Heterophobic interactions hinder consensus formation in sparse random networks

Alejandro Castro [1], Tuan Minh Pham [2,3], Ernesto Ortega [4], and David Machado [1,5,6,*]

[1] *Group of Complex Systems and Statistical Physics, Department of Theoretical Physics, University of Havana, Havana, Cuba*

[2] *Dutch Institute for Emergent Phenomena, Amsterdam 1090 GL, The Netherlands*

[3] *Institute of Physics, University of Amsterdam, Science Park 904, The Netherlands*

[4] *Complexity Science Hub Vienna, Metternichgasse 8, 1030 Vienna, Austria*

[5] *Dipartimento di Fisica, Sapienza Università di Roma, P.le Aldo Moro 5, 00185 Rome, Italy*

[6] *CNR-Nanotec, unità di Roma, 00185 Italy*

(Received 2 July 2025; accepted 24 November 2025; published 15 December 2025)

Heterophobic interactions, which drive individuals to be repelled from others with opposite opinions, play a role as important as homophilic ones in shaping many dynamical processes on social networks, such as opinion formation, social balance, or epidemic spreading. In this paper, we use belief propagation and Monte Carlo simulations on treelike signed graphs to predict that a sufficient propensity to heterophobia can impede a consensus that would otherwise emerge via a phase transition. As the strength of heterophobic interactions and the rationality of individuals with respect to social stress decrease, this transition changes from continuous to discontinuous, with a strong dependence on the initial conditions. The size of the parameter region where consensus can be reached from any initial condition decays as a power-law function of the number of discussed topics.

DOI: 10.1103/cxsr-dh25

I. INTRODUCTION

Opinion dynamics underlie a variety of emergent social phenomena, including consensus [1–3], polarization [4,5], fragmentation [6,7], emergence of leaders [8,9], or norm formations [10]. Extensive work has studied how network structure determines opinion outcomes [11–16]. However, social interaction between any two agents often depends on their opinion similarity [17–23]. Positively linked agents tend to align due to homophily, while negatively linked agents tend to disagree due to heterophobia—a repulsion from those who are different [24]. This type of interaction, also known as xenophobia in the literature [24,25], can drive society toward polarization [25]. The question of how consensus is affected by the joint action of homophilic and heterophobic ties, nevertheless, remains largely unanswered.

To address this question, we study a minimal spin model [22] that has been validated empirically [26]. In this model, agents interact via both homophilic and heterophobic couplings on random sparse graphs. The focus of Ref. [22] is on the emergence of triadic motifs, but not on consensus formation. To study the latter, we apply belief propagation (BP) to treelike graphs (i.e., those without short cycles), such as Erdős–Rényi or random regular [27–32]. Although BP is exact in this type of graphs, it can also be extended to networks with short loops [33–40]. Identifying the conditions for consensus to emerge via a disorder-to-order transition, we find that consensus is harder to reach when the number of discussed issues (denoted by g) increases. In particular, we show how the feasibility of consensus decreases as a power law of the number of discussed issues. Since this law is scale-free, there is no characteristic value of g that guarantees consensus, instead of having a maximum value of g as previously suggested in Refs. [1,41].

The rest of the article is organized as follows. In Sec. II, we presented the model as defined in Ref. [22], while in Sec. III we derive the corresponding BP equations. In Sec. IV, we show the phase diagrams predicted by BP for Erdős–Rényi and random regular graphs, and compare with the output of Monte Carlo simulations for large systems. Finally, the results are discussed in Sec. V. We include more information on other relevant analytical and technical details in the Appendixes.

II. MODEL

We consider a system of N individuals connected via a network $\mathcal{G}(V, E)$, where V denotes the set of vertices and E the set of edges. Each node $i \in V$ has k_i neighbors and opinions on g different "issues," forming a feature vector $\mathbf{s}_i$ of dimension g with $s_i^\mu \in \{-1, 1\}$, for $\mu = 1, 2, \cdots, g$. This captures, for instance, an election process in which a candidate is to be selected based on multiple topics, such as abortion, public versus private health care, climate change, gun control, etc., like the American National Election Survey data considered in Ref. [42].

Although real-world agents typically experience bounded rationality and finite attention spans, we assume that individuals distribute their cognitive attention evenly across all

*Contact author: david.machado@uniroma1.it

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

topics. This simplification allows us to isolate the purely structural effects of opinion dimensionality [22]. Note that, instead of treating heterophobia as quenched disorder that captures agents' intrinsic characteristic [43], here we focus on *annealed* disorder. Specifically, we allow the interactions to co-evolve dynamically with opinions: agents adjust their couplings $J_{i,j} = \text{sign}(\mathbf{s}_i \cdot \mathbf{s}_j)$ based on the opinion similarity, while $J_{i,j}$, in turn, affects the update of opinions through social pressure specified by a Hamiltonian. The model of Ref. [22] introduced the following *local* Hamiltonian with every node i having both homophilic and heterophobic interactions:

$$H^{(i)}(\mathbf{s}_i, \mathbf{s}_{\partial i}) = -\frac{\alpha}{g} \sum_{\substack{j \in \partial i: \\ J_{i,j}=1}} \mathbf{s}_i \cdot \mathbf{s}_j + \frac{1-\alpha}{g} \sum_{\substack{j \in \partial i: \\ J_{i,j}=-1}} \mathbf{s}_i \cdot \mathbf{s}_j, \tag{1}$$

where ∂i denotes the set of neighbors of i. The first sum in the Hamiltonian goes over all the nodes that have a positive coupling $J_{i,j} = 1$ with i, while the nodes in the second sum are those with a negative coupling $J_{i,j} = -1$. The parameter α, termed affinity, controls the propensity to focus on homophilic interactions, and $(1-\alpha)$ that to focus on heterophobic ones. As α decreases, the effect of heterophobic interaction becomes stronger. This local Hamiltonian (social stress) accounts for the influence that an agent i receives from interactions with its friends through the first term and from those with its enemies through the second term. Monte Carlo simulations of this model consist of a large number of steps, at each of which we select one agent and its attribute, both at random. Then, we propose a new attribute that is accepted or rejected using the Metropolis scheme [44] at a given temperature T. In a social context, T characterizes the average volatility of individuals: the higher T, the more likely an agent is to update its opinion, regardless of whether that update reduces social stress [45]. Since individuals are assumed to update opinions and social links on the same timescale, once the opinion has been updated, we recompute their couplings $J_{i,j} = \text{sign}(\mathbf{s}_i \cdot \mathbf{s}_j)$ before continuing the next step of the opinion dynamics.

III. BELIEF PROPAGATION EQUATIONS

The *marginal* probability $P(\mathbf{s}_i)$ of i's opinion vector is given by

$$P(\mathbf{s}_i) \propto \sum_{\mathbf{s}_{\partial i}} e^{\beta H^{(i)}(\mathbf{s}_i, \mathbf{s}_{\partial i})} \prod_{j \in \partial i} m_{j \to i}(\mathbf{s}_j), \tag{2}$$

where $\beta = T^{-1}$ represents the rationality of individuals with respect to social stress. The factors in the product are the marginal probabilities (called messages) of the opinion $\mathbf{s}_j$, where $j \in \partial i$, in a modified graph where the edge (i, j) is removed. Since the graph is treelike, in a scenario with finite correlation length, the removal of (i, j) effectively disconnects the nodes i and j and makes $m_{j \to i}(\mathbf{s}_j)$ independent of $\mathbf{s}_i$. As derived in Appendix A, these messages can be computed recursively using the following rule:

$$m_{j \to i}(\mathbf{s}_j) \propto \prod_{k \in \partial j \setminus i} \sum_{\mathbf{s}_k} m_{k \to j}(\mathbf{s}_k)\, e^{\frac{\beta}{2g}(\mathbf{s}_k \cdot \mathbf{s}_j)[2\alpha - 1 + \text{sgn}(\mathbf{s}_k \cdot \mathbf{s}_j)]}. \tag{3}$$

This set of equations can be solved iteratively to find the messages (if a stable fixed point of this iterative map exists). Once all these final BP's messages are obtained, we can compute $P(\mathbf{s}_i)$ using Eq. (2). To measure the level of agreement, we define the magnetization as follows:

$$m = \left\| \frac{1}{Ng} \sum_i \sum_{\mathbf{s}_i} P(\mathbf{s}_i)\, \mathbf{s}_i \right\|_2, \tag{4}$$

where $||\cdot||_2$ is the Euclidean norm. Note that $m = 0$ if the opinions $\mathbf{s}_i$ have no preferred direction, while $m > 0$ if the opinions are at least partially aligned along a given direction. When all people share the same opinions, a global consensus is achieved, so $m = 1$. At each temperature $T = \beta^{-1}$, there exists a critical value $\alpha_c(T)$ such that $m = 0$ (paramagnetic state) for $\alpha < \alpha_c(T)$, and $m > 0$ (ferromagnetic state) for $\alpha > \alpha_c(T)$. Therefore, $\alpha_c(T)$ marks a disorder-to-order transition. To find $\alpha_c(T)$ (with a precision equal to $d\alpha$), we start at some $\alpha_0 < \alpha_c(T)$ and measure $m(\alpha_0)$. As long as $m(\alpha_0) = 0$, we set $\alpha_0 - > \alpha_0 + d\alpha$ and repeat the process until $m(\alpha_0) > 0$.

Since BP does not converge for $\alpha < 0.5$, not even for a single run in a specific graph, we focus only on $\alpha \geqslant 0.5$. This lack of convergence usually indicates the existence of a frustrated phase, due to the high nonconvexity of the underlying solution space with multiple local minima [46]. Intuitively, at $\alpha < 0.5$, the system is more likely to be driven by the agents' tendency to separate their opinions from those of their neighbors. At low temperature and in the presence of long loops in the graph, individuals face inconsistencies in separating themselves from the rest, resulting in a frustrated phase with a rich phenomenology, such as the coexistence of opposite opinion clusters [47]. We left this scenario for future work.

IV. RESULTS

Figure 1 shows the dependence of consensus on the level of affinity in random regular graphs. As α increases beyond $\alpha_c(T)$, the system goes from disagreement to consensus. The higher the temperature T, the larger $\alpha_c(T)$. In a random regular graph, given that all links are equivalent, we can reduce Eq. (3) to a single equation for one message by initializing all messages $m_{j \to i}(\mathbf{s}_j)$ with the same (random or aligned) value. This further facilitates studying the dependence of the BP solution on the initial conditions. In essence, choosing an initial condition is assigning a message (a probability) to each of the 2^g different opinion vectors, for each edge (i, j) and in both directions $i \to j$ and $j \to i$. To simplify the process, we initially take all $m_{j \to i}(\mathbf{s}_j)$ equal to the same distribution $m_{\to}(\mathbf{s})$, regardless of i and j. We then set $m_{\to}(\mathbf{s})$ to depend only on the number $n = \sum_{\mu=1}^{g} \delta_{s^\mu,1}$ of positive components in the opinion vector. Therefore, to each $\mathbf{s}$ we assign the message:

$$m_{\to}(\mathbf{s}) \equiv m_{\to}(n) = p^n\, (1-p)^{g-n}, \tag{5}$$

where $p \in [0.51, 1]$ is a parameter that can be continuously varied from completely random configurations ($p = 0.5$) to a fully aligned one ($p = 1$).

