

# Phase flipping emerges from coupling opinion formation and message propagation

Yiwen Liu,<sup>1</sup> Charo I. del Genio<sup>2,3,4</sup>, Stefano Boccaletti,<sup>2,5,6</sup> Ming Tang<sup>2,1,7,\*</sup> and Shuguang Guan<sup>1,†</sup>

<sup>1</sup>*School of Physics, East China Normal University, Shanghai 200241, China*

<sup>2</sup>*Research Institute of Interdisciplinary Intelligent Science, Ningbo University of Technology, Ningbo 315104, China*

<sup>3</sup>*School of Mathematics, North University of China, Taiyuan 030051, China*

<sup>4</sup>*Institute of Smart Agriculture for Safe and Functional Foods and Supplements, Trakia University, Stara Zagora 6000, Bulgaria*

<sup>5</sup>*Sino-Europe Complexity Science Center, North University of China, Taiyuan 030051, China*

<sup>6</sup>*CNR-Institute of Complex Systems, Via Madonna del Piano 10, I-50019 Sesto Fiorentino, Italy*

<sup>7</sup>*Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China*



(Received 28 August 2025; accepted 12 December 2025; published 31 December 2025)

During decision-making processes, many real-world social systems exhibit phase flipping, a relatively unexplored phenomenon consisting in sudden macroscopic switches in the majority opinion of the system. Existing studies mostly treat consensus formation as an isolated process. However, in the human society, opinion dynamics is almost always influenced by external factors, such as the diffusion of news on social media. Here, we construct a noisy threshold voter-susceptible-infected-recovered coupled model with multiple timescales and study the synergistic coevolution dynamics of opinion formation and message diffusion. We show that the model has two mechanisms that account for macroscopic changes in the consensus state and demonstrate how, in small enough systems, one of the two mechanisms can cause the occurrence of phase flipping. Our work not only improves the understanding of phase flipping phenomena in real-world opinion systems, but also provides potential strategies to stabilize the collective opinion in a positive state.

DOI: [10.1103/xd9p-tvz1](https://doi.org/10.1103/xd9p-tvz1)

## I. INTRODUCTION

Network models, which are routinely used to study structure and dynamics of complex systems, often feature the occurrence of phase transitions [1–4]. Notable examples of processes that exhibit such behavior include epidemic spreading, percolation, and synchronization, as well as the general mechanisms that can affect the structure itself of a network [5–16]. Within the particular context of opinion dynamics, such transitions can lead to unexpected outcomes, such as the majority of a voting population suddenly switching their support from one candidate to the other. However, notwithstanding the long history of research on opinion formation, existing studies have mostly focused on characterizing the steady state of the dynamics and estimating the time needed to reach consensus [17]. Currently, only a limited number of studies have explored the mechanism underlying collective-opinion flipping. In the context of discrete opinions, they are mostly modifications and evolutions of the Kirman model, also known as the noisy voter model [18], which include both spontaneous and neighbor-induced state-switching

mechanisms [19,20]. Similarly, in continuous opinion models, random changes in the opinion of some agents can cause the macroscopic opinion state to fluctuate [21]. Also note that when the possibility of spontaneous state switching is removed, such models reduce to the classic voter models [22]. Thus, our understanding of such switching phenomena is still incomplete, as the underlying mechanism that drives both the formation of opinions and the potential for sudden shifts in the opinion of the majority has not yet been positively identified.

In the pursuit of an answer to this question, researchers have recently recognized the impact that modern means of mass communication have on the formation of public opinion [23,24]. In fact, when communications were mostly limited by the necessity of a certain degree of physical presence, the processes of information spreading and opinion formation occurred on the same timescale. However, with the advent of social media, a significant difference between the two has arisen, due to the fact that, while a round of message dissemination in an online environment lasts a time in the order of magnitude of a few hours [25,26], the formation of group opinion is still a relatively slower process. Thus, a more realistic approach is to consider the two contact processes as being separate, but influencing each other. Existing studies have shown that, in multiplex networks, coevolution dynamics lead to the emergence of critical phenomena that are not observed in single-layer cases [27–29]. Also, the use of an opinion-network coevolution framework has allowed researchers to explore the effects of polarization and echo chamber phenomena [30,31]. However, the dynamical collective-opinion transition and the effects of message dissemination have not

\*Contact author: tangminghan007@gmail.com

†Contact author: ssguan@phy.ecnu.edu.cn

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

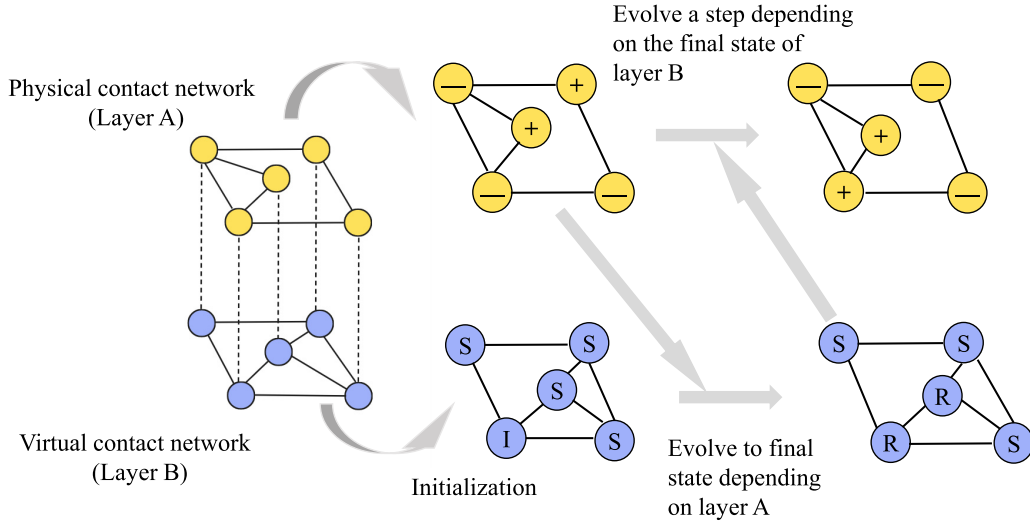


FIG. 1. Schematic representation of our coupled model on a duplex network. Every step of opinion updating on layer A is influenced by the steady state previously reached by the message dissemination process on layer B. In turn, the message dissemination dynamics evolves depending on the opinions of the nodes in layer A.

been completely elucidated. Additionally, the difference in timescale has been shown to have a sizable effect on coevolutionary dynamics, and to be able to affect the final state of the system [32–35]. However, a specific application of this idea to opinion dynamics is still missing. In this article, we introduce a noisy threshold voter model coupled to a Susceptible-Infected-Recovered (SIR) model, to study the coevolution of public opinion formation, represented by the former, and message dissemination, represented by the latter. Our approach takes into account the timescale difference between the two processes as well as the multiround feature of message spreading on social media. We study the bifurcations of the model and, using extensive numerical simulations and theoretical arguments, we show that it features a first-order transition and a bistable state, which disappear if the noise in the system is above a critical threshold. Also, we find two separate mechanisms that can induce opinion flipping and drive a social system to switch between two different consensus states.

## II. RESULTS

### A. Coupled model

To construct our model, we use a multiplex network [36], in which the same nodes are linked over more than one interaction layer, and each layer relates to an individual process affecting separate state variables. Specifically, we employ two layers, layer A and layer B.

As shown in Fig. 1, layer A represents the physical contact network, whose interactions are described by the noisy threshold voter model, with the nodes holding a “positive” (+) or a “negative” (−) opinion. This is a modification of the classic voter model, upon which two additional mechanisms are introduced. The first is a noise rate  $\nu$ , which allows individuals to spontaneously change their opinion. The other is a threshold factor  $m$ , such that an agent can switch their state, with a rate  $r$ , only if their opinion is shared by at most  $m$  neighbors of theirs [37]. Thus, the threshold  $m$  can be thought

of as being inversely related to the depth of conviction of an agent, as a smaller value of  $m$  means that nodes require more neighbors with opposite opinion to consider changing state. Note that the symmetry between the deterministic and stochastic evolutionary rules of this model allows one to consider the relative strength of the two mechanisms, determined by  $\frac{\nu}{r}$ , as an effective parameter.

The message propagation runs instead on layer B, which represents a generalized interaction network on which agents exchange information, and is thus not restricted to actual physical contacts. The dynamics follows the classic SIR model, in which individuals can be in one of three states, namely, susceptible ( $S$ ), infected ( $I$ ), and recovered ( $R$ ). In the context of message propagation, susceptible individuals can receive messages from infected neighbors, whereas recovered ones correspond to agents that have already received messages and have eventually lost interest in reposting them. The initial state of the system includes a small number of randomly chosen nodes that are already in state  $I$ , representing the message propagation sources, while all the other nodes are in state  $S$ . Susceptible nodes turn to state  $I$  with a rate  $\beta$  when in contact with an infected neighbor, while infected nodes recover with rate  $\mu$ . To account for the difference of timescale in the evolution of the dynamical processes, we assume that message propagation on layer B reaches equilibrium before nodes in layer A update their opinion states. In other words, every round of opinion updating is based on the steady state of the message dissemination. Thus, the evolution of our coupled model is as follows. First, the opinion states of the nodes in layer A are initialized by assigning the positive opinion to a randomly selected subset of nodes and the negative opinion to the rest. Then, a small number of nodes in layer B are put in the  $I$  state as initial spreaders, while the rest are put in state  $S$ . Next, the process of message spreading is allowed to reach steady state, after which the nodes in layer A undergo one pass of opinion update. The message spreading and opinion formation processes are then repeated until both layers reach a stable state, as measured by the fraction  $\rho^+$  of nodes with

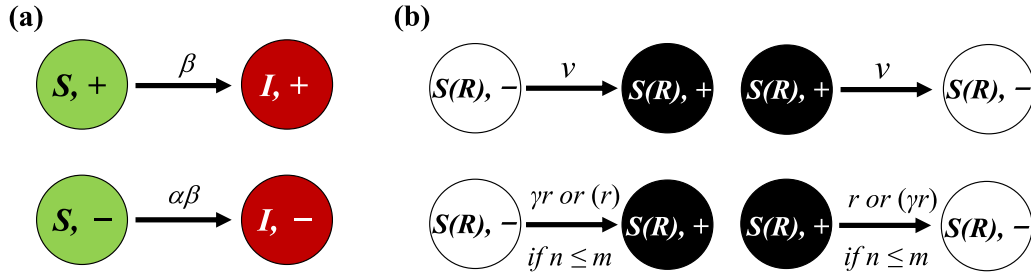


FIG. 2. Diagram of state transitions in the model. (a) In the process of message diffusion, susceptible nodes ( $S$ ) with positive opinion (+) receive a message and become infected with rate  $\beta$ , whereas those with negative opinion (−) receive it with reduced rate  $\alpha\beta$ . Infected nodes become recovered at a rate  $\mu$  regardless of their opinion states (not shown). (b) The opinion of a node in layer A flips to the opposite state spontaneously at rate  $\nu$ . In addition to spontaneous flipping, a node may flip its opinion when the number of its neighbors  $n$  holding the same opinion is not greater than the threshold  $m$ . In this case, the flipping rate depends on its current opinion state and on whether the node did receive a message in the previous round of message diffusion. The flipping rate for nodes that have received a message and hold negative opinion, and the one for nodes that have not received a message and hold positive opinion are both equal to  $r$ ; for nodes that have received a message and hold positive opinion, and that for nodes that have not received a message and hold negative opinion are both equal to  $\gamma r$ .

positive opinion in layer A and the fraction  $\rho^R$  of nodes with state  $R$  in layer B, which measures the overall coverage of message diffusion.

