

# Analytical and numerical study of the non-linear noisy voter model on complex networks

A. F. Peralta,<sup>1,a)</sup> A. Carro,<sup>2,3</sup> M. San Miguel,<sup>1</sup> and R. Toral<sup>1</sup>

<sup>1</sup>*IFISC (Institut de Física Interdisciplinar y Sistemas Complejos), Universitat de les Illes Balears-CSIC, 07122 Palma de Mallorca, Spain*

<sup>2</sup>*Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, OX2 6ED Oxford, UK*

<sup>3</sup>*Mathematical Institute, University of Oxford, OX2 6GG Oxford, UK*

(Received 19 March 2018; accepted 22 June 2018; published online 24 July 2018)

We study the noisy voter model using a specific non-linear dependence of the rates that takes into account collective interaction between individuals. The resulting model is solved exactly under the all-to-all coupling configuration and approximately in some random network environments. In the all-to-all setup, we find that the non-linear interactions induce *bona fide* phase transitions that, contrary to the linear version of the model, survive in the thermodynamic limit. The main effect of the complex network is to shift the transition lines and modify the finite-size dependence, a modification that can be captured with the introduction of an effective system size that decreases with the degree heterogeneity of the network. While a non-trivial finite-size dependence of the moments of the probability distribution is derived from our treatment, mean-field exponents are nevertheless obtained in the thermodynamic limit. These theoretical predictions are well confirmed by numerical simulations of the stochastic process. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5030112>

**Imitation models where individuals copy the actions or opinions of others are the basis to understand the transition to consensus and organized behavior in societies. The voter model incorporates the simplest mechanism of blind imitation and has become one of the accepted paradigms in this field. To make the model closer to reality, one needs to take into account the addition of spontaneous changes of state, the characteristics of the network of interactions, and the details of the imitation mechanism. All three ingredients have been considered in the present paper. The resulting non-equilibrium model turns out to offer a rich phenomenology, and its phase diagram includes tricritical points, catastrophes, and a non-trivial scaling behavior that can be analyzed using some tools of equilibrium statistical mechanics.**

## I. INTRODUCTION

The original voter model<sup>1,2</sup> implements, in its simplest way, the mechanism of random imitation by which a “voter” (an individual represented by a variable that can be in any of two possible states) adopts the state of one of its neighbors in a given network of interactions. The voter model displays stochastic dynamics with absorbing states that correspond to the collective consensus in each of the two possible states and that, from the point of view of Statistical Physics and Critical Phenomena, have very special characteristics.<sup>3</sup> Some studies<sup>4,5</sup> have examined in detail the robustness or generic behavior of the voter model dynamics by the introduction of nonlinear variations of the random imitation mechanism that, however, respect the existence of the two absorbing states of the stochastic dynamics. A different modification of the

voter model is the noisy voter<sup>6</sup> model, also called the Kirman model,<sup>7</sup> that has appeared under different names and contexts in the literature<sup>8–12</sup> and that adds to the random imitation mechanism spontaneous switches of state or noise. It was introduced, in the social sciences literature, as a simple stochastic process that could explain some experimental features observed in ant colonies, when ants have to choose between two symmetrical food sources, as well as some stylized facts—or non-Gaussian statistical properties—observed in financial markets, where traders have to decide whether to buy or sell a given stock. In both cases, noise is introduced to allow switches between the two collective consensus avoiding the existence of absorbing states. With its very simple behavioral rules, the noisy voter model is able to capture the emergence of herding and non-Gaussian statistical properties in those two different contexts.<sup>13,14</sup> Furthermore, the model presents an order-disorder transition as a function of the relative importance of the spontaneous switching with respect to the copying mechanism, the order parameter being the fraction of voters that share a common value for the state variable. This is known to be a finite-size noise induced transition,<sup>7</sup> with well characterized critical properties in the thermodynamic limit,<sup>6</sup> where the transition disappears—or rather, it occurs in the limit of vanishing spontaneous switching. Some modifications of the system size scaling of the copying mechanism have been proposed to preserve the existence of the order-disorder transition in the thermodynamic limit.<sup>13,14</sup> In this paper, we address the question of the robustness of the order-disorder transition of the noisy voter model under nonlinear modifications of the random imitation mechanism, emphasizing the system size scaling of its critical properties.

The voter model can be thought of as based on a dyadic, direct interaction between two neighboring voters so that the

<sup>a)</sup>Electronic mail: [afperalta@ifisc.uib-csic.es](mailto:afperalta@ifisc.uib-csic.es)

probability that a voter changes its state is proportional to the fraction of neighboring voters in the opposite state. It is clear that in some situations a more complex mechanism of group interaction can be relevant. This collective interaction introduces a nonlinearity in which the probability to change state is proportional to a power  $\alpha$  of the fraction of neighbors in the opposite state. Similar types of nonlinearities, with minor variations in some cases, as the ones introduced here in the noisy voter model have been previously considered in different contexts.<sup>5,15–26</sup> For example, the degree of nonlinearity,  $\alpha$ , is assumed in social impact theory<sup>18,19</sup> to be a positive real number measuring the nonlinear effect of local majorities, while the same parameter is termed “volatility” and is represented by the letter “ $a$ ” in problems of language competition dynamics.<sup>16,17</sup> In the interpretation of Refs. 5, 21, and 23 each individual interacts with a set of  $\alpha$  of its nearest neighbors, and therefore,  $\alpha > 1$  and it takes integer values. In order to allow for a detailed study of the robustness of the transition of the noisy voter model—that is, near  $\alpha = 1$ —we focus here on continuous values of  $\alpha$ .

We show in this paper that the noisy voter model is structurally unstable with respect to non-linearity so that its order-disorder transition becomes now well defined in the thermodynamic limit for any value of  $\alpha > 1$ . The resulting rich phase diagram, with the presence of tricriticality and catastrophe transitions, extends to the case of continuous  $\alpha$  and asymmetry in the transition rates, as some previous results recently presented in the literature<sup>21,23</sup> seem to display. Using analytical techniques and numerical simulations, we carry out a detailed analysis of the dependence with system size of the moments and maxima of the probability distribution of the order parameter. In particular, we focus on the all-to-all interaction and on highly heterogeneous networks of interactions, and we obtain the critical exponents of the finite-size scaling behavior. Furthermore, we prove that the system size scaling of the copying mechanism proposed in Refs. 13 and 14 for the noisy voter model is not applicable for nonlinear interactions. In the context of language competition,<sup>16,17</sup> our study can be understood as an analysis of the robustness of the transition between language coexistence and extinction under noisy perturbations. In this sense, we find that the transition from language extinction to language coexistence at  $\alpha = 1$  occurs now at a value modified by noise. Interestingly, we also find that there is a tricritical point<sup>27</sup> such that for  $\alpha > 5$  there is a range of parameters with a new phase of coexistence of three solutions: extinction of each of the two languages or language coexistence in which, in the absence of stochasticity, the initial condition determines the final state.

The paper is organized as follows: in Section II, we define the model, set the notation, and introduce the non-linear interaction term. In Section III, we solve the model analytically in an all-to-all scenario. This includes the identification of transition lines and the statistical properties of the global state of the system. Section IV deals with the characterization of the model in a network structure using the pair-approximation, which is capable of reproducing the main differences with respect to the all-to-all result. We end with a summary and conclusions in Section V.

## II. MODEL

We consider a population of  $N$  individuals located in the nodes of a given (undirected) network of interactions. Each node holds a time-dependent binary variable  $n_i = 0, 1$ , defining the state of individual  $i = 1, \dots, N$ . The meaning of this binary variable does not concern us in this paper, but typical interpretations include the optimistic/pessimistic state of a stock market broker,<sup>7</sup> the language A/B used by a speaker,<sup>16,17</sup> or the direction of the velocity right/left in a one-dimensional model of active particles.<sup>28</sup> The network of interactions is mapped onto the usual (symmetric) adjacency matrix of coefficients  $A_{ij} = 1$  if nodes  $i$  and  $j$  are connected and  $A_{ij} = 0$  otherwise. The degree of node  $i$  is its total number of connected nodes (also known as neighbors)  $k_i = \sum_j A_{ij}$ . An important characteristic of the network is its degree distribution,  $P_k \equiv N_k/N$ , where  $N_k$  is the number of nodes with degree  $k$ . The  $m$ -moment of the degree distribution is defined as  $\mu_m = \sum_k P_k k^m$ , with short notation  $\mu \equiv \mu_1$ .

The state of a node can change over time following the combination of a purely random, or noisy, effect and a herding, or copying, mechanism. They are both stochastic processes defined as follows:

1. The purely random effect takes into account that individuals can change their state independently of the state of others, with a rate  $a_0$  or  $a_1$ , depending on whether the node holds state 0 or 1, respectively. This process reflects an idiosyncratic or autonomous behavioral rule.
2. The herding mechanism considers that individuals can be persuaded by their neighbors to change their state. Hence, individual  $i$  in state  $n_i = 0$  can change the state to  $n_i = 1$  with a rate  $hf[\sigma_1(i)]$ . Here,  $\sigma_1(i)$  is the fraction of neighbors of  $i$  holding the opposite state 1,  $h$  is a parameter measuring the herding intensity, and  $f$  is a monotonically increasing function that fulfills  $f[0] = 0$  and  $f[1] = 1$ . Similarly, an individual in the state  $n_i = 1$  can change to state  $n_i = 0$  with a rate  $hf[\sigma_0(i)]$ , where  $\sigma_0(i) = 1 - \sigma_1(i)$ .

The total individual transition rates  $\pi_i^\pm$  result from the sum of these two complementary processes

$$\pi_i^+ \equiv \text{rate}(n_i = 0 \rightarrow n_i = 1) = a_0 + hf[\sigma_1(i)], \quad (1)$$

$$\pi_i^- \equiv \text{rate}(n_i = 1 \rightarrow n_i = 0) = a_1 + hf[\sigma_0(i)]. \quad (2)$$

In terms of the the adjacency matrix,  $\sigma_{0/1}(i)$  can be written as follows:

$$\sigma_1(i) = \frac{1}{k_i} \sum_{j=1}^N A_{ij} n_j, \quad \sigma_0(i) = 1 - \sigma_1(i). \quad (3)$$

The function  $f[\sigma]$  depicts the nature of the copying mechanism. For example, the traditional voter model uses a linear dependence  $f[\sigma] = \sigma$  which corresponds to the process in which a node copies the state of a randomly selected neighbor. We will focus, however, on a more general way to model the interactions, considering a non-linear function  $f[\sigma] = \sigma^\alpha$ . For  $\alpha > 1$  individuals are more resistant to follow the opinion

of the neighbors holding the opposite state, while the contrary is true for  $\alpha < 1$ , it is easier to copy the opposite state of a neighbor. Hence, a value of  $\alpha > 1$  (probability of imitation below random) can be representative of a situation of aversion to change, where a larger fraction of neighbors in the opposite state is needed to switch state compared to the purely random imitation of the voter model. In fact, values of  $\alpha > 1$  have been fitted in some problems of language competition.<sup>16</sup> On the other hand, values of  $\alpha < 1$ , corresponding to a probability of imitation above random or a situation of preference for change, have been considered in social impact theory.<sup>18</sup> The cases  $\alpha = 2, 3 \dots$  are equivalent to a process in which an individual changes state if and only if after checking repeatedly the state of  $\alpha$  randomly selected neighbors (and allowing for repetitions in the selection), all of them happen to be in the opposite state to the one held by the individual.<sup>5</sup>

The global state of the system can be characterized by the total number of nodes in state 1,  $n = \sum_{i=1}^N n_i$ , taking integer values  $n \in [0, N]$ , or by the intensive variable (also known as “magnetization”)  $m = 2n/N - 1$  that takes values in the range  $m \in [-1, +1]$ . The values  $m = -1$  and  $m = +1$  correspond to a situation where all the nodes hold the same state, respectively,  $n_i = 0$  and  $n_i = 1$ , while  $m = 0$  is the perfectly balanced case, where half of the nodes are in state  $n_i = 0$  and the other half in  $n_i = 1$ . One of the particularities of the noisy version of the voter model is that for  $a_0 > 0$  the system, even after long enough times, does not get stuck at  $n = 0$  nor at  $n = N$  for  $a_1 > 0$ , the two absorbing states of the noiseless voter model.