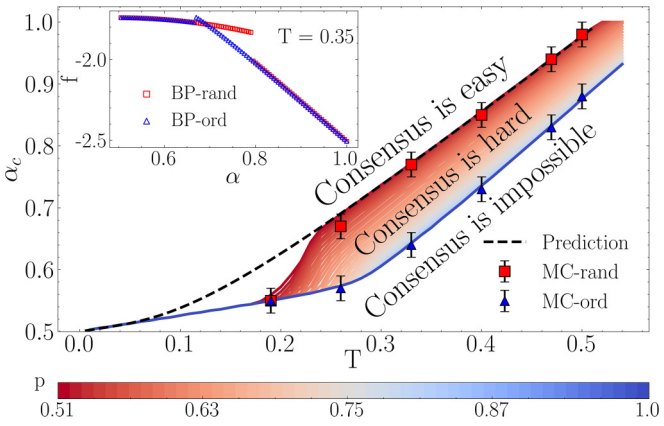


FIG. 1. Phase diagram of the model in random regular graphs with $g = c = 5$. The figure includes BP's predictions (solid lines), the results of Monte Carlo simulations (squares and triangles), and our analytic computation of the stability of the paramagnetic solution (dashed lines). Different colors represent BP's predictions for different initial conditions obtained using Eq. (5) to assign the initial value of the messages. Monte Carlo simulations are performed over systems with $N = 10^3$ and $N = 10^4$ individuals. When $p = 0.5$, the initial state has zero magnetization (no alignment of the opinions). When $p = 1.0$, the initial state is completely aligned ($m = 1$). BP's transition lines have an uncertainty of order $d\alpha = 0.001$, not shown since it is of the size of the lines themselves. The inserted graphic shows the two branches of BP's free-energy density for different α at $T = 0.35$. The blue triangles are obtained with completely aligned initial conditions and the red squares with random initial conditions.

In Fig. 1 we use colored lines to depict the transitions obtained by BP for different initial conditions, each given by a particular value of p. The disorder-to-order transition obtained with $p = 0.51$ is marked by the darkest tone of red, while that with $p = 1$ is the darkest blue. In the intermediate region, the system has a peculiar behavior. There, although consensus exists as an attractor for the dynamics, it would take a time that grows with the number N of agents for a system starting from a random configuration to reach consensus (see Appendix D). The larger the number of topics discussed, the more pronounced this dependence on initial conditions becomes.

The phase diagram in Fig. 1 is divided into three distinct regions. Above the darkest red line, it is easy to reach a consensus from any initial condition. Below the darkest blue line, it is impossible to achieve consensus, as it is not an attractor of the dynamics. There is a region in between, where it is hard but possible to attain consensus due to a strong dependence on the initial conditions.

The lines generated using $p \in (0.51, 1)$ continuously cover the area between these two extremes. Nevertheless, we observe that a small perturbation ($p \approx 0.63$) already shifts the transition far away from the results obtained with $p \sim 0.5$. However, the results for $p > 0.75$ remain very close to what is obtained with $p = 1$. This shows how, at a fixed temperature, a small perturbation in the right direction can facilitate consensus. In Fig. 1 we also verify that, regardless of the value of p, there exists a critical point $(T_c, \alpha_c(T_c))$ such that, for $T > T_c$, there are two different transitions, depending on whether we start from random or aligned initial conditions. Exactly at $(T_c, \alpha_c(T_c))$, all the colored lines merge into a single one, which remains unique for all $T < T_c$.

To better understand the phase transition, we show the free-energy density of the system, obtained using BP, in the inset of Fig. 1. At $T > T_c$, the free energy exhibits two distinct branches, confirming a first-order transition, typically accompanied by a strong dependence on initial conditions. The *true* free energy is always the minimum among the two branches and has a discontinuity in its derivative at the crossing point of the branches, where the actual thermodynamic transition takes place. For α beyond this point, the paramagnetic state is still stable but has a higher free energy (is metastable). As usual in first-order transitions, jumping from this state to the ordered regime requires crossing extensive free-energy barriers. According to Arrhenius law [46,48], this takes a time that grows exponentially with the system size, making consensus considerably hard to achieve when starting with random initial conditions.

However, at $T < T_c$ the free energy has only one branch and does not features any discontinuities in its derivative (see Appendix C). In other words, the nature of the phase transition changes from discontinuous at high T to continuous at low T, with the point $(T_c, \alpha_c(T_c))$ separating these two regimes.

We next compare BP's predictions with Monte Carlo simulations for two system sizes ($N = 10^3$ and $N = 10^4$) and two different initial conditions: (i) globally aligned opinions (blue points) or (ii) iid random opinions at each node (red points). In Appendix D, we demonstrate a strong dependence of the results on the initial conditions. A strong dependence on the system size is also observed for some specific values of T and α. To determine the location of the phase transition for each temperature and each type of initial condition, we ran simulations for different values of α and observed the temporal evolution of the magnetization (see Figs. 6 and 7 in Appendix D). If the results for $N = 10^3$ and $N = 10^4$ were different, then we labeled that value of α as *uncertain*. By comparing the results for these two system-sizes, we reduce chance of misclassifications due to finite-size effects. The reader can access the data and graphics in the repository of Ref. [49].

For the rest of the values of α, we looked at the final magnetization obtained with both system sizes. All values of α for which the dynamics arrives at $m \sim 0$ for large times were classified as inside the paramagnetic phase for that specific initial condition. Analogously, if the final magnetization is different from zero, the value of α is considered to be inside the ferromagnetic phase. The *uncertain* values lie inside the error bars. The critical values of α_c obtained by BP with $p = 1.0$ and $p = 0.5$, respectively, are in excellent agreement with simulations.

We complete our analysis of the phase transition by *approximately* computing the transition from disorder to order. Our prediction is represented by a dashed black line in Fig. 1 and is obtained from the stability analysis of BP's disordered solution. Specifically, we apply a small perturbation to the paramagnetic fixed point and expand Eq. (3) to the first order in powers of the perturbation. We then study whether the perturbed messages will return to the paramagnetic fixed-point solution, or, instead, will depart from it. In the latter case, the fixed point is linearly unstable, as detailed in Appendix B.

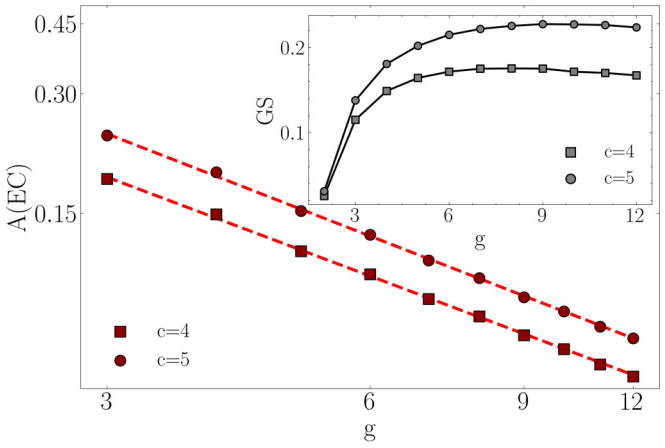


FIG. 2. Area of easy consensus A(EC) vs the length g of the opinion vectors on a logarithmic scale in random regular graphs with fixed connectivity $c = 4$ (square) and $c = 5$ (circle). Red dashed lines are power-law fits to the data points: $\mathrm{A(EC)} = \Phi(c)\, g^{-\gamma(c)}$. The fitted parameters depend on the connectivity. They are: $\Phi(4) = 0.459(8)$, $\Phi(5) = 0.61(1)$ and $\gamma(4) = 0.83(1)$, $\gamma(5) = 0.85(1)$. The inserted graphic shows the gap's size GS defined in Eq. (6) as a function of g.

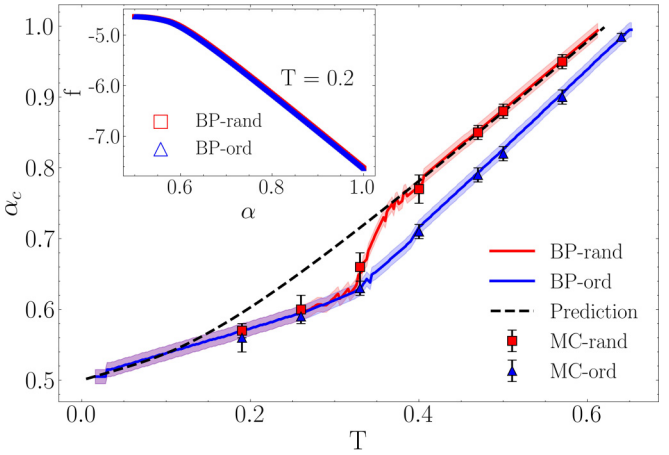


FIG. 3. Phase diagram of the model in Erdős-Rényi graphs. The figure includes BP's predictions (solid lines), the results of Monte Carlo simulations (squares and triangles), and our approximation of the stability of the paramagnetic solution (dashed lines). Monte Carlo simulations were carried out for systems with $N = 10^4$, $N = 5 \times 10^5$, and $N = 10^5$ individuals, $g = 3$, and average connectivity $\kappa = 3$. The red points and lines are obtained by initializing all the opinions uniformly at random. The blue points and lines are instead the result of initializing the opinions aligned toward a given direction. The colored shadows represent the uncertainty in the determination of BP's transition lines, which is of order $d\alpha = 0.01$ in this case. The inserted graphic shows BP's free-energy density as a function of α at $T = 0.2$. The blue triangles are obtained with completely aligned initial conditions, and the red squares with random initial conditions.

Expanding to the first order in the perturbation is just an upper bound, since subsequent powers could destroy the stability of this fixed point. It is not surprising that this works well only with first-order phase transitions, for large α and T. Our method can be directly applied to other graph ensembles like Erdős-Rényi, as we will show later. Since we derive an analytical expression for the stability line, its numerical evaluation is considerably faster than solving the BP equations or performing Monte Carlo simulations.