To model the interplay between the two processes, we introduce a mechanism to break the symmetry of the states in layer A. Specifically, without loss of generality, we make the reception of messages on layer B be inhibited by holding a negative opinion in layer A. Thus, as shown in Fig. 2(a), nodes of states  $+S$  and  $-S$  can transition to states  $+I$  and  $-I$  with rates  $\beta$  and  $\alpha\beta$ , respectively. Here,  $\alpha$  is a parameter ranging from 0 to 1, characterizing the influence of one's own opinion on receptivity to messages. Thus, the smaller  $\alpha$  is, the less likely nodes with negative opinion are to receive and approve messages. Infected nodes lose interest in forwarding messages at rate  $\mu$  regardless of the opinion they hold.

At the same time, messages that individuals receive from their neighbors may influence their opinion switching. Again without loss of generality, we assume that nodes that have received a message are less likely to abandon a positive opinion than those who have not. Vice versa, nodes who have not received a message are less likely to abandon a negative opinion than those who have. Thus, if the threshold condition is met, nodes in states  $-R$  and  $+R$  change opinion state with rates  $r$  and  $\gamma r$ , respectively, whereas nodes in states  $+S$  and  $-S$  switch with rates  $r$  and  $\gamma r$ , as illustrated in Fig. 2(b). Here,  $\gamma$  is a parameter ranging from 0 to 1, characterizing the influence of message spreading on opinion formation. Note that no nodes in layer B are in state  $I$ , as they are first made to reach the steady state of the message propagation dynamics. Also note that the noise-induced opinion switching mechanism, which occurs independently of the threshold condition, is not affected by message propagation, and its rate is always  $\nu$ .

### B. Bistable state and critical noise

To study the behavior of our model, we start with a mean-field approximation (for details, see Appendix A), which yields two equations for the values of the order parameters  $\rho^+$  and  $\rho^R$  at steady state,

$$\rho^R = 1 - \rho^+ e^{-\frac{\beta \langle k_B \rangle \rho^R}{\mu}} - \rho^- e^{-\frac{\alpha \beta \langle k_B \rangle \rho^R}{\mu}} \quad (1)$$

and

$$\begin{aligned} \frac{d}{dt} \rho^+ = & \left( \nu + \sum_{k_A} p_{k_A} F(k_A, \rho^+) \gamma r \right) \rho^{-S} \\ & + \left( \nu + \sum_{k_A} p_{k_A} F(k_A, \rho^+) r \right) \rho^{-R} \\ & - \left( \nu + \sum_{k_A} p_{k_A} F(k_A, \rho^-) r \right) \rho^{+S} \\ & - \left( \nu + \sum_{k_A} p_{k_A} F(k_A, \rho^-) \gamma r \right) \rho^{+R} = 0. \quad (2) \end{aligned}$$

Here,  $\rho^-$  is the total fraction of nodes that hold a negative opinion,  $\langle k_B \rangle$  is the average degree of the nodes in layer B, the sums are over all degrees  $k_A$  of the nodes in layer A,  $p_{k_A}$  is the degree distribution of the nodes in layer A,  $\rho^{\pm S}$  and  $\rho^{\pm R}$  are the fractions of nodes in states  $\pm S$  and  $\pm R$ , respectively, and the function  $F$  is the binomial cumulative distribution function with  $k_A$  trials,  $m$  successes, and success probability  $1 - x$ ,

$$F(k_A, x) = \sum_{j=0}^m \binom{k_A}{j} (1-x)^j x^{k_A-j}. \quad (3)$$

Thus, Eq. (1) is a self-consistent equation for  $\rho^R$ , given the initial condition  $\rho^+$ , whereas the first equality of Eq. (2) describes the time evolution of  $\rho^+$  with the initial conditions given by the steady state of the dynamics in layer B. Equating this to 0, as in the second equality of Eq. (2), yields the condition for the steady state.

A bifurcation analysis of the two equations above shows that the number of stable states of the system depends on the value of the parameter  $\beta$ . So, for example, when  $\beta = 0.18$ , the solution curves of Eqs. (1) and (2) have a single intersection, at a low value of  $\rho^+ \approx 0.1$  and an infinitesimal value of  $\rho^R$ , as shown in Fig. 3(a). This indicates that, in this regime, a large majority of people ends up holding the negative opinion and, at the same time, the message cannot spread through the

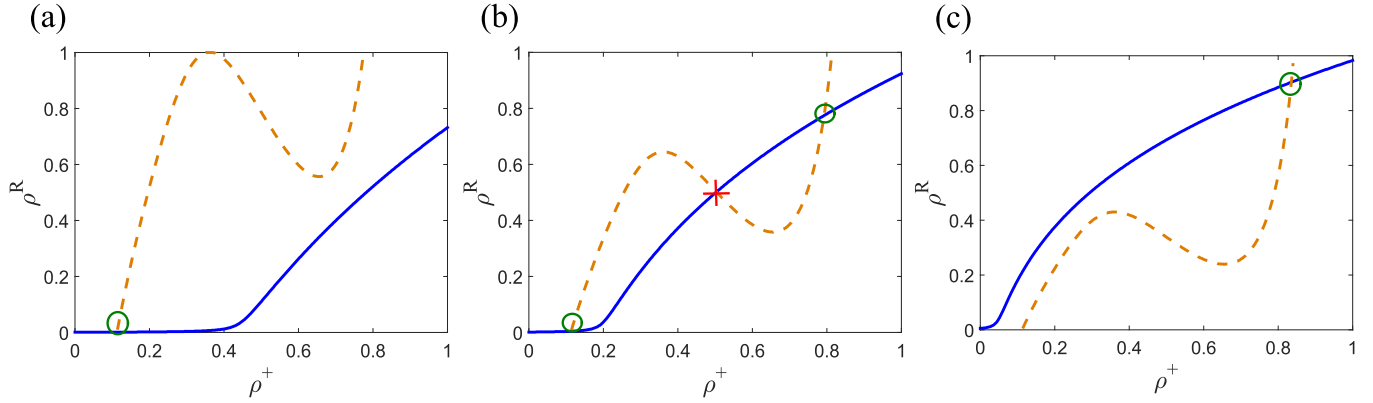


FIG. 3. The number of steady states of the model depends on the message transmission rate  $\beta$ . (a) For  $\beta = 0.18$ , there is only one intersection (green circle) of the solution of Eq. (1) (blue solid line) and that of Eq. (2) (brown dashed line). (b) For  $\beta = 0.28$ , there are three intersections, representing two stable states (green circles) and one unstable one (red cross). (c) For  $\beta = 0.42$ , the number of steady states returns to one (green circle). In all panels, the solutions are obtained for  $\alpha = 0.2$ ,  $\gamma = 0.1$ ,  $\mu = 1.0$ ,  $r = 1.0$ ,  $m = 4$ , and  $\nu = 0.14$ , and both layers are random regular networks with  $k = 10$ .

population. When  $\beta$  increases to 0.28, the curves intersect in three points, as illustrated in Fig. 3(b). However, not all three solutions correspond to the stable states. In fact, due to the collaborative interplay between message dissemination and positive opinion, for a state to be stable, it has to correspond to a point where  $\rho^R$  increases monotonically with  $\rho^+$ , and vice versa. Thus, only the two points with coordinates approximately (0.1, 0) and (0.8, 0.8) are stable. Finally, if we increase  $\beta$  even more to 0.42, the two curves have once more only one intersection, corresponding to a state where approximately 90% of the agents have been reached by the message, and more than 80% hold a positive opinion, as shown in Fig. 3(c). Further stability considerations are discussed in Appendix B.

Based on the analysis above, which shows how the number of stable states changes with the message transmission rate  $\beta$ , we expect a discontinuous transition to happen around a critical value  $\beta = \beta_c$ . To verify the accuracy of this prediction, we performed numerical simulations on a duplex network with random regular layers of size  $N = 10^4$  and degree  $k = 10$ . The results, illustrated in Fig. 4, show that the variation of  $\rho^+$  with  $\beta$  depends on the specific value of the noise parameter  $\nu$ . As shown in Fig. 4(a), for low enough values of  $\nu$ , the positive-opinion density  $\rho^+$  remains low until  $\beta$  increases to a forward critical value  $\beta_c^F$ , where a first-order transition occurs and  $\rho^+$  jumps up to a high value. As one, then, decreases  $\beta$ , a hysteresis cycle appears, and  $\rho^+$  remains high until  $\beta$  reaches a backward critical value  $\beta_c^B < \beta_c^F$ . Thus, for low noise, two stable states exist, one with low-positive-opinion density (LP) and one with high-positive-opinion density (HP). However, the size of the hysteresis loop decreases with the strength of the noise, and the transition undergoes a crossover to a continuous one when  $\nu$  is above a critical noise value  $\nu_c$ , in good agreement with the theoretical results.