The aim of this paper is to study the stationary statistical properties of the magnetization, namely, its moments  $\langle m^k \rangle$  and the location of the maxima of its probability distribution as a function of the parameters  $a_0, a_1, h, \alpha$ , as well as the population size  $N$  and the structure of the network of interactions, with special attention to the role of the non-linear parameter  $\alpha$ . In order to do that, in Section III, we will develop analytical results that will be compared with those coming from numerical simulations of the stochastic process using well-known techniques.<sup>49</sup>

### III. ALL-TO-ALL SOLUTION

In the all-to-all scenario, all the nodes are equivalent, and the only relevant variable is  $n$ . In this sense, we do not consider the individual rates Eqs. (1) and (2) but the global rates  $\pi^\pm(n) \equiv \text{rate}(n \rightarrow n \pm 1)$  at which a change of state in any of the nodes takes place. The individual fraction  $\sigma_1(i)$  is replaced by the global fraction  $n/N$  [while  $\sigma_0(i)$  by  $(N - n)/N$ ] and thus the global rates read as follows:

$$\pi^+(n) = (N - n) \left[ a_0 + h \left( \frac{n}{N} \right)^\alpha \right], \quad (4)$$

$$\pi^-(n) = n \left[ a_1 + h \left( \frac{N - n}{N} \right)^\alpha \right]. \quad (5)$$

The approach is exact only when the network is fully connected, namely, when each node is connected to all the other nodes. The transition rates Eqs. (4) and (5) define a Markovian stochastic process, which is of the one-step type. The process can be fully described by means of the probability  $P(n, t)$  to

have  $n$  individuals in state 1 at time  $t$ , which obeys the master equation:<sup>31</sup>

$$\frac{\partial P(n, t)}{\partial t} = (E^+ - 1) [\pi^-(n)P] + (E^- - 1) [\pi^+(n)P], \quad (6)$$

where we have introduced the step operators  $E^\pm$ , defined to act on an arbitrary function  $g(n)$  as  $E^\pm[g(n)] = g(n \pm 1)$ . One can derive a continuous version of this equation, known as Fokker-Planck equation, through a systematic expansion in  $N$ . Rewriting Eq. (6) in terms of the magnetization  $m$  and expanding it in power series of  $\Delta m = \pm 2/N$  to second order, one finds

$$\frac{\partial P(m, t)}{\partial t} = - \frac{\partial}{\partial m} [F(m)P] + \frac{1}{N} \frac{\partial^2}{\partial m^2} [D(m)P], \quad (7)$$

where  $P(m, t)$  is now a probability density function (pdf), related to the discrete probability function as  $P(n, t) = P(m, t) \cdot |dm/dn| = P(m, t) \cdot 2/N$ . The functions  $F(m) = [\pi^+ - \pi^-]2/N$  and  $D(m) = [\pi^+ + \pi^-]2/N$  are usually called drift and diffusion, respectively, and, for this model, are given by

$$F(m) = a_0 - a_1 - (a_0 + a_1)m + 2^{-\alpha} h(1 - m^2)[(1 + m)^{\alpha-1} - (1 - m)^{\alpha-1}], \quad (8)$$

$$D(m) = a_0 + a_1 - (a_0 - a_1)m + 2^{-\alpha} h(1 - m^2)[(1 + m)^{\alpha-1} + (1 - m)^{\alpha-1}]. \quad (9)$$

It is also possible to describe the evolution of the system in terms of a Langevin stochastic differential equation for the trajectories  $m(t)$ , which, within the Itô convention, reads

$$\frac{dm(t)}{dt} = F(m) + N^{-1/2} \sqrt{2D(m)} \cdot \xi(t), \quad (10)$$

where  $\xi(t)$  is a Gaussian white noise with zero mean  $\langle \xi(t) \rangle = 0$  and correlations  $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$ .

### A. Modes and transitions

In general, it is difficult to find an analytical solution of the master equation (6) or the Fokker-Planck equation (7). There is, however, a general solution  $P_{\text{st}}(m)$  of the latter<sup>50</sup> in the stationary state ( $\partial/\partial t P(m, t) = 0$ ) that can be written in the exponential form

$$P_{\text{st}}(m) = \mathcal{Z}^{-1} \cdot \exp[-NV(m)], \quad (11)$$

$$V(m) = \int^m \frac{-F(z) + D'(z)/N}{D(z)} dz, \quad (12)$$

where  $V(m)$  is called the “potential function” and  $\mathcal{Z}$  is a normalization constant.

Although the solution Eqs. (11) and (12) is a complicated expression given the drift and diffusion functions Eqs. (8) and (9), it is possible to portray the shape of the stationary distribution  $P_{\text{st}}(m)$ . In particular, in this subsection, we will focus on characterizing the modes, or values of  $m_*$  for which  $P_{\text{st}}(m_*)$  is a maximum, and the transitions where modes appear or disappear as a function of the parameters of the model.

The condition of local maximum for the mode is  $P'_{\text{st}}(m_*) = 0$  (condition of local extreme) and  $P''_{\text{st}}(m_*) < 0$ , but

note that modes can also be located at the boundary values  $m_* = \pm 1$ . In the symmetric case  $a_0 = a_1$ , there is a trivial extreme  $m_* = 0$  which corresponds to the perfectly balanced case. Its dynamical stability under perturbations comes determined by the second derivative of  $P_{st}(m)$  or, alternatively, of  $V(m)$

$$V''(m_* = 0) = \frac{2^\alpha}{2^\alpha \varepsilon + 1} [\varepsilon - \varepsilon_c(N)], \quad (13)$$

where we define the noise-herding intensity ratio parameter  $\varepsilon \equiv (a_0 + a_1)/2h$ , and  $\varepsilon_c(N)$  reads

$$\varepsilon_c(N) = 2^{-\alpha} \left( \alpha - 1 + \frac{\alpha(3 - \alpha)}{N} \right), \quad (14)$$

$$\varepsilon_c \equiv \varepsilon_c(\infty) = 2^{-\alpha} (\alpha - 1). \quad (15)$$

Then, for  $\varepsilon < \varepsilon_c(N)$ ,  $m_* = 0$  corresponds to a minimum of the probability distribution, while for  $\varepsilon > \varepsilon_c(N)$  to a maximum. Note that in the thermodynamic limit  $N \rightarrow \infty$ , the transition between maximum and minimum occurs at a finite positive value of  $\varepsilon$  for  $\alpha > 1$ , while for  $\alpha < 1$  there is no transition since  $\varepsilon_c(\infty) < 0$  and  $m_* = 0$  is always a stable solution. For the particular value  $\alpha = 1$ , the linear model, the transition is only a finite-size effect since  $\varepsilon_c(N) = 1/N$  and, in the thermodynamic limit, the maximum is located at  $m_* = 0$  for all values of  $\varepsilon > 0$ .

Consequently, only in the non-linear regime  $\alpha > 1$ , the parameters  $a_0$ ,  $a_1$ , and  $h$  take values of the same order near the transition, and we can safely disregard the term  $D'(z)/N$  in front of  $F(z)$  in the numerator of Eq. (12), in the limit of large  $N$ . In this limit, that we assume in the remaining of this subsection, the extrema of the distribution fulfill the condition  $F(m_*) = 0$  while the stability comes determined by the sign of the derivative  $F'(m_*)$ . Since it is not possible to find a closed solution of this equation, we will expand the drift function in power-series to  $O(m^7)$

$$\begin{aligned} F(m) = & h\Delta\varepsilon + 2h(\varepsilon_c - \varepsilon)m + \frac{h}{3}2^{-\alpha}(\alpha - 5)(\alpha - 1)\alpha m^3 \\ & + \frac{h}{15}2^{-2-\alpha}(\alpha - 9)(\alpha - 3)(\alpha - 2)(\alpha - 1)\alpha m^5 \\ & + O(m^7), \end{aligned} \quad (16)$$

where we define the parameter  $\Delta\varepsilon = (a_0 - a_1)/h$  as a measure of the asymmetry in the rates for switching states  $0 \rightarrow 1$  and  $1 \rightarrow 0$ . This expression coincides with the derivative of the potential function found by Vazquez *et al.*<sup>17</sup> in their study of the noiseless version ( $\varepsilon = \Delta\varepsilon = 0$ ) of the nonlinear model in the unbiased case.<sup>51</sup> In principle, the solution (16) using the power-series expansion will be valid only for  $m_* \approx 0$ , but it is enough to identify additional extrema and transitions. We now find the extrema and their stability distinguishing between the symmetric  $a_0 = a_1$  and asymmetric  $a_0 \neq a_1$  cases.

### 1. Symmetric case

Using the expansion Eq. (16), there are five (maybe complex) solutions of  $F(m_*) = 0$ . When  $a_0 = a_1$ , they are as follows: the trivial one  $m_* = 0$  and other four roots  $\pm m_*^\pm$ ,

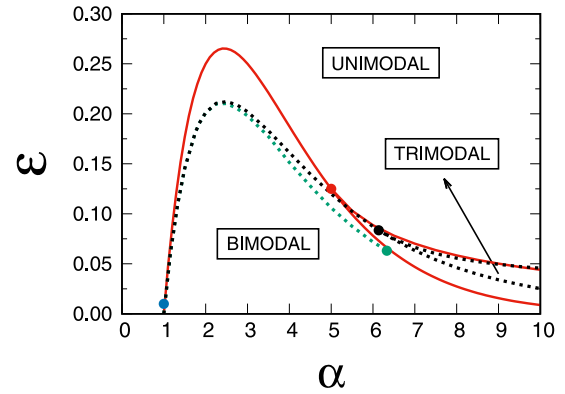


FIG. 1. Phase diagram of the symmetric case  $\Delta\varepsilon = 0$ . The dot (blue) at  $\alpha = 1$  is the finite-size transition point  $\varepsilon_c(N)$ , Eq. (14) for  $N = 100$ . The solid line corresponds to the all-to-all setting, while the dotted line to a 15-regular network (black) and a scale free network  $P_{5 \leq k \leq 966} \sim k^{-2.34}$  with  $\mu \approx 15$  (green), see Section IV. The trimodal region is delimited by the transition lines  $\varepsilon_c$  and  $\varepsilon_t$ , which for the all-to-all network correspond to expressions Eqs. (14) and (18), with a tricritical point (red) at  $\alpha = 5$ ,  $\varepsilon = 1/8$ . The tricritical point for the 15-regular network (black) is at  $\alpha = 6.14$ ,  $\varepsilon = 0.084$  and for the scale free network (green) at  $\alpha = 6.33$ ,  $\varepsilon = 0.063$  (the trimodal region of the scale free network is removed for clarity in the figure).

$\pm m_*^\pm$  obtained from

$$\begin{aligned} (m_*^\pm)^2 \approx & \frac{10(\alpha - 5)}{(\alpha - 2)(\alpha - 3)(\alpha - 9)}, \quad (17) \\ & \left[ \pm \sqrt{1 - \frac{6(\alpha - 2)(\alpha - 3)(9 - \alpha)}{\alpha(\alpha - 5)^2} \left( \frac{\varepsilon - \varepsilon_c}{\varepsilon_c} \right)} - 1 \right]. \end{aligned}$$

The acceptable maxima of the probability distribution  $P_{st}(m)$  must correspond to values of  $m_*$  which are real and inside the interval  $[-1, 1]$ . The four roots given by Eq. (17) are real or imaginary depending on the values of the parameters  $\varepsilon$  and  $\alpha$ :

1. In the range  $1 < \alpha < 5$ , for  $\varepsilon < \varepsilon_c$  the pair of solutions  $\pm m_*^+$  are real and correspond to probability maxima, while for  $\varepsilon > \varepsilon_c$  all four roots are imaginary. Then, in this range, the line  $\varepsilon_c$  divides unimodal (a single maximum of the probability distribution) from bimodal (two real maxima) regimes, see Fig. 1 and top panels of Fig. 2.
2. In the range  $\alpha > 5$ , for  $\varepsilon < \varepsilon_c$  the pair of solutions  $\pm m_*^+$  are real and correspond to probability maxima; for  $\varepsilon_c < \varepsilon < \varepsilon_t$  all four roots are real,  $\pm m_*^-$  correspond to probability maxima, and  $\pm m_*^+$  to minima; while for  $\varepsilon > \varepsilon_t$  all four are imaginary, where  $\varepsilon_t$  reads

$$\varepsilon_t \approx \varepsilon_c \cdot \left( 1 + \frac{5\alpha(\alpha - 5)^2}{6(\alpha - 2)(\alpha - 3)(9 - \alpha)} \right). \quad (18)$$

Consequently, for  $\alpha > 5$ , the line  $\varepsilon_c$  divides bimodal from trimodal regimes, while  $\varepsilon_t$  divides trimodal from unimodal regimes, see Fig. 1 and the top panels of Fig. 3. The two lines meet at the tricritical point  $\alpha = 5$  and  $\varepsilon = 2^{-3} = 0.125$ .

