{2}$$

Given two agents i and j, the product of their expressed opinions $v_i(s)$ and $v_j(s)$ is weighted with an exponentially decaying function in time. This entails that the more recent history has a larger influence than the distant past in determining the

sign and the strength of the interactions between the two agents. The rate of decay of this function is described by the constant γ; its inverse $\tau_\gamma = 1/\gamma$ thus determines a memory time scale, or more specifically, the length of the memory span. In this way our society uses the past history to interpret the instantaneous inputs that it receives from outside. The prefactor J_0 sets an overall scale for the strength of the interactions, while the N^{-1} scaling of couplings with system size N is to ensure that a large system limit of the dynamics exists. A detailed study of the behaviour of this model under the influence of constant or periodic external information can be found in [15]. In the case of constant external information, we observed that at stationarity the u_i are approximately Gaussian distributed $u_i \sim \mathcal{N}(\langle u_i \rangle, \sigma_u)$, with:

$$\langle u_i \rangle = \sum_{j(\neq i)} J_{ij} \langle v_j \rangle + I_i \tag{3}$$

$$\sigma_u^2 = \sigma^2/2 , \tag{4}$$

and

$$\langle v_i \rangle = \mathrm{erf}\left(\frac{\langle u_i \rangle}{\sqrt{1 + 2\sigma_u^2}} \right) \tag{5}$$

where σ^2 is the noise variance. In the scenarios which will be analysed in the present paper, the couplings keep evolving in time while the external signal changes. However, in order to *probe* the evolving memory of the society at any given point in time, we will use a protocol which determines the information processing capabilities of a system using couplings that are frozen at the values they attained at that moment in time. We will, for instance, see that after a long exposure to a random sequence of external news the couplings will reach statistical stationarity, entailing that information processing capabilities of the society become statistically stationary as well. In this paper we will use the Gaussian approximation used in [15] to obtain Eq. (3), (4) and (5) to analyse the society under the effect of different kinds of random external signals. Firstly we will study the behaviour of the society under the influence of a finite number of different news items presented at random. Then we will study a society influenced by an infinite sequence of different news items and we will focus on how many of these can be actually remembered.

3. Random presentation of a finite number of patterns

In our previous paper [15] on the model described by Eqs. (1)–(2), we examined the society under the effect of sequential periodic external news. In particular, the signal $I_i(t)$ as perceived by agent i was constructed in terms of different pieces of news, each switched on for a period Δ_0, with the different news items appearing periodically in an ordered sequence. Each piece of news $\boldsymbol{I}^\mu = I_0 \boldsymbol{\xi}^\mu = I_0(\xi_i^\mu)$, labelled by $\mu \in \{1, \ldots, p\}$ is characterized by its amplitude I_0, which was chosen to be uniform (i.e. μ independent) in [15], and a vector $\boldsymbol{\xi}^\mu$, with components ξ_i^μ describing the agent-dependent idiosyncratic perceptions of a news item. For example $\xi_i^\mu = +1$ means that the agent i has a positive attitude towards the piece of information μ. In [15] the ξ_i^μ were modelled as unbiased, binary i.i.d. random variables uniformly chosen from $\{\pm 1\}$. The resulting expression of the signal was:

$$I_i(t) = I_0 \xi_i^{\mu(t)} , \tag{6}$$

with

$$\mu(t) = 1 + \left\lfloor \frac{t}{\Delta_0} \right\rfloor \mod p , \tag{7}$$

In that case the order of presentation determined which information was most clearly remembered. More realistically, however, we expect to see news appearing in a less regular fashion, with some signals more frequent or stronger than others. In this section we show how the society responds when news items may have a variable strength or are received in a random order, each returning with a given probability. In this case we will have a finite set of possible news labelled by $\mu \in \{1, \ldots, p\}$ appearing with probabilities $\{\pi_1, \ldots, \pi_p\}$. The signal will be made of a sequence of patterns $I_0^{\mu_k(t)} \xi_i^{\mu_k(t)}$ switched on for a time Δ_0, each one labelled by a randomly chosen index $\mu_k(t) \in \{1, \ldots, p\}$, where k counts the presentation times:

$$I_i(t) = I_0^{\mu_k(t)} \xi_i^{\mu_k(t)} \qquad \text{for} \qquad (k-1)\Delta_0 < t < k\Delta_0 , \tag{8}$$

$$\mu_k(t) = \mu \qquad\qquad \text{with} \quad \mathrm{Prob}(\mu_k = \mu) = \pi_\mu . \tag{9}$$

Unlike in the previous paper, the strengths I_0^μ of the different news items are now μ-dependent. When the I_0^μ are sufficiently large, the agents' opinions will quickly align with the information received, so that u_i becomes very large with its sign determined by ξ_i^μ, and consequently $v_i = \mathrm{erf}(u_i) \simeq \xi_i^\mu$ almost immediately after the signal $\boldsymbol{\xi}^\mu$ is switched on. This corresponds to a situation in which the external information is disruptive and captures the attention of the whole society. In this case we are able to calculate the couplings J_{ij}, proceeding as in [15] for periodically presented patterns, splitting the integral over intervals of length Δ_0, during which $v_i(t) \sim \xi_i^\mu$. We adopt a convention different from that

used in [15], counting "backward" in time, and using $k = 1$ to label the final presentation period, $k = 2$ the previous one, and so forth. This gives

$$J_{ij}(t) = \frac{J_0}{N} \sum_{\mu=1}^{p} \xi_i^{\mu} \xi_j^{\mu} \sum_{k=1}^{N_p} w_k \, \delta_{\mu,\mu_k} \ , \qquad (10)$$

with N_p denoting the total number of news presentation periods up to time t, i.e. $t = N_p \Delta_0$, and

$$w_k = (1 - e^{-\gamma \Delta_0}) e^{-\gamma(k-1)\Delta_0} \ . \qquad (11)$$

The k-sum in Eq. (10) is a large weighted sum of Bernoulli random variables δ_{μ,μ_k}. It can be evaluated as a sum of averages by appeal to the Law of Large Numbers, giving

$$J_{ij} = \frac{J_0}{N}(1 - e^{-\gamma \Delta_0 N_p}) \sum_{\mu=1}^{p} \pi_{\mu} \xi_i^{\mu} \xi_j^{\mu} \ . \qquad (12)$$

We will be mostly dealing with the limit $\gamma \Delta_0 \ll 1 \ll \gamma \Delta_0 N_p$ for which $\sum_{k=1}^{N_p} w_k = 1 - e^{-\gamma \Delta_0 N_p} \simeq 1$, and for which it can be shown that corrections to Eq. (12) are negligible, being smaller than the dominant contribution by a factor $\mathcal{O}(\sqrt{\gamma_0 \Delta_0})$. If $\pi_{\mu} = 1/p$ for $\mu = 1, \dots, p$, couplings are (apart from a factor $1/p$) equal to Hebb–Hopfield couplings [28,34] used in the well known Hopfield model of associative neural network. This similarity was already discussed in [15], suggesting that our network of agents can behave in a manner analogous to a network of neurons, in which firing patterns are stored in the couplings. Instead of neural patterns we will have *opinion patterns* $\boldsymbol{\xi}^{\mu}$ that under suitable conditions can be stored and recalled by the society by means of a reshaping of individuals' interactions. For periodic news presentation, we showed in [15] that at stationarity each opinion pattern contributed to the couplings with a weight depending on the presentation order of the corresponding signal. Instead in Eq. (12) we see that, when news items are presented in a random order, after long time each pattern contributes to the couplings with a weight which depends only on the probability of appearance of its corresponding signal. In the context of this paper we will only analyse the scenario in which a very large number of signal presentations occur. This allow us to neglect the effect of the order of presentation of the patterns on each realization of the dynamics, which may be important when only few patterns are shown.

A measure of the similarity between the stored opinion patterns and a given piece of news μ is the overlap between the pattern μ and the system state

$$m^{\mu}(t) = \frac{1}{N} \sum_i \xi_i^{\mu} \langle v_i(t) \rangle \ . \qquad (13)$$

The overlap m^{μ} will be close to one during the presentation of the signal contribution μ and close to zero when other signal contributions are presented. Most importantly, when the signal is totally removed, the value of m^{μ} tells us whether the society's opinions remain aligned with one of the previously presented news.

To investigate this question, we evaluate the m^{μ} in a stationary regime with fixed external signal I_i, after couplings are frozen at their values given by Eq. (12). Exploiting the fact that the $u_i(t)$ in the stationary regime are Gaussian, with means and variances defined in terms of Eqs. (3), (4) and (5), we obtain

$$m^{\mu} = \frac{1}{N} \sum_i \xi_i^{\mu} \, \mathrm{erf}\left(\frac{\langle u_i \rangle}{\sqrt{1 + 2\sigma_u^2}} \right) = \frac{1}{N} \sum_i \xi_i^{\mu} \mathrm{erf}\left(\frac{\sum_j J_{ij}\langle v_j \rangle + I_i}{\sqrt{1 + \sigma^2}} \right)$$
$$= \frac{1}{N} \sum_i \xi_i^{\mu} \, \mathrm{erf}\left(\frac{J_0 \sum_{\nu} \xi_i^{\nu} \pi_{\nu} m^{\nu} + I_i}{\sqrt{1 + \sigma^2}} \right) \ , \qquad (14)$$

i.e., a set of self-consistency equations for the overlaps $\{m^{\mu}\}$, $\mu = 1, \dots, p$. Here we have used the approximation $1 - e^{-\gamma \Delta_0 N_p} \simeq 1$ valid in the large N_p limit as discussed above. For large N, the fixed point equations can be expressed in terms of averages over the statistics of the ξ_i^{μ} and the I_i by appeal to the LLN, giving

$$m^{\mu} = \left\langle\!\!\left\langle \xi^{\mu} \mathrm{erf}\left(\frac{J_0 \sum_{\nu} \xi^{\nu} \pi_{\nu} m^{\nu} + I}{\sqrt{1 + \sigma^2}} \right) \right\rangle\!\!\right\rangle_{\{\xi^{\mu}\},I} \ , \qquad (15)$$

in which the double angled brackets denote an average over the distribution of the $\{\xi^{\mu}\}$ and the external stationary signals I. In much of what follows we will be interested in so-called retrieval solutions in the absence of external signals, for which it is assumed that the system state is aligned with only a single opinion pattern, so that the vector $\mathbf{m}$ of overlaps has only a single non-vanishing component, $\mathbf{m} = (0, \dots, m^{\mu}, \dots, 0)$, for some $\mu \in \{1, \dots, p\}$, for which the system of self-consistency Eqs. (13) simplifies to

$$m^{\mu} = \mathrm{erf}\left(\frac{J_0 \pi_{\mu} m^{\mu}}{\sqrt{1 + \sigma^2}} \right) \ . \qquad (16)$$
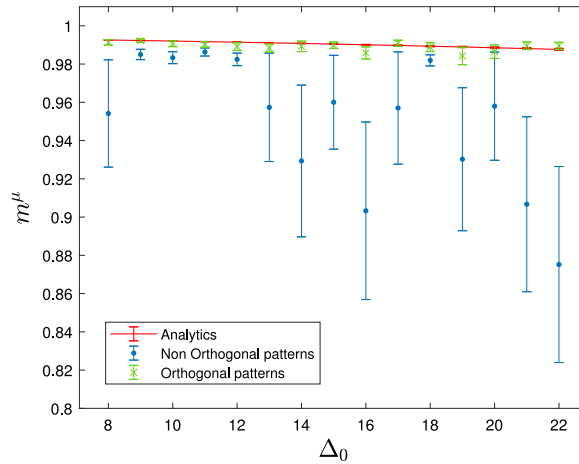
Fig. 1. The figure shows the comparison between the analytical (full line) and simulated (dots) value of m^μ as a function of the presentation time length Δ_0, for $J_0 = 6$, $I_0 = 10$ and $\pi_\mu = 1/3$ $\forall \mu$.