At each temperature, we observe a finite gap $\alpha_c^{\mathrm{rand}}(T) - \alpha_c^{\mathrm{ord}}(T) > 0$ due to the existence of two different transitions, one for random ($p = 0.51$) and another for aligned ($p = 1$) initial conditions. We can quantify the difficulty of reaching a consensus, even when it is possible, by introducing two definitions. Let the area of easy consensus [A(EC)] be the area above the darkest red line, where consensus is easy, and the area of possible consensus [A(PC)] be the entire area above the darkest blue line, where consensus is possible. Figure 2 shows that a large number of discussed issues hinders the achievement of consensus. In particular, the area A(EC) tends to zero when g increases. For each connectivity c, the decay of this quantity follows a power-law function: $\mathrm{A(EC)} = \Phi(c) g^{-\gamma(c)}$, whose parameters depend on c. For both $c = 4$ and $c = 5$, the exponent γ is smaller than one, signaling the scale-free nature of the decay, i.e., the absence of any characteristic lengths. Although we have no rigorous argument for this observation, we can intuitively understand it from the following argument. For a large system size N, the density of configurations concentrates in the region $m \sim 0$, and this effect is stronger when g is increased (see Appendix G). The dynamics needs to visit a larger number of configurations to abandon the region $m \sim 0$ when g grows, and the paramagnetic state becomes more stable.

The difference A(PC) − A(EC) gives the area where consensus is possible, but hard to reach if starting from random initial conditions. Normalizing this difference by A(PC), we have the following measure of the gap's size:

$$\mathrm{GS} = \frac{\mathrm{A(PC)} - \mathrm{A(EC)}}{\mathrm{A(PC)}}. \tag{6}$$

The behavior of this GS is shown in the inserted graphic in Fig. 2. This region grows with the number of topics until it stabilizes above $g \sim 8$. For completeness, we show that A(PC) does not follow a power-law behavior in Appendix E.

Finally, in Fig. 3 we confirm the validity of our results, including the approximate computation for the stability of the paramagnetic solution, in Erdős-Rényi random graphs. The transition lines were obtained following the same methodology described above. The phenomenology is essentially the same as in random regular graphs. The inserted graphic shows that for low $T < T_c$ the free energy is continuously differentiable. To complete this analogy, we include a figure for the free energy at high $T > T_c$ in Appendix F.

There are some numerical challenges to obtaining accurate results for Erdős-Rényi graphs. They arise from the fact that the computational cost of running BP increases exponentially with the number of neighbors and with the number of discussed issues g. As the connectivity is not uniform, we used a standard population dynamics algorithm [50,51] to sample the distribution of BP's messages. This method allows making predictions for the infinite size limit while keeping a modest computational cost. Furthermore, to reduce a strong finite-size effects of Monte Carlo simulations in this case, we used three large system sizes ($N = 10^4$, $N = 5 \times 10^4$, and $N = 10^5$). The reader can access the data and graphics in the repository

of Ref. [49]. Nevertheless, we find good agreement between BP's predictions and Monte Carlo simulations.

V. DISCUSSION

In summary, we demonstrate the impact of heterophobic interactions on opinion formation in sparse network topologies. The results deliver a general message: the presence of such interactions significantly complicate the achievement of consensus, even when this state exists as a fixed point of the dynamics. The phase diagrams for random regular and Erdős-Rényi graphs (Figs. 1 and 3, respectively), both contain regions where an uninformed initialization converges to consensus only at very long times. In this case, typical of first-order phase transitions, the time to reach consensus grows exponentially with the system size.

As our results confirm (see Fig. 2), the situation gets worse with increasing the number of issues discussed g: the area where consensus can be reached from any initial condition decays to zero as a power-law function of g. Moreover, we find that consensus can emerge via both discontinuous and continuous phase transitions. Such a mixed-order phase transition has recently been observed in other opinion dynamics models [21,52,53]. Here, we confirm that this picture remains qualitatively unaltered when moving from random regular to Erdős-Rényi graphs. Its validity for other topologies must be evaluated in future investigations.

We can also use BP to obtain other observables, such as the total level of social stress in the whole society and the fraction of positive links. Preliminary results for these observables are presented in Appendix C, highlighting the need for further investigation to understand their nontrivial behaviors.

The treelike networks considered in the present paper do not have many triangles. It would be interesting to study the effect of heterophobia in networks having short loops using generalized BP [54]. Moreover, the definition of the interactions could be modified to add more realistic details. For example, we could include heterogeneous interactions and/or consider the effects of adding bias to individual opinions. One could also consider diverse timescales between opinion-forming and relation-forming processes. It is also possible to study other related models, such as the Axelrod model [55] or the integration of structural balance and status hierarchies [21].

ACKNOWLEDGMENTS

We thank Alejandro Lage Castellanos for his useful comments. T.M.P. was supported by the Dutch Institute for Emergent Phenomena cluster at the University of Amsterdam under the Research Priority Area Emergent Phenomena in Society: Polarisation, Segregation, and Inequality. This study was conducted using the DARIAH HPC cluster at CNR-NANOTEC in Lecce, funded by the "MUR PON Ricerca e Innovazione 2014-2020" project, Code No. PIR01_00022.

DATA AVAILABILITY

The data that support the findings of this article are openly available on github and Zenodo [49].

APPENDIX A: DERIVATION OF BELIEF PROPAGATION EQUATIONS

To write the belief propagation equations that characterize the system in a given graph $\mathcal{G}(V, E)$, we start from the full Hamiltonian,

$$H = -\frac{\alpha}{g} \sum_{\substack{(i,j)\in E:\\ J_{i,j}=1}} \mathbf{s}_i \cdot \mathbf{s}_j + \frac{1-\alpha}{g} \sum_{\substack{(i,j)\in E:\\ J_{i,j}=-1}} \mathbf{s}_i \cdot \mathbf{s}_j, \tag{A1}$$

where J_{ij}, as said in the main text, is the sign of the product $\mathbf{s}_i \cdot \mathbf{s}_j$. The first sum covers all edges of the graph that have a positive coupling $J_{i,j} = 1$, while the second sum goes over all edges with a negative coupling $J_{i,j} = 1$. Although this way of writing the Hamiltonian makes the physical interpretation easier, it will be useful here to recast everything as a more traditional sum over all the interacting pairs in the system:

$$\begin{aligned} H = &-\frac{\alpha}{2g} \sum_{(i,j)\in E} (1 + J_{ij})(\mathbf{s}_i \cdot \mathbf{s}_j) \\ &+ \frac{1-\alpha}{2g} \sum_{(i,j)\in E} (1 - J_{ij})(\mathbf{s}_i \cdot \mathbf{s}_j). \end{aligned} \tag{A2}$$

Both forms are equivalent, and we can use the explicit form of J_{ij} and reorganize the terms to get

$$H = \frac{1-2\alpha}{2g} \sum_{(i,j)\in E} (\mathbf{s}_i \cdot \mathbf{s}_j) - \frac{1}{2g} \sum_{(i,j)\in E} |\mathbf{s}_i \cdot \mathbf{s}_j|, \tag{A3}$$

where $|\mathbf{s}_i \cdot \mathbf{s}_j| = J_{ij}(\mathbf{s}_i \cdot \mathbf{s}_j) = \mathrm{sgn}(\mathbf{s}_i \cdot \mathbf{s}_j)\,(\mathbf{s}_i \cdot \mathbf{s}_j)$ is the absolute value of the dot product.

Thus, the partition function Z is

$$\begin{aligned} Z &= \sum_{\underline{\mathbf{s}}} e^{-\beta H(\underline{\mathbf{s}})} \\ &= \sum_{\underline{\mathbf{s}}} \prod_{(i,j)\in E} \exp\left\{\frac{\beta}{2g}[(2\alpha-1)(\mathbf{s}_i \cdot \mathbf{s}_j) + |\mathbf{s}_i \cdot \mathbf{s}_j|]\right\}, \\ Z &= \sum_{\underline{\mathbf{s}}} \prod_{(i,j)\in E} \exp\left\{\frac{\beta}{2g}(\mathbf{s}_i \cdot \mathbf{s}_j)[(2\alpha-1) + \mathrm{sgn}(\mathbf{s}_i \cdot \mathbf{s}_j)]\right\}. \end{aligned} \tag{A4}$$

The Boltzmann-Gibbs distribution of the different configurations is

$$\begin{aligned} P(\underline{\mathbf{s}}) &= \frac{1}{Z} \prod_{(i,j)\in E} \exp\left\{\frac{\beta}{2g}(\mathbf{s}_i \cdot \mathbf{s}_j)[(2\alpha-1) + \mathrm{sgn}(\mathbf{s}_i \cdot \mathbf{s}_j)]\right\} \\ &= \frac{1}{Z} \prod_{(i,j)\in E} \varphi_{ij}(\mathbf{s}_i, \mathbf{s}_j), \end{aligned} \tag{A5}$$

Once we have identified the shape of the pairwise factor $\varphi_{ij}(\mathbf{s}_i, \mathbf{s}_j)$, it is straightforward to obtain the BP equations by following the standard procedure [28,56–58]:

$$\begin{aligned} m_{j\to i}(\mathbf{s}_j) \propto &\prod_{k\in\partial j\setminus i} \sum_{\mathbf{s}_k} m_{k\to j}(\mathbf{s}_k) \\ &\times \exp\left\{\frac{\beta}{2g}(\mathbf{s}_k \cdot \mathbf{s}_j)[2\alpha - 1 + \mathrm{sgn}(\mathbf{s}_k \cdot \mathbf{s}_j)]\right\}. \end{aligned} \tag{A6}$$

APPENDIX B: ANALYTIC PREDICTION FOR SOME CRITICAL LINES

When the messages are initially set as random, we expect BP to stay with zero magnetization unless that disordered configuration is not a fixed point of the algorithm. To locate this transition, we study the stability of the disordered solution under small perturbations. This can be done in all the ensembles where the graphs are locally treelike.