Note that the expressions for the roots Eq. (17) and the transition line Eq. (18) are based on the expansion Eq. (16), and consequently, they are an approximation, hence the  $\approx$  symbol used in those formulas. In fact, we have found that



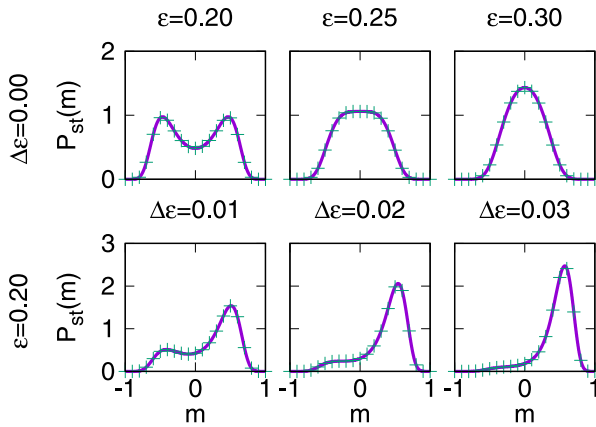


FIG. 2. Probability distribution  $P_{st}(m)$ , for different values of the parameters with fixed  $\alpha = 2$ . In the symmetric case  $\Delta\epsilon = 0$  (top panels), the transition point is at  $\epsilon_c = 0.25$  where the distribution switches from symmetric bimodal to unimodal. In the bottom panels, we identify the transition from an asymmetric bimodal (with two maxima, one of them absolute and the other local) to a unimodal distribution occurring at a value of the asymmetry parameter  $\Delta\epsilon_a = 0.017$  for  $\epsilon = 0.2$  (see the corresponding points in Fig. 5). Dots correspond to numerical simulations of the process defined by the rates Eqs. (4) and (5), while lines are the function Eq. (11).

they remain accurate only for  $\alpha < 7$ . Despite that the classification of maxima/minima is completely general and additional extrema or transitions are not observed in a numerical analysis using the exact expression for the potential function Eqs. (11), (8), and (9).

It is also important to mention that for  $1 < \alpha < 5$ , the line  $\epsilon = \epsilon_c$  corresponds to a classical (Landau) second order phase transition from the “ordered” phase  $m_* \neq 0$  to the “disordered” one<sup>52</sup>  $m_* = 0$ , with scaling  $|m_*| \sim (\epsilon_c - \epsilon)^\beta$  for  $\epsilon \leq \epsilon_c$  with  $\beta = 1/2$  and  $m_* = 0$  for  $\epsilon \geq \epsilon_c$ . This line is delimited by two special degenerate points,  $\alpha = 1$  and  $\alpha = 5$ :

- For  $\alpha = 1$ , all the terms of the expansion Eq. (16) vanish and  $P_{st}(m)$  is completely flat at the transition. It is only at  $\epsilon = \epsilon_c = 0$  (noiseless voter model) that the probability distribution is a sum of two delta-functions at  $m = \pm 1$ , leading to  $m_* = \pm 1$  if  $\epsilon = 0$  or  $m_* = 0$  if  $\epsilon > 0$ . This discontinuity

survives in the finite  $N$  limit, where the result is  $m_* = \pm 1$  if  $\epsilon < \epsilon_c(N)$  or  $m_* = 0$  if  $\epsilon > \epsilon_c(N)$ . Formally, we still have a scaling of the form  $|m_*| \sim (\epsilon_c(N) - \epsilon)^\beta$ , but now  $\beta = 0$ .

- For  $\alpha = 5$ , the first three terms of the expansion vanish and  $P_{st}(m)$  is flatter than for the classical second order transition, with a scaling  $|m_*| \sim (\epsilon_c - \epsilon)^\beta$ ,  $\beta = 1/4$ .

The condition that isolates the classical second order line from the degenerate points is that we can disregard the  $O(m^5)$  term of Eq. (16) in front of the  $O(m^3)$ , which leads to

$$|\epsilon - \epsilon_c| \ll \epsilon_c \cdot \left| \frac{\alpha(\alpha - 5)^2}{(\alpha - 2)(\alpha - 3)(9 - \alpha)} \right|. \quad (19)$$

For  $\alpha > 5$  an increase of  $\epsilon$  leads first to a transition from a bimodal to a trimodal distribution at  $\epsilon = \epsilon_c$ , implying a discontinuity of the location of the absolute maximum from  $|m_*| > 0$  to  $m_* = 0$ . This corresponds to a first-order transition. The local maxima at  $|m_*| > 0$  remain up to  $\epsilon = \epsilon_t$  where they disappear and the distribution is unimodal. The absolute maximum of the distribution is at  $|m_*| > 0$  for  $\epsilon < \epsilon_M$  and at  $m_* = 0$  for  $\epsilon > \epsilon_M$ , being  $\epsilon_M$  the Maxwell point where the potential takes the same value at the maxima,  $|m_*| > 0$  and  $m_* = 0$ .

A summary of the equation of state  $m(\epsilon)$  for different values of  $\alpha$  in the symmetric case is presented in Fig. 4.

## 2. Asymmetric case

When  $a_0 \neq a_1$ , there are two main effects: (i) the locations of the modes shift and (ii) one of the members of the previous symmetrical solution pair  $\pm m_*^\pm$  is promoted to be more likely than its corresponding partner. If the asymmetry parameter  $|\Delta\epsilon|$  is large enough, bimodality and trimodality may even be destroyed, inducing new transitions in the  $(a_0, a_1, \alpha)$  or, alternatively, in the  $(\epsilon, \Delta\epsilon, \alpha)$  space of parameters. These transitions are known as catastrophes.<sup>32,33</sup> For fixed  $\alpha$ , the projection  $(\epsilon_a, \Delta\epsilon_a)$  of the transition region in the plane  $(\epsilon, \Delta\epsilon)$  must fulfill the transition conditions  $F(m_*) = 0$  and  $F'(m_*) = 0$ , which in the approximation Eq. (16) and in

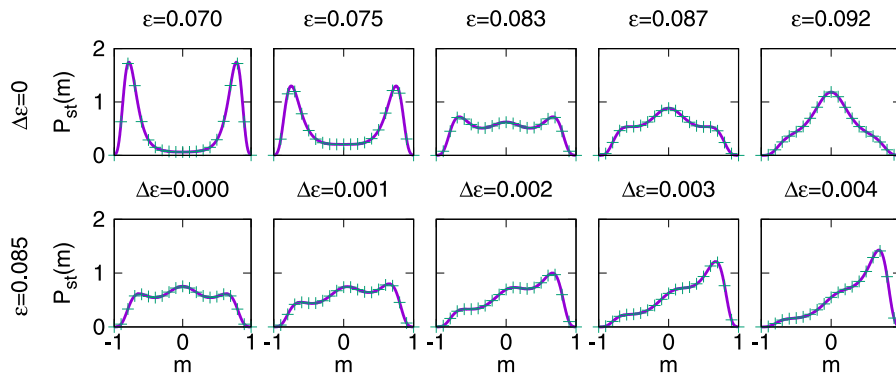


FIG. 3. Probability distribution  $P_{st}(m)$ , for different values of the parameters with fixed  $\alpha = 6$ . In the symmetric case  $\Delta\epsilon = 0$  (top panels), the transition points are  $\epsilon_c = 0.075$  (from bimodal to trimodal) and  $\epsilon_t = 0.086$  (from trimodal to unimodal). In the bottom panels, we fix  $\epsilon = 0.085$  and identify the transition from a trimodal to a bimodal distribution occurring at a value of the asymmetry parameter  $\Delta\epsilon_{a,1} = 0.00082 \dots$ , and the transition from bimodal to unimodal occurring at  $\Delta\epsilon_{a,2} = 0.00280 \dots$  (see the corresponding points in Fig. 5). Dots correspond to numerical simulations of the process defined by the rates Eqs. (4) and (5), while lines are the function Eq. (11).

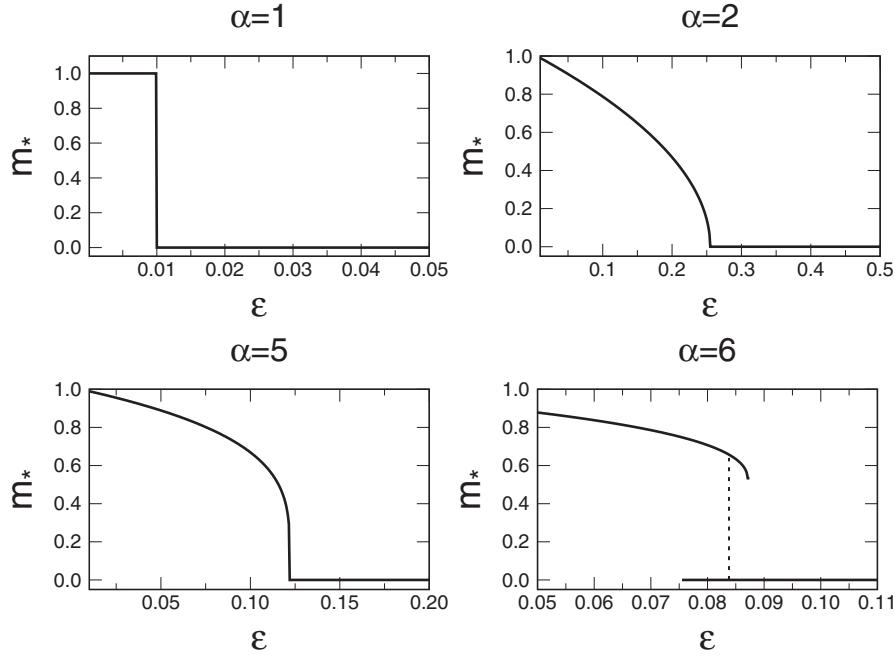


FIG. 4. Summary of the equation of state for different values of  $\alpha$  in the symmetric case  $\Delta\epsilon = 0$ . The lines come from the determination of the maxima  $m_*$  of the exact potential Eq. (12) for  $N = 100$ . The vertical dashed line for  $\alpha = 6$  is  $\epsilon_M = 0.0838$ , the Maxwell point where the potential takes the same value at the two maxima,  $m_* > 0$  and  $m_* = 0$ .

parametric representation read

$$\frac{\epsilon_a}{\epsilon_c} \approx 1 + \frac{\alpha(\alpha-5)}{2} m_*^2 + \frac{\alpha(\alpha-2)(\alpha-3)(\alpha-9)}{24} m_*^4, \quad (20)$$

$$\frac{\Delta\epsilon_a}{\epsilon_c} \approx \frac{2\alpha(\alpha-5)}{3} m_*^3 + \frac{\alpha(\alpha-2)(\alpha-3)(\alpha-9)}{15} m_*^5.$$

The transition line in the plane  $(\epsilon_a, \Delta\epsilon_a)$  is obtained by elimination of the parameter  $m_*$ . As indicated in Fig. 5, there could be several transition lines.

For  $1 < \alpha < 5$ , and when the condition Eq. (19) is fulfilled, we can disregard the terms of order  $O(m^4)$  or higher in Eqs. (20). The resulting transition line corresponds to the well known *Cusp Catastrophe*<sup>33</sup> with geometry  $\epsilon_c - \epsilon_a \sim |\Delta\epsilon_a|^{2/3}$  (see the left panel of Fig. 5). The degenerate points  $\alpha = 1, 5$  also display cusp-like transition curves with different geometries:  $\epsilon_c - \epsilon_a \sim |\Delta\epsilon_a|^{4/5}$  for  $\alpha = 5$  and  $\epsilon_c - \epsilon_a \sim |\Delta\epsilon_a|$  for  $\alpha = 1$ . In the bottom panels of Fig. 2, we plot the probability distribution for  $\alpha = 2$ ,  $\epsilon = 0.2$  as the parameter  $\Delta\epsilon$  crosses the transition line at  $\Delta\epsilon_a = 0.017$  depicted in Fig. 5.