The solution of Eq. (16) can be compared to the value of m^μ estimated by simulating the model dynamics described in Eqs. (1), (2), (8), using presentation probabilities π_μ for the different patterns. Equations of motion are integrated using a simple Euler algorithm. During the dynamics the couplings evolve and reach a plateau corresponding to their stationary value. When this happens we stop the dynamics and we freeze the couplings at their current value. We then switch on each signal pattern in turn for a brief period, sufficient to have the society aligned with the patterns, after which the signal is removed. We then wait until the value of m^μ converges to a "stationary" level around which, due to the noise, it oscillates with a small variance. The final average value of m^μ is evaluated averaging over 700 values of $m^\mu(t)$ at stationarity. Results are then averaged over 10 simulations performed with different random pattern realizations.

In the simulation we assume that $I_0^\mu = I_0$ and – initially – for $\pi_\mu = 1/p$ for all μ. In this way all the patterns have the same probability of appearance and the same strength. Other details of how simulations are performed can be found in Appendix C. In Fig. 1 we compare the results for m^μ as a function of the presentation period Δ_0 as obtained from simulations with the analytical predictions. For the simulations we use systems of size $N = 100$, and $p = 3$ for the number of patterns which, as argued in [15], is representative of a system at low loading. As Δ_0 becomes large, results for random patterns begin to deviate from the theoretical predictions. This is mainly due to the effect that random patterns have non-zero mutual overlaps in systems of finite size N, which is in contrast to assumptions used in the derivation of the fixed-point equation for retrieval solutions in the thermodynamic limit $N \to \infty$. Indeed, using a set of mutually orthogonal patterns in the simulation, i.e. random patterns extracted with the condition that their scalar product is exactly null, we observe a much better agreement between analytical predictions and results from simulations.

To better understand this aspect we must carefully think of the role played by non-orthogonality: given two positively correlated patterns, the memory of the first gets reinforced while the second is presented and vice versa. When two patterns are anti-correlated, the opposite happens and their memory is instead weakened. This phenomenon is amplified when Δ_0 is large because the same patterns have more time to get mutually reinforced or weakened as the case may be. The creation of this pattern imbalance may lead to some patterns having a smaller overlap than others, or even being forgotten, which explains the deviation from predictions by numerical data for non-orthogonal patterns.

When the signal strength I_0^μ and the probability π_μ vary between patterns, the behaviour of the society is non trivially affected. The influence of varying these two parameters is investigated in the following by performing two different numerical experiments. In the first experiment, described in Section 3.1, we show how the recovery of a piece of news depends on the probability of its appearance and on its strength in the news stream. In the second experiment, described in Section 3.2, we investigate the recovery of patterns when their parameters are chosen at random, such to observe how the probability of a piece of news to be embedded in the society depends on the properties of the *other* news also presented in the same dynamics.

3.1. Storing random news with different strengths and frequencies of appearance

In the previous experiment we analysed the behaviour of a society under the effect of news items presented randomly with equal probability. With the aim of modelling a more realistic scenario we consider now an external signal made of random news which have a different probability of being presented to the society as well as different intensities. To start this new analysis we will assign to the first pattern a probability π_1 and a signal strength I_0^1 and equal parameters to the other two with $\pi_2 = \pi_3 = (1 - \pi_1)/2$ and $I_0^2 = I_0^3$. In this experiment we will follow the same simulation protocols as before, except that we do not average the value of m^μ over several realizations. Rather than recording average overlaps, we
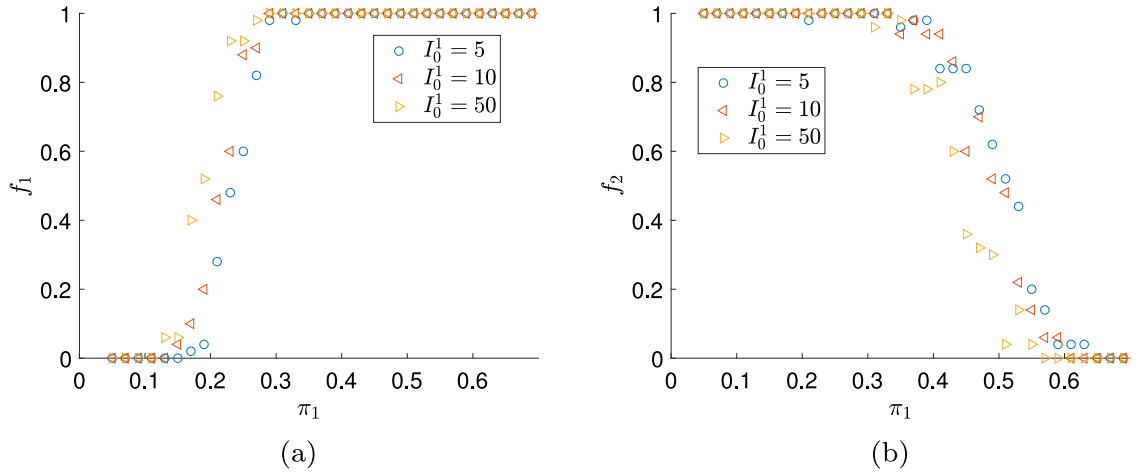
Fig. 2. (a) Fraction of instances in which pattern one is recovered as a function of its probability of appearance. (b) Same for pattern two. The quantities f_1 and f_2 have been calculated over 50 tests and then normalized on the interval $[0, 1]$. The simulations have been performed using the following parameters: $I_0^2 = I_0^3 = 10, J_0 = 8, \Delta_0 = 10$.
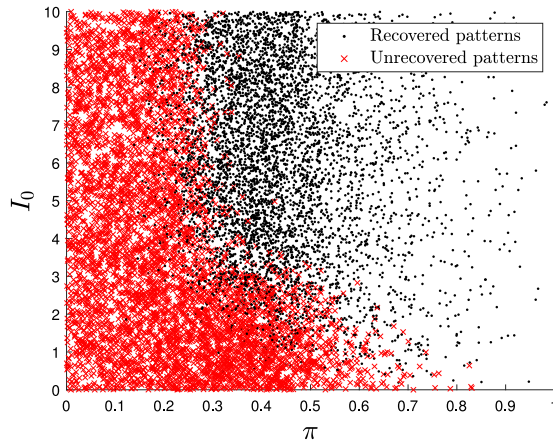

Fig. 3. Black and red dots, corresponding to recovered and unrecovered patterns, are plotted as a function of their signal strength and probability of presentation. The parameters of the simulation are $N = 100, J_0 = 8$ and $\Delta_0 = 17$.

measure the fraction of instances f_μ in which m^μ exceeds a critical value $m_c = 0.4$. This value is chosen to be significantly larger than the value of random mutual overlaps between different patterns (see Appendix C for other simulations details.)

In Fig. 2 we show f_1 and f_2 as a function of π_1 for different values of I_0^1 keeping $I_0^2 = 10$ fixed. For all signal strengths, the fraction f_1 increases steeply as the pattern probability π_1 increases. We expect that for $N \to \infty$ the curve presents a discontinuous transition between 0 and 1. As anticipated, when the probability π_1 grows beyond 0.33 (the point where the three patterns are equally probable), the pattern $\mu = 1$ is always recovered, while the other two are progressively less frequently remembered. However the most interesting effect is linked to the signal strength. In a real society we expect that the intensities I_0^μ of external information are not equal for all pieces of news but that some are significantly weaker or stronger than others. In Fig. 2 we show the effect of this variability in our model. The recovery of the first pattern becomes easier for $I_0^1 > I_0^2 = I_0^3$ and more difficult for $I_0^1 < I_0^{2,3}$. This result suggests that even a rarer piece of news, if particularly intense and disruptive, can leave a strong mark on the society.

In Fig. 2 we can also notice that the increased strength of pattern one results in a more difficult recovery of pattern 2 even if its signal strength is not changed. This effect suggests that the possibility for a piece of news to impress the society depends not only on its own immediate impact but also on the impact of other news that is present in the news stream. This means that a piece of news strong enough to be remembered on its own can be forgotten when presented alongside stronger or more frequent information. This can explain, for example, how news about politics that would be remembered when broadcast alone may be forgotten when presented with other shocking news, such as those about terror-attacks or earthquakes.

3.2. Storing random news with random intensities and frequencies of appearance

In the previous experiment we have shown that, when we fix the parameters of two pieces of information, the third one presented in the same experiment is more likely to be recovered if presented more often or if it is presented with bigger intensity. This gives us only a partial information on how our society reacts to random news. To complete this picture, we investigated the recovery of opinion patterns using a broad range of parameters. Using the protocol of Section 3.1, we now measure the recovery of patterns in a society where the signal strengths and probabilities of appearance are chosen uniformly at random in the intervals $\pi_\mu \in [0, 1]$, with $\sum_{\mu=1}^p \pi_\mu = 1$ and $I_0^\mu \in [0, 10]$. As before we take $p = 3$. The total number of samples of triplets taken is 3600. In Fig. 3 we plot a map of recovered (black dots) and unrecovered (red crosses) patterns as a function of their probabilities of appearance and their strengths. For weak signals the patterns will not be remembered by the society even when they appear very often. For larger signal strengths, there is a separation between recovered and unrecovered patterns around $\pi = 0.3$.

We will call recovery zone the region of the graph where pattern recovery is predominant and will refer to the complementary region as non-recovery zone. However Fig. 3 shows that some of the patterns around the boundary between the two zones do not behave as expected, e.g. some news around $I_0^1 = 5$ and $\pi = 0.2$ are recovered, while some others around $I_0^1 = 5$ and $\pi = 0.4$ are not. More generally, the separation between the two zones is not sharp and we observe a large transition area. A smooth transition is to be expected due to finite size effects, as discussed earlier. In fact, we present patterns only for a finite number of times in any realization of the dynamics, giving rise to fluctuations in their effective frequency of appearance. However, other aspects may play a role, broadening even further the transition area visible in Fig. 3. For instance as also discussed in the previous subsection the relative strength (and in this case the relative probability of appearances) of concomitant news in the recent history can result in a systematically more frequent recovery of news that would be otherwise unrecovered and vice versa. An attempt to disentangle these two aspects is presented in Appendix A.

4. Extensively many patterns and society's storage capacity

Real-life news are divulged with different intensity and characterized by different frequency of appearances. Moreover, it is also conceivable that the number of different news the society receives should not be considered to be very small in comparison to system size, as assumed when analysing the situation of a finite number of external signals that we have considered so far.