Let us summarize the results and give the details of the computation later. The disordered solution stops being stable at the line defined as follows:

$$\frac{Q(g,\beta,\alpha)}{R(g,\beta,\alpha)} = \frac{g}{2}\frac{1+\langle\gamma\rangle}{\langle\gamma\rangle}, \tag{B1}$$

where $\gamma = c-1$ is the reduced connectivity of a node, which is distributed according to $\mathbb{P}(\gamma)$, i.e., $\langle\gamma\rangle = \sum_{\gamma=0}^{\infty}\gamma\,\mathbb{P}(\gamma)$. The shape of the functions Q and R depends on the parity of g. For even $g = 2m$ we have

$$R(2m,\beta,\alpha) = e^{-\beta(\alpha-1)}A_{2m}[e^{\beta(\alpha-1)/m}] + e^{\beta\alpha}A_{2m}[e^{-\beta\alpha/m}] + \binom{2m}{m}, \tag{B2}$$

$$Q(2m,\beta,\alpha) = e^{-\beta(\alpha-1)}e^{\beta(\alpha-1)/m}A'_{2m}[e^{\beta(\alpha-1)/m}] - e^{\beta\alpha}e^{-\beta\alpha/m}A'_{2m}[e^{-\beta\alpha/m}] + 2me^{\beta\alpha}A_{2m}[e^{-\beta\alpha/m}] + m\binom{2m}{m}, \tag{B3}$$

where, in terms of the ordinary hypergeometric function ${}_2F_1$, we have

$$A_{2m}(z) = (1+z)^{2m} - \frac{m+1}{m}\binom{2m}{m-1}z^m(1+z)^{2m}\,{}_2F_1(2m+1,m;m+1;-z), \tag{B4}$$

$$A'_{2m}(z) = \frac{2m}{1+z}A_{2m}(z) - (m+1)\binom{2m}{m-1}\frac{z^{m-1}}{1+z}. \tag{B5}$$

However, for odd $g = 2m+1$ we obtain

$$R(2m+1,\beta,\alpha) = e^{-\beta(\alpha-1)}A_{2m+1}[e^{2\beta(\alpha-1)/(2m+1)}] + e^{\beta\alpha}A_{2m+1}[e^{-2\beta\alpha/(2m+1)}], \tag{B6}$$

$$\begin{aligned} Q(2m+1,\beta,\alpha) = {} & e^{-\beta(\alpha-1)}e^{2\beta(\alpha-1)/(2m+1)}A'_{2m+1}[e^{2\beta(\alpha-1)/(2m+1)}] \\ & -e^{\beta\alpha}e^{-2\beta\alpha/(2m+1)}A'_{2m}[e^{-2\beta\alpha/(2m+1)}] + (2m+1)e^{\beta\alpha}A_{2m+1}[e^{-2\beta\alpha/(2m+1)}], \end{aligned} \tag{B7}$$

where

$$A_{2m+1}(z) = (1+z)^{2m+1} - \binom{2m+1}{m}z^{m+1}(1+z)^{2m+1}\,{}_2F_1(2m+2,m+1;m+2;-z), \tag{B8}$$

$$A'_{2m+1}(z) = \frac{2m+1}{1+z}A_{2m+1}(z) - (m+1)\binom{2m+1}{m}\frac{z^m}{1+z}. \tag{B9}$$

Derivation

Now let us proceed with the details of the computations needed to get our result in Eq. (B1). First, we rewrite the Eq. (A6) considering explicitly the normalization factor $\mathcal{Z}_{j\to i}$:

$$m_{j\to i}(\mathbf{s}_j) = \frac{1}{\mathcal{Z}_{j\to i}}\prod_{k\in\partial j\setminus i}\sum_{\mathbf{s}_k} m_{k\to j}(\mathbf{s}_k)\,\exp\left\{\frac{\beta}{2g}(\mathbf{s}_k\cdot\mathbf{s}_j)[2\alpha-1+\mathrm{sgn}(\mathbf{s}_k\cdot\mathbf{s}_j)]\right\}, \tag{B10}$$

$$\mathcal{Z}_{j\to i} = \sum_{\mathbf{s_j}}\prod_{k\in\partial j\setminus i}\sum_{\mathbf{s}_k} m_{k\to j}(\mathbf{s}_k)\,\exp\left\{\frac{\beta}{2g}(\mathbf{s}_k\cdot\mathbf{s}_j)[2\alpha-1+\mathrm{sgn}(\mathbf{s}_k\cdot\mathbf{s}_j)]\right\}. \tag{B11}$$

In the region where there is no preferred orientation, the structure of the BP solutions corresponds to

$$m_{j\to i}(\mathbf{s}_j) = \left(\frac{1}{2}\right)^g, \quad \forall\mathbf{s_j}. \tag{B12}$$

To find the place where $m_{j\to i}(\mathbf{s}_j)$ stops being a solution of the equations, we make a small perturbation to each term of the product in Eq. (B12):

$$m_{j\to i}(\mathbf{s}_j) = \prod_{l=1}^{g}\left(\frac{1+s_j^l\mu_{j\to i}^l}{2}\right). \tag{B13}$$

Here, s_j^l is the lth component of $\mathbf{s}_j$ and $\mu_{j\to i}^l$ is the average value of that lth component. Notice that the distribution in Eq. (B13) is still normalized, and we assume that it remains factorized in the different components. In the completely disordered solution, we have $\mu_{j\to i}^l = 0$. Thus, this will be the small parameter in the perturbation we are making.

There is no reason to believe that one direction should be preferred. Thus, we expect a sort of spontaneous symmetry breaking in which the system randomly chooses to orient itself into one of the directions. To fix ideas, let us say that the preferred direction will have $s_j^l = 1$ for all $l = 1, \ldots, g$. In that case $\mu_{j\to i}$ is the same for all components and Eq.(B13) transforms into

$$m_{j\to i}(\mathbf{s}_j) = \left(\frac{1+\mu_{j\to i}}{2}\right)^{n_{j\to i}} \left(\frac{1-\mu_{j\to i}}{2}\right)^{g-n_{j\to i}}, \tag{B14}$$

where $n_{j\to i} = \sum_{l=1}^{g} \delta(s_j^l, 1)$ is the number of components of $\mathbf{s}_j$ that point upward.

By inserting Eq. (B14) into Eq. (B10) and averaging over the graph ensemble, one gets

$$m_{\text{new}}(\mathbf{s}^*) = \sum_{\gamma=0}^{\infty} P(\gamma)\, \frac{N_\gamma(g,\beta,\alpha)}{Z_\gamma(g,\beta,\alpha)}. \tag{B15}$$

The numerator $N_\gamma(g,\beta,\alpha)$ and the denominator $Z_\gamma(g,\beta,\alpha)$ are given by

$$N_\gamma(g,\beta,\alpha) = \prod_{k=1}^{\gamma} \sum_{n_k=0}^{g} \binom{g}{n_k} \left(\frac{1+\mu}{2}\right)^{n_k} \left(\frac{1-\mu}{2}\right)^{g-n_k} \exp\left\{\frac{\beta}{2g}(2n_k - g)[2\alpha - 1 + \text{sgn}(2n_k - g)]\right\}, \tag{B16}$$

$$Z_\gamma(g,\beta,\alpha) = \sum_{n_j=0}^{g} \binom{g}{n_j} \prod_{k=1}^{\gamma} \sum_{n_k^+=0}^{n_j} \binom{n_j}{n_k^+} \sum_{n_k^-=0}^{g-n_j} \binom{g-n_j}{n_k^-} \left(\frac{1+\mu}{2}\right)^{n_k^+ + g - n_j - n_k^-} \left(\frac{1-\mu}{2}\right)^{n_j + n_k^- - n_k^+}$$
$$\times \exp\left\{\frac{\beta}{2g}(2n_k^+ + 2n_k^- - g)[2\alpha - 1 + \text{sgn}(2n_k^+ + 2n_k^- - g)]\right\}. \tag{B17}$$

In Eq. (B15), we have chosen $n_{j\to i} = g$ and, therefore, all the components of the vector $\mathbf{s}^*$ are equal to one. The message $m_{\text{new}}(\mathbf{s}^*)$ is the one obtained after applying the update rule to the perturbed messages. The reader should recall that we are studying the stability of the fixed point $\hat{m}(\mathbf{s}) = (1/2)^g$ for all $\mathbf{s}$. We need to understand whether, after applying the small perturbation, $m_{\text{new}}(\mathbf{s}^*)$ goes back in the direction of the fixed point or, on the contrary, goes away from it.

Thus, the idea is to expand $m_{\text{new}}(\mathbf{s}^*)$ around $\hat{m}(\mathbf{s}) = (1/2)^g$ in powers of μ. The fixed point is unstable if

$$\left|\left(\frac{1+\mu}{2}\right)^g - \left(\frac{1}{2}\right)^g\right| \leqslant \left|m_{\text{new}}(\mathbf{s}^*) - \left(\frac{1}{2}\right)^g\right|. \tag{B18}$$

Expanding the left-hand side of this inequality in powers of μ, keeping only the first-order terms and taking the equality, one gets

$$\left(\frac{1}{2}\right)^g (1 + g\mu) = \left|m_{\text{new}}(\mathbf{s}^*) - \left(\frac{1}{2}\right)^g\right|. \tag{B19}$$

We will estimate the stability line of the paramagnetic state by finding the solutions of Eq. (B19). To help the reader follow the computation, let us interpret Eqs. (B16) and (B17). Notice that we are giving an expression for a message between two sites, say i and j. We are using all the incoming messages from all the other neighbors of j. The numerator (B16) is a product of sums, and each factor in the product is the contribution of one of the $\gamma = c - 1$ neighbors. They all become equivalent because we are assuming all the messages to be equal.

The sum inside the product in Eq. (B16) goes over all the possible configurations of one of these neighbors. To fix ideas, let us say that the neighbor has an opinion $\mathbf{s}_k$. Since what is important is the scalar product between the opinion $\mathbf{s}_j$ and the opinion of the neighbor $\mathbf{s}_k$, we care about how many of the g components of $\mathbf{s}_k$ are equal to the corresponding components of $\mathbf{s}_j$. That number is what we call n_k in Eq. (B16). The sum over all the configurations of $\mathbf{s}_k$ is substituted for a sum over all the values of n_k, with a combinatorial factor that counts the configurations with the same scalar product $\mathbf{s}_j \cdot \mathbf{s}_k = 2n_k - g$. In Eq. (B16) we also used Eq. (B14) for the messages, remembering that we set all the components of $\mathbf{s}_j$ to one.

The denominator in Eq. (B17) is just the normalization of the numerator [Eq. (B16)]. We consider all the values of $\mathbf{s}_j$ by summing over n_j, which is the number of positive components of $\mathbf{s}_j$. Inside the outer sum, the structure is the same. The integer n_k^+ is the number of components of $\mathbf{s}_k$ whose value is "1" and that are also equal to the corresponding components of $\mathbf{s}_j$. We must have $n_k^+ \leqslant n_j$. However, n_k^- is the number of components of $\mathbf{s}_k$ whose value is "−1" and that are also equal to the corresponding components of $\mathbf{s}_j$. We have $n_k^- \leqslant g - n_j$. Therefore, the scalar product is $\mathbf{s}_j \cdot \mathbf{s}_k = 2n_k^+ + 2n_k^- - g$. The total number of components of $\mathbf{s}_k$ whose value is "1" is $n_k^+ + g - n_j - n_k^-$, and the other $n_j + n_k^- - n_k^+$ are "−1."