For  $\alpha > 5$  one has to keep all the terms of Eqs. (20). In this case, trimodality, together with bimodality and unimodality, is also possible and the transition curves correspond to the so-called *Butterfly Catastrophe*.<sup>33</sup> Its geometry is similar to three cusps, whose intersections delimit the trimodal regime (see the right panel of Fig. 5, corresponding to  $\alpha = 6$ ). The two tips at the top divide unimodal from bimodal zones, while the tip at the bottom separates trimodal from bimodal. In the bottom panel of Fig. 2, we plot the probability distribution for  $\alpha = 6$ ,  $\epsilon = 0.085$  as the parameter  $\Delta\epsilon$  crosses the transition lines at  $\Delta\epsilon_{a,1} = 0.00082\dots$  and  $\Delta\epsilon_{a,2} = 0.00280\dots$ , as depicted in Fig. 5.

## B. Fluctuations and $N$ -dependence

Fluctuations, including those due to the finite-size of the system, play a very important role in the original version of the noisy voter model,  $\alpha = 1$ . This is mainly because, as we mentioned earlier, the transition is noise-induced and finite-size. In this case, with  $\alpha = 1$ , and if we approach the critical point  $\epsilon_c(N) = 1/N$  by keeping  $\epsilon \sim 1/N$  and  $\Delta\epsilon \sim 1/N$ , the expression Eq. (12) for the potential can be simplified to

$$V(m) = \left(\frac{1}{N} - \epsilon\right) \log(1 - m^2) + \frac{\Delta\epsilon}{2} \log\left(\frac{1 - m}{1 + m}\right). \quad (21)$$

If we conveniently redefine the rates Eqs. (4) and (5) such that we are always in the regime  $\epsilon \sim 1/N$  and  $\Delta\epsilon \sim 1/N$  (for example, considering  $h/N \equiv h^*$  with  $h^* = O(1)$  being a new parameter), then the distribution  $P_{st}(m)$  is  $N$ -independent, which is one of the interesting properties of the model with rescaled parameters widely discussed in the literature.<sup>13,14,34</sup> We will, however, show that this is not the case for the non-linear version of the model and that the  $N$ -independence property is not a robust result of the noisy voter model.

Let us expand the potential Eq. (12) in power series of  $m$ , around the critical point  $\epsilon \sim \epsilon_c$ ,  $\Delta\epsilon \sim 0$ , to find

$$V(m) = -c_1 m + c_2 m^2 + c_4 m^4 + c_6 m^6 + \dots, \quad (22)$$

with coefficients

$$c_1 = \frac{2^\alpha}{2^\alpha} \Delta\epsilon \equiv c \cdot \Delta\epsilon, \quad (23)$$

$$c_2 = \frac{2^\alpha}{2^\alpha} (\epsilon - \epsilon_c) \equiv c \cdot (\epsilon - \epsilon_c), \quad (24)$$

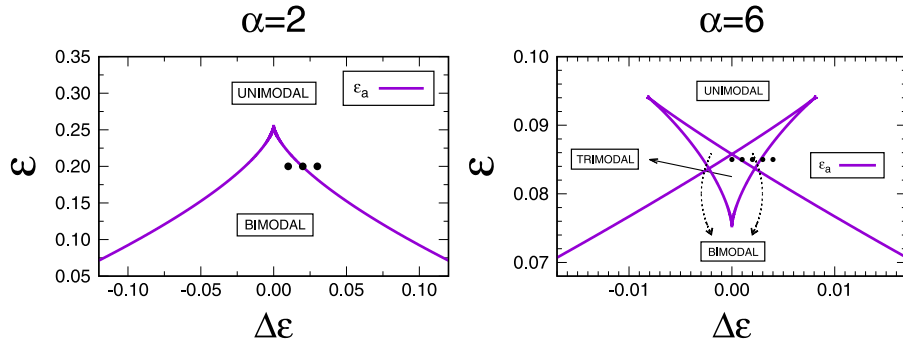


FIG. 5. Phase diagram of the asymmetric case  $\Delta\epsilon \neq 0$ . The left panel corresponds to  $\alpha = 2$  and the right one to  $\alpha = 6$ . The transition lines correspond to the expressions Eqs. (20). The dots are the points whose probability distributions are displayed in the bottom panels of Figs. 2 and 3.

$$c_4 = \frac{(\alpha - 1)(5 - \alpha)}{24}, \quad (25)$$

$$c_6 = \frac{(\alpha - 1)(\alpha - 3)(\alpha - 4)(17 - \alpha)}{720}. \quad (26)$$

Note that we have only kept the first order of each coefficient and only some of the coefficients. In doing so, we have neglected terms of order  $1/N$ ,  $m^2(\epsilon - \epsilon_c)^2$ ,  $m^2\Delta\epsilon^2$ ,  $m^3\Delta\epsilon$ , etc. If we now assume that  $1 < \alpha < 5$ , then we can also neglect the order  $m^6$  of the potential, and an arbitrary  $k$ -moment of the distribution  $P_{st}(m)$  will read

$$\langle m^k \rangle_{st} = \frac{\int_{-1}^1 dm m^k e^{N(c_1 m - c_2 m^2 - c_4 m^4)}}{\int_{-1}^1 dm e^{N(c_1 m - c_2 m^2 - c_4 m^4)}}. \quad (27)$$

If we make the change of variables  $z = N^{1/4}m$ , the leading term can be rewritten as follows:

$$\langle m^k \rangle_{st} = N^{-k/4} \phi_k [N^{1/2}(\epsilon - \epsilon_c), N^{3/4}\Delta\epsilon], \quad (28)$$

with the scaling function

$$\phi_k[x, y] = \frac{\int_{-\infty}^{\infty} dz z^k e^{cyz - cxz^2 - c_4 z^4}}{\int_{-\infty}^{\infty} dz e^{cyz - cxz^2 - c_4 z^4}}. \quad (29)$$

Note that this function  $\phi_k[x, y]$  is independent of  $N$ , and all the  $N$ -dependence is displayed explicitly in the expression Eq. (28). The case  $\alpha = 5$  is a degenerate case where  $c_4 = 0$ , and

we have to keep the order  $m^6$  of the potential, this leads to a different scaling function:

$$\langle m^k \rangle_{st} = N^{-k/6} \phi_k [N^{2/3}(\epsilon - \epsilon_c), N^{5/6}\Delta\epsilon], \quad (30)$$

with

$$\phi_k[x, y] = \frac{\int_{-\infty}^{\infty} dz z^k e^{cyz - cxz^2 - c_6 z^6}}{\int_{-\infty}^{\infty} dz e^{cyz - cxz^2 - c_6 z^6}}. \quad (31)$$

If we write the special case  $\alpha = 1$  in this form, we obtain  $\langle m^k \rangle = \phi_k[N\epsilon, N\Delta\epsilon]$ , where now the scaling function uses the expression Eq. (21) of the potential with all the orders of  $m$  included. In this case of  $\alpha = 1$ , a rescaling of the parameters of the model  $\epsilon$ ,  $\Delta\epsilon$  in an appropriate way, will make all the moments  $\langle m^k \rangle$   $N$ -independent. However, for  $\alpha \neq 1$ , even if we rescale the parameters of the model to be close enough to the critical point, the moments  $\langle m^k \rangle$  vanish as  $N^{-k/4}$  for  $1 < \alpha < 5$  and as  $N^{-k/6}$  for  $\alpha = 5$ . It is also important to mention that the width of the critical region where these scaling properties hold is wider for the non-linear case  $\epsilon - \epsilon_c \sim N^{-1/2}$  (for  $1 < \alpha < 5$ ) compared to  $\epsilon - \epsilon_c \sim N^{-1}$  for  $\alpha = 1$ . In Fig. 6, we check the scaling form of the stationary variance of the magnetization  $\sigma_{st}^2[m]$  in the cases  $\alpha = 2$  and  $\alpha = 5$ .

#### IV. COMPLEX NETWORKS

So far, we have considered the case of an all-to-all interaction, in which each node is a neighbor of every other node. Given the application of the model to problems of consensus formation in populations, it seems important to consider in

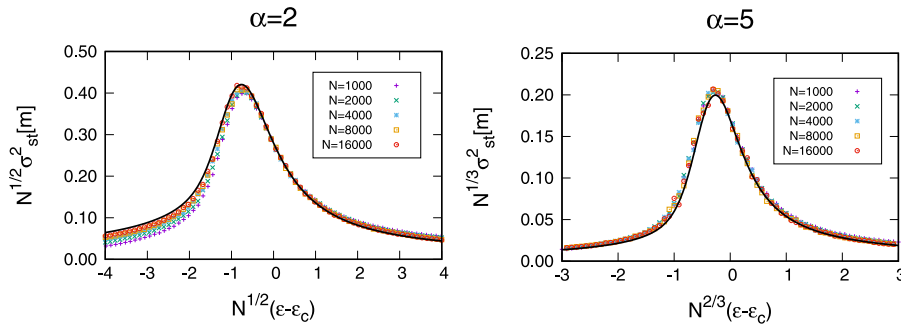


FIG. 6. Variance of the magnetization  $\sigma_{st}^2[m] \equiv \langle m^2 \rangle_{st} - \langle m \rangle_{st}^2$  as a function of  $\epsilon - \epsilon_c$ , rescaled with the correct power of system size  $N$ , in order to check the scaling relations Eqs. (28) and (30). Both panels are for the symmetric case  $\Delta\epsilon = 0$ , the left one with  $\alpha = 2$  and the right one with  $\alpha = 5$ . Dots correspond to Monte Carlo simulations of different system sizes, averaged over  $10^5$  Monte Carlo steps, while the solid black line is the scaling function Eqs. (29) and (31).

detail the role the network of interactions might have on the transitions identified in Section III. In this way, we are able to go beyond the global coupling approach and consider in some detail the effect of the locality of the interactions.

There are several theoretical approaches under which the network structure can be considered.<sup>34–40</sup> In this work, we focus on a particular version of the so-called *pair approximation* as developed in Refs. 17 and 41. In this approach, one defines the network related variables  $\mathbf{n}_k$  and  $\rho$  as, respectively, the number of nodes with degree  $k$  in state 1 (not to be confused with  $n_i$ , which is the node state variable) and the density of active links, i.e., the ratio between the number of links connecting nodes in different states and the total number of links  $\mu N/2$ . Additionally, it is also convenient to define the variable  $n_L = \sum_{i=1}^N k_i n_i$ , which is the number of links coming out of nodes in state 1. The intensive versions of the variables  $\mathbf{n}_k$  and  $n_L$  are  $m_k = 2\mathbf{n}_k/N_k - 1$  and  $m_L = 2n_L/N\mu - 1$  (also known as *link magnetization*). We follow closely the method of Vazquez and Eguiluz<sup>41</sup> in their study of the linear noiseless voter model adapted to the presence of the noise term and non-linear rates of our model, and we refer to that work for a more detailed explanation of the approach. Similar, but not identical, treatments have been developed by Diakonova *et al.*<sup>12</sup> in their study of the linear noisy voter model, and by Jędrzejewski<sup>23</sup> and by Min and San Miguel<sup>20</sup> in particular versions of nonlinear voter models.

For simplicity, we will restrict ourselves to the symmetric case  $\Delta\varepsilon = 0$  and  $a_0 = a_1 \equiv a$ , such that  $m_k = m_L = 0$  is always a fixed point of the dynamics. We start by finding an approximate mean-field time evolution equation for  $\rho$ :

$$\begin{aligned} \frac{d\rho}{dt} = & \sum_k \frac{P_k}{2} \sum_{q=0}^k P_0(k, q) \left[ a + h \left( \frac{q}{k} \right)^\alpha \right] \frac{\Delta\rho_{k,q}}{\Delta t} \\ & - \sum_k \frac{P_k}{2} \sum_{q=0}^k P_1(k, q) \left[ a + h \left( \frac{k-q}{k} \right)^\alpha \right] \frac{\Delta\rho_{k,q}}{\Delta t}. \end{aligned} \quad (32)$$

Here,  $P_{0/1}(k, q)$  is the probability of selecting a node within the population in state 0/1 with degree  $k$  and that has  $q$  neighbors in state 1,  $\Delta\rho_{k,q} = 2(k-2q)/\mu N$  is the change in  $\rho$  when a node with  $(k, q)$  changes from 0 to 1 [being  $-\Delta\rho_{k,q}$  the corresponding change in  $\rho$  when a  $(k, q)$  node changes from 1 to 0], and  $\Delta t = 1/N$  is the elementary time increment in Monte Carlo steps.<sup>49</sup> In this way, the right-hand side of Eq. (32) is simply the probability of selecting nodes  $(0/1, k, q)$ , times the rate at which those nodes change state, times the corresponding change in  $\rho$ , summed over all the possible values of  $k, q$ . The so-called pair approximation assumes  $P_{0/1}(k, q)$  to be binomial, with a single event probability  $p_0 = \rho$  and  $p_1 = 1 - \rho$ , i.e.,  $P_0(k, q) \approx \binom{k}{q} \rho^q (1 - \rho)^{k-q}$ ,  $P_1(k, q) \approx \binom{k}{q} (1 - \rho)^q \rho^{k-q}$ . Under this approximation, Eq. (32) reduces to

$$\frac{d\rho}{dt} = 2a(1 - 2\rho) + \frac{2h}{\mu} \sum_k P_k \left\langle (k - 2q) \left( \frac{q}{k} \right)^\alpha \right\rangle_0, \quad (33)$$

where  $\langle \dots \rangle_{0/1}$  is an average over  $P_{0/1}(k, q)$ . In the case of directed networks,<sup>29</sup> the density of  $0 \rightarrow 1$  links  $\rho_{01}$  is not the

same as the density of  $1 \rightarrow 0$  links  $\rho_{10}$ , and one would need to obtain an evolution equation for each. While such a generalization is beyond the scope of this contribution, we point the interested reader to Ref. 53, where the necessary steps are described.