We present here the analysis of a society hit by a stream of different news items, with each pattern μ presented only once. For each presentation, a pattern is generated choosing each component ξ_i^μ uniformly at random from $\{\pm 1\}$. In particular, we will investigate the number of associated opinion patterns that the society can effectively remember, or the *storage capacity* of the society, whose interactions have been shaped by such a history of news. In order to store a large number of patterns, proportional to the number N of agents of the system, the society needs to receive a number of news items of order N *within* the memory time $\tau_\gamma = 1/\gamma$. For this reason we need to consider a scaling of the memory decay rate of the form $\gamma = \gamma_0/N$, as it is needed to ensure that an extensive number of patterns remains within "memory range". In the initial formulation of the model (Eq. (2)) we used an explicit scaling factor $1/N$ in the couplings in order to get a meaningful theory in the thermodynamic limit $N \to \infty$. In the present case, there is no need to introduce such an explicit scaling factor by hand, as the correct scaling follows automatically from the scaling of γ. Adopting such a scaling, the couplings J_{ij} are thus seen to take the form

$$J_{ij} = \frac{J_0 \gamma_0}{N} \int_0^t \mathrm{d}s\, v_i(s) v_j(s) e^{-\gamma_0(t-s)/N} \ . \tag{17}$$

We should note that this scaling is correct in the limit of a large number of agents (thermodynamic limit), only when the society receives a series of different random news. We consider an external signal made of a sequence of news labelled by $\mu = 1, 2, 3, \ldots, p$, with $p = \alpha_{max}N$. Given that we are eventually considering a history of infinitely many patterns presented, i.e. $\alpha_{max} \to \infty$, it is convenient to count them in reverse order, with $\mu = 1$ indicating the most recent pattern received and $\alpha_{max}N$ indicating the oldest one. We should remark that α_{max} tending to infinity does not mean that the society is able to store an infinite number of patterns, but only that the number of patterns presented is very large. The time dependent external signal will thus be of the form

$$I(s) = I_0 \xi_i^\mu \qquad (\mu - 1)\Delta_0 < t - s < \mu \Delta_0 \ . \tag{18}$$

As before, we assume signal intensities to be large. Each signal pattern μ is presented for a time Δ_0, during which the society's opinions are well aligned with the signal with $v_i(s) \simeq \xi_i^\mu$. We can use this to evaluate the integral in Eq. (17), splitting it in pieces of length Δ_0 and calculating them separately as described in [15]. The resulting couplings for small $\gamma_0 \Delta_0$ are:

$$J_{ij} = \frac{J_0 \tilde{\gamma}_0}{N} \sum_{\mu=1}^{\alpha_{max}N} \xi_i^\mu \xi_j^\mu e^{-\frac{(\mu-1)\tilde{\gamma}_0}{N}} + o(\tilde{\gamma}_0^2) \ . \tag{19}$$

where we introduced $\tilde{\gamma}_0 \equiv \gamma_0 \Delta_0$. We will use α_c to denote the storage capacity of the society, i.e., $\alpha_c N$ is the maximum number of patterns that the society is able to recall. The concept of storage capacity was introduced in neural networks and first computed for the Hopfield model in [35]. Our couplings in Eq. (19) are in fact reminiscent of a particular type of neural network model, the Hopfield model with forgetful memory [36–39]. In the present paper, we investigate the storage capacity of a society in the noiseless limit of our model (1). In this case, the dynamics of the u_i is known to be governed by a Lyapunov function [40,41], provided the interaction matrix is symmetric and the function v_j describing the expressed opinions as functions of the preference fields u_j are monotone increasing functions of their argument. Both conditions met in our case. Following [42,43] we can locate the minima of the Lyapunov function, and thus the attractors of the dynamics, by taking the zero-temperature limit of the free-energy of a system with the Lyapunov function as its energy function. The analysis can be found in Appendix B. We will describe the collective properties of the society using three order parameters:

1. the overlap between the society state and a given opinion pattern presented by the signal (taken to be pattern μ)

$$m = \frac{1}{N} \sum_i \xi_i^{\mu} \overline{\langle v_i \rangle} , \qquad (20)$$

   where angled brackets denote a thermal average and over-bars an average over the disorder embodied by the other opinion patterns $\nu (\neq \mu)$ embedded in the society. For a finite number of patterns, this parameters corresponds to the one defined in Eq. (13).

2. the mean of averaged squared opinions, which are the off-diagonal elements of the replica symmetric matrix of Edwards–Anderson order parameters (see Appendix B)

$$q = \frac{1}{N} \sum_i \overline{\langle v_i \rangle^2} , \qquad (21)$$

3. a susceptibility-type parameter

$$C = \beta(q_d - q) = \frac{\beta}{N} \sum_i \left(\overline{\langle v_i^2 \rangle} - \overline{\langle v_i \rangle^2} \right) , \qquad (22)$$

   where the q_d are diagonal elements of the replica symmetric matrix of Edwards–Anderson order parameters.

Making use of replica theory (details of the calculations in Appendix B) we obtain the following three fixed point equations, the solutions of which self-consistently determine these three order parameters and their dependence on the parameters characterizing the system:

$$
\begin{aligned}
m &= \left\langle\!\!\left\langle \xi^{\mu} \hat{v} \right\rangle\!\!\right\rangle , \\
C &= \frac{1}{J_0 \sqrt{r}} \left\langle\!\!\left\langle z \hat{v} \right\rangle\!\!\right\rangle , \\
q &= \left\langle\!\!\left\langle \hat{v}^2 \right\rangle\!\!\right\rangle ,
\end{aligned} \qquad (23)
$$

where

$$\hat{v}(\xi^{\mu}, I, z) = g\left(m\xi^{\mu} J_0 \tilde{\gamma}_0 e^{-\tilde{\gamma}_0 \alpha} + J_0 \sqrt{r} z + I - J_0 \left(1 + \frac{\ln(1 - J_0 \tilde{\gamma}_0 C)}{J_0 \tilde{\gamma}_0 C} \right) \hat{v} \right) , \qquad (24)$$

with

$$r = \frac{q}{J_0 C} \left[\frac{1}{1 - J_0 \tilde{\gamma}_0 C} + \frac{\ln(1 - J_0 \tilde{\gamma}_0 C)}{J_0 \tilde{\gamma}_0 C} \right] , \qquad (25)$$

and where z is a normally distributed random variable of zero-mean and unit variance. The double angled brackets denote an average over the site-random variables, i.e ξ^{μ} and I, and over the Gaussian random variable z. We note that the equations above have been obtained under the assumption that during every presentation time the society is perfectly aligned with the signal patterns presented i.e. in the large I_0 limit. The I appearing in Eq. (24) is not the signal described by Eq. (18) that appeared during the history of pattern presentations to the society, but a node-dependent perceived signal to which agents of the society may be subjected at the end of an extensive number of such pattern presentations. In the following subsections we will study two different retrieval scenarios, one is the spontaneous retrieval of patterns ξ^{μ} obtained setting I to 0, the other is the retrieval of old patterns when one of them is presented again, though in the form of a weak randomly distorted version of the original.

In both cases the problem of solving z-dependent fixed-point equations within the set (23) of fixed-point equations is avoided by transforming the Gaussian z- distribution into a $\hat{v}$-distribution as done in [43] and then taking the averages respect to the $\hat{v}$-distribution.
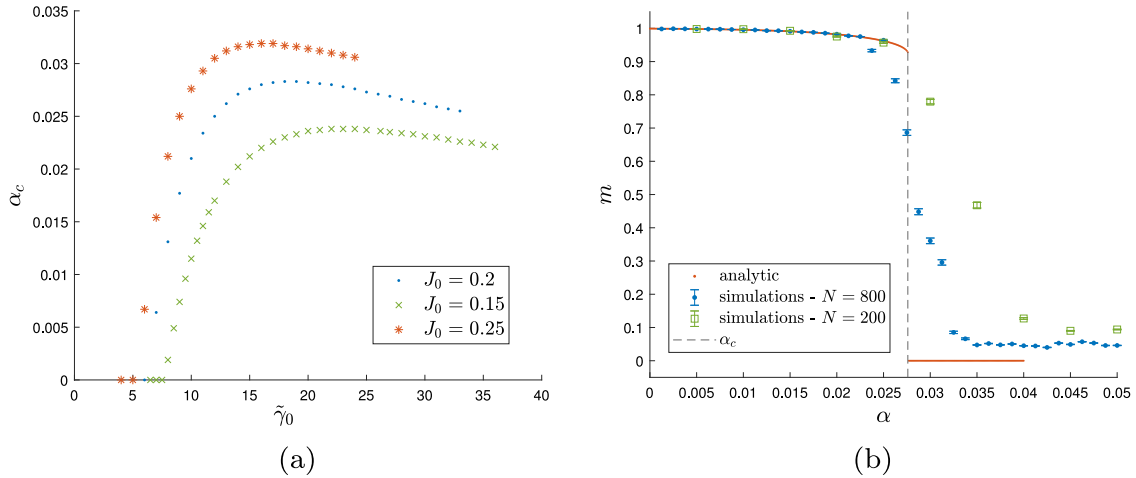
Fig. 4. Spontaneous retrieval of patterns in a society with an infinite history of news. (a) Critical capacity of the society as a function of $\tilde{\gamma}_0$ for different values of J_0. (b) Comparison between analytical solutions of m for different values of α at $\tilde{\gamma}_0 = 15$ and $J_0 = 0.2$ and simulations of societies of different sizes with Gaussian noise with mean 0 and variance $\sigma^2 = 0.01$. The dashed line corresponds to the analytical critical capacity.

4.1. Spontaneous retrieval

We solved Eqs. (23) numerically for $I = 0$ and different values of α and $\tilde{\gamma}_0$. We expect this system to have two solutions when $\alpha \leq \alpha_c$, one with $m = 0$ and one with $m > 0$, the latter disappearing for $\alpha > \alpha_c$. We thus computed the storage capacity α_c as the threshold of α after which the system has a unique solution with $m = 0$. In Fig. 4(a) the values of α_c found for different values of J_0 are plotted as a function of $\tilde{\gamma}_0$. Whatever the value of J_0, the storage capacity is zero when $\tilde{\gamma}_0$ is too small; it becomes non-zero at a critical value of $\tilde{\gamma}_0$ and grows until it reaches a peak and then decreases again. A small value of $\tilde{\gamma}_0$ corresponds to a long memory time or very short pattern presentation periods. This results in a society being exposed to too many patterns within its memory time, which interfere with each other and make the retrieval of these memories impossible. Conversely, when $\tilde{\gamma}_0$ is too large, the memory of the patterns fades too quickly for a large number of patterns to be remembered. These phenomena are similar to those observed in [36–39] in the context of forgetful Hopfield networks.

In Fig. 4(b) we show an example of how the solution jumps from $m \sim 1$ to $m = 0$ when $\alpha > \alpha_c$. The numerical solution is compared to the values of m found using the simulated dynamics at low temperature for the model described in Eq. (1). In the simulations we present a sequence of $p = N$ signal patterns to a society of $N = 800$ and $N = 200$ individuals. We record the values of m in absence of signal as described at the end of Section 3 and we averaged them over 50 simulations with different patterns realizations. Other simulation details can be found in Appendix C. Fig. 4(b) shows the values obtained as a function of $\alpha = \mu/N$, which are in good agreement with the theory. Given that we have a finite N, the transition from one solution to the other at $\alpha = \alpha_c$ is not abrupt, but is rounded and becomes steeper with increasing N. For the same reason there is a non-zero overlap of order $\mathcal{O}(1/\sqrt{N})$ between the patterns, which results in a residual m for $\alpha > \alpha_c$.