Now, we can expand both equations in powers of μ. Let us take the numerator [Eq. (B16)] and keep only the terms up to first order:

$$N_\gamma(g,\beta,\alpha) = 2^{-\gamma g}\prod_{k=1}^{\gamma}\sum_{n_k=0}^{g}\binom{g}{n_k}(1+(2n_k-g)\mu)\exp\left\{\frac{\beta}{2g}(2n_k-g)[2\alpha-1+\mathrm{sgn}(2n_k-g)]\right\}+o(\mu),$$

$$N_\gamma(g,\beta,\alpha) = 2^{-\gamma g}\prod_{k=1}^{\gamma}[\,(1-g\mu)R(g,\beta,\alpha)+2\mu Q(g,\beta,\alpha)\,]+o(\mu),$$

$$N_\gamma(g,\beta,\alpha) = 2^{-\gamma g}[(1-\gamma g\mu)R(g,\beta,\alpha)+2\mu Q(g,\beta,\alpha)]^{\gamma}+o(\mu),$$

$$N_\gamma(g,\beta,\alpha) = 2^{-\gamma g}(R(g,\beta,\alpha))^{\gamma-1}[\,(1-\gamma\, g\,\mu)R(g,\beta,\alpha)+2\mu Q(g,\beta,\alpha)\,]+o(\mu), \tag{B20}$$

where we defined the sums

$$R(g,\beta,\alpha) = \sum_{n_k=0}^{g}\binom{g}{n_k}\exp\left\{\frac{\beta}{2g}(2n_k-g)[2\alpha-1+\mathrm{sgn}(2n_k-g)]\right\}, \tag{B21}$$

$$Q(g,\beta,\alpha) = \sum_{n_k=0}^{g}\binom{g}{n_k}n_k\,\exp\left\{\frac{\beta}{2g}(2n_k-g)[2\alpha-1+\mathrm{sgn}(2n_k-g)]\right\}. \tag{B22}$$

To compute $R(g,\beta,\alpha)$ and $Q(g,\beta,\alpha)$, we need to deal with the "sign" function inside the exponential. The straightforward choice is to split the sum into two or possibly three parts: when the sign is negative, positive, or zero.

When g is even, the sign can be zero. Let us take $m = g/2$ and write

$$R(2m,\beta,\alpha) = \sum_{n_k=0}^{m-1}\binom{2m}{n_k}\exp\left\{\frac{\beta}{m}(n_k-m)(\alpha-1)\right\}+\binom{2m}{m}+\sum_{n_k=m+1}^{2m}\binom{2m}{n_k}\exp\left\{\frac{\beta}{m}(n_k-m)\alpha\right\},$$

$$R(2m,\beta,\alpha) = e^{-\beta(\alpha-1)}\sum_{n_k=0}^{m-1}\binom{2m}{n_k}\exp\left\{\frac{\beta(\alpha-1)}{m}\,n_k\right\}+\binom{2m}{m}+e^{\beta\,\alpha}\sum_{n_k=0}^{m-1}\binom{2m}{n_k}\exp\left\{-\frac{\beta\,\alpha}{m}\,n_k\right\}. \tag{B23}$$

Then, we identify that

$$R(2m,\beta,\alpha) = e^{-\beta(\alpha-1)}A_{2m}[e^{\beta(\alpha-1)/m}]+e^{\beta\alpha}A_{2m}[e^{-\beta\alpha/m}]+\binom{2m}{m}, \tag{B24}$$

where

$$A_{2m}[z] = \sum_{n=0}^{m-1}\binom{2m}{n}z^n. \tag{B25}$$

To compute this last sum, we take the derivative of Eq. (B25) with respect to z,

$$A'_{2m}[z] = \sum_{n=0}^{m-1}n\binom{2m}{n}z^{n-1} = \sum_{n=0}^{m-2}\frac{(2m)!\,z^n}{n!(2m-n-1)!} = \sum_{n=0}^{m-1}(2m-n)\frac{(2m)!\,z^n}{n!(2m-n)!}-(m+1)\frac{(2m)!\,z^{m-1}}{(m-1)!(m+1)!},$$

$$A'_{2m}[z] = 2mA_{2m}[z]-zA'_{2m}[z]-(m+1)\binom{2m}{m-1}z^{m-1}, \tag{B26}$$

from where

$$(1+z)^{2m}\frac{d}{dz}[(1+z)^{-2m}A_{2m}[z]] = -(m+1)\binom{2m}{m-1}\frac{z^{m-1}}{1+z},$$

$$(1+z)^{2m}A_{2m}[z] = (1+z)^{2m}-(m+1)\binom{2m}{m-1}(1+z)^{2m}\int_0^x dt\,\frac{t^{m-1}}{(1+t)^{2m+1}}, \tag{B27}$$

where we used that $A_{2m}[0] = 1$.

Taking into account now the following integral expression of the ordinary hypergeometric function:

$${}_2F_1(a,b;c;z) = \frac{1}{B(b,c-b)}\int_0^1 x^{b-1}(1-x)^{-b-1}(1-zx)^{-a}dx,\quad \mathrm{Re}[c]>\mathrm{Re}[b]>0, \tag{B28}$$

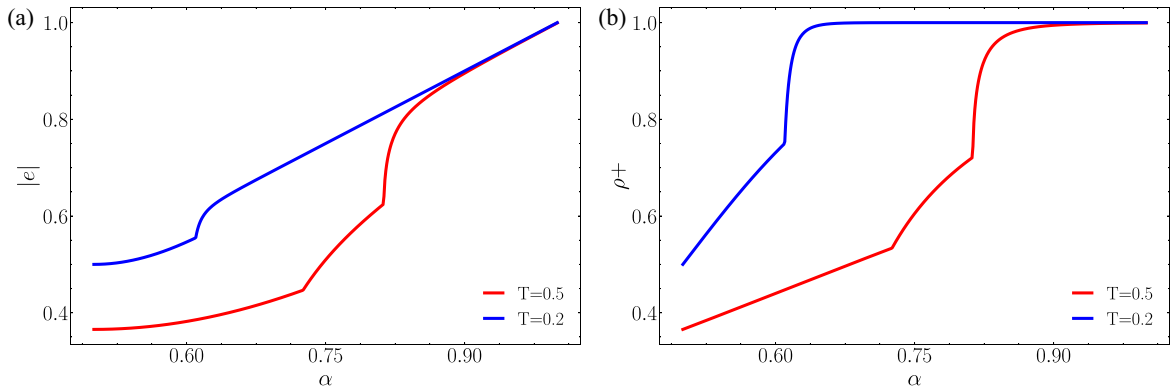


FIG. 4. Belief Propagation results for the absolute value of the energy density $|e|$ and the fraction of positive links ρ_+ in random regular graphs with connectivity $c = 3$. The lines are the outcome of BP's convergence process for a given value of α and the temperature T. All the messages are initially aligned toward a given direction. The panels contain results $g = 2$ and two temperatures, $T = 1/5$ and $T = 1/2$.

where $B(b, c-b)$ is Euler's β function, we get Eq. (B4):

$$A_{2m}(z) = (1+z)^{2m} - \frac{m+1}{m}\binom{2m}{m-1} z^m (1+z)^{2m}\, {}_2F_1(2m+1, m; m+1; -z). \tag{B29}$$

Once we see how our computations are connected to the hypergeometric function, the rest is analogous. It is not difficult to see that for even g the sum $Q(g, \beta, \alpha)$ is then related to the derivative of $A_{2m}[z]$ [see Eqs. (B3) and (B5)]. When g is odd, the procedure is the same with only small modifications that are reflected in the results presented in Eqs. (B6), (B7), (B8), and (B9).

When we insert everything back into the Eq. (B15) and keep expanding in powers of μ to the first order, we get

$$1 + g\mu = 1 + \langle\gamma\rangle\left(\frac{2Q(g,\beta,\alpha)}{R(g,\beta,\alpha)} - g\right)\mu + o(\mu), \tag{B30}$$

where $\langle\gamma\rangle = \sum_{\gamma=0}^{\infty} \gamma\, \mathbb{P}(\gamma)$. Therefore, we arrive at Eq. (B1):

$$\frac{Q(g,\beta,\alpha)}{R(g,\beta,\alpha)} = \frac{g}{2}\frac{1+\langle\gamma\rangle}{\langle\gamma\rangle}. \tag{B31}$$

APPENDIX C: CAPTURING THE TRANSITION WITH BELIEF PROPAGATION

There are 2×2^g equations like Eq. (A6) for each edge in the interaction graph. We can solve them by iterating until the convergence of the messages $m_{i\to j}(\mathbf{s}_j)$. If our graph is locally treelike, then we should expect a good agreement between the results of this belief propagation (BP) algorithm and the Monte Carlo simulations. Since we suspect the presence of a transition between two different states, one where the opinion vectors $\mathbf{s}_i$ do not have any preferential orientation, and other where they exhibit a spontaneous alignment toward a preferred direction, we will compute the absolute value of the energy density $|e|$, where e is also known as the level of social stress per pair of agents, the fraction of positive links ρ_+, and the magnetization m. Their definitions are

$$\begin{aligned}
|e| &= -\langle H\rangle = \frac{1}{L}\sum_{\langle ij\rangle}\sum_{\mathbf{s}_i}\sum_{\mathbf{s}_j} P(\mathbf{s}_i,\mathbf{s}_j)\,\frac{\mathbf{s}_i\cdot\mathbf{s}_j}{2G}[2\alpha - 1 \\
&\quad + \mathrm{sgn}(\mathbf{s}_i\cdot\mathbf{s}_j)], \\
\rho_+ &= \frac{1}{L}\sum_{\langle ij\rangle}\sum_{\mathbf{s}_i}\sum_{\mathbf{s}_j} P(\mathbf{s}_i,\mathbf{s}_j)\,\Theta(\mathbf{s}_i\cdot\mathbf{s}_j), \\
m &= \left|\left|\frac{1}{Ng}\sum_i\sum_{\mathbf{s}_i} P(\mathbf{s}_i)\mathbf{s}_i\right|\right|_2,
\end{aligned} \tag{C1}$$

where $P(\mathbf{s}_i)$ and $P(\mathbf{s}_i, \mathbf{s}_j)$ can be computed using BP or Monte Carlo, L is the number of edges in the graph and N the number of vertices, $\Theta(x)$ is one only if $x > 0$ and gives zero otherwise, and $||\,\mathbf{v}\,||_2$ is the Euclidean norm of the vector $\mathbf{v}$.

It can be easily checked that the maximum value of $|e|$ depends on α. For $\alpha = 1/2$ we have

$$|e| = \frac{1}{L}\sum_{\langle ij\rangle}\frac{\mathbb{E}_{ij}[\,|\mathbf{s}_i\cdot\mathbf{s}_j|\,]}{2G} \leqslant \frac{1}{2},$$

where $\mathbb{E}_{ij}[\,\cdot\,] \equiv \sum_{\mathbf{s}_i}\sum_{\mathbf{s}_j} P(\mathbf{s}_i,\mathbf{s}_j)\,[\,\cdot\,]$. For $\alpha = 1$ we have

$$|e| = \frac{1}{L}\sum_{\langle ij\rangle}\left(\frac{\mathbb{E}_{ij}[\mathbf{s}_i\cdot\mathbf{s}_j]}{2G} + \frac{\mathbb{E}_{ij}[\,|\mathbf{s}_i\cdot\mathbf{s}_j|\,]}{2G}\right) \leqslant 1.$$

In Fig. 4 we show the values of $|e|$ and ρ_+ for different values of α around the transition. The calculations were performed for random regular graphs. It is important to note that, in general, the computational cost of finding the fixed point of BP equations increases exponentially in g. However, when one chooses the same initial values for all the messages in a random regular graph with homogeneous interactions, the set of BP equations reduces to only one equivalent equation. Therefore, the numerics can be directly done using only one equation, and it is very fast.