Note that Eq. (33) is applicable for an arbitrary  $\alpha$  and that, in particular, for integer  $\alpha$ , the right-hand side contains powers of  $\rho$  up to order  $\alpha + 1$ . It also depends on the mean degree  $\mu$  and the first negative moments of the degree distribution  $\mu_{-1}, \dots, \mu_{-\alpha+1}$ , e.g., for  $\alpha = 1$  it is<sup>12</sup>

$$\frac{d\rho}{dt} = -4\rho^2 \frac{\mu - 1}{\mu} + 2\rho \left( \frac{\mu - 2}{\mu} - 2a \right) + 2a, \quad (34)$$

and for  $\alpha = 2$ ,

$$\frac{d\rho}{dt} = 2a(1 - 2\rho) + \frac{2h\rho}{\mu} (\omega_1 + \omega_2\rho - 2\omega_3\rho^2), \quad (35)$$

with  $\omega_1 = 1 - 2\mu_{-1}$ ,  $\omega_2 = \mu - 7 + 6\mu_{-1}$ , and  $\omega_3 = \mu - 3 + 2\mu_{-1}$  (as no node of the network is allowed to be isolated, the degree distribution satisfies  $P_0 = 0$  and the  $m$ -moment  $\mu_m$  is defined also for negative  $m$ , as required in the previous formulas). Although we have not been able to provide a rigorous proof, for all analyzed cases it turns out that the stationary state  $\frac{d\rho}{dt} = 0$  admits a physical (stable) solution  $\rho = \xi \in [0, 1]$ . For  $\alpha = 2$ ,  $\xi$  is the root of the cubic equation

$$\varepsilon(1 - 2\xi) + \frac{\xi}{\mu} (\omega_1 + \omega_2\xi - 2\omega_3\xi^2) = 0, \quad (36)$$

satisfying  $\xi \in [0, 1]$ . See Fig. 7 for an example of the dependence of  $\xi$  with  $\varepsilon = a/h$ . It is possible to recover the all-to-all limit by taking a degree distribution  $P_k = \delta(k - \mu)$  such that all averages can be written as  $\mu_m = \mu^m$ , and taking the limit  $\mu \rightarrow \infty$ . In this limiting case, it is found that  $\xi = 1/2$ .

Next, we derive a master equation for the link-magnetization variable  $m_L$ . A single update  $0 \rightarrow 1$  of a node with degree  $k$  implies a change  $m_L \rightarrow m_L + \Delta_k$ , with  $\Delta_k = \frac{2k}{\mu N}$ , whereas an update  $1 \rightarrow 0$  of a node with degree  $k$  changes

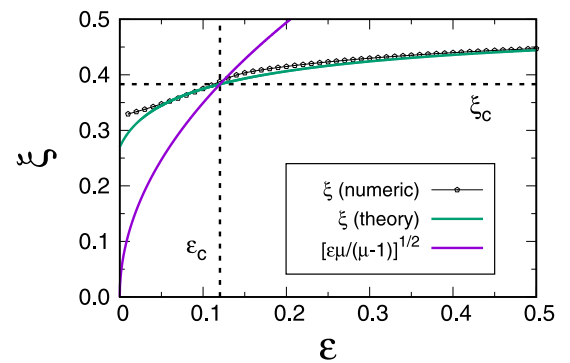


FIG. 7. Stationary solution  $\rho = \xi$  from Eq. (36), corresponding to  $\alpha = 2$ , as a function of  $\varepsilon$  for  $\mu = 5.54$ ,  $\mu_{-1} = 0.35$  and  $\omega_1 = 0.30$ ,  $\omega_2 = 0.64$ ,  $\omega_3 = 3.24$  (green line). The black line-point is the numerical results of  $\xi(\text{numeric}) \equiv \langle \rho \rangle_{\text{st}} / \langle 1 - m^2 \rangle_{\text{st}}$  for an ensemble of 100 scale free networks with the specified degree moments. In purple  $\sqrt{\frac{\varepsilon\mu}{\mu-1}}$ , the crossing of the two curves is the critical point  $\varepsilon_c = 0.120$ ,  $\xi_c = 0.383$ .



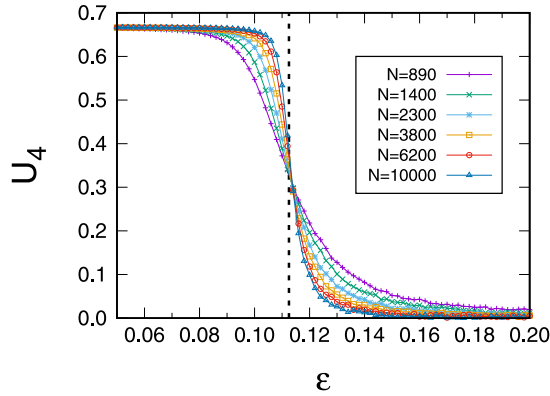


FIG. 8. Binder cumulant as a function of  $\varepsilon$  for a 5-regular network. Dots (joined by lines as a guide to the eye) correspond to Monte Carlo simulations of different system sizes, averaged over  $10^6$  Monte Carlo steps, while the vertical dashed black line is the prediction  $\varepsilon_c = 0.1125$  of the pair-approximation.

$m_L \rightarrow m_L - \Delta_k$ . The master equation for  $P(m_L, t)$  then reads

$$\frac{\partial P(m_L, t)}{\partial t} = \sum_k (E_{m_L}^{\Delta_k} - 1) [\pi_k^- P] + (E_{m_L}^{-\Delta_k} - 1) [\pi_k^+ P], \quad (37)$$

where  $E_{m_L}^\ell$  is the step operator acting on an arbitrary function  $g(m_L)$  as  $E_{m_L}^\ell[g(m_L)] = g(m_L + \ell)$ , and  $\pi_k^\pm$  are the rates of the proposed processes in the  $m_L$  variable:

$$\pi_k^+ = (N_k - \mathbf{n}_k) \left[ a + h \left\langle \left( \frac{q}{k} \right)^\alpha \right\rangle_0 \right], \quad (38)$$

$$\pi_k^- = \mathbf{n}_k \left[ a + h \left\langle \left( \frac{k-q}{k} \right)^\alpha \right\rangle_1 \right]. \quad (39)$$

The single event probabilities  $p_{0/1}$  are different as  $m_L \neq 0$  during the dynamical evolution. They are calculated as the ratio between the number of active links and the number of links coming out from nodes in state 0/1, i.e.,  $p_0 = \frac{\rho}{1-m_L}$  and  $1-p_1 = \frac{\rho}{1+m_L}$ . If we expand the master equation in powers

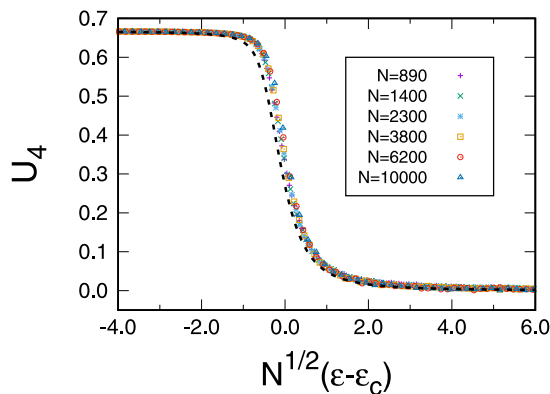


FIG. 9. The same data of Fig. 8 are plotted versus the rescaled variable  $N^{1/2}(\varepsilon - \varepsilon_c)$ . The good collapse of the data validates the proposed scaling law  $U_4(\varepsilon, N) = u(N^{1/2}(\varepsilon - \varepsilon_c))$ , being  $u(x)$  the scaling function. The theoretical prediction (dotted line) is obtained from Eq. (29) using the coefficients  $c = 1.602$  and  $c_4 = 0.093$ .

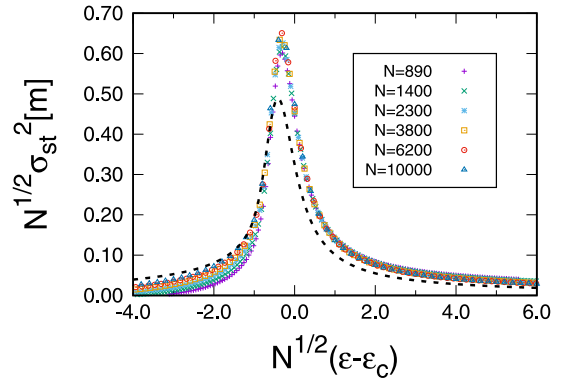


FIG. 10. Rescaled variance of the magnetization as a function of  $N^{1/2}(\varepsilon - \varepsilon_c)$  for a 5-regular network. Dots correspond to Monte Carlo simulations of different system sizes, averaged over  $10^6$  Monte Carlo steps, while the dashed black line is the prediction of the pair-approximation.

of  $\Delta_k$  to second order, we find

$$\frac{\partial P(m_L, t)}{\partial t} = -\frac{\partial}{\partial m_L} [F_L(m_L)P] + \frac{1}{N} \frac{\partial^2}{\partial m_L^2} [D_L(m_L)P], \quad (40)$$

with drift  $F_L = \sum_k \frac{2k}{\mu N} [\pi_k^+ - \pi_k^-]$  and diffusion  $D_L = \sum_k \frac{2k^2}{\mu^2 N} [\pi_k^+ + \pi_k^-]$ .

In order for Eq. (37) to be consistent,  $\pi_k^\pm$  should depend only on  $m_L$ . Following Refs. 38, 39, and 41, we will make an approximation based on an adiabatic elimination that allows us to relate variables  $m_k \approx m_L$ ,  $\rho \approx \xi(1 - m_L^2)$  as functions of  $m_L$  only. Performing the averages on Eqs. (38) and (39) using the known moments of the binomial distributions  $P_{0/1}(k, q)$  and replacing the above mentioned relations resulting from the adiabatic approximation, we obtain specific formulas for the drift and diffusion in terms of the  $m_L$  variable. For example, if we take  $\alpha = 2$ , we obtain

$$F_L = -2am_L + 2h\xi^2 \frac{\mu - 1}{\mu} m_L(1 - m_L^2), \quad (41)$$

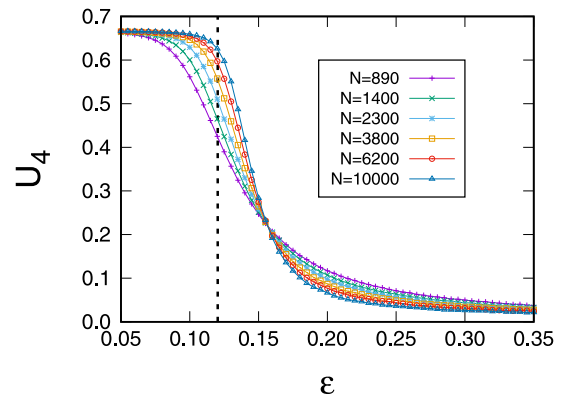


FIG. 11. Binder cumulant as a function of  $\varepsilon$  for a scale-free network with  $\lambda = 2.3$ . Dots (joined by lines as a guide to the eye) correspond to Monte Carlo simulations of different system sizes, averaged over  $10^3$  Monte Carlo steps and an ensemble of 100 networks, while the dashed black line is the prediction  $\varepsilon_c = 0.120$  of the pair approximation.