4.2. Noisy signal

In this last subsection we study the ability of a society to retrieve patterns of news after having been exposed to an infinite series of such news items. In particular, we focus on what happens when one of these pieces of news is presented again to the society, albeit at low intensity and distorted by some amount of noise. How will the society respond? Will it develop a reaction in line with the new pattern *as presented*, or will the interpersonal relations formed in response to the society's history of news exposures allow it to recognize and retrieve the information in its "pure" form as previously stored? The answer depends not only on the strength of the signal presented and on how much it is distorted by the noise, but also on how long ago in the past the society had been exposed to the original un-distorted version of it. In other words, it depends on how well the memory of the original opinion pattern is still embedded in the system. In order to model this situation, we consider a noisy signal of the form:

$$I_i = \widetilde{I}_0 \xi_i^\mu + \sigma_I z_i \tag{26}$$

where ξ^μ is one of the news presented in the society's history, $\widetilde{I}_0$ the amplitude of the new signal to be presented and $\sigma_I z_i$ is a Gaussian noise with mean 0 and variance σ_I^2.

First we comment on the average behaviour of the society. Fig. 5 shows the overlap m with the original pattern for different values of σ_I and $\widetilde{I}_0$ at $\tilde{\gamma}_0 = 10$ and $\alpha = 0.012$. In panel (a) it is visible how the retrieval of the original opinion
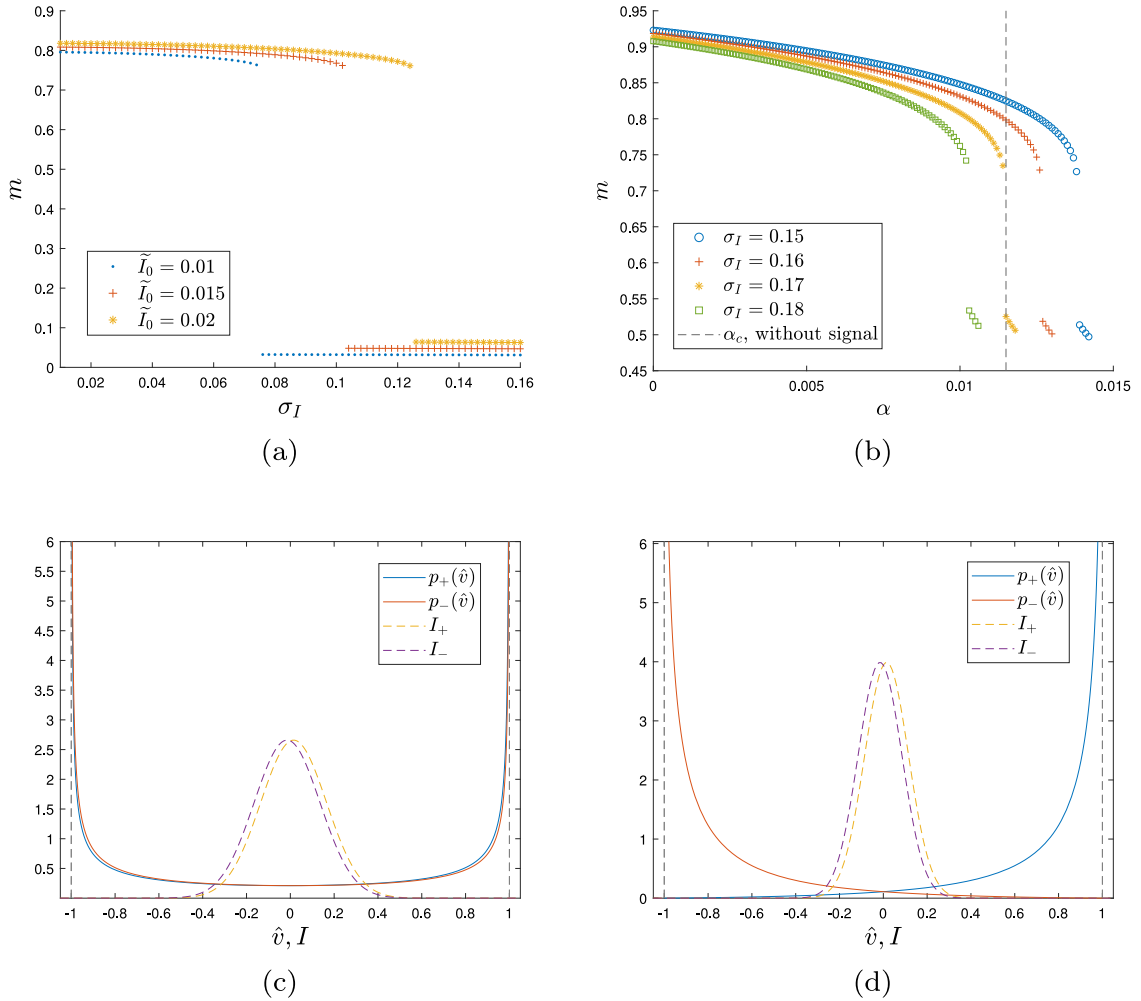
Fig. 5. Recovery of patterns through a noisy signal in a society which has received an infinite history of news. (a) Solution of m as a function of the level of noise of the applied signal, for different values of $\widetilde{I}_0$, given $J_0 = 0.15$, $\tilde{\gamma}_0 = 10$ and $\alpha = 0.012$. (b) Solution of m as a function of α for different level of noise and fixed $\widetilde{I}_0 = 0.2$, $J_0 = 0.15$ and $\tilde{\gamma}_0 = 10$. The dashed line indicates the value of α_c in absence of the noisy signal. (c) Plot of the distribution of expressed opinions $p_\pm(\hat{v})$ (corresponding to $p(\hat{v})$ at $\xi^\mu = \pm 1$), for $\alpha = 0.012$, $J_0 = 0.15$, $\widetilde{I}_0 = 0.015$ and $\sigma_I = 0.15$. The dashed lines correspond to the plots of the noisy signals with average $\pm \widetilde{I}_0$. (d) Same plot with $\sigma_I = 0.1$.

pattern becomes more difficult as σ_I grows, until the pattern is no longer recovered. Naturally, as $\widetilde{I}_0$ gets bigger, the point at which this happens moves towards larger values of σ_I. Conversely, when the noise is too disruptive the pattern is mostly not retrieved but the presence of the signal still guarantees a non zero value of m.

We show in panel (b) how the distortion of the signal also changes the maximal 'age' α_c of a news item that the society is able to retrieve when exposed to a distorted version of it. While we use the same symbol as for the storage capacity, i.e. the maximum number of opinion patterns a society can *spontaneously* retrieve, these quantities are obviously not strictly the same. In particular in absence of distortion, the presence of a signal aligned with one news already seen produces a larger α_c than the storage capacity of spontaneous retrieval. In other words, the presence of $\widetilde{I}_0$ allows the retrieval of opinion patterns that would have not been recovered spontaneously. When some amount of distortion is also present, how far in the past these opinion patterns can have been presented and still be recalled, clearly depends on the level σ_I of distortion of the pure memories. Fig. 5(b) shows how the value of α_c at which the solution for m jumps between a high and a low value, depends on σ_I when the distortion is added to a signal of strength $\widetilde{I}_0 = 0.2$. When the noise increases, the degree of similarity with an item originally stored decreases implying that older opinion patterns, which are less strongly embedded, cannot be recalled any longer; this is confirmed by the fact that the corresponding α_c decreases with increasing level σ_I of pattern distortion.

These results on the average behaviour can be complemented by looking at the distribution of expressed opinions, which reveal the emergence of an interesting behaviour. Figs. 5(c) and 5(d) show how the distributions of expressed opinions in response to a distorted incoming piece of news, change in a very non trivial way with its noise level σ_I. The

functions I_+ and I_- show the distribution of the external fields corresponding to positive and negative values of ξ_i^μ, with

$$I_\pm = \pm\widetilde{I_0} + \sigma_I z_i . \tag{27}$$

The function $p_+(\hat{v})$ is the distribution of expressed opinions of all agents i which originally received the signal $I = I_0\xi_i^\mu$ with $\xi_i^\mu = +1$, while $p_-(\hat{v})$ is the same for $\xi_i^\mu = -1$. When expressed opinions have the same sign as the original signal we can conclude that the society has successfully retrieved the corresponding news.

Panels (c) and (d) show two scenarios in which the signal is in both cases very distorted and the distributions of I_+ and I_- barely differ, being just slightly shifted versions of each other, yet the response of the society is very different. Panel (c) shows that in presence of a very large distortion $p_+(\hat{v})$ and $p_-(\hat{v})$ are approximately symmetrical and almost identical. This means that the society is not able to retrieve the original opinion pattern and individuals' opinions are *uncorrelated* with the original un-distorted opinion pattern the society had been exposed to in the past.

A small reduction of the noise level or the amount of distortion of an original news item leads to a completely different behaviour in terms of the distributions of expressed opinions. The results corresponding to this scenario are shown in panel (d), with distributions of expressed opinions being strongly asymmetric and biased in the direction of the original *un-distorted* signal with, $p_+(\hat{v})$ $(p_-(\hat{v}))$ peaked at 1 (-1) only. In this case individual's opinions are almost perfectly aligned with the originally presented information, despite the fact that the society is exposed to an external signal which has very little resemblance with it.

The important message delivered by these results is that even a weak and distorted signal can trigger the society to retrieve old stored memories. The society described by our model is able to remember information from the past, even when exposed to very noisy versions of it. The collective memory that emerges in this context connects distorted pieces of information to their clear un-distorted versions embedded in the collective memory. The maximum amount of noise tolerated for this to happen depends on the signal strength $\widetilde{I_0}$ and the value of α, *i.e.* how far in the past an original news item has impacted the society. In a real society shaped by a history of intense events, like terror attacks, earthquakes, or other crises, a new signal which resembles one of the old memories, even if only barely, can trigger in the population the same reactions it had shown in the past. These results open the way to design social experiments to understand how collective memory is formed and activated or re-activated by new information, and to test whether this mechanism can explain strong and pronounced collective responses to news items that do not appear to be particularly disruptive.