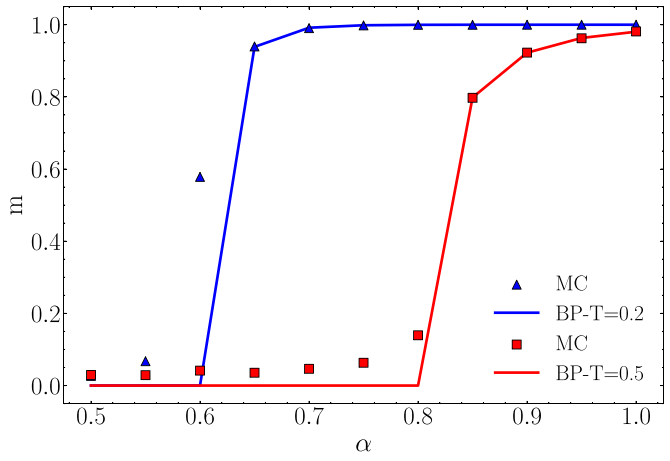


FIG. 5. Comparison between Monte Carlo's and BP's magnetization in random regular graphs with connectivity $c = 3$. The number of individual opinions is $g = 2$ in all cases. Each point is the result of averaging $s = 50$ Monte Carlo simulations in systems of size $N = 1000$. All simulations started with all the spins aligned toward a given direction ($m(t = 0) = 1$). The lines are the outcome of BP's convergence for some values of α and the temperature T. All the messages are set initially to favor the same direction chosen for Monte Carlo's initial conditions. The figure contains results for two temperatures, $T = 1/5$ and $T = 1/2$.

Although the energy and fraction of positive links show some small changes in behavior around specific values of α, it is not possible to see the extent of what is happening in the system. Therefore, we also computed the magnetization in Fig. 5, where we find a good agreement between Monte Carlo simulations and the results of the BP equations. The figure also shows the existence of a sharp temperature-dependent change in the system's behavior around a specific value of α: below a certain α there is no preferred orientation, and above that α the system aligns.

The sharp transitions observed in the magnetization coincide with discontinuities in the derivative with respect to *alpha* of the energy and the fraction of positive links. Notice that at $T = 0.2$ there is only one discontinuity in the derivative (blue curves in Fig. 4). However, at $T = 0.5$ both the energy and the fraction of positive links display two discontinuities in their derivatives with respect to α. The one located at the larger α is explained by the order-disorder transition signaled by the magnetization in Fig. 5. The other one, at a smaller α, does not seem to be related to this transition, and further investigation is needed to understand its origin.

APPENDIX D: MONTE CARLO DYNAMICS

The results obtained with BP suggest the relevance of the initial conditions in the system's behavior. In particular, the

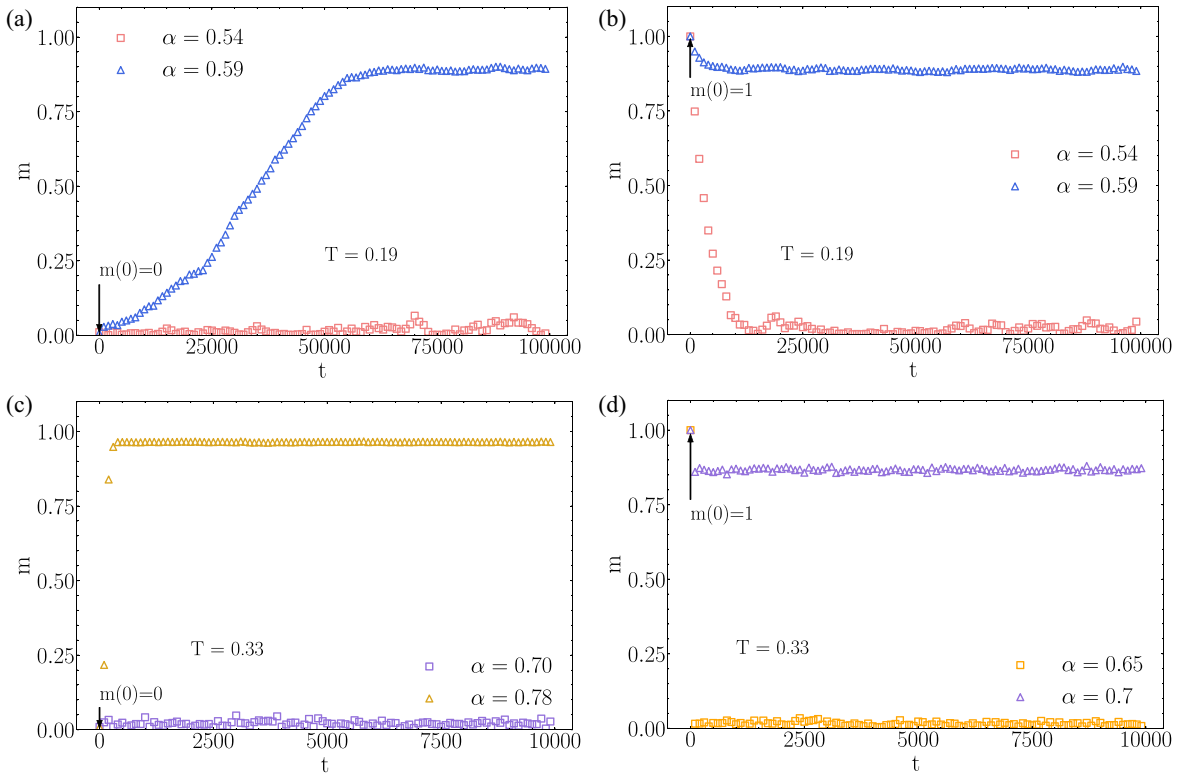


FIG. 6. Monte Carlo dynamics in random regular graphs with connectivity $c = 4$. The length of the opinions of each individual is $g = 4$. The upper panels show the average of 5 Monte Carlo runs at a low temperature ($T = 0.19$). The bottom panels show the average of 20 Monte Carlo runs at a higher temperature ($T = 0.33$). The simulations in the left panels start with a random initial condition, while the ones in the right panels start completely aligned toward a given direction. System size is $N = 10\,000$. The time unit is a Monte Carlo step, consisting of N individual Metropolis steps. Arrows indicate the starting points of the trajectories.

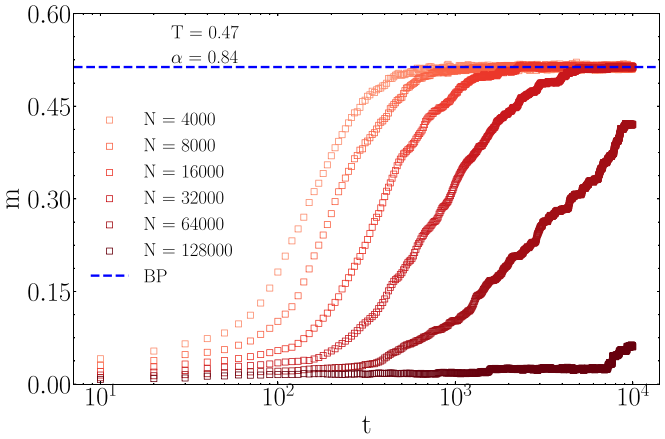


FIG. 7. Monte Carlo dynamics in Erdős-Rényi graphs with average connectivity $\kappa = 3$ across different system sizes. The figure shows the average of 200 Monte Carlo runs for fixed temperature T and α, with random initial conditions. The length of the opinions of each individual is $g = 3$. The time unit is a Monte Carlo step, consisting of N individual Metropolis steps. The horizontal line marks the magnetization of the ferromagnetic fixed point obtained with BP.

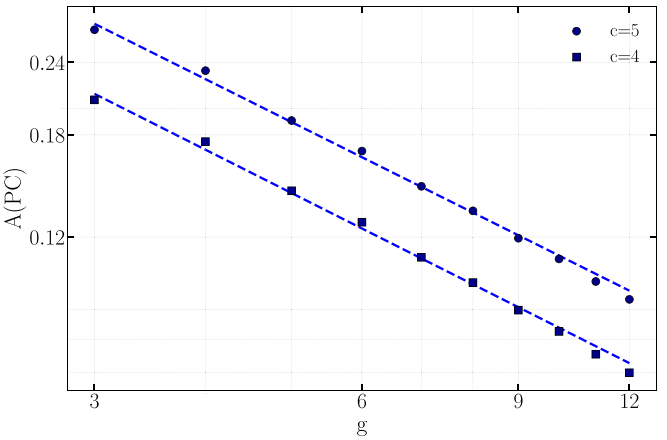


FIG. 8. Area of possible consensus A(PC) vs the length g of the opinion vectors on a logarithmic scale with fixed connectivity $c = 4$ (squares) and $c = 5$ (circles). Blue dashed lines are fitted to the data points using a power law $\mathrm{A(EC)} = \Phi(c)\, g^{-\gamma(c)}$. The optimal parameters depend on the connectivity. They are: $\Phi(4) = 0.494(15)$, $\Phi(5) = 0.61(2)$, and $\gamma(4) = 0.771(18)$, $\gamma(5) = 0.764(18)$.

location of the transition strongly depends on the initial conditions. We show in Fig. 6 that this is also the case for Monte Carlo simulations. There, the upper panels show Monte Carlo runs performed at a low temperature ($T = 0.19$), while the lower panels are done for a high temperature ($T = 0.33$). The left column presents results obtained from a random initial condition with zero magnetization [$m(0) = 0$], and the right column shows those obtained from a completely aligned initial condition [$m(0) = 1$].

For $T = 0.19$ (top panels), the steady states are independent of the initial magnetization. When $\alpha = 0.59$ (blue points), the Monte Carlo dynamics leads the system to the same final nonzero magnetization ($0.8 < m < 0.9$) in both panels. Instead, for $\alpha = 0.54$ (red points) and independently of the initial configuration, the steady-state magnetization remains close to zero.

However, for $T = 0.33$ (bottom panels), the steady states *do* depend on the initial condition. In the bottom-left panel, the system is not magnetized at any time for $\alpha = 0.70$ when initiated in a random configuration [$m(0) = 0$]. In the bottom-right, the dynamics for the same $\alpha = 0.70$ remains magnetized if initiated at $m(0) = 1$. This simple example illustrates the dependence on the initial conditions near the first-order phase transition described in the main text.