The mean-field theory also accurately predicts the forward and backward critical transmission rates and the location of the critical noise  $\nu_c$ , as shown in the phase diagram of Fig. 4(b). The values of the order parameters  $\rho^+$  and  $\rho^R$  at steady state for each choice of  $\beta$  and  $\nu$  are shown in Figs. 4(c) and 4(d), illustrating that for parameters that correspond to the LP or the HP phases, the system does indeed evolve to a

low-positive-opinion or a high-positive-opinion state. Conversely, for parameters within the bistable phase region, the system evolves to either stable state depending on the initial conditions. The right-hand and left-hand spinodals that surround the bistable region in the phase diagram are first-order transition lines, corresponding, respectively, to  $\beta_c^F$  and  $\beta_c^B$  for each  $\nu$ . For details on the effect of the coupling parameters  $\alpha$  and  $\gamma$  on the transition curves, see Appendix C.

### C. Phase flipping

Studying the time evolution of the system, one can observe an interesting phenomenon. As illustrated in Fig. 5(a), when the system is within the bistable region, the fraction of nodes with positive opinion  $\rho^+$  repeatedly flips between the LP phase and the HP one, with a bimodal probability distribution, shown in Fig. 5(b), that is consistent with this observation. This kind of switching is often called “phase flipping,” and a typical example of this behavior that has been widely observed and studied occurs in the spontaneous recovery model (SRM) [38–41]. In a way similar to the noisy threshold voter model, active nodes in the SRM switch to the inactive state due to internal or external failures. However, inactive nodes can recover spontaneously and become active again at a fixed rate or after a fixed time interval, making the switching between the two states asymmetric. In contrast, in our model, the transition between positive and negative opinions is symmetric with respect to its direction. It is therefore natural to wonder whether the spontaneous recovery mechanism is not the only factor that can lead to phase flipping.

To formulate a hypothesis about this, we start by noting that, unlike the SRM, which describes an isolated dynamical process, our coupled model characterizes the coevolutionary dynamics of opinion formation and message propagation. Thus, the first question to answer is which of the two processes has a stronger effect on the phase flipping. Due to the different evolutionary timescales between opinion dynamics and message diffusion, we answer this question by assuming that the primary driver of switches in the macroscopic state of the system is the opinion dynamics process, in spite of the synergetic coupling effects between the two.

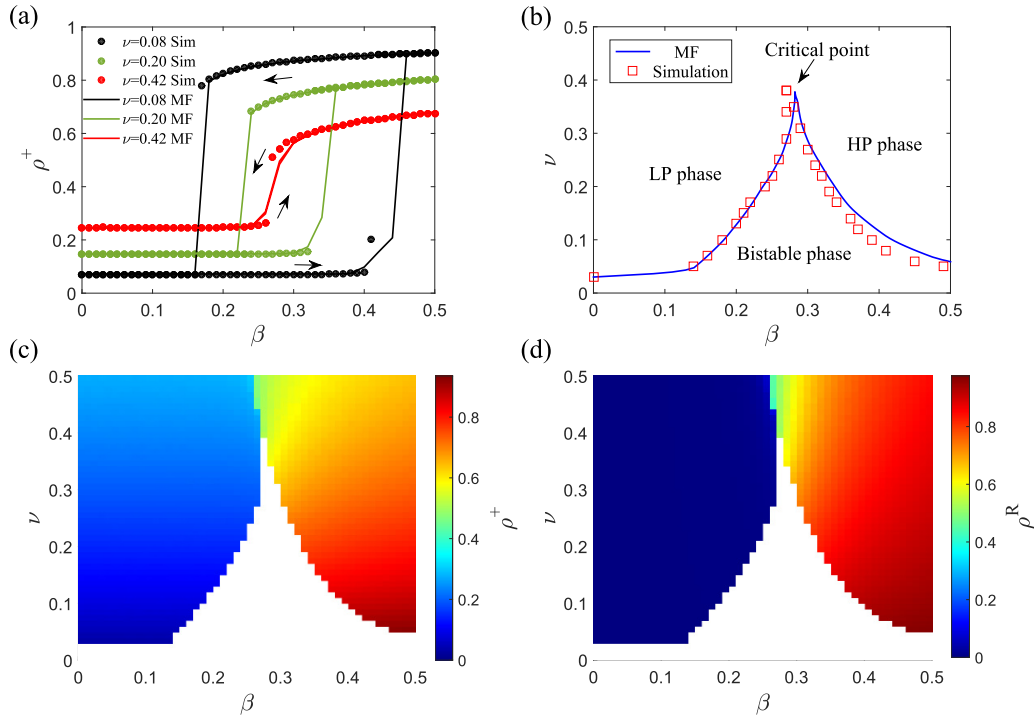


FIG. 4. The order of the transition in our coupled model depends on the intensity of the noise. (a) The fraction of nodes holding a positive opinion  $\rho^+$  undergoes a first-order transition for noise intensities  $\nu$  smaller than critical value  $\nu_c$ . When the noise exceeds  $\nu_c$ , the hysteresis loop disappears. Simulation results (colored dots) are compared with mean-field theory predictions (solid lines). (b) The mean-field theory predicts quite accurately the phase diagram of the system. Here, the positive opinion is held by a minority of nodes in the “LP phase” and by a majority of nodes in the “HP phase.” The “bistable phase” represents a region from where the system evolves either toward the LP phase or the HP one depending on the initial conditions. (c), (d) The order parameters  $\rho^+$  and  $\rho^-$ , obtained from numerical simulations, are consistent with the phase diagram of panel (b), with the bistable phase indicated in white. In all panels above, the results are the average of ten realizations of a bilayer regular random network with  $N = 10^4$  nodes,  $\alpha = 0.2$ ,  $\gamma = 0.1$ ,  $\mu = 1.0$ ,  $r = 1.0$ ,  $m = 4$ , and  $k = 10$ .

Next, we have already shown that, when the system is in the bistable phase, it evolves toward the LP phase or the HP one depending on the initial conditions. Thus, we conjecture that, in finite-size systems, the fluctuations during the process may cause the system to deviate from its current stable state and transition to the other one. Specifically, stochastic fluctuations could push the system toward a state that matches what the expected behavior would be for different values of the parameters, which, in turn, would correspond to conditions that evolve toward the other stable state. The next question, then, is what the relevant parameters are.

To answer this, note that, according to Eq. (2), the evolution of the opinion dynamics is determined jointly by  $\nu$ ,  $\gamma$ , and  $r$ . However, while  $\nu$  controls the opinion flipping of all the nodes,  $\gamma$  and  $r$  only affect the strength of the neighbor-induced flipping mechanism. Now note that in our model, when the system is in the LP phase, positive-opinion nodes have a high probability of satisfying the threshold condition, even though they are a very small minority; conversely, negative-opinion nodes, which constitute the vast majority of the network, have a very small probability of satisfying the threshold. Thus, only a very small fraction of nodes have the potential to flip their opinion by interacting with their neighbors when the system is in the LP phase. A mirror argument, with the same conclusion, holds when the system is in HP phase. Therefore, we conclude that the relevant fluctuations that may induce phase flipping are those that change the effective value of  $\nu$ , rather than  $\gamma$

and  $r$ . To validate this consideration and further investigate the underlying mechanism of phase flipping, we analyze the relation between the fluctuations of  $\nu$  and the dynamics of phase flipping. Due to the stochasticity of numerical simulations, the fraction of nodes changing their opinion states at any given time is not equal to  $\nu$  exactly, but fluctuates around it. Then, it is convenient to indicate with  $\nu^+$  the effective rate for the nodes that convert from a negative opinion to a positive one via noise-induced flipping, and with  $\nu^-$  the one for the nodes that convert from a positive opinion to a negative one via the same mechanism. This allows us to rewrite Eq. (2) as

$$\begin{aligned} \frac{d}{dt}\rho^+ = & \left( \nu^+ + \sum_{k_A} p_{k_A} F(k_A, \rho^+) \gamma r \right) \rho^{-S} \\ & + \left( \nu^+ + \sum_{k_A} p_{k_A} F(k_A, \rho^+) r \right) \rho^{-R} \\ & - \left( \nu^- + \sum_{k_A} p_{k_A} F(k_A, \rho^-) r \right) \rho^{+S} \\ & - \left( \nu^- + \sum_{k_A} p_{k_A} F(k_A, \rho^-) \gamma r \right) \rho^{+R} = 0 \end{aligned} \quad (4)$$