5. Conclusions

In this paper we studied how exposure to strong external information can shape the behaviour of a society within a model of opinion dynamics introduced in our previous paper [15] that produces interactions between individual agents based on their recent history of mutual agreement or disagreement. In this way people tend to agree with others if they have a history of predominant of mutual agreement, whereas they will be more likely to disagree with them in case of a history of predominant recent disagreement. This mechanism gives rise to a collective memory effect by which a society can remember past configurations of opinions. As mentioned initially, the opinion configurations we are interested in are those produced in reaction to external news. In the present paper we studied the effect of random news and how properties, such as their frequency of appearance, their strength or their relative intensity with respect to other news, determine whether the corresponding opinion patterns can be remembered or not. Unsurprisingly, we have seen that opinions in reaction to strong and frequent news are more easily remembered. However, we observe that even rare signals can have a deep impact on the society if they are powerful enough. Moreover, we showed that the memory formation of an opinion pattern depends not only on its own strength and frequency of appearance, but also on the characteristics of the other news presented within the memory span of a society. In fact, a piece of news which has a sufficient strength to be remembered if presented along with news items which have a similar amplitude, can be forgotten if were instead presented along with stronger news. A different aspect of the model behaviour is the proportion of consecutive news that a society can remember when a very long sequence of information is presented. Using techniques borrowed from statistical mechanics, we found this proportion analytically and confirmed it using simulations. The results which we found are compatible with the behaviour of models appearing in the literature of forgetful Hopfield networks [36–39]. The last result concerns how noisy information is perceived by the society. We showed that if a piece of news was presented in the recent history of a society, and it is confronted with distorted version of it, the society is able to reconnect the distorted version with the clean memory it corresponds to, recalling the originally stored memory of it. This implies that the collective memory of a society is able to produce interesting effects that modulate the collective perception of current news or affairs, and it could be used to understand attitudes of people towards external events that may be only weakly correlated with past ones. This means that a shocking event that hits a society can change its perception of future similar events.

In the future we would like to extend our work by adding other interesting features to our model. We know for example that in the real world not all people receive the same news or perceive them with the same strength, so we aim to take this into account. Another possible improvement consists in personalizing the intensity of the pressure of the society for different agents. In this way some people will be more influenced by their peers and others less. Another possibility is to consider multiple opinions related to a set of different issues that may be relevant at the same time, and to have interactions depending on the history of such sets of opinions on a range of different topics. Apart from the addition of further ingredients to the model, a possible research direction would be the validation of our predictions with data from the real world. While we foresee some difficulties in finding readily available data suitable for this purpose, a way to circumvent this problem could be to design social experiments to collect the necessary data.

CRediT authorship contribution statement

Gioia Boschi: Conceptualization, Methodology, Software, Writing - original draft. **Chiara Cammarota:** Conceptualization, Methodology, Writing - review & editing, Supervision. **Reimer Kühn:** Conceptualization, Methodology, Writing - review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Patterns mutual influence

To inspect the effect of mutual interactions between patterns of news concomitantly presented to the society, we studied several societal dynamics where random news are presented in triplets with random probability of appearance and random intensity. In Fig. 3 we report the information on the final recovery or non-recovery of the patterns from the society, as a function of the probability and the intensity with which they were presented. In the graph we can notice a smooth transition region between a recovery zone and a non-recovery zone with some patterns within the recovery zone remaining un-recovered and, conversely, some patterns in the non-recovery zone apparently being recovered. Apart from an expected finite size effect, a factor which can also contribute to this phenomenon is the influence of the relative strengths and the relative probabilities of patterns presented during the same run of the dynamics. In fact, in Section 3.1, Fig. 2(b) we saw that, when there is one pattern in a triplet of patterns with a stronger amplitude, the other two are recovered less frequently than the strong pattern in question. In order to understand if this is a reason for anomalously recovered or un-recovered patterns, we propose to isolate outliers in the recovery and non-recovery zone which we identify as those patterns that have an anomalous behaviour when compared to their neighbours in the parameter space of frequency of occurrence and intensity, *i.e.* they are points corresponding to recovered patterns which are predominantly surrounded by those corresponding to un-recovered patterns and vice versa. Fig. A.6(a) shows a selection of points μ which have at least 50% of neighbours μ' with opposite retrieval outcome (and opposite colour in our representation) within distance $d = \sqrt{(I_0^\mu - I_0^{\mu'})^2 + ((\pi_\mu - \pi_{\mu'}) \cdot 10)^2} = 0.2$. We multiplied the difference in probability to a factor 10 in order to account for the difference in scale between the probability which is randomly drawn between 0 and 1, and the signal strength which is randomly drawn between 0 and 10.

To investigate whether mutual interactions between patterns influences recovery, we evaluate where the anomalous patterns are located relative to the other patterns in within their triplet which have the opposite recovery state. This means for example, that given a pattern that is unexpectedly recovered, we want to know how it is located relative to the un-recovered patterns presented within the same run of the dynamics. We thus calculate the average coordinate difference between the anomalously recovered point and the patterns in their triplets with opposite behaviour. Given a recovered pattern $\mu = 1$, assuming that both the patterns $\mu = 2$ and $\mu = 3$ are un-recovered, we calculate the coordinate difference in the I_0 direction as $\Delta_{I_0} = \frac{I_0^1 - I_0^2}{2} + \frac{I_0^1 - I_0^3}{2}$. If only $\mu = 2$ is un-recovered then $\Delta_{I_0} = I_0^1 - I_0^2$. Analogous formulae apply to the difference in the π direction. We thus replot the points in Fig. A.6(a) as a function of the coordinate differences.[1]

Interestingly, we notice that recovered (unrecovered) patterns typically fall into their 'natural' or expected top-right (bottom-left) regions when plotted in terms of their new coordinates. In other words, unexpected recovery (or non-recovery) of patterns, as highlighted in the right panel, can be explained by the fact that they were appearing together with particularly weak (strong) or less (more) frequent patterns, which affected their likelihood of retrieval despite their own features.

Additional evidence about the mutual influence of news retrievals can be obtained by looking directly at the statistics of news, which belong to the triplets of patterns that are anomalously retrieved or unretrieved by the society. Figs. A.6(c) and A.6(d) show in yellow respectively the anomalously recovered and unrecovered patterns of Fig. A.6(a), accompanied only by the patterns presented in the same run of the dynamics. The accompanying patterns are presented in red when they are not recovered and in black when they are recovered. In Fig. A.6(c) we observe that anomalously recovered

---

[1] Given that we are looking at the relative location of patterns behaving differently in the same triplet, we do not plot the points in dynamics where all the three patterns are recovered (little circles in Fig. A.6(a)). Dynamics in which all the three patterns are not recovered are not present in the simulations performed.
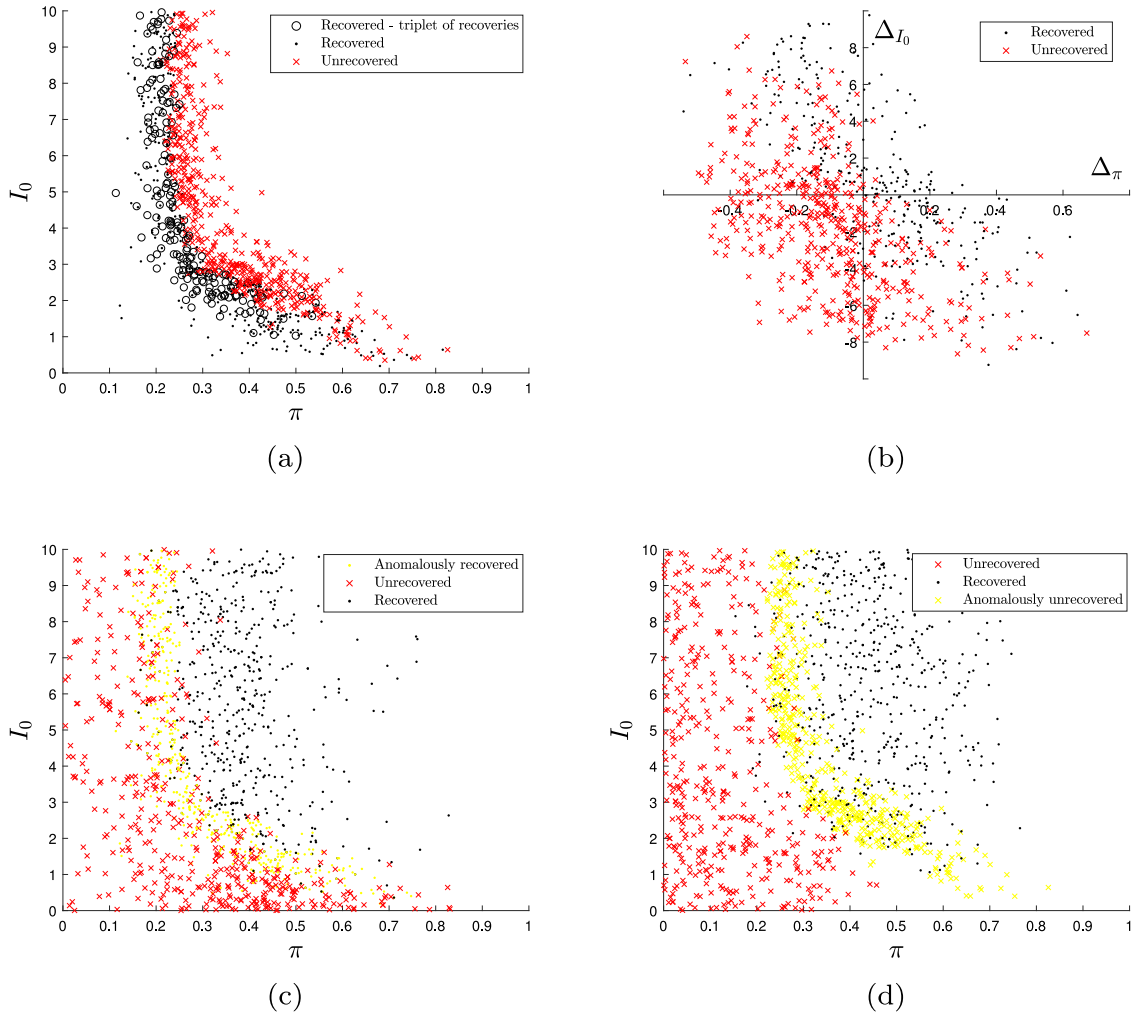
Fig. A.6. (a) A selection of points from Fig. 3 which have more than 50% of neighbours of the opposite colour within a circle of radius of 0.05, plotted as a function of their signal strength and probability of presentation. Black markers correspond to recovered patterns and red markers to unrecovered patterns. Circles highlight patterns belonging to dynamics in which all three patterns presented have been recovered. Panel (b) shows a subset of the data of (a), excluding those that belong to triplets of patterns that are *all* recovered. Points are plotted as a function of their average distance from the patterns in the same run of the dynamics which have a different recovery state. (c) Yellow dots represent the recovered patterns in panel (a) while the red crosses and black dots represent respectively the unrecovered and recovered patterns in a run of the dynamics that gives rise to the yellow points. (d) Same as (c) but this time yellow crosses indicate unrecovered patterns from panel (a). The parameters of the simulation are $N = 100$, $J_0 = 8$ and $\Delta_0 = 17$.

patterns are often accompanied by other recovered patterns with medium probability of appearance, which rarely reaches $\pi > 0.5$. Conversely, anomalously unrecovered patterns are accompanied by other recovered patterns with probability of appearance often in the range $0.5 - 0.8$, as visible in Fig. A.6(d). These observations give additional evidence to the hypothesis that very frequent news can hinder the ability of other news to be retrieved. Moreover, we note that anomalously recovered patterns often appear concomitantly with very weak news with medium probability of appearance, see the dense cluster of unrecovered patterns with π around $0.3 - 0.6$ and I_0 between 0 and 1 in Fig. A.6(c). The same region of parameters is totally empty in Fig. A.6(d), which contain information about triplets of anomalously unretrieved patterns. This last observation supports the idea that the retrieval of a pattern is enhanced when accompanying patterns, which appear often in the dynamics, are weak enough.