To explore the effects of metastability, we ran Monte Carlo simulations with different system sizes in the region where the paramagnetic state is indeed metastable. Figure 7 presents the dynamics at a fixed temperature T and α corresponding to this region, using random initial conditions [$m(0) = 0$]. When the system size N is small enough ($N \leqslant 32\,000$ in the figure), the system escapes the metastable state and arrives in the ferromagnetic ground state at $t < 10\,000$. Indeed, we checked that the final magnetization of the Monte Carlo simulations for all sizes $N \leqslant 32\,000$ corresponds to the ferromagnetic solution of BP (horizontal line in the figure).

As N increases, the transient becomes longer, indicating a significant slowing down of the dynamics. For $N = 128\,000$, the system barely leaves the paramagnetic state for $t < 10\,000$. As we argue in the main text, the time needed to reach the ferromagnetic state is related to the crossing of extensive free-energy barriers and grows exponentially following an Arrhenius law [46,48]. Figure 7 qualitatively corroborates this idea, since the curves for different system sizes continue showing big shifts to the right, even when we use a logarithmic scale for the time.

APPENDIX E: BEHAVIOR OF THE AREA OF POSSIBLE CONSENSUS WHEN THE NUMBER OF DISCUSSED ISSUES INCREASES

In the main text, we identify two relevant regions of the model's phase diagram. We denoted by A(EC) the area where consensus is easy to find, regardless of the initial conditions. We showed that A(EC) decreases when the number of discussed issues (g) increases, following a power law. However, A(PC) is the entire area where consensus is possible. For completeness, in this Appendix we include Fig. 8, illustrating the dependence of A(PC) on the value of g. The area where consensus is possible also decays with g, but this cannot be strictly described using a power law. For small g, the points in Fig. 8, displayed in logarithmic scale, are close to a straight line representing the best power-law fit to the points. For large g, instead, the points start to deviate from this power law.

Two things must be commented on here. First, the value of A(PC) is not as relevant as the value of A(EC). Only the latter represents the region where consensus can be reached from a random initial condition, i.e., A(EC) is more suitable as a quantitative measure of consensus feasibility. Second, the fact that A(PC) is not a power law is responsible for the saturation of the gap size (see Fig. 2 in the main text) to a constant value other than one. Therefore, we can say that all the relevant

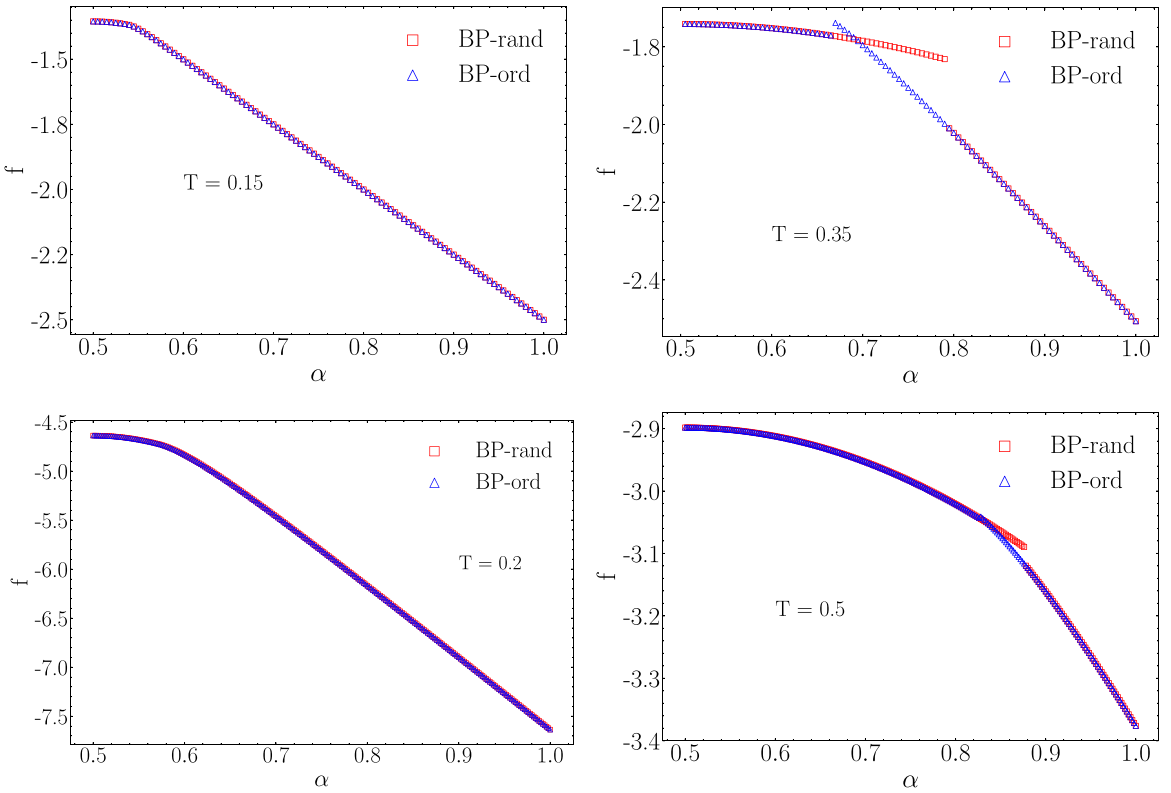


FIG. 9. Free-energy density computed with BP vs α for random regular (top panels) and Erdős-Rényi (bottom panels) graphs. Red squares are obtained with random initial conditions, while blue triangles are obtained with fully aligned initial conditions. The temperatures are $T = 0.15$ (top-left) and $T = 0.35$ (top-right) for random regular graphs. For Erdős-Rényi graphs we selected $T = 0.2$ (bottom-left) and $T = 0.5$ (bottom-right).

information can be extracted from Fig. 2 in the main text, although Fig. 8 is still a good complement.

APPENDIX F: COEXISTENCE OF DISCONTINUOUS AND CONTINUOUS PHASE TRANSITIONS

The model defined in the main text displays several interesting features. One of them is the coexistence, for the same number of discussed issues g and the same random graph of interactions, of continuous and discontinuous phase transitions. For high temperatures, the order-disorder transition (ferromagnetic state to paramagnetic state) is discontinuous. For low temperatures, instead, it is continuous.

Figure 9 shows that this phenomenon is qualitatively the same in both families of random graphs: random regular and Erdős-Rényi. For low temperatures (left panels), the free-energy density f computed with BP is indeed a continuously differentiable function of α. For high temperatures (right panels), we observe two distinct branches of the free energy. When two possible values coexist, the true f is always the minimum. The resulting function has a discontinuity in its derivative. Thus, the transition is discontinuous.

The existence of different fixed points for the same value of α is enough to explain the strong dependence of the dynamics on the initial conditions. For example, in the top-right panel of Fig. 9 we see that around $\alpha \sim 0.8$ we have two distinct stable states (one marked with red points and the other with blue points). The one with the larger free energy (the paramagnetic state) is metastable, while the other (ferromagnetic state) is the ground state. Even though the ground state is accessible and consensus exists, when the dynamics is initialized close to the paramagnetic state it will remain there for a time that grows exponentially with the system size. Since everything is explained by looking at the possible fixed points, the fact that BP gives the appropriate description in equilibrium is enough to predict the behavior of Monte Carlo simulations.

APPENDIX G: CONCENTRATION OF THE DENSITY OF CONFIGURATIONS AROUND $m = 0$

The existence of a stable paramagnetic state, even in the region where the ferromagnetic solution is the ground state, is related to the fact that most configurations of the system have a magnetization around $m = 0$. To show it, it is useful to start with only one issue to discuss ($g = 1$), in which case the problem is equivalent to an Ising model with homogeneous interactions (ferromagnetic for $\alpha > 0.5$ and antiferromagnetic for $\alpha < 0.5$). Therefore, the magnetization is directly related to the nonnegative integer N_+, which counts the number of variables with value $s_i = 1$. Using that $N_+ = N(1 + m)/2$, we can write the density of configurations with magnetization m for large system size N as

$$\mathcal{D}(m) = \frac{1}{2^N}\frac{N!}{[N(1+m)/2]!\,[N(1-m)/2]!} \sim \frac{1}{\sqrt{2\pi N}}\frac{1}{2^N}(\sqrt{(1+m)/2}\,)^{-N(1+m)-1}(\sqrt{(1-m)/2}\,)^{-N(1-m)-1},$$
$$\mathcal{D}(m) \sim \sqrt{\frac{2}{\pi N}}(\sqrt{1+m}\,)^{-N(1+m)-1}(\sqrt{1-m}\,)^{-N(1-m)-1} \sim \sqrt{\frac{2}{\pi N}}[(1+m)^{(1+m)/2}\,(1-m)^{(1-m)/2}]^{-N}, \tag{G1}$$

where we applied Stirling's approximation for the factorial and the fact that $N(1 \pm m) + 1 \sim N(1 \pm m)$ for $N \gg 1$.

The maximum of $\mathcal{D}(m)$ is reached at $m = 0$ independently of the value of N, where the density is equal to $\mathcal{D}(0) = \sqrt{2/\pi N}$. Notice that this implies the inequality $(1+m)^{(1+m)/2}\,(1-m)^{(1-m)/2} \geqslant 1$ for all values of m, so that the density $\mathcal{D}(m)$ can be rewritten as

$$\mathcal{D}(m) \sim \sqrt{\frac{2}{\pi N}}\,[q(m)]^{-N}, \tag{G2}$$

with $q(m) > 1$ for all $m \neq 0$, increasing monotonically when m grows away from zero and with its minimum at $q(0) = 1$.

In particular, this means that the fraction of configurations with a magnetization larger than m has the form

$$\bar{F}_1(m) = \int_m^1 dx\,\mathcal{D}(x) \leqslant \sqrt{\frac{2}{\pi N}}\,(1-m)\,[q(m)]^{-N}, \tag{G3}$$

since $q(x) \geqslant q(m)$ for all $x \in [m, 1]$, and then we have $[q(x)]^{-N} \leqslant [q(m)]^{-N}$ in that interval. Recalling that $q(m) > 1$ for any positive m, we conclude that the fraction of configurations with magnetization larger than m decreases exponentially when the system size increases. Although this is written for the case $m > 0$, the reader should note that something analogous is valid for $m < 0$.