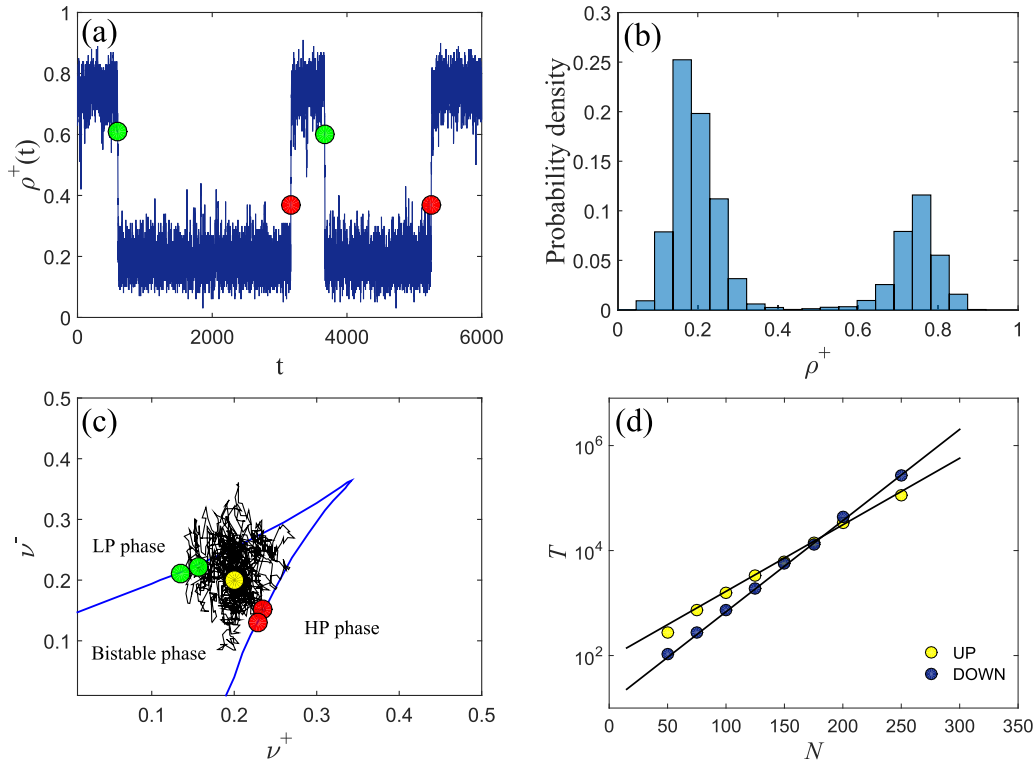


FIG. 5. Phase flipping in finite systems. (a) The fraction of nodes holding a positive opinion,  $\rho^+$ , repeatedly switches between a majority of approximately 0.75 to a minority of approximately 0.2. The moments when the switches occur are marked by colored dots. (b) As a consequence of the phase flipping described in panel (a), the probability distribution of  $\rho^+$  is bimodal. (c) Mapping the simulation trajectory of panel (a) onto the phase diagram on the  $(\nu^+, \nu^-)$  plane shows that the points at which the system flips from positive majority to negative majority (green dots) correspond to the trajectory  $(\nu_\lambda^+(t), \nu_\lambda^-(t))$  (black line) crossing the upper spinodal into the LP phase. Similarly, the points at which the system flips to positive majority (red dots) correspond to the trajectory crossing the lower spinodal into the HP phase. The yellow dot corresponds to the nominal value  $\nu = \nu^+ = \nu^- = 0.2$  chosen for the simulations. The green dots and the red ones represent the position of  $(\nu_\lambda^+(t), \nu_\lambda^-(t))$  on the plane at the times that are marked by the same colored dots in panel (a). Note that the black line only shows the trajectory  $(\nu_\lambda^+(t), \nu_\lambda^-(t))$  from  $t = 3000$  to the first subsequent phase flipping, which occurs at  $t = 3167.5$ . (d) The average permanence times in the HP phase (yellow dots, “UP”) and in the LP phase (blue dots, “DOWN”) scale exponentially with system size, in agreement with Eqs. (5) and (6). The results shown are averages over 1000 independent simulations. For all panels, the system is a two-layer random regular network with  $N = 100$  nodes of degree 10,  $\nu = 0.2$ ,  $\beta = 0.3$ ,  $\alpha = 0.2$ ,  $\gamma = 0.1$ ,  $\mu = 1.0$ ,  $r = 1.0$ , and  $m = 4$ .

and to construct a phase diagram in the  $(\nu^+, \nu^-)$  plane. The diagram, shown in Fig. 5(c), features two spinodals delimiting the region of bistability between the LP phase and the HP one and suggests that the inequality between  $\nu^+$  and  $\nu^-$  may induce the system to undergo flipping transitions. To map the simulation trajectory onto this phase diagram, we define  $\nu_\lambda^+(t)$  as the time-averaged fraction of nodes converting via noise from the negative opinion to the positive one in the time interval  $[t - \lambda, t]$  and indicate with  $\nu_\lambda^-(t)$  the time-averaged fraction of nodes undergoing the opposite switch in the same time interval. Thus,  $\nu_\lambda^+(t)$  and  $\nu_\lambda^-(t)$  can be viewed as effective noise parameters within each time interval. Here,  $\lambda$  is a parameter that, to give meaningful results, should correspond to the system relaxation time. For a network of 100 nodes, this is  $\lambda = 5$  [38]. This construction shows that the points at which the system flips from majority-positive to majority-negative opinion lie on the upper spinodal, whereas those corresponding to the opposite flipping lie on the lower one. This indicates that when the system switches from the HP phase to the LP one, its trajectory in the  $(\nu_\lambda^+, \nu_\lambda^-)$  plane crosses the spinodal from the bistable region into the LP one, whereas when it flips

from the LP phase to the HP one, the trajectory crosses the other spinodal into the HP region.

Additionally, we can compute the expected permanence time in the LP phase and in the HP one, which we indicate with  $T_{\text{down}}$  and  $T_{\text{up}}$ , respectively, in the assumption of the existence of two critical values  $\nu_s^+$  and  $\nu_s^-$  such that the system moves to the LP phase whenever  $\nu^- > \nu_s^-$ , and it switches to the HP phase whenever  $\nu^+ > \nu_s^+$ . The expected permanence times in the two phases, then, are given by the following equations (see Appendix D for details):

$$T_{\text{down}}(N) \approx e^{\frac{N(1-\langle\rho^+\rangle_{\text{low}})(\nu_s^+-\nu^+)^2}{2\nu^+(1-\nu^+)}} \quad (5)$$

$$T_{\text{up}}(N) \approx e^{\frac{N\langle\rho^+\rangle_{\text{high}}(\nu_s^--\nu^-)^2}{2\nu^-(1-\nu^-)}} \quad (6)$$

Here,  $\langle\rho^+\rangle_{\text{low}}$  and  $\langle\rho^+\rangle_{\text{high}}$  represent the average fraction of positive-opinion nodes in the LP phase and in the HP one, respectively. Because the strength of the stochastic fluctuations ultimately decreases with system size, the two equations above predict that the average time that the system spends in one of the two phases increases exponentially with the number

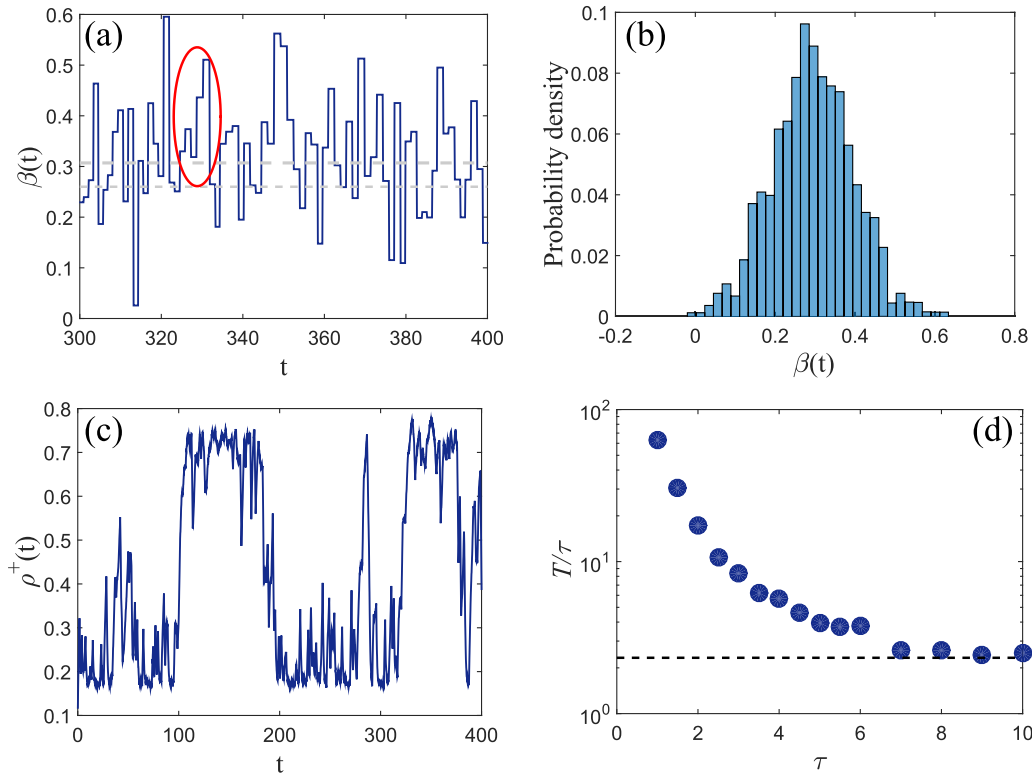


FIG. 6. Macroscopic state transitions can be induced by varying the transmission rate. (a) The transmission rate  $\beta$ , which is extracted at the end of each time interval of size  $\tau = 1.4$ , occasionally increases above the forward critical value  $\beta_c^F = 0.307$  (upper gray dashed line) or decreases below the backward critical value  $\beta_c^B = 0.260$  (lower gray dashed line). As an illustrative example, in the period from  $t = 324.8$  to  $t = 331.8$ ,  $\beta$  exceeds  $\beta_c^F$  for a duration of  $5\tau$ , highlighted by a red ellipse. This corresponds to a switch from the LP phase to the HP one, as seen in panel (c). (b) The probability density of  $\beta$  is Gaussian. (c) The value of  $\rho^+$  shows several switches between the LP phase and the HP one, due to the fluctuations of  $\beta$ . (d) As  $\tau$  increases, the ratio  $\frac{T}{\tau}$  converges to the value given by Eq. (7). The results are averaged over one 100 independent simulations of a two-layer random regular network with  $N = 10^4$  nodes of degree 10,  $\nu = 0.25$ ,  $\alpha = 0.2$ ,  $\gamma = 0.1$ ,  $m = 4$ ,  $\mu = 1$ ,  $r = 1$ ,  $\bar{\beta} = 0.28$ , and  $\sigma^2 = 8.71 \times 10^{-3}$ .

of nodes  $N$ . This is confirmed by numerical simulations, as shown in Fig. 5(d), and indicates that phase flipping can only be consistently observed in small enough systems.

#### D. Effect of message transmission rate

The mechanism discussed in the previous subsection can explain how macroscopic changes in system state can be caused by internal fluctuations in the rate of spontaneous opinion switching. However, fluctuations of this kind are inhibited in real-world systems with large size. Here, we show how such flipping phenomena can also be driven by external fluctuations in the message transmission rate  $\beta$ . To explore this possibility, we simulate the model changing the value of  $\beta$  at every time step. Specifically, we extract  $\beta$  from a normal distribution with mean  $\bar{\beta}$  and variance  $\sigma^2$ , as shown in Figs. 6(a) and 6(b). Thus, the evolution of the system can be considered to happen at a fixed transmission rate  $\bar{\beta}$  affected by small random perturbations.

Because of the different timescales between the processes of message spreading and opinion formation, the transient fluctuations of  $\beta$  do not necessarily cause a persistent change in the collective opinion as measured by  $\rho^+$ . However, if  $\beta$  exceeds the forward critical transmission rate  $\beta_c^F$  for long enough, or if it remains smaller than the backward critical rate

$\beta_c^B$  for an equally sufficient time, then the system switches to the HP phase or to the LP one, respectively. This is clear from the values of  $\rho^+$  along the simulations, illustrated in Fig. 6(c), which show repeated switches between the two phases.