Unfortunately, even in the simple case of societal dynamics involving only three patterns discussed here, these phenomena do not establish quantitative causal relations on the possibility to retrieve or not retrieve a pattern, although it becomes clear that fluctuations of retrieval probabilities cannot be simply ascribed to finite-size effects. We have, however, collected indirect evidence that the relative strength and probability of the patterns in the system can influence the likelihood of their recovery. In particular, a pattern which has a probability and/or a strength which is small in magnitude can be recovered if the other patterns in the same dynamics are weaker in probability or signal strength. Similarly,

patterns in a triplet of patterns with medium/high probability and/or high signal strength can end up being unexpectedly unrecovered.

Appendix B. Replica calculations for an infinite flow of patterns

In this Appendix we derive the system of Eqs. (23), which defines the order parameters of the model in the noiseless limit of Eq. (1). The society described by this model is subject to an infinite sequence of external signals representing exposure to an infinite history of news items (Eq. (18)), as introduced in Section 4. Using replica calculations, we obtain the free-energy of the model and we derive and solve the corresponding fixed-point equations in the zero-temperature limit. The calculation follows standard reasoning. For further details, we refer to [43].

B.1. Replica approach

We start assuming that the asymptotic state of our system is macroscopically correlated with at most one of the presented patterns, taken to be pattern μ. We want to obtain the average of the free energy of the system taken over the randomness in all but pattern $\mu = \alpha N$. This is done using replica trick, for which we can write the free energy in the form

$$f(\beta) = -\frac{1}{\beta} \lim_{N \to \infty} \lim_{n \to 0} (Nn)^{-1} \ln \overline{Z_N^n} \tag{B.1}$$

where β is the reciprocal of the thermodynamic temperature which tends to infinity in the noiseless limit, while n is the number of replica and

$$Z_N^n = \int \prod_{i\sigma} dv_{i\sigma} \exp\left\{-\beta \sum_{\sigma=1}^n \mathcal{H}(\{\boldsymbol{v}_\sigma\})\right\} . \tag{B.2}$$

is the replicated partition function. One can write the Hamiltonian of a replica σ appearing in this expression as follows

$$\mathcal{H}(\{\boldsymbol{v}_\sigma\}) = -\frac{N J_0 \tilde{\gamma}_0}{2} m_\sigma^{\mu 2} e^{-\tilde{\gamma}_0 \alpha} - \frac{J_0 \tilde{\gamma}_0}{2} \sum_{\nu(\neq\mu)} X_\sigma^{\nu 2} e^{-\tilde{\gamma}_0 \nu/N}$$

$$+ \frac{J_0}{2}\left(1 - e^{-\tilde{\gamma}_0 \alpha_{\max}}\right) \sum_i v_{i\sigma}^2 - \sum_i I_i v_{i\sigma} + \sum_i G(v_{i\sigma}) . \tag{B.3}$$

Here

$$m_\sigma^\mu = \frac{1}{N} \sum_i \xi_i^\mu v_{i\sigma}$$

$$X_\sigma^\nu = \frac{1}{\sqrt{N}} \sum_i \xi_i^\nu v_{i\sigma} \quad , \qquad \nu(\neq \mu) ,$$

and $G(v)$ is the integrated inverse input–output relation

$$G(v) = \int^v dv' g^{-1}(v') . \tag{B.4}$$

The disorder due to the patterns $\nu(\neq \mu)$ appears only through the X_σ^ν, which for any fixed configuration $\{\boldsymbol{v}_\sigma\}$ are Gaussian random variables of zero mean and covariance

$$\langle X_\sigma^\nu X_{\sigma'}^{\nu'} \rangle = \frac{\delta_{\nu\nu'}}{N} \sum_i v_{i\sigma} v_{i\sigma'} = \delta_{\nu\nu'} q_{\sigma\sigma'} . \tag{B.5}$$

The average over the patterns disorder can therefore be performed as a Gaussian integral, resulting in

$$\overline{Z_N^n} = \int \prod_{i\sigma} dv_{i\sigma} \exp\left\{ N\left[\beta \frac{J_0 \tilde{\gamma}_0}{2} \sum_\sigma m_\sigma^{\mu 2} e^{-\tilde{\gamma}_0 \alpha} + \right.\right. \tag{B.6}$$

$$\left. -\frac{1}{2}\int_0^{\alpha_{\max}} dx \, \text{tr} \ln\left(\mathbb{I} - \beta J_0 \tilde{\gamma}_0 e^{-\tilde{\gamma}_0 x} Q\right) \right]$$

$$\left. -\frac{\beta J_0}{2}\left(1 - e^{-\tilde{\gamma}_0 \alpha_{\max}}\right) \sum_{i,\sigma} v_{i\sigma}^2 + \beta \sum_{i,\sigma} I_i v_{i\sigma} - \beta \sum_{i\sigma} G(v_{i\sigma}) \right\}$$

in which Q is a matrix with elements $q_{\sigma\sigma'}$. The standard procedure now is to introduce the overlaps m_σ^μ and Edwards–Anderson order parameters $q_{\sigma\sigma'}$ as integration variables, using conjugate variables for Fourier representations of δ-functions that enforce the definition of these order parameters. This results in

$$
\begin{aligned}
\langle Z_N^n \rangle = \int \prod_\sigma \frac{\mathrm{d}m_\sigma^\mu \mathrm{d}\hat{m}_\sigma^\mu}{2\pi/N} \prod_{\sigma\sigma'} \frac{\mathrm{d}q_{\sigma\sigma'} \mathrm{d}\hat{q}_{\sigma\sigma'}}{2\pi/N} \exp\Bigg\{ N\Bigg[& \beta\frac{J_0\tilde{\gamma}_0}{2} \sum_\sigma m_\sigma^{\mu 2} \, \mathrm{e}^{-\tilde{\gamma}_0\alpha} \\
& -\frac{1}{2}\int_0^{\alpha_{\max}} \mathrm{d}x \, \mathrm{tr}\ln\Big(\mathbb{I} - \beta J_0\tilde{\gamma}_0 \, \mathrm{e}^{-\tilde{\gamma}_0 x} Q\Big) \\
& -\sum_\sigma \mathrm{i}\hat{m}_\sigma m_\sigma - \sum_{\sigma\sigma'} \mathrm{i}\hat{q}_{\sigma\sigma'} q_{\sigma\sigma'} \\
& +\frac{1}{N}\sum_i \ln\int \prod_\sigma dv_\sigma \exp\Bigg\{ \sum_\sigma \mathrm{i}\hat{m}_\sigma \xi_i^\mu v_\sigma + \sum_{\sigma\sigma'} \mathrm{i}\hat{q}_{\sigma\sigma'} v_\sigma v_{\sigma'} \\
& -\frac{\beta J_0}{2}\Big(1 - \mathrm{e}^{-\tilde{\gamma}_0\alpha_{\max}}\Big)\sum_\sigma v_\sigma^2 + \beta I_i \sum_\sigma v_\sigma - \beta \sum_\sigma G(v_\sigma) \Bigg\} \Bigg] \Bigg\} .
\end{aligned}
\tag{B.7}
$$

This is now of a form that can be evaluated by the saddle-point method. In the thermodynamic limit, the (empirical) average over single site free energies $\frac{1}{N}\sum_i \ln\int \prod_\sigma dv_\sigma \exp\{\dots\}$ appearing in the last two lines of (B.7) will, by the law of large numbers, converge to a joint average over all forms of on-site randomness present in the average, i.e., the I_i, and the ξ_i^μ.

The stationarity requirement on the exponent w.r.t. variations of the conjugate variables gives two self-consistency equations

$$
m_\sigma^\mu = \Big\langle\!\!\Big\langle \xi^\mu \langle v_\sigma\rangle \Big\rangle\!\!\Big\rangle \quad , \qquad q_{\sigma\sigma'} = \Big\langle\!\!\Big\langle \langle v_\sigma v_{\sigma'}\rangle \Big\rangle\!\!\Big\rangle ,
\tag{B.8}
$$

while the stationarity requirement w.r.t. the order parameters m_σ^μ and $q_{\sigma\sigma'}$ results in

$$
\mathrm{i}\hat{m}_\sigma^\mu = \beta J_0\tilde{\gamma}_0 m_\sigma^\mu \, \mathrm{e}^{-\tilde{\gamma}_0\alpha} ,
\tag{B.9}
$$

$$
\mathrm{i}\hat{q}_{\sigma\sigma'} = \frac{\beta J_0\tilde{\gamma}_0}{2}\int_0^{\alpha_{\max}} \mathrm{d}x \, \mathrm{e}^{-\tilde{\gamma}_0 x}\Big(\mathbb{I} - \beta J_0\tilde{\gamma}_0 \, \mathrm{e}^{-\tilde{\gamma}_0 x} Q\Big)^{-1}_{\sigma\sigma'} .
\tag{B.10}
$$

Inner averages in (B.8) denote thermal averages over effective i-dependent replicated single-site Hamiltonians defined via (B.7); after inserting the values of the conjugate order parameters as defined via Eqs. (B.9) and (B.10), it is seen to take the form

$$
\begin{aligned}
H_{\mathrm{eff}}^{(i)}(\{v_\sigma\}) = & -J_0\tilde{\gamma}_0 \, \mathrm{e}^{-\tilde{\gamma}_0\alpha} \sum_\sigma \xi_i^\mu m_\sigma^\mu v_\sigma \\
& -\frac{J_0\tilde{\gamma}_0}{2}\sum_{\sigma\sigma'}\int_0^{\alpha_{\max}} \mathrm{d}x \, \mathrm{e}^{-\tilde{\gamma}_0 x}\Big(\mathbb{I} - \beta J_0\tilde{\gamma}_0 \, \mathrm{e}^{-\tilde{\gamma}_0 x} Q\Big)^{-1}_{\sigma\sigma'} v_\sigma v_{\sigma'} \\
& +\frac{J_0}{2}\Big(1 - \mathrm{e}^{-\tilde{\gamma}_0\alpha_{\max}}\Big)\sum_\sigma v_\sigma^2 - I_i \sum_\sigma v_\sigma + \sum_\sigma G(v_\sigma)
\end{aligned}
\tag{B.11}
$$

The i dependence originates from the on-site disorder in (B.7), and the outer average denoted by double angle brackets in (B.8) denotes an average over the joint distribution of this remaining on-site disorder.