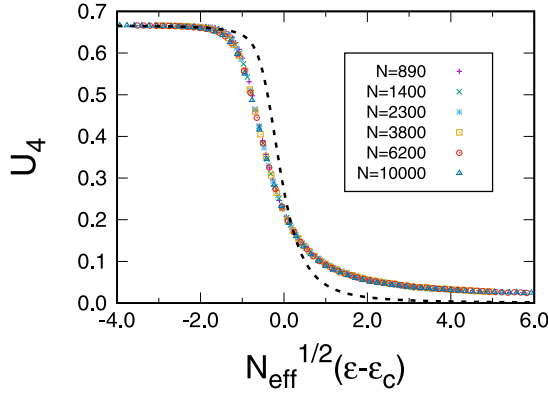


FIG. 12. The same data of Fig. 11 are plotted versus the rescaled variable  $N_{\text{eff}}(\epsilon - \epsilon_c)^{1/2}$ . The good collapse of the data validates the proposed scaling law  $U_4(\epsilon, N) = u(N_{\text{eff}}^{1/2}(\epsilon - \epsilon_c))$ , being  $u(x)$  the scaling function. The theoretical prediction (dotted line) is obtained from Eq. (29) using instead of  $c$ ,  $c_4$  the rescaled coefficients  $\bar{c} = 1.84$  and  $\bar{c}_4 = 0.11$ .

$$D_L = \frac{\mu_2}{\mu^2} \left[ 2a + 2h\xi \frac{\mu + (\mu_2 - \mu)\xi}{\mu_2} (1 - m_L^2) \right], \quad (42)$$

which coincide with Eqs. (8) and (9) in the all-to-all limit  $\mu_m = \mu^m$ ,  $\mu \rightarrow \infty$ ,  $\xi = 1/2$ . It is now straightforward to repeat the analysis of Section III with those new drift and diffusion terms Eqs. (41) and (42) or their equivalent for other values of  $\alpha$ . The critical point is determined by the condition  $\epsilon_c = \frac{\mu - 1}{\mu} \xi_c^2$  where the dependence of  $\xi_c$  with  $\epsilon_c$  is given by Eq. (36) or its equivalent for other values of  $\alpha$ . Note that  $\epsilon_c$  is always smaller, for finite  $\mu$ , than the all-to-all result  $\epsilon_c = 1/4$  (see Fig. 7). For the particular case of a 15-regular lattice, we find that the tricritical point moves to a larger value  $\alpha = 6.14$ , with corresponding  $\epsilon = 0.084$ , and the trimodal region shrinks with respect to the one found in the all-to-all setup (see Fig. 1). For  $z$ -regular lattices with  $z \leq 5$ , the trimodal region disappears altogether. In general, multiple stability and tricritical points are also possible in the case

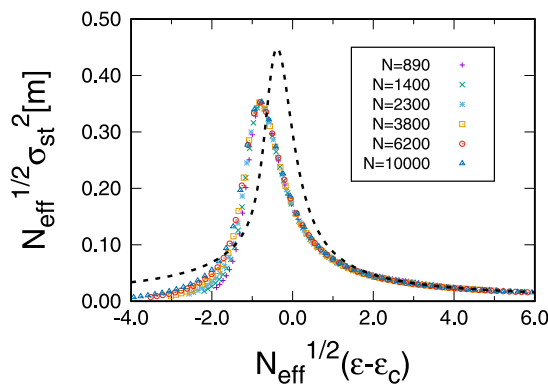


FIG. 13. Rescaled variance of the magnetization as a function of  $\epsilon$  in the same case than in Fig. 11, showing the validity of the scaling law  $\sigma^2[m](\epsilon, N) = N_{\text{eff}}^{-1/2} v(N_{\text{eff}}^{1/2}(\epsilon - \epsilon_c))$ . The theoretical prediction (dotted line) is obtained from Eq. (29) using instead of  $c$ ,  $c_4$  the rescaled coefficients  $\bar{c} = 1.84$  and  $\bar{c}_4 = 0.11$ .

of other complex networks and we leave for future work a detailed analysis of the phase diagram for different network types.

We now compare the theoretical results with numerical simulations of the stochastic process as defined by the individual rates Eqs. (1) and (2). We consider a quadratic interaction  $\alpha = 2$  and two network configurations: (i) a  $z$ -regular random network, where each node is randomly connected to exactly  $z$  neighbors and hence  $\mu = z$  and  $\mu_m = \mu^m$ ; and (ii) a scale-free network, with a power-law degree distribution  $P_k \sim k^{-\lambda}$ . Both networks have been generated with the configuration model as detailed in Ref. 42.

In the  $z$ -regular network, for which we take  $z = 5$ , the previous analysis leads to  $\epsilon_c = 0.1125$ ,  $\xi_c = 0.3899$  and coefficients of the potential [see Eqs. (22) and (24)]  $c = 1.602$ ,  $c_4 = 0.093$ , while the all-to-all solution is  $\epsilon_c = 0.25$ ,  $\xi_c = 0.5$  and  $c = 1$ ,  $c_4 = 0.125$ . To compare the predictions of the pair approximation with the numerical simulations in the steady state, we have chosen two particular combinations of moments: the Binder cumulant,<sup>30,43</sup> defined as  $U_4 = 1 - \frac{\langle m^4 \rangle}{3\langle m^2 \rangle^2}$ , and the variance  $\sigma^2[m]$ . As shown in Fig. 8, the theory agrees very well with the simulations. According to the finite-size scaling relation Eq. (28) for  $\Delta\epsilon = 0$ , the dependence on system size is  $U_4(\epsilon, N) = u(N_{\text{eff}}^{1/2}(\epsilon - \epsilon_c))$ , being  $u(x)$  the scaling function. This means that the cumulant curves for different values of  $N$  intersect at the critical point of the infinite system  $\epsilon_c$ . From the numerical data, we obtain  $\epsilon_c = 0.114 \pm 0.002$  which is compatible with the theoretical prediction. A plot of  $U_4(\epsilon, N)$  vs  $N_{\text{eff}}^{1/2}(\epsilon - \epsilon_c)$  (see Fig. 9) indicates a good data collapse in a single function  $u(x)$ . For the variance, we obtain the scaling law  $\sigma^2[m](\epsilon, N) = N_{\text{eff}}^{-1/2} v(N_{\text{eff}}^{1/2}(\epsilon - \epsilon_c))$ , being  $v(x)$  the scaling function, which is checked in Fig. 10. The comparison in Figs. 9 and 10 indicates that the scaling functions  $u(x)$  and  $v(x)$  obtained analytically capture reasonably well the overall shape of the numerical results.

In the scale-free network, we have chosen  $\lambda = 2.3$ , and measured directly  $\mu \approx 5.54$ ,  $\mu_{-1} \approx 0.35$ , while the second moment depends on system size  $N$  as  $\mu_2(N) \sim N^b$  with  $b = 0.64$ . The theoretical results are  $\epsilon_c = 0.120$  and  $\xi_c = 0.383$ . The analysis of the Binder cumulant obtained numerically (see Fig. 11) yields  $\epsilon_c = 0.157$ , while the all-to-all prediction is  $\epsilon_c = 0.25$ . The finite-size scaling of the scale-free network is somehow tricky. In our theoretical analysis, the critical point  $\epsilon_c$  depends only on  $\mu$  and  $\mu_{-1}$ , which tend to finite values as  $N \rightarrow \infty$ , but the second moment  $\mu_2$  diverges as  $N^b$ . Under the assumption  $\mu_2 \gg \mu$ , which is true in general for scale-free networks, the diffusion scales as  $D_L \propto \mu^2/\mu_2$  [see, for example, Eq. (42) for  $\alpha = 2$ ], which eventually leads to coefficients  $c \propto \mu^2/\mu_2$  and  $c_4 \propto \mu^2/\mu_2$ . One can still keep a scaling law as in Eq. (28) if one replaces the size  $N$  by an effective system size defined as  $N_{\text{eff}} = N\mu^2/\mu_2$ , in such a way that the dependence of  $c$  and  $c_4$  on  $\mu_2$  is absorbed by  $N_{\text{eff}}$  and the scaling law becomes  $U_4(\epsilon, N) = u(N_{\text{eff}}^{1/2}(\epsilon - \epsilon_c))$ , being  $u(x)$  the scaling function. This scaling law is confirmed by the data collapse presented in Fig. 12. The theoretical scaling function obtained from Eq. (28) replacing  $c$  and  $c_4$  by their rescaled versions  $\bar{c} = c\mu_2/\mu^2 = 1.84$  and

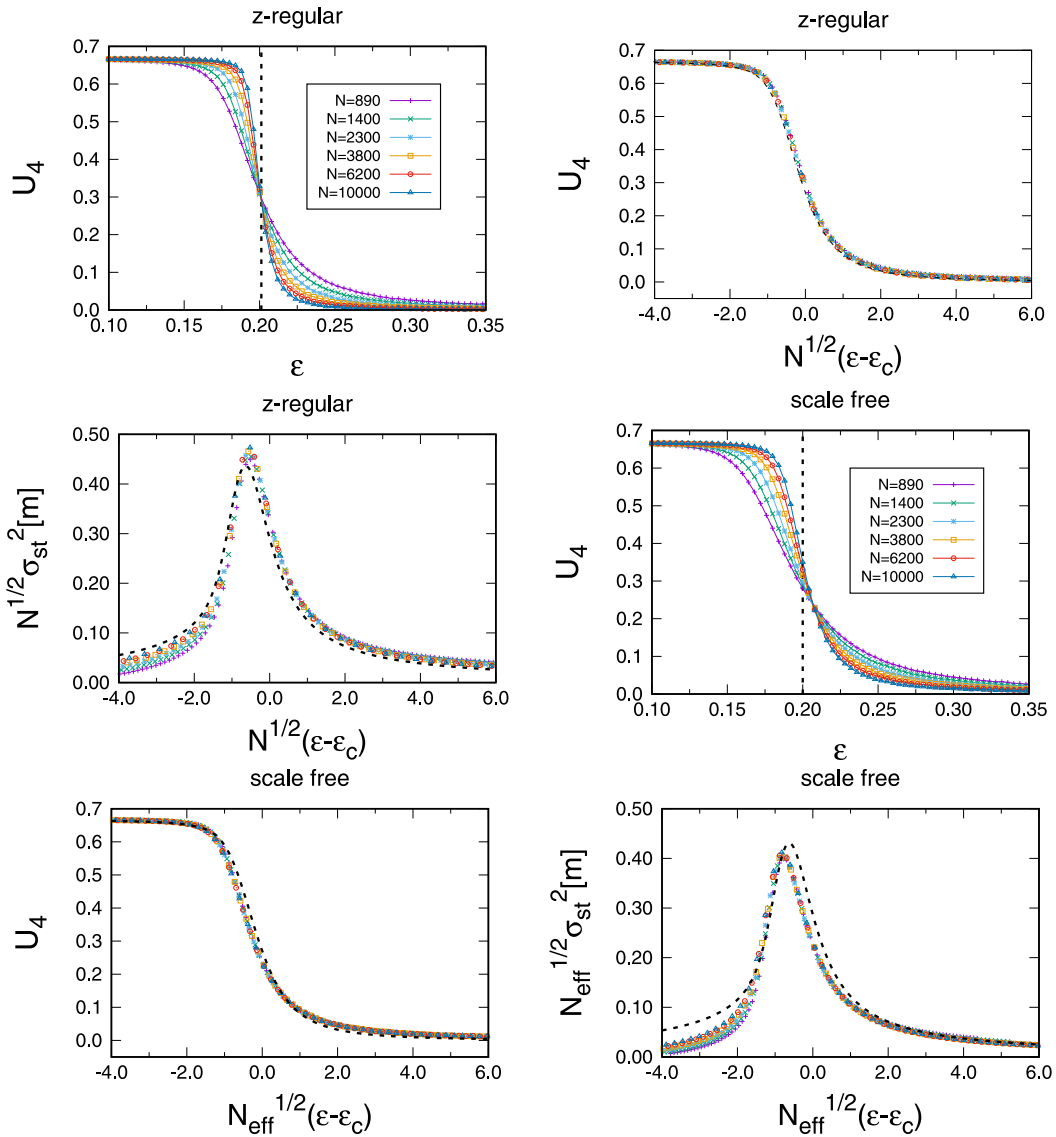


FIG. 14. Numerical and theoretical results for the 15-regular network and a scale free network with  $P_{k \geq 6} \sim k^{-2.51}$  and  $\mu \approx 14.79$ ,  $\mu_{-1} \approx 0.11$ . Dots correspond to numerical simulations of different system sizes averaged over  $10^6$  Monte Carlo steps and an ensemble of 100 networks, while the dashed black lines are the prediction of the pair-approximation which are:  $\epsilon_c = 0.201$ ,  $c = 1.15$  and  $c_4 = 0.12$  for the 15-regular network and  $\epsilon_c = 0.200$ ,  $\bar{c} = 1.19$  and  $\bar{c}_4 = 0.12$  for the scale free network.