The permanence time of the system in either of the two majority phases depends on both the interval  $\tau$  between updates of  $\beta$  and on the probability that the value of  $\beta$  lies beyond the corresponding critical one. Thus, for large  $\tau$ , the permanence time in the LP phase,  $T_{\text{low}}$ , and that in the HP phase,  $T_{\text{high}}$ , can be expressed as

$$T_{\text{low}}(\tau) = \tau \left( \frac{1}{2} - \frac{1}{2} \operatorname{erf} \left( \frac{\beta_c^F - \bar{\beta}}{\sqrt{2}\sigma} \right) \right)^{-1} \quad (7)$$

and

$$T_{\text{high}}(\tau) = \tau \left( \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{\beta_c^B - \bar{\beta}}{\sqrt{2}\sigma} \right) \right)^{-1}, \quad (8)$$

where  $\operatorname{erf}$  is the error function. A direct measurement of these times via numerical simulations confirms the correctness of these equations, as illustrated in Fig. 6(d).

These findings make it naturally interesting to explore the effect of the interplay between the two types of fluctuations. To do so, we simulated a smaller system, with  $N = 100$  nodes, to ensure that the two mechanisms described above can

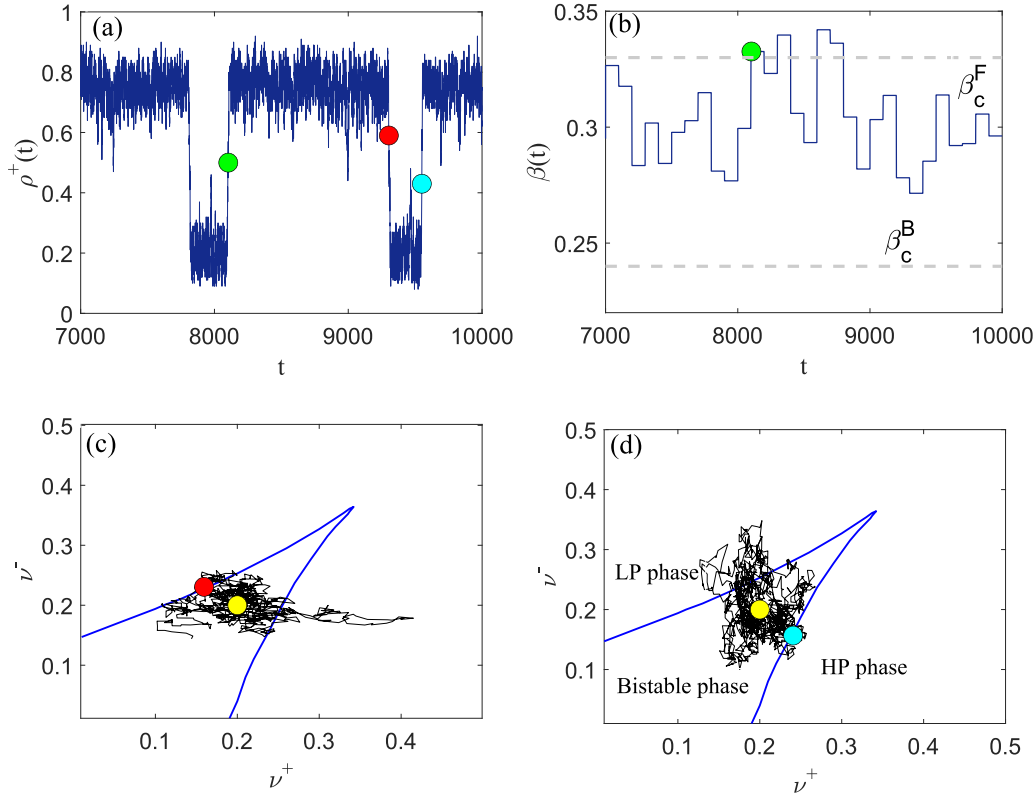


FIG. 7. Interplay of the fluctuation-driven mechanisms on phase flipping. (a) The fraction of nodes holding a positive opinion,  $\rho^+$ , repeatedly switches between a majority of approximately 0.8 to a minority of approximately 0.2. The moments when the switches occur are marked by colored dots. (b) In the time period between  $t = 7000$  and  $t = 10000$ , the transmission rate  $\beta$  occasionally increases above the forward critical value  $\beta_c^F$  (upper gray dashed line), but never decreases below the backward critical value  $\beta_c^B$  (lower gray dashed line). The values of  $\beta$  follow a normal distribution with mean 0.3 and variance  $2.25 \times 10^{-4}$ . The sampling interval is  $\tau = 100$  and the green circle indicates the value of  $\beta$  at the moment of the phase flipping shown in green in panel (a). (c) Mapping the simulation trajectory of panel (a) from  $t = 9200$  to  $t = 9303.3$  onto the phase diagram on the  $(\nu^+, \nu^-)$  plane shows that during this interval, the system is within the HP phase and the bistable region. At  $t = 9303.3$ , the trajectory reaches the demarcation boundary (red circle), indicating an imminent transition toward the LP phase. (d) The simulation trajectory from  $t = 9400$  to  $t = 9550$  shows that the system lies within the LP phase and the bistable phase. At  $t = 9550$ , it intersects the lower boundary (cyan circle), signaling an impending phase flipping toward the HP phase. For all panels, the system is a two-layer random regular network with  $N = 100$  nodes of degree 10,  $\nu = 0.2$ ,  $\beta = 0.3$ ,  $\alpha = 0.2$ ,  $\gamma = 0.1$ ,  $\mu = 1.0$ ,  $r = 1.0$ , and  $m = 4$ . The yellow circles indicate the nominal values of  $\nu = \nu^+ = \nu^- = 0.2$  chosen for the simulations.

coexist, and analyzed the temporal evolution of the system, disentangling the contributions of each mechanism.

The results, illustrated in Fig. 7, show that three phase flips occur within the time interval from  $t = 7000$  to  $t = 10000$ . The first one is due to  $\beta(t)$  exceeding the forward critical threshold  $\beta_c^F$ . The next two, however, are induced by intrinsic fluctuations of the parameter  $\nu$  as the trajectory  $(\nu_\lambda^+(t), \nu_\lambda^-(t))$  crosses the phase demarcation boundary. This can be seen from the trajectory of  $(\nu_\lambda^+(t), \nu_\lambda^-(t))$  on the  $(\nu^+, \nu^-)$  phase diagram in correspondence of the last two phase flips, which we plot in panels (c) and (d) of the figure. It is worth to note that if the two mechanisms operate simultaneously and exert opposing effects, for instance, when  $\beta(t)$  exceeds  $\beta_c^F$  but the trajectory  $(\nu_\lambda^+(t), \nu_\lambda^-(t))$  is in the LP region, they may interfere with each other, potentially suppressing the phase flipping from the LP phase to the HP phase.

The effect of the heterogeneity of the network structure is further explored in Appendix E.

### III. DISCUSSION

In summary, we proposed a noisy threshold voter-SIR coupled model that takes into account the interplay between message spreading and opinion formation, and features differences in the characteristic timescales of the two constituent dynamical processes. Unlike the classic noisy voter model, in which there is no steady state, our coupled model exhibits a bistable phase with hysteresis loops that disappear if the noise is greater than a critical value. Conversely, when the noise is low and individuals are not very likely to switch opinions autonomously, two message transmission rates appear that determine a forward transition and a backward one between states with definite and opposite opinion majorities. Note that our results assume that the timescale of opinion formation is much larger than that of the message propagation. In most real-world cases, this is a reasonable assumption. Nonetheless, it is one that is not necessarily always satisfied. In the future, we plan to explore the role of timescale difference in

the coevolution dynamics, as well as the effect of network topology and memory effects on final steady states of the model. These are particularly relevant questions, since recent studies indicate that in complex contagion processes, such as in the presence of social reinforcement, the ultimate effects of network structure depend strongly on the specific model used [42,43].

A detailed analysis of the dynamics of opinion switching reveals that, in its current form, our model has two intrinsically different mechanisms that can induce macroscopic changes in the consensus state of the system. The first mechanism is a spontaneous symmetry breaking between  $v^+$  and  $v^-$  that can be induced by stochastic fluctuations in small enough systems, causing a phase flipping. It is worth noting that the possibility of phase flipping means that the system can transition from a full-functioning state to a nonfunctional one without the control parameters changing. Moreover, the fact that the phase flipping can be caused by fluctuations in the noise parameter suggests that the lack of depth in the conviction of individuals may constitute a more serious problem than the peer pressure of their social contacts for the goal of controlling the public opinion.

Additionally, macroscopic changes in the consensus state can be caused by fluctuations in the rate of message transmission. However, since these fluctuations originate from the external environment, in general their distribution varies with time, and their macroscopic effects can be observed regardless of the size of the system. This implies that, in a real-world situation, the rise and decline in the attention given to a particular type of news can lead to an abrupt and drastic shift in the collective opinion.

Additionally, the evolution of the system in the bistable region depends on the initial distribution of opinions among individuals. This confirms the effectiveness of strategies that have been adopted by some governments, consisting in the leveraging of so-called bot farms to create a constant stream of posts on social media with the aim of fostering and maintaining a desired general attitude of the public.

#### ACKNOWLEDGMENTS

M.T. acknowledges support from the National Natural Science Foundation of China (Grants No. 12231012 and No. 42461144209), and from the Science and Technology Commission of Shanghai Municipality (Grant No. 22DZ2229004). C.I.D.G. acknowledges support from the Bulgarian Ministry of Education and Science under Project No. BG-RRP-2.004-0006-C02.

#### DATA AVAILABILITY

The data that support the findings of this article are openly available [45].

#### APPENDIX A: MEAN-FIELD THEORY

We start by analyzing the message spreading process, which takes place on layer B, and we indicate the six possible joint states of the nodes as  $\pm S$ ,  $\pm I$ , and  $\pm R$ . On average, the

time evolutions of the densities of nodes in  $+S$ ,  $+I$ , and  $+R$  states can be written as

$$\frac{d}{dt}\rho^{+S}(t) = -\beta\langle k_B\rangle\rho^I(t)\rho^{+S}(t), \quad (\text{A1})$$

$$\frac{d}{dt}\rho^{+I}(t) = -\mu\rho^{+I}(t) + \beta\langle k_B\rangle\rho^I(t)\rho^{+S}(t), \quad (\text{A2})$$

$$\frac{d}{dt}\rho^{+R}(t) = \mu\rho^{+I}(t), \quad (\text{A3})$$

where  $\langle k_B\rangle$  is the average degree of the nodes in layer B,  $\rho^{\pm S}$ ,  $\rho^{\pm I}$ , and  $\rho^{\pm R}$  are the fractions of nodes in state  $\pm S$ ,  $\pm I$ , and  $\pm R$ , respectively, and  $\rho^S$ ,  $\rho^I$ , and  $\rho^R$  are the fractions of nodes in states  $S$ ,  $I$ , and  $R$ , respectively, regardless of their opinion.