B.2. Replica symmetry

Assuming replica symmetry (RS) for the solutions of the fixed point equations, with

$$
m_\sigma^\mu = m \quad , \quad q_{\sigma\sigma} = q_d \quad , \quad \text{and} \quad q_{\sigma\sigma'} = q \ \ \text{for} \ \ \sigma \neq \sigma' ,
\tag{B.12}
$$

one can decouple the replica and take the $n \to 0$-limit of the theory as required. Using the shorthand $a(x) = \beta J_0\tilde{\gamma}_0 \, \mathrm{e}^{-\tilde{\gamma}_0 x}$, we notice that the term coupling replica in the effective single-site Hamiltonian $H_{\mathrm{eff}}^{(i)}(\{v_\sigma\})$ in Eq. (B.11) is

$$
\frac{1}{2}\sum_{\sigma\sigma'}\int_0^{\alpha_{\max}} \mathrm{d}x \, a(x)\Big(\mathbb{I} - a(x)Q\Big)^{-1}_{\sigma\sigma'} v_\sigma v_{\sigma'}
\tag{B.13}
$$

$$
= \frac{1}{2}\sum_{\sigma\sigma'}\int_0^{\alpha_{\max}} \mathrm{d}x \, a(x)\Big(\big(1 - a(x)(q_d - q)\big)\mathbb{I} - a(x)q\mathbf{1}\Big)^{-1}_{\sigma\sigma'} v_\sigma v_{\sigma'}
$$

in which $\mathbb{I}$ is the $n \times n$ unit matrix, while $\mathbf{1}$ is the $n \times n$ with all elements equal to 1. Using the algebra of these matrices one finds

$$(A\mathbb{I} - B\mathbf{1})^{-1} = \frac{1}{A}\mathbb{I} + \frac{B}{A(A + Bn)}\mathbf{1} \quad \longrightarrow \quad \frac{1}{A}\mathbb{I} + \frac{B}{A^2}\mathbf{1}\,, \qquad \text{as } n \to 0\,.$$

We thus get

$$\frac{1}{2}\sum_{\sigma\sigma'}\int_0^{\alpha_{\max}}\mathrm{d}x\,a(x)\Big(\big(1 - a(x)(q_d - q)\big)\mathbb{I} - a(x)q\mathbf{1}\Big)^{-1}_{\sigma\sigma'}v_\sigma v_{\sigma'}$$

$$= \frac{1}{2}\int_0^{\alpha_{\max}}\mathrm{d}x\,\frac{a(x)}{1 - a(x)(q_d - q)}\sum_\sigma v_\sigma^2$$

$$+ \frac{1}{2}\int_0^{\alpha_{\max}}\mathrm{d}x\,\frac{a(x)^2 q}{\big(1 - a(x)(q_d - q)\big)^2}\Big(\sum_\sigma v_\sigma\Big)^2 \tag{B.14}$$

Exploiting the fact that $a'(x) = -\tilde\gamma_0 a(x)$, we find (assuming $\alpha_{\max} \to \infty$) the first integral to give

$$\frac{1}{2}\int_0^\infty \mathrm{d}x\,\frac{a(x)}{1 - a(x)(q_d - q)} = -\frac{1}{2}\frac{\ln\big(1 - \beta J_0\tilde\gamma_0(q_d - q)\big)}{\tilde\gamma_0(q_d - q)} \tag{B.15}$$

To evaluate the second integral, use

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{a(x)}{1 - a(x)(q_d - q)}\right] = \frac{-\tilde\gamma_0 a(x)}{1 - a(x)(q_d - q)} - \frac{a(x)^2\tilde\gamma_0(q_d - q)}{\big(1 - a(x)(q_d - q)\big)^2}\,, \tag{B.16}$$

hence

$$\frac{1}{2}\int_0^\infty \mathrm{d}x\,\frac{a(x)^2 q}{\big(1 - a(x)(q_d - q)\big)^2}$$

$$= \frac{1}{2}\frac{q}{\tilde\gamma_0(q_d - q)}\left[\frac{\beta J_0\tilde\gamma_0}{1 - \beta J_0\tilde\gamma_0(q_d - q)} + \frac{\ln\big(1 - \beta J_0\tilde\gamma_0(q_d - q)\big)}{(q_d - q)}\right]$$

$$= \frac{1}{2}\frac{\beta J_0 q}{q_d - q}\left[\frac{1}{\big(1 - \beta J_0\tilde\gamma_0(q_d - q)\big)} + \frac{\ln\big(1 - \beta J_0\tilde\gamma_0(q_d - q)\big)}{\beta J_0\tilde\gamma_0(q_d - q)}\right] \tag{B.17}$$

Although it is not obvious at this point, this integral is positive by construction.

It is expected that $C = \beta(q_d - q)$ will remain finite in the $\beta \to \infty$-limit we are interested in.

This gives

$$\frac{1}{2}\int_0^\infty \mathrm{d}x\,\frac{a(x)}{1 - a(x)(q_d - q)} = -\frac{\beta J_0}{2}\frac{\ln\big(1 - J_0\tilde\gamma_0 C\big)}{J_0\tilde\gamma_0 C}$$

and

$$\frac{1}{2}\int_0^\infty \mathrm{d}x\,\frac{a(x)^2 q}{\big(1 - a(x)(q_d - q)\big)^2} = \frac{1}{2}\frac{(\beta J_0)^2 q}{J_0 C}\left[\frac{1}{1 - J_0\tilde\gamma_0 C} + \frac{\ln\big(1 - J_0\tilde\gamma_0 C\big)}{J_0\tilde\gamma_0 C}\right]$$

$$\equiv \frac{1}{2}(\beta J_0)^2 r \tag{B.18}$$

The coupling between replica through a complete square embodied in the second contribution in (B.14) is then dealt with in the usual way by Gaussian linearization.

The fixed point equations for m, q and C finally take the form

$$m = \left\langle\!\!\left\langle \xi^\mu \langle v\rangle \right\rangle\!\!\right\rangle\,, \quad C = \frac{1}{J_0\sqrt{r}}\left\langle\!\!\left\langle z\,\langle v\rangle \right\rangle\!\!\right\rangle\,, \quad q = \left\langle\!\!\left\langle \langle v\rangle^2 \right\rangle\!\!\right\rangle\,, \tag{B.19}$$

in which r is defined through Eq. (B.18):

$$r = \frac{q}{J_0 C}\left[\frac{1}{1 - J_0\tilde\gamma_0 C} + \frac{\ln\big(1 - J_0\tilde\gamma_0 C\big)}{J_0\tilde\gamma_0 C}\right]\,. \tag{B.20}$$

The inner average in Eq. (B.19) is now a thermal average over effective RS single-site Hamiltonians of the form

$$H_{RS} = -\Big(\xi^\mu m\tilde J_0 e^{-\gamma_0\Delta_0\alpha} + J_0\sqrt{r}\,z + I\Big)v + \frac{J_0}{2}\left(1 + \frac{\ln(1 - J_0\tilde\gamma_0 C)}{J_0\tilde\gamma_0 C}\right)v^2 + G(v) \tag{B.21}$$

while outer averages are over the on-site disorder *and* over the additional zero-mean unit-variance Gaussian z appearing in H_{RS}. In the limit for $\beta \to \infty$, given a generic function F the inner averages can be written as:

$$\langle F(v) \rangle = F(\hat{v}) \tag{B.22}$$

where $\hat{v}$ is the value of v which minimizes H_{RS}

$$\hat{v}(\xi^\mu, I, z) = g\left(m\xi^\mu J_0 \tilde{\gamma}_0 e^{-\tilde{\gamma}_0 \alpha} + J_0 \sqrt{r} z + I - J_0\left(1 + \frac{\ln(1 - J_0 \tilde{\gamma}_0 C)}{J_0 \tilde{\gamma}_0 C}\right) \hat{v} \right). \tag{B.23}$$

We are thus left with the problem of solving a z-dependent fixed-point equation, Eq. (B.23), *within* the system (B.19) of fixed-point equations, which is avoided by transforming the Gaussian z-distribution into a $\hat{v}$-distribution and then taking the averages respect to $\hat{v}$. For a vanishing signal, i.e. $I = 0$, the resulting $p(\hat{v})$ will be given by

$$
\begin{aligned}
p(\hat{v}) &= \frac{e^{-z^2/2}}{\sqrt{2\pi}} \frac{dz}{d\hat{v}}, \\
&= \frac{e^{-z^2/2}}{J_0 \sqrt{2\pi r}} \left[\frac{\sqrt{\pi}}{2} e^{\mathrm{erf}^{-1}(\hat{v})^2} + J_0\left(1 + \frac{\ln(1 - J_0 \tilde{\gamma}_0 C)}{J_0 \tilde{\gamma}_0 C}\right) \right],
\end{aligned}
\tag{B.24}
$$

with:

$$z = \frac{1}{J_0 \sqrt{r}} \left[\mathrm{erf}^{-1}(\hat{v}) - m\xi^\mu J_0 \tilde{\gamma}_0 e^{-\tilde{\gamma}_0 \alpha} - I + J_0 \hat{v}\left(1 + \frac{\ln(1 - J_0 \tilde{\gamma}_0 C)}{J_0 \tilde{\gamma}_0 C}\right) \right]. \tag{B.25}$$

If instead we consider a Gaussian signal, as described in Eq. (26), the value of $\hat{v}$ can be written as:

$$\hat{v}(\xi^\mu, z, z') = g\left(\xi^\mu(m J_0 \tilde{\gamma}_0 e^{-\tilde{\gamma}_0 \alpha} + \tilde{I}_0) + J_0 \sqrt{r} z + \sigma_I z' - J_0\left(1 + \frac{\ln(1 - J_0 \tilde{\gamma}_0 C)}{J_0 \tilde{\gamma}_0 C}\right) \hat{v} \right). \tag{B.26}$$

in which z, z' are two independent standard Gaussians. Using the fact that the sum of Gaussians is itself Gaussian (with variances given by the sum of variances), we can write

$$\hat{v}(\xi^\mu, z) = g\left(\xi^\mu(m J_0 \tilde{\gamma}_0 e^{-\tilde{\gamma}_0 \alpha} + \tilde{I}_0) + \sqrt{J_0^2 r + \sigma_I^2}\, z - J_0\left(1 + \frac{\ln(1 - J_0 \tilde{\gamma}_0 C)}{J_0 \tilde{\gamma}_0 C}\right) \hat{v} \right). \tag{B.27}$$

so

$$z = \frac{1}{\sqrt{J_0^2 r + \sigma_I^2}} \left[\mathrm{erf}^{-1}(\hat{v}) - \xi^\mu(m J_0 \tilde{\gamma}_0 e^{-\tilde{\gamma}_0 \alpha} + \tilde{I}_0) + J_0 \hat{v}\left(1 + \frac{\ln(1 - J_0 \tilde{\gamma}_0 C)}{J_0 \tilde{\gamma}_0 C}\right) \right]. \tag{B.28}$$

and

$$p(\hat{v}) = \frac{e^{-z^2/2}}{\sqrt{2\pi} \sqrt{J_0^2 r + \sigma_I^2}} \left[\frac{\sqrt{\pi}}{2} e^{\mathrm{erf}^{-1}(\hat{v})^2} + J_0\left(1 + \frac{\ln(1 - J_0 \tilde{\gamma}_0 C)}{J_0 \tilde{\gamma}_0 C}\right) \right]. \tag{B.29}$$

Appendix C. Methods

C.1. Simulations

In this paper we simulate the dynamics of the society described by Eq. (1) in order to measure the spontaneous retrieval of opinion patterns induced by the presentation of different kind of external signals.

In order to do this, we discretize the Eq. (1) with a time step $dt = 0.1$, where smaller dt were seen not to significantly change the results. At each time step we use Euler method to calculate the preference of each agent u_i, and the couplings J_{ij}, discretizing with the same time step its differential equation (obtained differentiating Eq. (2)):

$$\dot{J}_{ij}(t) = \gamma \left[\frac{J_0}{N} v_i(t) v_j(t) - J_{ij}(t) \right]. \tag{C.1}$$

Even if the Euler method is a simple integration method, it is accurate enough for our purposes.