$\bar{c}_4 = c_4 \mu_2 / \mu^2 = 0.11$  yields a reasonable qualitative agreement with the numerical data. A similar scaling law holds for the variance  $\sigma^2[m](\epsilon, N) = N_{\text{eff}}^{-1/2} v(N_{\text{eff}}^{1/2}(\epsilon - \epsilon_c))$ , the validity of which is checked in Fig. 13. While the numerical results indeed follow the proposed scaling law, the theory can only offer a qualitative agreement for the scaling function.

There are at least three possible sources of the quantitative discrepancy between the scaling function derived from the theory and the numerical results: (i) the scale-free network has no cut-off so that the maximum degree scales as  $k_{\text{max}} \sim N^{1/(\lambda-1)}$  and there will be degree-degree correlations which are neglected in the pair-approximation; (ii) the truncation of expansion Eq. (40) may not converge rapidly with  $\Delta_k$  for the scale free network, producing anomalous diffusion; and (iii) the binomial ansatz for the probabilities  $P_{0/1}(k, q)$

fails in two different ways for low degrees  $k$ . First, the single event probabilities  $p_0$  and  $p_1$  depend also on the degree (see Ref. 40) and, second, the shape of the distribution may differ from being binomial (see Ref. 36). In order to elucidate which one is more important, we repeated the analysis for highly connected networks, which are equally affected by the error sources (i) and (ii) compared to the previous low average degree networks. The corresponding results are presented in Fig. 14. We explored a 15-regular network, and a scale free network with  $P_{k \geq 6} \sim k^{-2.51}$  and average degree  $\mu \approx 14.79$ . As one can appreciate in the figure, the theoretical predictions are greatly improved in this case, which indicates that the third cause is the most significant one. There are still some discrepancies for the scale free network, which are caused by the other two possible error sources mentioned above.

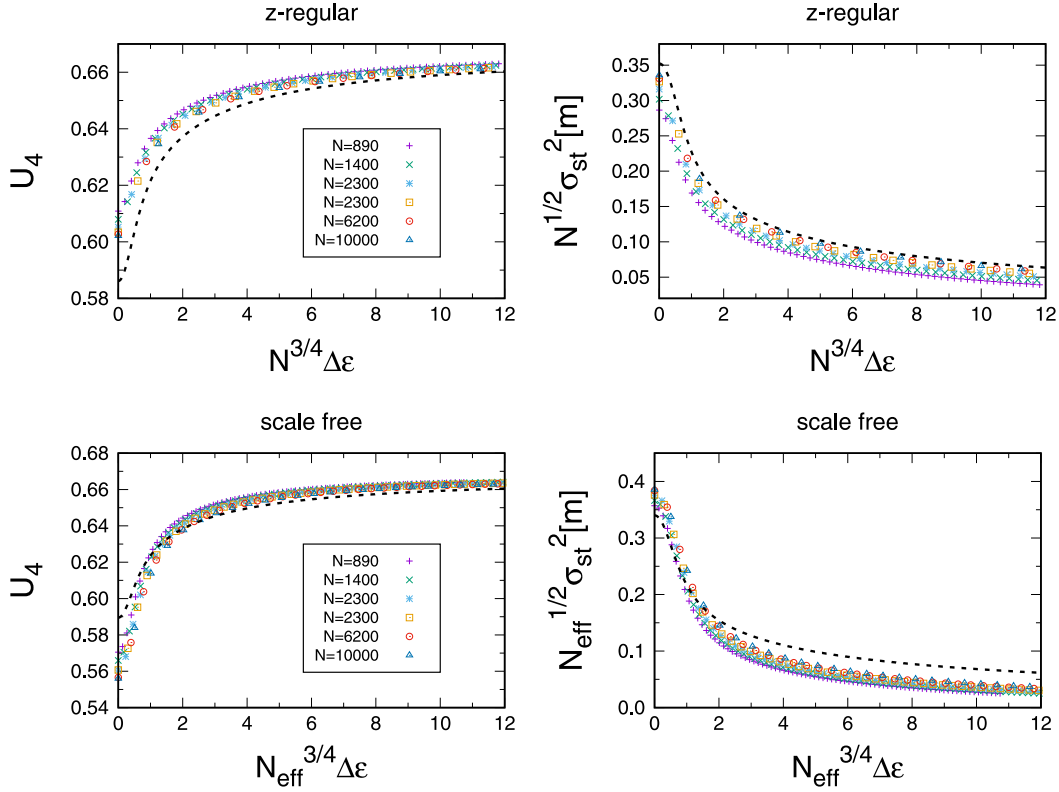


FIG. 15. Binder cumulant and rescaled magnetization variance as a function of the rescaled asymmetry parameter  $N^{3/4} \Delta \epsilon$  for the z-regular network, and  $N_{\text{eff}}^{3/4} \Delta \epsilon$  for the scale free network. The simulation and network specifications, together with the theoretical results, are the same of those of Fig. 14. The value of  $\epsilon$  is taken to be below the critical value,  $\epsilon = \epsilon_c - 1/N^{1/2}$  for the 15-regular network and  $\epsilon = \epsilon_c - 1/N_{\text{eff}}^{1/2}$  for the scale free network, so that as the asymmetry increases the system crosses the catastrophe lines. The corresponding scaling functions are calculated using Eq. (29) taking into account that  $x = -1$  is fixed and the variable in this case is the asymmetry  $y$ .

In Fig. 15, we additionally checked if the theoretical analysis is also valid for the asymmetric case. For the highly connected networks, we take values of  $\epsilon \leq \epsilon_c$  below the critical point previously calculated in Fig. 14 and vary the asymmetry parameter  $\Delta \epsilon$  checking the scaling law Eq. (28). The results suggest that the theory is equally accurate for the asymmetric case, with a very different tendency of the moments with respect to  $\Delta \epsilon$  before and after the catastrophe.

As in the scale-free network we have  $N_{\text{eff}} \sim N^{1-b}$ , the scaling laws Eq. (28) can be written in terms of the physical system size  $N$  as

$$\langle m^k \rangle_{\text{st}} = N^{-k\beta/\bar{\nu}} \tilde{\phi}_k [N^{1/\bar{\nu}}(\epsilon - \epsilon_c), N^{\delta\beta/\bar{\nu}} \Delta \epsilon], \quad (43)$$

with an appropriate scaling function  $\tilde{\phi}_k$  and  $\beta = 1/2$ ,  $\delta = 3$ ,  $1/\bar{\nu} = (1-b)/2$ . In the thermodynamic limit, one recovers the Landau theory mean-field exponents  $\langle |m| \rangle \sim (\epsilon_c - \epsilon)^{1/2}$  for  $\Delta \epsilon = 0$ , and  $\langle |m| \rangle \sim (\Delta \epsilon)^{1/3}$  for  $\epsilon = \epsilon_c$ . For the normalized fluctuations (the equivalent of the “magnetic susceptibility”)  $\chi = N[\langle m^2 \rangle_{\text{st}} - \langle |m| \rangle_{\text{st}}^2]$ , we find the scaling behavior  $\chi(\epsilon, N) = N^{\gamma/\bar{\nu}} \tilde{\chi}[N^{1/\bar{\nu}}(\epsilon - \epsilon_c), N^{\delta\beta/\bar{\nu}} \Delta \epsilon]$  with a scaling function  $\tilde{\chi}$  and  $\gamma/\bar{\nu} = 1 - 2\beta/\bar{\nu}$ , and  $\gamma = 1$ . Again, in the thermodynamic limit, we find the mean-field result  $\chi \sim |\epsilon - \epsilon_c|^{-1}$ . Therefore, although the critical exponents of the transition are those given by the mean-field theory, the analysis in terms of the finite system size  $N$  yields a network-dependent exponent  $1/\bar{\nu} = (1-b)/2$ . We have

computed  $b = 0.36(1)$  for the scale free network with tail exponent  $\lambda = 2.43$ , yielding  $1/\bar{\nu} = 0.160(5)$  and  $\gamma/\bar{\nu} = 1 - 2\beta/\bar{\nu} = 0.680(5)$  in perfect agreement with the numerical fits  $\beta/\bar{\nu} = 0.157(6)$  and  $\gamma/\bar{\nu} = 0.691(9)$  reported in Ref. 23.

## V. SUMMARY AND CONCLUSIONS

In this paper, we have studied the impact of a non-linear herding mechanism in the noisy voter model, with the consideration of a parameter  $\alpha$ , the “volatility” in the context of the Abrams-Strogatz model,<sup>16</sup> that measures the resistance of individuals to copy the state of a neighbor in the opposite state or the effect of complex, collective interactions.  $\alpha$  is taken as a real parameter that can take any non-negative value. The solution in the all-to-all scenario, where all voters are neighbors of each other, shows that the well studied bimodal-unimodal finite-size transition of the noisy voter model turns into a classical second-order transition, with a critical point  $\epsilon_c = 2^{-\alpha}(\alpha - 1) \neq 0$  for  $\alpha > 1$ , no transition for  $\alpha < 1$ , and a finite-size transition  $\epsilon_c(N) = 1/N$  for the traditional linear case  $\alpha = 1$ . In the strong non-linear regime  $\alpha > 5$ , the transition  $\epsilon_c$  separates bimodal from trimodal states, with the appearance of a new first order transition line  $\epsilon_t > \epsilon_c$  that separates trimodal from unimodal states. This implies that this non-linear mechanism is able to generate tristability and the point where the tristability parameter region begins is a

tricritical point  $\alpha = 5$ ,  $\varepsilon = \varepsilon_c = \varepsilon_t$ , in accordance with the results of Ref. 21, valid for integer  $\alpha$ . The asymmetry in the rates of the model  $\Delta\varepsilon \neq 0$  is capable of destroying bimodality and trimodality at what is known as cusp and butterfly catastrophes, respectively.

The existence of a *bona-fide* transition that remains in the thermodynamic limit offers a convenient solution to the  $N$ -dependence of the transition without the need to resort to a  $N$ -dependent rescaling of the model parameters,<sup>14</sup> also known as non-extensive formulation. Although this formulation partially solves this problem, it makes the model structurally unstable in such a way that simple perturbations such as the inclusion of agents that do not change state,<sup>45–47</sup> an external signal of information,<sup>44</sup> or a network structure, can drastically change the properties to be Gaussian.<sup>34</sup> Similar conclusions were already reported in the context of the voter model<sup>48</sup> (without spontaneous switching). The addition of a third state or a small perturbation of the functional form of the rates of the voter model drastically changes the dynamics and the ordering mechanism. These structural changes of the voter model are specially relevant in the context of agent-based modeling of language competition.<sup>17</sup> In this context, the inclusion of a third state represents the role of bilingual speakers in the evolution of the number of speakers of a particular language.

The  $N$ -dependence of the non-linear model in the presence of noise is dramatically different from the linear case. The aforementioned non-extensive formulation of the linear model leads to  $N$ -independent statistics of the magnetization  $\langle m^k \rangle \sim N^0$ . For the non-linear case, however, we have shown that rescaling the parameters leads to  $\langle m^k \rangle \sim N^{-k/4}$  for  $1 < \alpha < 5$  at the critical point  $\varepsilon_c$  and  $\langle m^k \rangle \sim N^{-k/6}$  at the tricritical point for  $\alpha = 5$ , i.e., vanishing moments in the thermodynamic limit  $N \rightarrow \infty$ . The critical region where these scaling laws are valid, however, is wider for the non-linear case making them more robust against perturbations.