Similarly, for nodes holding a negative opinion, the evolution equations are

$$\frac{d}{dt}\rho^{-S}(t) = -\alpha\beta\langle k_B\rangle\rho^I(t)\rho^{-S}(t), \quad (\text{A4})$$

$$\frac{d}{dt}\rho^{-I}(t) = -\mu\rho^{-I}(t) + \alpha\beta\langle k_B\rangle\rho^I(t)\rho^{-S}(t), \quad (\text{A5})$$

$$\frac{d}{dt}\rho^{-R}(t) = \mu\rho^{-I}(t). \quad (\text{A6})$$

From Eqs. (A1) and (A4), it follows that

$$\frac{1}{\rho^{-S}(t)}\frac{d}{dt}\rho^{-S}(t) = \alpha\frac{1}{\rho^{+S}(t)}\frac{d}{dt}\rho^{+S}(t). \quad (\text{A7})$$

Integrating this equation, we obtain

$$\rho^{-S}(t) = \rho^{-S}(0)e^{\alpha\ln\frac{\rho^{+S}(t)}{\rho^{+S}(0)}}. \quad (\text{A8})$$

When the message spreading process reaches its final state, no node is in the  $I$  state, so that  $\rho^I = 0$ . To obtain the final density of nodes in the  $R$  state, we need to find the relationship between  $\rho^S$  and  $\rho^R$ . To do so, we start from Eqs. (A1), (A3), (A4), and (A6), to find the evolution equations for  $\rho^S$  and  $\rho^R$ ,

$$\frac{d}{dt}\rho^S(t) = -\beta\langle k_B\rangle\rho^I(t)(\rho^{+S}(t) + \alpha\rho^{-S}(t)) \quad (\text{A9})$$

and

$$\frac{d}{dt}\rho^R(t) = \mu\rho^I(t). \quad (\text{A10})$$

Then, using these, we obtain

$$\frac{1}{\rho^{+S}(t) + \alpha\rho^{-S}(t)}\frac{d}{dt}\rho^S(t) = -\frac{\beta\langle k_B\rangle}{\mu}\frac{d}{dt}\rho^R(t). \quad (\text{A11})$$

However,  $\rho^S(t) = \rho^{-S}(t) + \rho^{+S}(t)$ . Substituting this expression in the equation above, and using Eq. (A7), we get

$$\frac{1}{\rho^{+S}(t)}\frac{d}{dt}\rho^{+S}(t) = -\frac{\beta\langle k_B\rangle}{\mu}\frac{d}{dt}\rho^R(t). \quad (\text{A12})$$

Integrating this equation, we obtain

$$\rho^{+S}(t) = \rho^{+S}(0)e^{-\frac{\beta\langle k_B\rangle\rho^R(t)}{\mu}}. \quad (\text{A13})$$

Combining Eqs. (A8) and (A13), we obtain a self-consistent equation for the steady-state fraction of nodes in the  $R$  state

$$\begin{aligned} \rho^R &= 1 - \rho^{+S} - \rho^{-S} \\ &= 1 - \rho^{+S}(0)e^{-\frac{\beta\langle k_B\rangle\rho^R}{\mu}} - \rho^{-S}(0)e^{-\frac{\alpha\beta\langle k_B\rangle\rho^R}{\mu}}, \end{aligned} \quad (\text{A14})$$

where  $\rho^{\pm S}$  is the steady-state fraction of nodes in the  $\pm S$  state. In the thermodynamic limit, assuming that the initial fraction of spreaders is very small, we can approximate the steady-state fractions of nodes with positive or negative opinion,  $\rho^+$  and  $\rho^-$ , with the initial fractions of susceptible nodes, obtaining

$$\rho^R = 1 - \rho^+ e^{-\frac{\beta \langle k_B \rangle \rho^R}{\mu}} - \rho^- e^{-\frac{\alpha \beta \langle k_B \rangle \rho^R}{\mu}}, \quad (\text{A15})$$

which is Eq. (1).

Next, we analyze the process of opinion formation on layer A. When the message spreading on layer B reaches its final state, the node can be in states  $\pm S$  and  $\pm R$ . Then, the opinion of the nodes in layer A evolves on the basis of their state in layer B, and the evolution equation for the fraction of nodes holding positive opinion is

$$\begin{aligned} \frac{d}{dt} \rho^+ = & \left( v + \sum_{k_A} p_{k_A} F(k_A, \rho^+) \gamma r \right) \rho^{-S} \\ & + \left( v + \sum_{k_A} p_{k_A} F(k_A, \rho^+) r \right) \rho^{-R} \\ & - \left( v + \sum_{k_A} p_{k_A} F(k_A, \rho^-) r \right) \rho^{+S} \\ & - \left( v + \sum_{k_A} p_{k_A} F(k_A, \rho^-) \gamma r \right) \rho^{+R}, \end{aligned} \quad (\text{A16})$$

where  $p_{k_A}$  is the degree distribution in layer A, and  $F(k_A, \rho^+)$  and  $F(k_A, \rho^-)$  are the probabilities that, given a node with

degree  $k$ , the numbers of its neighbors with negative or positive opinion are less than or equal to the threshold  $m$ . In the assumption of a well-mixed system with no degree-degree correlations, this is

$$F(k_A, x) = \sum_{j=0}^m \binom{k_A}{j} (1-x)^j x^{k_A-j}, \quad (\text{A17})$$

which is Eq. (3). At steady state, the densities of nodes in each state do not change. Thus, the final equation is obtained by imposing that  $\frac{d}{dt} \rho^+$  vanish in Eq. (A16), which results in Eq. (2).

## APPENDIX B: STABILITY OF THE COUPLED MODEL

Equations (1) and (2) characterize the evolution of our coupled model under a mean-field approximation. Here, we prove the stability of their equilibrium point by finding an appropriate Lyapunov function  $V$ .

To do so, we first note that  $V$  must be a function of  $\rho^+$  and  $\rho^R$ . Then, when taking its derivative with respect to time, we get

$$\begin{aligned} \frac{d}{dt} V(\rho^+, \rho^R) &= \frac{\partial}{\partial \rho^+} V \frac{d}{dt} \rho^+ + \frac{\partial}{\partial \rho^R} V \frac{d}{dt} \rho^R \\ &= \frac{\partial}{\partial \rho^+} V \frac{d}{dt} \rho^+ + \frac{\partial}{\partial \rho^+} V \frac{d}{d \rho^R} \rho^+ \frac{d}{d \rho^+} \rho^R \frac{d}{dt} \rho^+ \\ &= 2 \frac{\partial}{\partial \rho^+} V \frac{d}{dt} \rho^+. \end{aligned} \quad (\text{B1})$$

This shows that it would be useful to find a form for  $V$  that related its functional dependence on  $\rho^+$  with that of the evolution of  $\rho^+$  in time. To reach this goal, we exploit Eqs. (2) and (3) and assume  $\frac{\partial}{\partial \rho^+} V = -\frac{d}{dt} \rho^+$ , obtaining the form

$$\begin{aligned} \frac{\partial}{\partial \rho^+} V = & -v + 2v\rho^+ - \sum_{k_A} p_{k_A} \sum_{j=0}^m \binom{k_A}{j} (1-\rho^+)^{j+1} (\rho^+)^{k_A-j} r \left[ 1 + (\gamma-1) e^{-\frac{\alpha \beta \langle k_B \rangle \rho^R}{\mu}} \right] \\ & + \sum_{k_A} p_{k_A} \sum_{j=0}^m \binom{k_A}{j} (1-\rho^+)^{k_A-j} (\rho^+)^{j+1} r \left[ \gamma + (1-\gamma) e^{-\frac{\beta \langle k_B \rangle \rho^R}{\mu}} \right]. \end{aligned} \quad (\text{B2})$$

Next, we use the binomial theorem to expand the powers of  $(1-\rho^+)$ , obtaining

$$\begin{aligned} \frac{\partial}{\partial \rho^+} V = & -v + 2v\rho^+ - \sum_{k_A} p_{k_A} \sum_{j=0}^m \binom{k_A}{j} \sum_{l=0}^{j+1} \binom{j+1}{l} (-1)^l (\rho^+)^{l+k_A-j} r \left[ 1 + (\gamma-1) e^{-\frac{\alpha \beta \langle k_B \rangle \rho^R}{\mu}} \right] \\ & + \sum_{k_A} p_{k_A} \sum_{j=0}^m \binom{k_A}{j} \sum_{l=0}^{k_A-j} \binom{k_A-j}{l} (-1)^l (\rho^+)^{l+j+1} r \left[ \gamma + (1-\gamma) e^{-\frac{\beta \langle k_B \rangle \rho^R}{\mu}} \right]. \end{aligned} \quad (\text{B3})$$

Finally, to reach the form we are aiming for, we take the indefinite integral of the equation above, resulting in

$$\begin{aligned} V(\rho^+, \rho^R) = & -v\rho^+(1-\rho^+) - \sum_{k_A} p_{k_A} \sum_{j=0}^m \binom{k_A}{j} \sum_{l=0}^{j+1} \binom{j+1}{l} (-1)^l \frac{(\rho^+)^{l+k_A-j+1}}{l+k_A-j+1} r \left[ 1 + (\gamma-1) e^{-\frac{\alpha \beta \langle k_B \rangle \rho^R}{\mu}} \right] \\ & + \sum_{k_A} p_{k_A} \sum_{j=0}^m \binom{k_A}{j} \sum_{l=0}^{k_A-j} \binom{k_A-j}{l} (-1)^l \frac{(\rho^+)^{l+j+2}}{l+j+2} r \left[ \gamma + (1-\gamma) e^{-\frac{\beta \langle k_B \rangle \rho^R}{\mu}} \right]. \end{aligned} \quad (\text{B4})$$