Now, returning to the cases with more than one issue to be discussed, we have opinion vectors $\mathbf{s}_i$ of length $g > 1$. Nevertheless, we can ask an equivalent question to understand how many configurations are close to one of the 2^g possible aligned configurations. Given that they are all equivalent, we can choose one. For example, we can examine whether the average opinion vector associated with a configuration ($\langle \mathbf{s} \rangle = \sum_i \mathbf{s}_i/N$) points toward the specific direction that has all its components equal to one ($\mathbf{s}^\mu = 1, \forall \mu = 1, \ldots, g$). Therefore, we define that $\langle \mathbf{s} \rangle$ is within a distance $1 - m$ from this specific direction if the absolute value of each of its components is greater than m. The fraction of configurations with such opinion vectors is simply

$$\bar{F}_g(m) \leqslant \left[\sqrt{\frac{2}{\pi N}}\,(1-m)\,[q(m)]^{-N}\right]^g = \left(\frac{2}{\pi N}\right)^{g/2}(1-m)^g\,[q(m)]^{-gN}. \tag{G4}$$

As before, the function in Eq. (G4) decreases exponentially when N increases. Furthermore, it is important to note that the dominating factor $[q(m)]^{-gN}$ also decreases exponentially when g increases. We conclude that increasing g reinforces the concentration of the density of configurations around $m = 0$. This is key to understanding why the area of the phase diagram where consensus is easy to find shrinks when we increase g. In that case, the dynamics must visit a larger number of configurations to abandon the region $m \sim 0$.

[1] C. Castellano, S. Fortunato, and V. Loreto, Statistical physics of social dynamics, Rev. Mod. Phys. **81**, 591 (2009).

[2] M. Jusup, P. Holme, K. Kanazawa, M. Takayasu, I. Romić, Z. Wang, S. Geček, T. Lipić, B. Podobnik, L. Wang, W. Luo, T. Klanjšček, J. Fan, S. Boccaletti, and M. Perc, Social physics, Phys. Rep. **948**, 1 (2022).

[3] S. Redner, Reality-inspired voter models: A mini-review, C. R. Phys. **20**, 275 (2019).

[4] P. DiMaggio, J. Evans, and B. Bryson, Have American's social attitudes become more polarized? Am. J. Sociol. **102**, 690 (1996).

[5] M. P. Fiorina and S. J. Abrams, Political polarization in the American public, Annu. Rev. Polit. Sci. **11**, 563 (2008).

[6] M. Granovetter, Threshold models of collective behavior, Am. J. Sociol. **83**, 1420 (1978).

[7] J. P Gleeson, Cascades on correlated and modular random networks, Phys. Rev. E **77**, 046117 (2008).

[8] E. Bonabeau, G. Theraulaz, and J.-L. Deneubourg, Phase diagram of a model of self-organizing hierarchies, Phys. A: Stat. Mech. Appl. **217**, 373 (1995).

[9] D. Mistry, Q. Zhang, N. Perra, and A. Baronchelli, Committed activists and the resilience of crowd dynamics, Phys. Rev. E **92**, 042805 (2015).

[10] N. E. Friedkin, Norm formation in social influence networks, Soc. Netw. **23**, 167 (2001).

[11] V. Sood and S. Redner, Voter model on heterogeneous graphs, Phys. Rev. Lett. **94**, 178701 (2005).

[12] P. Holme and M. E. J. Newman, Nonequilibrium phase transition in the coevolution of networks and opinions, Phys. Rev. E **74**, 056108 (2006).

[13] F. Vazquez, V. M. Eguíluz, and M. S. Miguel, Generic absorbing transition in coevolution dynamics, Phys. Rev. Lett. **100**, 108702 (2008).

[14] R. Durrett, J. P. Gleeson, A. L. Lloyd, P. J. Mucha, F. Shi, D. Sivakoff, J. E. S. Socolar, and C. Varghese, Graph fission in an evolving voter model, Proc. Natl. Acad. Sci. USA **109**, 3682 (2012).

[15] 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).

[16] T. Pham, S. Redner, L. Waldorp, J. Armas, and H. L. J. van der Maas, Polarisation in increasingly connected societies, arXiv:2503.24098.

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

[18] T. M. Pham, A. C. Alexander, J. Korbel, R. Hanel, and S. Thurner, Balance and fragmentation in societies with homophily and social balance, Sci. Rep. **11**, 17188 (2021).

[19] P. J. Górski, K. Bochenina, J. A. Hołyst, and R. M. D' Souza, Homophily based on few attributes can impede structural balance, Phys. Rev. Lett. **125**, 078302 (2020).

[20] P. J. Górski, C. Atkisson, and J. A. Hołyst, A general model for how attributes can reduce polarization in social groups, Net. Sci. **11**, 536 (2023).

[21] P. J Górski, A. Sulik, G. Andres, G. Vaccario, and J. A Hołyst, Coexistence of balance and hierarchies: An ego perspective to explain empirical networks, PNAS Nexus **4**, pgaf130 (2025).

[22] T. M. Pham, J. Korbel, R. Hanel, and S. Thurner, Empirical social triad statistics can be explained with dyadic homophylic interactions, Proc. Natl. Acad. Sci. USA **119**, e2121103119 (2022).

[23] J. Korbel, S. D. Lindner, T. M. Pham, R. Hanel, and S. Thurner, Homophily-based social group formation in a spin glass self-assembly framework, Phys. Rev. Lett. **130**, 057401 (2023).

[24] A. Flache, M. Mäs, T. Feliciani, E. Chattoe-Brown, G. Deffuant, S. Huet, and J. Lorenz, Models of social influence: Towards the next frontiers, JASSS **20**, 2 (2017).

[25] M. W. Macy, J. A. Kitts, A. Flache, and S. Benard, Polarization in dynamic networks: A Hopfield model of emergent structure, in *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers* (National Academies Press, Washington, DC, 2003), p. 162.

[26] M. Galesic, H. Olsson, T. Pham, J. Sorger, and S. Thurner, Experimental evidence confirms that triadic social balance can be achieved through dyadic interactions, npj Complex. **2**, 1 (2025).

[27] Y. Kabashima and D. Saad, Belief propagation vs TAP for decoding corrupted messages, Europhys. Lett. **44**, 668 (1998).

[28] J. S Yedidia, W. T Freeman, and Y. Weiss, Constructing free-energy approximations and generalized belief propagation algorithms, Inf. Theory IEEE Trans. **51**, 2282 (2005).

[29] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, Burlington, MA, 1988).

[30] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, Factor graphs and the sum-product algorithm, IEEE Trans. Inf. Theory **47**, 498 (2001).

[31] J. S Yedidia, W. Freeman, and Y. Weiss, Generalized belief propagation, in *Advances in Neural Information Processing Systems*, edited by T. Leen, T. Dietterich, and V. Tresp (MIT Press, Cambridge, MA, 2000), Vol. 13, pp. 689–695.

[32] J. S Yedidia, W. T Freeman, and Y. Weiss, Understanding belief propagation and its generalizations, in *Exploring Artificial Intelligence in the New Millennium* (Morgan Kaufmann, Burlington, MA, 2003), pp. 239–269.

[33] T. Rizzo, A. Lage-Castellanos, R. Mulet, and F. Ricci-Tersenghi, Replica cluster variational method, J. Stat. Phys. **139**, 375 (2010).

[34] F. L. Metz, I. Neri, and D. Bollé, Spectra of sparse regular graphs with loops, Phys. Rev. E **84**, 055101(R) (2011).

[35] D. Bollé, F. L. Metz, and I. Neri, On the spectra of large sparse graphs with cycles, in *Spectral Analysis, Differential Equations and Mathematical Physics: A Festschrift in Honor of Fritz Gesztesy's 60th Birthday*, edited by H. Holden *et al.* (American Mathematical Society, Providence, RI, 2013), pp. 35–58.

[36] A. Kirkley, G. T. Cantwell, and M. E. J. Newman, Belief propagation for networks with loops, Sci. Adv. **7**, eabf1211 (2021).

[37] G. T. Cantwell, A. Kirkley, and F. Radicchi, Heterogeneous message passing for heterogeneous networks, Phys. Rev. E **108**, 034310 (2023).

[38] Y. Wang, Y. E. Zhang, F. Pan, and P. Zhang, Tensor network message passing, Phys. Rev. Lett. **132**, 117401 (2024).

[39] Grover E. C. Guzman, P. F. Stadler, and A. Fujita, Cavity approach for the approximation of spectral density of graphs with heterogeneous structures, Phys. Rev. E **109**, 034303 (2024).

[40] T. M. Pham, T. Peron, and F. L. Metz, Effects of clustering heterogeneity on the spectral density of sparse networks, Phys. Rev. E **110**, 054307 (2024).

[41] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, Mixing beliefs among interacting agents, Adv. Complex Syst. **03**, 87 (2000).

[42] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, Emergence of polarized ideological opinions in multidimensional topic spaces, Phys. Rev. X **11**, 011012 (2021).

[43] A. Jedrzejewski and K. Sznajd-Weron, Person-situation debate revisited: Phase transitions with quenched and annealed disorders, Entropy **19**, 415 (2017).

[44] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. **21**, 1087 (1953).

[45] D. B. Bahr and E. Passerini, Statistical mechanics of opinion formation and collective behavior: Micro-sociology, J. Math. Soc. **23**, 1 (1998).

[46] M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, Oxford, UK, 2009).

[47] A. Bizyaeva, A. Franci, and N. E. Leonard, Multitopic belief formation through bifurcations over signed social networks, IEEE Trans. Autom. Control **70**, 5082 (2025).

[48] S. Arrhenius, Über die dissociationswärme und den einfluss der temperatur auf den dissociationsgrad der elektrolyte, Z. Phys. Chem. **4U**, 96 (1889).

[49] See https://github.com/d4v1d-cub/Homophily_RandGraphs and https://doi.org/10.5281/zenodo.15789890.

[50] M. Mézard and G. Parisi, The Bethe lattice spin glass revisited, Eur. Phys. J. B **20**, 217 (2001).

[51] A. Mozeika and A. C. C. Coolen, Dynamical replica analysis of processes on finitely connected random graphs: II Dynamics in the griffiths phase of the diluted ising ferromagnet, J. Phys. A: Math. Theor. **42**, 195006 (2009).

[52] J. Korbel, S. Havlin, and S. Thurner, Microscopic origin of abrupt mixed-order phase transitions, Nat. Commun. **16**, 2628 (2025).

[53] M. G. Noudehi, A. Kargaran, N. Azimi-Tafreshi, and G. R. Jafari, Second- to first-order phase transition: Coevolutionary versus structural balance, Phys. Rev. E **106**, 044303 (2022).

[54] A. Lage-Castellanos and R. Mulet, Zero temperature solutions of the Edwards-Anderson model in random Husimi lattices, Eur. Phys. J. B **65**, 117 (2008).

[55] L. Pedraza, S. Pinto, J. P. Pinasco, and P. Balenzuela, Analytical approach to the axelrod model based on similarity vectors, Phys. Rev. E **103**, 012307 (2021).

[56] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian, Clusters of solutions and replica symmetry breaking in random k-satisfiability, J. Stat. Mech. (2008) P04004.

[57] F. R. -Tersenghi and G. Semerjian, On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms, J. Stat. Mech. (2009) P09001.

[58] L. Zdeborová and F. Krzakala, Generalization of the cavity method for adiabatic evolution of Gibbs states, Phys. Rev. B **81**, 224205 (2010).