We have also checked the role of the network of interactions in the presented results of the model, making use of the pair approximation. It would indeed be possible to develop higher accuracy theoretical methods at the cost of simplicity and analytical tractability.<sup>35,36,40</sup> In general, we conclude that the critical point  $\varepsilon_c$  is lowered compared to the all-to-all solution, depending on the mean degree  $\mu$  and the first negative moments of the degree distribution  $\mu_{-1}, \dots, \mu_{-\alpha+1}$  (for  $\alpha$  integer). The  $N$ -dependent scaling laws are also changed for networks whose second moment of the degree distribution  $\mu_2$  changes dramatically with system size, e.g., scale-free networks  $\mu_2 \sim N^b$  with  $0 < b < 1$ . In this case, we have shown that one can define an effective system size  $N_{\text{eff}} = N\mu^2/\mu_2 \sim N^{1-b}$  for which the all-to-all scaling laws become valid. In the limit of highly heterogeneous networks,  $\lambda \rightarrow 2$  and  $b \rightarrow 0$ , statistics may be  $N$ -independent or with a very weak dependence on system size.

## ACKNOWLEDGMENTS

We acknowledge financial support from Agencia Estatal de Investigación (AEI, Spain) and Fondo Europeo de Desarrollo Regional under Project ESoTECoS Grant No.

FIS2015-63628-C2-2-R (AEI/FEDER,UE) and the Spanish State Research Agency, through the María de Maeztu Program for Units of Excellence in R&D (MDM-2017-0711). A.F.P. acknowledges support by the Formación de Profesorado Universitario (FPU) program of Ministerio de Educación, Cultura y Deportes (MECD) (Spain).

- <sup>1</sup>P. Clifford and A. Sudbury, "A model for spatial conflict," *Biometrika* **60**, 581 (1973).
- <sup>2</sup>R. A. Holley and T. M. Liggett, "Theorems for weakly interacting infinite systems and the voter model," *Ann. Probab.* **3**, 643 (1975).
- <sup>3</sup>O. Al Hammal, H. Chaté, I. Dornic, and M. A. Muñoz, "Langevin description of critical phenomena with two symmetric absorbing states," *Phys. Rev. Lett.* **94**, 230601 (2005).
- <sup>4</sup>F. Vazquez and C. López, "Systems with two symmetric absorbing states: Relating the microscopic dynamics with the macroscopic behavior," *Phys. Rev. E* **78**, 061127 (2008).
- <sup>5</sup>C. Castellano, M. A. Muñoz, and R. Pastor-Satorras, "Nonlinear  $q$ -voter model," *Phys. Rev. E* **80**, 041129 (2009).
- <sup>6</sup>B. L. Granovsky and N. Madras, "The noisy voter model," *Stoch. Proc. Appl.* **55**, 23–43 (1995).
- <sup>7</sup>A. Kirman, "Ants, rationality, and recruitment," *Quart. J. Econ.* **108**, 137–156 (1993).
- <sup>8</sup>P. A. P. Moran, "Random processes in genetics," in *Mathematical Proceedings of the Cambridge Philosophical Society* (Cambridge University Press, 1958), Vol. 54, No. 1.
- <sup>9</sup>J. L. Lebowitz and H. Saleu, "Percolation in strongly correlated systems," *Physica A* **138**, 194 (1986).
- <sup>10</sup>K. Fichtorn, E. Gulari, and R. Ziff, "Noise-induced bistability in a Monte Carlo surface-reaction model," *Phys. Rev. Lett.* **63**, 1527 (1989).
- <sup>11</sup>D. Considine, S. Redner, and H. Takayasu, "Comment on 'Noise-induced bistability in a Monte Carlo surface-reaction model'," *Phys. Rev. Lett.* **63**, 2857 (1989).
- <sup>12</sup>M. Diakonova, V. M. Eguiluz, and M. San Miguel, "Noise in coevolving networks," *Phys. Rev. E* **92**, 032803 (2015).
- <sup>13</sup>S. Alfaro, T. Lux, and F. Wagner, "Estimation of agent-based models: The case of an asymmetric Herding model," *Comput. Econ.* **26**, 19–49 (2005).
- <sup>14</sup>S. Alfaro, T. Lux, and F. Wagner, "Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach," *J. Econ. Dyn. Control* **32**, 101–136 (2008).
- <sup>15</sup>F. Schweitzer and L. Behera, "Nonlinear voter models: The transition from invasion to coexistence," *Eur. Phys. J. B* **67**, 301–318 (2009).
- <sup>16</sup>D. M. Abrams and S. H. Strogatz, "Modelling the dynamics of language death," *Nature* **424**, 900 (2003).
- <sup>17</sup>F. Vazquez, X. Castelló, and M. San Miguel, "Agent based models of language competition: Macroscopic descriptions and order-disorder transitions," *J. Stat. Mech. Theory Exp.* **2010**(04), P04007 (2010).
- <sup>18</sup>A. Nowak, J. Szamrej, and B. Latané, "From private attitude to public opinion: A dynamic theory of social impact," *Psychol. Rev.* **97**, 362 (1990).
- <sup>19</sup>D. Nettle, "Using social impact theory to simulate language change," *Lingua* **108**, 95 (1999).
- <sup>20</sup>B. Min and M. San Miguel, "Fragmentation transitions in a coevolving nonlinear voter model," *Sci. Rep.* **7**, 12864 (2017).
- <sup>21</sup>P. Nyczka, K. Sznajd-Weron, and J. Cislo, "Phase transitions in the  $q$ -voter model with two types of stochastic driving," *Phys. Rev. E* **86**, 011105 (2012).
- <sup>22</sup>P. Nyczka and K. Sznajd-Weron, "Anticonformity or independence? Insights from statistical physics," *J. Stat. Phys.* **151**, 174 (2013).
- <sup>23</sup>A. Jędrzejewski, "Pair approximation for the  $q$ -voter model with independence on complex networks," *Phys. Rev. E* **95**, 012307 (2017).
- <sup>24</sup>P. Moretti, S. Liu, C. Castellano, and R. Pastor-Satorras, "Mean-field analysis of the  $q$ -voter model on networks," *J. Stat. Phys.* **151**, 113 (2013).
- <sup>25</sup>R. Lambiotte and S. Redner, "Dynamics of non-conservative voters," *EPL* **82**, 18007 (2008).
- <sup>26</sup>J. Molofsky, R. Durrett, J. Dushoff, D. Griffeth, and S. Levine, "Local frequency dependence and global coexistence," *Theor. Popul. Biol.* **55**, 270 (1999).
- <sup>27</sup>K. Huang, *Statistical Mechanics* (John Wiley & Sons, New York, 1963).
- <sup>28</sup>D. Escaff, R. Toral, C. Van den Broeck, and K. Lindenberg, "A continuous-time persistent random walk model for flocking," *Chaos* **28**, 075507 (2018).



- <sup>29</sup>M. A. Serrano, K. Klemm, F. Vazquez, V. M. Eguiluz, and M. San Miguel, "Conservation laws for voter-like models on random directed networks," *J. Stat. Mech. Theory Exp.* **2009**(10), P10024 (2009).
- <sup>30</sup>R. Toral and P. Colet, *Stochastic Numerical Methods: An Introduction for Students and Scientists* (Wiley, 2014).
- <sup>31</sup>N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland Physics Publishing, Amsterdam, 2007).
- <sup>32</sup>E. C. Zeeman, "Catastrophe theory," *Sci. Am.* **234**, 65–83 (1976).
- <sup>33</sup>R. Gilmore, *Catastrophe Theory for Scientists and Engineers* (Wiley, New York, 1981).
- <sup>34</sup>S. Alfarano and M. Milaković, "Network structure and  $N$ -dependence in agent-based herding models," *J. Econ. Dyn. Control* **33**, 78–92 (2009).
- <sup>35</sup>J. P. Gleeson, "High-accuracy approximation of binary-state dynamics on networks," *Phys. Rev. Lett.* **107**, 068701 (2011).
- <sup>36</sup>J. P. Gleeson, "Binary-state dynamics on complex networks: Pair approximation and beyond," *Phys. Rev. X* **3**, 021004 (2013).
- <sup>37</sup>A. Carro, R. Toral, and M. San Miguel, "The noisy voter model on complex networks," *Sci. Rep.* **6**, 24775 (2016).
- <sup>38</sup>A. F. Peralta, R. Toral, A. Carro, and M. San Miguel, "Stochastic pair approximation treatment of the noisy voter model," [arXiv:1807.04615](https://arxiv.org/abs/1807.04615).
- <sup>39</sup>V. Sood, T. Antal, and S. Redner, "Voter models on heterogeneous networks," *Phys. Rev. E* **77**, 041121 (2008).
- <sup>40</sup>E. Pugliese and C. Castellano, "Heterogeneous pair approximation for voter models on networks," *EPL* **88**, 58004 (2009).
- <sup>41</sup>F. Vazquez and V. M. Eguiluz, "Analytical solution of the voter model on uncorrelated networks," *New J. Phys.* **10**, 063011 (2008).
- <sup>42</sup>M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, "Generation of uncorrelated random scale-free networks," *Phys. Rev. E* **71**, 027103 (2005).
- <sup>43</sup>K. Binder, "Finite-size scaling analysis of Ising model block distribution functions," *Z. Phys. B Condens. Matter* **43**, 119 (1981).
- <sup>44</sup>A. Carro, R. Toral, and M. San Miguel, "Markets, herding and response to external information," *PLoS ONE* **10**(7), e0133287 (2015).
- <sup>45</sup>N. Khalil, M. San Miguel, and R. Toral, "Zealots in the mean-field noisy voter model," *Phys. Rev. E* **97**, 0123101 (2018).
- <sup>46</sup>A. Kononovicius and V. Gontis, "Control of the socio-economic systems using herding interactions," *Physica A* **405**, 80 (2014).
- <sup>47</sup>A. Kononovicius and V. Gontis, "Herding interactions as an opportunity to prevent extreme events in financial markets," *Eur. Phys. J. B* **88**, 189 (2015).
- <sup>48</sup>X. Castelló, V. M. Eguiluz, and M. San Miguel, "Ordering dynamics with two non-excluding options: Bilingualism in language competition," *New J. Phys.* **8**, 308 (2006).
- <sup>49</sup>The simulations proceed by selecting randomly one of the  $N$  nodes, say  $i$ . If  $n_i = 0$ , the transition  $n_i \rightarrow 1$  occurs with a probability  $\pi_i^+/C$ , and if  $n_i = 1$ , the transition  $n_i \rightarrow 0$  occurs with probability  $\pi_i^-/C$  with  $C = \max(a_0, a_1) + h$ .  $N$  node selections constitute one Monte Carlo step.<sup>30</sup>
- <sup>50</sup>There is also a recurrence relation for the master equation in the stationary state:  $P_{st}(n) = \frac{\pi^+(n-1)}{\pi^-(n)} P_{st}(n-1)$ , but for a general value of  $\alpha$ , it does not seem to be possible to reduce this expression to an analytically tractable function.
- <sup>51</sup>There is a misprint in Ref. 17 and the factor  $(\alpha - 9)$  is missing in formula (10).
- <sup>52</sup>In some interpretations of the model, the ordered phase corresponds to social consensus where a large fraction of the population holds the same option for the variable  $n_i$ .
- <sup>53</sup>In the case of directed networks, there is an equation for  $\rho_{01}$  and another one for  $\rho_{10}$ , which are formally similar to Eq. (32). The difference is that the changes of these quantities are  $\Delta\rho_{k,q;k_{in},q_{in}}^{01} = 2(k_{in} - q_{in} - q)/\mu N$  and  $\Delta\rho_{k,q;k_{in},q_{in}}^{10} = 2(k - q - q_{in})/\mu N$ , which depend on the in-degree  $k_i^{in} = \sum_j A_{ji}$  which is different from the out-degree  $k_i = \sum_j A_{ij}$ . The single event probabilities of  $P_{0/1}(k, q)$  are in this case  $p_0 = \rho_{01}$  and  $p_1 = 1 - \rho_{10}$ , and we have to define the binomial  $P_{0/1}^{in}(k_{in}, q_{in})$  with single event probabilities  $p_0^{in} = \rho_{10}$  and  $p_1^{in} = 1 - \rho_{01}$ . The corresponding Eq. (32) will need to be averaged additionally by the in-degree distribution  $P_{k_{in}}^{in}$  (or by  $P_{k_{in}|k}^{in}$  if there are  $k_{in} - k$  correlations) and  $P_{0/1}^{in}(k_{in}, q_{in})$ . Obviously, this treatment is valid only if reciprocal links are negligible; otherwise, the method is significantly more involved.