There are many parameters to be set up for these simulations. In Section 3 we simulate a society which receives $p = 3$ patterns in a random order. The size of the system is set to be $N = 100$, the strength of the interactions $J_0 = 6$ and the memory factor $\gamma = 10^{-3}$. Other parameters, as the time length of the external signal and the amplitude I_0 of the polarizing signal, change in different simulations and are indicated under the corresponding figures. The number of agents $N = 100$ is a good approximation of the thermodynamic limit at finite p. Although these numbers seem small, they produce results that are representative of the $N \to \infty$ limit with $p \ll N$.

In Section [4] we perform the same kind of simulations to describe a society which receives an infinite number of external news. In this case J_0 is fixed to 0.2, a value at which was possible to obtain analytical results to be compared to the simulations. The signal strength was chosen to be $I_0 = 10$ while the size of the system explored where $N = 200$ and $N = 800$. The value of $\tilde{\gamma}_0$ is chosen to be 15, close to the point at which the corresponding critical capacity is maximum. The value of $\tilde{\gamma}$ is varied through different simulations, with only the value of Δ_0 changing and γ_0 always kept at 1. Through all the paper we used a low noise level, with variance $\sigma^2 = 0.01$, to ensure that non-trivial collective states can emerge. All simulations start with random initial conditions $u_i \sim \mathcal{N}(0, \sigma^2/2)$ which would be the equilibrium distribution in a non-interacting system without external signal.

C.2. Numerical solutions

Numerical solutions of Eq. [(23)] are obtained iteratively. The double angle bracket are evaluated as an average over $\xi^\mu = \pm 1$ and an integral over $p(\hat{v})$ (Eq. [(B.24)]).

We would like to find a solution close to $m = 1$ so we start with an initial guess of $m \sim 1$, $c \sim 0$ and $q \sim 1$, and we iterate as follow:

$$m(i+1) = \rho m(i) + (1-\rho)\frac{1}{2}\left[\int_{-1}^{1} d\hat{v}\hat{v}p_+(\hat{v}) - \int_{-1}^{1} d\hat{v}\hat{v}p_-(\hat{v})\right],$$

$$C(i+1) = \rho C(i) + (1-\rho)\frac{1}{2J_0\sqrt{r}}\left[\int_{-1}^{1} d\hat{v}zp_+(\hat{v}) + \int_{-1}^{1} d\hat{v}zp_-(\hat{v})\right],$$

$$q(i+1) = \rho q(i) + (1-\rho)\frac{1}{2}\left[\int_{-1}^{1} d\hat{v}\hat{v}^2p_+(\hat{v}) + \int_{-1}^{1} d\hat{v}\hat{v}^2p_-(\hat{v})\right], \tag{C.2}$$

where z is expressed as a function of $\hat{v}$ as described by Eq. [(B.25)]. The notation $p_\pm(\hat{v})$ indicates the distribution $p(\hat{v})$ evaluated at $\xi^\mu = \pm 1$. We consider the iterative algorithm converged when the difference between the values found at step $i+1$ and i is smaller than 10^{-9}. Numerical solutions of the set are found for different values of α and $\tilde{\gamma}_0$. For each couple of parameters, the iterative algorithm is initialized in the solution found in the closest point in the parameter space.

We should finally notice that for

$$J_0\left(1 + \frac{\log(1 - J_0\tilde{\gamma}_0 C)}{J_0\tilde{\gamma}_0}\right) < 0 \tag{C.3}$$

the slope of the error function appearing in Eq. [(24)] induces the solution of $\hat{v}$ to be discontinuous. In particular $\hat{v}$ is negative from -1 to $-v_0$, jumps to a positive value v_0 which must be found numerically, and remains positive from v_0 to 1. In this case the integrals appearing in Eq. [(C.2)] must be split and taken from -1 to $-v_0$ and from v_0 to 1. For the parameters used in this paper the condition in Eq. [(C.3)] is not met, so this split is not necessary.

References

[1] C. Castellano, S. Fortunato, V. Loreto, Statistical physics of social dynamics, Rev. Modern Phys. 81 (2) (2009) 591.
[2] A. Flache, M. Mäs, T. Feliciani, E. Chattoe-Brown, G. Deffuant, S. Huet, J. Lorenz, Models of social influence: Towards the next frontiers, J. Artif. Soc. Soc. Simul. 20 (4) (2017).
[3] A. Baronchelli, The emergence of consensus: a primer, R. Soc. Open Sci. 5 (2) (2018) 172189.
[4] S. Galam, Y. Gefen, Y. Shapir, Sociophysics: A new approach of sociological collective behaviour. I. mean-behaviour description of a strike, J. Math. Sociol. 9 (1) (1982) 1–13.
[5] S. Galam, S. Moscovici, Towards a theory of collective phenomena: Consensus and attitude changes in groups, Eur. J. Soc. Psychol. 21 (1) (1991) 49–74.
[6] C. Borghesi, J.-P. Bouchaud, Of songs and men: a model for multiple choice with herding, Qual. Quant. 41 (4) (2007) 557–568.
[7] P. Clifford, A. Sudbury, A model for spatial conflict, Biometrika 60 (3) (1973) 581–588.
[8] R.A. Holley, T.M. Liggett, Ergodic theorems for weakly interacting infinite systems and the voter model, Ann. Probab. (1975) 643–663.
[9] S. Galam, Minority opinion spreading in random geometry, Eur. Phys. J. B 25 (4) (2002) 403–406.
[10] A.C. Martins, S. Galam, Building up of individual inflexibility in opinion dynamics, Phys. Rev. E 87 (4) (2013) 042807.
[11] A. Sîrbu, V. Loreto, V.D. Servedio, F. Tria, Opinion dynamics: models, extensions and external effects, in: Participatory Sensing, Opinions and Collective Awareness, Springer, 2017, pp. 363–401.
[12] T.V. Martins, M. Pineda, R. Toral, Mass media and repulsive interactions in continuous-opinion dynamics, Europhys. Lett. 91 (4) (2010) 48003.
[13] T. Carletti, D. Fanelli, S. Grolli, A. Guarino, How to make an efficient propaganda, Europhys. Lett. 74 (2) (2006) 222.
[14] W. Quattrociocchi, G. Caldarelli, A. Scala, Opinion dynamics on interacting networks: media competition and social influence, Sci. Rep. 4 (2014) 4938.
[15] G. Boschi, C. Cammarota, R. Kühn, Opinion dynamics with emergent collective memory: a society shaped by its own past, Physica A 558 (2020) 124909.
[16] M. Mäs, A. Flache, D. Helbing, Individualization as driving force of clustering phenomena in humans, PLoS Comput. Biol. 6 (10) (2010) e1000959.
[17] M. Mäs, A. Flache, J.A. Kitts, Cultural integration and differentiation in groups and organizations, in: Perspectives on Culture and Agent-Based Simulations, Springer, 2014, pp. 71–90.
[18] A. Sîrbu, V. Loreto, V.D. Servedio, F. Tria, Opinion dynamics with disagreement and modulated information, J. Stat. Phys. 151 (1–2) (2013) 218–237.

[19] K. Sznajd-Weron, M. Tabiszewski, A.M. Timpanaro, Phase transition in the Sznajd model with independence, Europhys. Lett. 96 (4) (2011) 48002.

[20] A. Radillo-Díaz, L.A. Pérez, M. del Castillo-Mussot, Axelrod models of social influence with cultural repulsion, Phys. Rev. E 80 (6) (2009) 066107.

[21] R. Axelrod, The dissemination of culture: A model with local convergence and global polarization, J. Conflict Resolut. 41 (2) (1997) 203–226.

[22] G. Deffuant, D. Neau, F. Amblard, G. Weisbuch, Mixing beliefs among interacting agents, Adv. Complex Syst. 3 (01n04) (2000) 87–98.

[23] M.W. Macy, J.A. Kitts, A. Flache, S. Benard, Polarization in dynamic networks: A hopfield model of emergent structure, Dy. Soc. Netw. Model. Anal. (2003) 162–173.

[24] A. Flache, M.W. Macy, Small worlds and cultural polarization, J. Math. Sociol. 35 (1–3) (2011) 146–176.

[25] N.P. Mark, Culture and competition: Homophily and distancing explanations for cultural niches, Am. Sociol. Rev. (2003) 319–345.

[26] S. Huet, G. Deffuant, Openness leads to opinion stability and narrowness to volatility, Adv. Complex Syst. 13 (03) (2010) 405–423.

[27] W. Jager, F. Amblard, Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change, Comput. Math. Organ. Theory 10 (4) (2005) 295–303.

[28] J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. 79 (8) (1982) 2554–2558.

[29] L. Dall'Asta, C. Castellano, Effective surface-tension in the noise-reduced voter model, Europhys. Lett. 77 (6) (2007) 60005.

[30] A. Jedrzejewski, K. Sznajd-Weron, Impact of memory on opinion dynamics, Physica A 505 (2018) 306–315.

[31] H.-U. Stark, C.J. Tessone, F. Schweitzer, Decelerating microdynamics can accelerate macrodynamics in the voter model, Phys. Rev. Lett. 101 (1) (2008) 018701.

[32] R. García-Gavilanes, A. Mollgaard, M. Tsvetkova, T. Yasseri, The memory remains: Understanding collective memory in the digital age, Sci. Adv. 3 (4) (2017) e1602368.

[33] S. Mariano, I. Morărescu, R. Postoyan, L. Zaccarian, A hybrid model of opinion dynamics with memory-based connectivity, IEEE Control Syst. Lett. 4 (3) (2020) 644–649.

[34] D.O. Hebb, The Organization of Behavior: a Neuropsychological Theory, J. Wiley; Chapman & Hall, 1949.

[35] D.J. Amit, H. Gutfreund, H. Sompolinsky, Statistical mechanics of neural networks near saturation, Ann. Physics 173 (1) (1987) 30–67.

[36] J. Nadal, G. Toulouse, J. Changeux, S. Dehaene, Networks of formal neurons and memory palimpsests, Europhys. Lett. 1 (10) (1986) 535.

[37] G. Parisi, A memory which forgets, J. Phys. A: Math. Gen. 19 (10) (1986) L617.

[38] J.L. van Hemmen, G. Keller, R. Kühn, Forgetful memories, Europhys. Lett. 5 (7) (1988) 663.

[39] E. Marinari, Forgetting memories and their attractiveness, Neural Comput. 31 (3) (2019) 503–516.

[40] M.A. Cohen, S. Grossberg, Absolute stability of global pattern formation and parallel memory storage by competitive neural networks, IEEE Trans. Syst. Man Cybern. (5) (1983) 815–826.

[41] J.J. Hopfield, Neurons with graded response have collective computational properties like those of two-state neurons, Proc. Natl. Acad. Sci. 81 (10) (1984) 3088–3092.

[42] R. Kühn, S. Bös, J.L. van Hemmen, Statistical mechanics for networks of graded-response neurons, Phys. Rev. A 43 (4) (1991) 2084.

[43] R. Kühn, S. Bös, Statistical mechanics for neural networks with continuous-time dynamics, J. Phys. A: Math. Gen. 26 (4) (1993) 831.