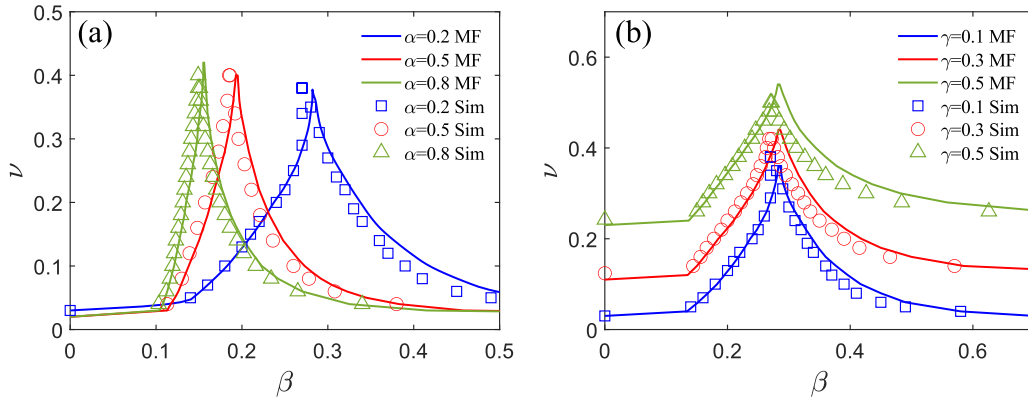


FIG. 8. The coupling parameters affect the critical transmission rates differently. (a) For a fixed  $\gamma = 0.1$ , higher values of  $\alpha$  shift the transition curves toward lower values of  $\beta$ . (b) For a fixed  $\alpha = 0.2$ , higher values of  $\gamma$  shift the transition curves symmetrically toward higher values of  $\nu$ . In both panels, the symbols correspond to simulation results for a two-layer random regular network of  $N = 10^4$  nodes with degree 10,  $\mu = 1.0$ ,  $r = 1.0$ , and  $m = 4$ , and solid lines represent the predictions of the mean-field theory.

The function  $V$  in the equation above is a valid Lyapunov function that satisfies  $\frac{\partial}{\partial \rho^+} V = -\frac{d}{dt} \rho^+$ . In turn, from Eq. (B1), this yields

$$\frac{d}{dt} V(\rho^+, \rho^R) = -2 \left( \frac{\partial}{\partial \rho^+} V \right)^2 < 0, \quad (\text{B5})$$

which demonstrates that in our model there is no oscillatory behavior, as instead observed in failure-recovery models [39,40] and in coevolution dynamics with asymmetric interactions [34,44]. Note that the phase flipping in our model is markedly different from the oscillatory behaviors just mentioned. In fact, the former consists in the stochastic switching between two stable solutions in a bistable region, whereas the latter indicates that the system has a deterministic closed orbit. In addition, the phase flipping is induced by stochastic fluctuations and it requires small systems sizes to happen, while the rapid decay and growth dynamics in oscillatory behavior are periodic and can occur also in large-scale systems.

### APPENDIX C: IMPACT OF COUPLING ON CRITICAL RATES

In Sec. II B, we discuss the phase diagram of our model in the  $(\beta, \nu)$  plane for some specific values of the coupling parameters, namely,  $\alpha = 0.2$  and  $\gamma = 0.1$ . Here, we study the impact of  $\alpha$  and  $\gamma$  on the critical transmission rates and on the critical noise.

The curves delimiting the three regions of the phase diagrams, shown in Fig. 8(a), shift toward lower values of  $\beta$  as  $\alpha$  increases. This indicates that both the forward and backward critical transmission rates,  $\beta_c^F$  and  $\beta_c^B$ , decrease when  $\alpha$  takes higher values. This can be explained by noting that, as  $\alpha$  increases, nodes in a  $-S$  state are more likely to transition to  $-R$ . In turn, because of the synergistic interplay between message spreading and opinion formation, compared to the  $-S$  state, nodes in  $-R$  have a higher probability of switching opinion via neighbor pressure. Thus, larger values of  $\alpha$  cause larger fractions of nodes with negative opinion to switch opinion, increasing the likelihood of the system to reach a positive-opinion consensus, and, as a result, both  $\beta_c^F$  and  $\beta_c^B$  decrease.

The main effect of  $\gamma$ , instead, is to shift the transition curves in the phase diagram toward higher values of  $\nu$ , as illustrated in Fig. 8(b). Thus, when  $\gamma$  increases, the region occupied by the bistable phase grows. This means that, for a fixed noise parameter  $\nu$ , the forward critical transmission rate  $\beta_c^F$  increases, while the backward one  $\beta_c^B$  decreases.

This is consistent with the mechanisms of our model. In fact, when the system is in the LP phase, only a small number of nodes with positive opinion satisfy the threshold condition with large probability, whereas the number of nodes with negative opinion that satisfy it is negligible. Thus, the main effect of an increase of  $\gamma$  is to increase the rate at which nodes in the  $+R$  state switch to  $-R$ , which, in turn, increases the probability that the system remains in the LP phase. Conversely, when the system is in the HP phase, the threshold condition is satisfied with high probability only by a small number of nodes with negative opinion, whereas those with positive opinion that satisfy it can be neglected. Then, in this case, the main effect of increasing  $\gamma$  is to increase the probability that nodes in the  $-S$  state switch to  $+S$ , increasing the likelihood of the system to remain in the HP phase.

Taken together, both cases also imply that the critical noise  $\nu_c$  increases with  $\gamma$ .

### APPENDIX D: PERMANENCE TIMES BETWEEN FLIPPINGS

To estimate the times  $T_{\text{up}}$  and  $T_{\text{down}}$  for which the system is expected to remain in the HP phase or in the LP one before flipping, we follow the approach in Ref. [38].

First, we note that the values of  $\nu^-$  at which the system switches to the LP phase are not much different from each other, as shown by the green dots in Fig. 5(c). Thus, we assume that, if the system is in the HP phase, it will flip to the LP one as long as the value of  $\nu_\lambda^-$  is greater than a typical value  $\nu_s^-$ . But then, for every node with positive opinion, the possibility of switching state in a unit time can be considered as an independent experiment with success probability  $\nu^-$ . Thus, we can write

$$\nu_\lambda^- = \frac{\lambda X}{n}, \quad (\text{D1})$$

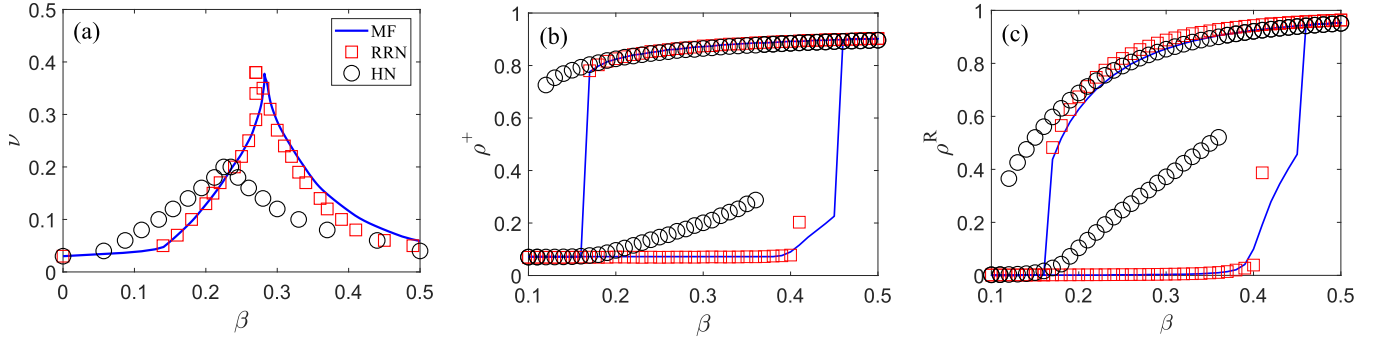


FIG. 9. Heterogeneity affects critical transmission rates and opinion distribution at steady state. The phase diagram of the model on the  $(\beta, v)$  plane [panel (a)], and the behavior of  $\rho^+$  [panel (b)] and  $\rho^R$  [panel (c)] as the message transmission rate  $\beta$  is changed show that the critical values of the forward and backward transmission rates decrease when the network is more heterogeneous. At the same time, a significant increase of  $\rho^+$  and  $\rho^R$  is observed as  $\beta$  is increased, together with a strong deviation from the mean-field results. In all panels, MF refers to the results of the mean-field theory, RRN to simulations using random regular networks, and HN to simulations using heterogeneous networks. The parameters used are  $\alpha = 0.2$ ,  $\gamma = 0.1$ ,  $\mu = 1.0$ ,  $r = 1.0$ , and  $m = 4$ . In heterogeneous networks, the threshold condition is set to be the fraction of neighboring nodes of the same opinion being less than or equal to 0.4.

where  $X$  is the number of successful trials in a unit time and  $n$  is the number of trials in a time interval of duration  $\lambda$ . Here,  $X$  follows a binomial distribution with success probability  $v^-$  and number of trials  $m = \frac{n}{\lambda}$ . Thus, the expectation value of  $X$  is  $mv^-$  and its variance is  $mv^-(1 - v^-)$ .

In turn, this means that also  $v_\lambda^-$  follows a binomial distribution. In particular, its expectation value is

$$E[v_\lambda^-] = \frac{\lambda}{n} E[X] = \frac{\lambda}{n} mv^- = \frac{\lambda}{n} \frac{n}{\lambda} v^- = v^- \quad (\text{D2})$$

and its variance is

$$\begin{aligned} \text{Var}[v_\lambda^-] &= \frac{\lambda^2}{n^2} \text{Var}[X] = \frac{\lambda^2}{n^2} mv^-(1 - v^-) \\ &= \frac{\lambda^2}{n^2} \frac{n}{\lambda} v^-(1 - v^-) = \frac{\lambda}{n} v^-(1 - v^-). \end{aligned} \quad (\text{D3})$$

However, in the limit of a large number of trials, we can write

$$n = N \langle \rho^+ \rangle_{\text{high}} \lambda, \quad (\text{D4})$$

and the binomial distribution approaches a normal one, so that

$$f(v_\lambda^-) \sim e^{-\frac{n(v_\lambda^- - v^-)^2}{2\lambda v^-(1 - v^-)}} \sim e^{-\frac{N(\rho^+)_{\text{high}}(v_\lambda^- - v^-)^2}{2v^-(1 - v^-)}}. \quad (\text{D5})$$

Thus, the probability that  $v_\lambda^-$  equals the typical transition value is

$$f(v_s^-) \sim e^{-\frac{N(\rho^+)_{\text{high}}(v_s^- - v^-)^2}{2v^-(1 - v^-)}}. \quad (\text{D6})$$

Finally, the expected time that the system spends in the HP phase is inversely proportional to this probability, as it will flip to the LP phase if the trajectory leads the value of  $v_\lambda^-$  to reach  $v_s^-$ . Thus, we conclude that

$$T_{\text{up}}(N) \approx e^{\frac{N(\rho^+)_{\text{high}}(v_s^- - v^-)^2}{2v^-(1 - v^-)}}, \quad (\text{D7})$$

which is Eq. (6).

The same steps can be followed to derive the expected time that the system spends in the LP phase, starting from the approximation of a critical value of  $v^+$  for this opposite

flipping. The result is

$$T_{\text{down}}(N) \approx e^{\frac{N(1 - (\rho^+)_{\text{low}})(v_s^+ - v^+)^2}{2v^+(1 - v^+)}} \quad (\text{D8})$$

which is Eq. (5).

## APPENDIX E: EFFECT OF NETWORK STRUCTURE

By eliminating the influence of node heterogeneity on the dynamics, the use of duplex networks with random regular layers allows one to more clearly attribute the phenomena observed on such networks, such as phase transitions, to the nonlinear interactions inherent to the model used, rather than to structural properties of the underlying network. Furthermore, the mean-field approximation, which is based on the assumption of homogeneous mixing, typically provides relatively accurate predictions of macroscopic behavior on random regular networks (RRN). This allows for a direct and clear comparison between the results of stochastic simulations and the analytical theory, thereby validating the basic logical self-consistency of the model.

Nonetheless, it is natural to wonder about the effect of a different network structure on the dynamics. Thus, to describe how our model is affected by structural heterogeneity, we simulated it on a two-layer Barabási-Albert (BA) network with mean degree 10 and power-law exponent 3. Our results, shown in Fig. 9, indicate that the increased network heterogeneity leads to a reduction in both the forward and backward critical message transmission rates,  $\beta_c^F$  and  $\beta_c^B$ . Furthermore, due to the vanishing epidemic threshold on BA networks in the thermodynamic limit, message propagation can persist even when the majority of nodes holds a negative opinion. Also, owing to the synergistic interaction between positive-opinion formation and message dissemination, we observe a significant increase in both  $\rho^+$  and  $\rho^R$  as the message transmission rate increases, before the occurrence of a first-order phase transition. Finally, degree heterogeneity leads to variations in how nodes of different degrees meet the threshold condition, in turn influencing the dynamical behavior of the coupled model.

- [1] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47 (2002).
- [2] M. E. J. Newman, Structure and function of complex networks, *SIAM Rev.* **45**, 167 (2003).
- [3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Complex networks: Structure and dynamics, *Phys. Rep.* **424**, 175 (2006).
- [4] S. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Critical phenomena in complex networks, *Rev. Mod. Phys.* **80**, 1275 (2008).
- [5] A. Pikovski, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, UK, 2001).
- [6] S. Boccaletti, J. Kurths, G. Osipov, D. L. Valladares, and C. S. Zhou, The synchronization of chaotic systems, *Phys. Rep.* **366**, 1 (2002).
- [7] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Synchronization in complex networks, *Phys. Rep.* **469**, 93 (2008).
- [8] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, *Rev. Mod. Phys.* **87**, 925 (2015).
- [9] S. Boccaletti, A. N. Pisarchik, C. I. del Genio, and A. Amann, *Synchronization: From Coupled Systems to Complex Networks* (Cambridge University Press, Cambridge, UK, 2018).
- [10] Z. H. Lin, M. Feng, M. Tang, Z. Liu, C. Xu, P. M. Hui, and Y.-C. Lai, Non-Markovian recovery makes complex networks more resilient against large-scale failures, *Nat. Commun.* **11**, 2490 (2020).
- [11] M. Li, R.-R. Liu, L. Lü, M.-B. Hu, S. Xu, and Y.-C. Zhang, Percolation on complex networks: Theory and application, *Phys. Rep.* **907**, 1 (2021).
- [12] X. Liu, D. Li, M. Ma, B. K. Szymanski, H. E. Stanley, and J. Gao, Network resilience, *Phys. Rep.* **971**, 1 (2022).
- [13] O. Artime, M. Grassia, M. De Domenico, J. P. Gleeson, H. Makse, G. Mangioni, M. Perc, and F. Radicchi, Robustness and resilience of complex networks, *Nat. Rev. Phys.* **6**, 114 (2024).
- [14] Y. Qi, S. Zhuang, X. Yu, Z. Zhao, W. Sun, Z. Li, J. Qiu, Y. Shi, F. Liu, C. I. del Genio, *et al.*, Multi-dimensional multi-option opinion dynamics leads to the emergence of clusters in social networks, *Chaos Solitons Fractals* **192**, 115983 (2025).
- [15] C. I. del Genio, J. Gómez-Gardeñes, I. Bonamassa, and S. Boccaletti, Synchronization in networks with multiple interaction layers, *Sci. Adv.* **2**, e1601679 (2016).
- [16] C. I. del Genio, S. Faci-Lazaro, J. Gómez-Gardeñes, and S. Boccaletti, Mean-field nature of synchronization stability in networks with multiple interaction layers, *Commun. Phys.* **5**, 121 (2022).
- [17] C. Castellano, S. Fortunato, and V. Loreto, Statistical physics of social dynamics, *Rev. Mod. Phys.* **81**, 591 (2009).
- [18] A. Kirman, Ants, rationality, and recruitment, *Q. J. Econ.* **108**, 137 (1993).
- [19] B. L. Granovsky and N. Madras, The noisy voter model, *Stoch. Proc. Appl.* **55**, 23 (1995).
- [20] A. Carro, R. Toral, and M. San Miguel, The noisy voter model on complex networks, *Sci. Rep.* **6**, 24775 (2016).
- [21] M. Pineda, R. Toral, and E. Hernandez-García, Noisy continuous-opinion dynamics, *J. Stat. Mech.* (2009) P08001.
- [22] R. A. Holley and T. M. Liggett, Ergodic theorems for weakly interacting infinite systems and the voter model, *Ann. Probab.* **3**, 643 (1975).
- [23] R. M. Bond, C. J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle, and J. Fowler, A 61-million-person experiment in social influence and political mobilization, *Nature (London)* **489**, 295 (2012).
- [24] W. Lee, S.-G. Yang, and B. J. Kim, The effect of media on opinion formation, *Physica A* **595**, 127075 (2022).
- [25] Z.-K. Zhang, C. Liu, X.-X. Zhan, X. Lu, C.-X. Zhang, and Y.-C. Zhang, Dynamics of information diffusion and its applications on complex networks, *Phys. Rep.* **651**, 1 (2016).
- [26] S. Vosoughi, D. Roy, and S. Aral, The spread of true and false news online, *Science* **359**, 1146 (2018).
- [27] M. De Domenico, C. Granell, M. Porter, and A. Arenas, The physics of spreading processes in multilayer networks, *Nat. Phys.* **12**, 901 (2016).
- [28] W. Wang, Q. H. Liu, J. Liang, Y. Hu, and T. Zhou, Coevolution spreading in complex networks, *Phys. Rep.* **820**, 1 (2019).
- [29] M. De Domenico, More is different in real-world multilayer networks, *Nat. Phys.* **19**, 1247 (2023).
- [30] F. 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).
- [31] J. Liu, S. Huang, N. M. Aden, N. F. Johnson, and C. Song, Emergence of polarization in coevolving networks, *Phys. Rev. Lett.* **130**, 037401 (2023).
- [32] H. Wang, C. Chen, B. Qu, D. Li, and S. Havlin, Epidemic mitigation via awareness propagation in communication networks: The role of time scales, *New J. Phys.* **19**, 073039 (2017).
- [33] P. C. V. da Silva, F. Velásquez-Rojas, C. Connaughton, F. Vazquez, Y. Moreno, and F. A. Rodrigues, Epidemic spreading with awareness and different timescales in multiplex networks, *Phys. Rev. E* **100**, 032313 (2019).
- [34] P. C. Ventura, Y. Moreno, and F. A. Rodrigues, Role of time scale in the spreading of asymmetrically interacting diseases, *Phys. Rev. Res.* **3**, 013146 (2021).
- [35] C.-R. Cai, N.-N. Liu, X. Chang, and X.-S. Liu, Physical images of relative timescales in coevolution dynamics, *Phys. Rev. Res.* **5**, 033220 (2023).
- [36] S. Boccaletti, G. Bianconi, R. Criado, C. I. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, The structure and dynamics of multilayer networks, *Phys. Rep.* **544**, 1 (2014).
- [37] M. Granovetter and R. Soong, Threshold models of diffusion and collective behavior, *J. Math. Sociol.* **9**, 165 (1983).
- [38] A. Majdandzic, B. Podobnik, S. V. Buldyrev, D. Y. Kenett, S. Havlin, and H. E. Stanley, Spontaneous recovery in dynamical networks, *Nat. Phys.* **10**, 34 (2014).
- [39] L. D. Valdez, M. A. Di Muro, and L. A. Braunstein, Failure-recovery model with competition between failures in complex networks: A dynamical approach, *J. Stat. Mech.* (2016) 093402.
- [40] L. Böttcher, M. Luković, J. Nagler, S. Havlin, and H. J. Herrmann, Failure and recovery in dynamical networks, *Sci. Rep.* **7**, 41729 (2017).

- [41] L. Böttcher, J. Nagler, and H. J. Herrmann, Critical behaviors in contagion dynamics, *Phys. Rev. Lett.* **118**, 088301 (2017).
- [42] D. Eckles, E. Mossel, M. A. Rahimian, and S. Sen, Long ties accelerate noisy threshold-based contagions, *Nat. Hum. Behav.* **8**, 1057 (2024).
- [43] A. Wan, C. Riedl, and D. Lazer, Diffusion of complex contagions is shaped by a trade-off between reach and reinforcement, *Proc. Natl. Acad. Sci. USA* **122**, e2422892122 (2025).
- [44] M. M. Danziger, I. Bonamassa, S. Boccaletti, and S. Havlin, Dynamic interdependence and competition in multilayer networks, *Nat. Phys.* **15**, 178 (2019).
- [45] Y. Liu, Phase-flipping-data, Version v1.0, (2025), <https://github.com/Lywww13/Phase-flipping-data>.