



PAPER • OPEN ACCESS

## Stochastic pair approximation treatment of the noisy voter model

To cite this article: A F Peralta *et al* 2018 *New J. Phys.* **20** 103045

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the [collection](#) - download the first chapter of every title for free.



## PAPER

## Stochastic pair approximation treatment of the noisy voter model

A F Peralta<sup>1</sup> , A Carro<sup>2,3</sup> , M San Miguel<sup>1</sup> and R Toral<sup>1</sup> <sup>1</sup> IFISC, Instituto de Física Interdisciplinar y Sistemas Complejos (UIB-CSIC), Campus Universitat de les Illes Balears, E-07122-Palma de Mallorca, Spain<sup>2</sup> Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, OX2 6ED, United Kingdom<sup>3</sup> Mathematical Institute, University of Oxford, OX2 6GG, United KingdomE-mail: [afperalta@ifisc.uib-csic.es](mailto:afperalta@ifisc.uib-csic.es)**Keywords:** pair approximation, voter model, system-size expansionRECEIVED  
18 July 2018REVISED  
22 September 2018ACCEPTED FOR PUBLICATION  
12 October 2018PUBLISHED  
31 October 2018

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

**Abstract**

We present a full stochastic description of the pair approximation scheme to study binary-state dynamics on heterogeneous networks. Within this general approach, we obtain a set of equations for the dynamical correlations, fluctuations and finite-size effects, as well as for the temporal evolution of all relevant variables. We test this scheme for a prototypical model of opinion dynamics known as the noisy voter model that has a finite-size critical point. Using a closure approach based on a system size expansion around a stochastic dynamical attractor we obtain very accurate results, as compared with numerical simulations, for stationary and time-dependent quantities whether below, within or above the critical region. We also show that finite-size effects in complex networks cannot be captured, as often suggested, by merely replacing the actual system size  $N$  by an effective network dependent size  $N_{\text{eff}}$ .

**1. Introduction**

From classical problems in statistical physics [1, 2] to questions in biology and ecology [3–5], and over to the spreading of opinions and diseases in social systems [6–10], stochastic binary-state models have been widely used to study the emergence of collective phenomena in systems of stochastically interacting components. In general, these components are modeled as binary-state variables—spin up or down—sitting at the nodes of a network whose links represent the possible interactions among them. While initial research focused on the limiting cases of a well-mixed population, where each of the components is allowed to interact with any other, and regular lattice structures, later works turned to more complex and heterogeneous topologies [11–14]. A most important insight derived from these more recent works is that the macroscopic dynamics of the system can be greatly affected by the particular topology of the underlying network. In the case of systems with critical behavior, different network characteristics have been shown to have a significant impact on the critical values of the model parameters [15–17], such as the critical temperature of the Ising model [18–20] and the epidemic threshold in models of infectious disease transmission [21–25]. Thus, the identification of the particular network characteristics that have an impact on the dynamics of these models, as well as the quantification of their effect, are of paramount importance.

The first theoretical treatments introduced for the study of stochastic, binary-state dynamics on networks relied on a global-state approach [2, 26], i.e. they focused on a single variable—for example, the number of nodes in one of the two possible states—assumed to represent the whole state of the system. In order to write a master equation for this global-state variable, some approximation is required to move from the individual particle transition rates defining the model to some effective transition rates depending only on the chosen global variable. Within this global-state approach, the *effective-field* approximation assumes that all nodes have the same rate of switching to the other state, such that local densities can be replaced by their global equivalents and all details of the underlying network are completely lost. As a consequence, results are only accurate for highly connected homogeneous networks, closer to fully connected topologies.

More recent theoretical treatments have proposed a node-state approach, departing from a master equation for the  $N$ -node probability distribution and, thus, taking into account the state of each individual node. From this master equation, general evolution equations for the moments and correlations of the relevant dynamical variables can be derived. Depending on the particular functional form of the individual particle transition rates, this system of dynamical equations can be open—an infinite hierarchy of equations, each of them depending on higher-order correlation functions—or closed. For open systems of equations, different schemes have been proposed for their closure: a *mean-field* approximation [25, 27, 28], in which all moments are replaced by products of individual averages, and a *van Kampen* type system size expansion [29, 30]. Other variants of the moments expansion have been previously considered in the literature with different types of moment closure assumptions [31–34], see also the review [35] of these techniques. Even when the system of equations is closed, further approximations are required to deal with the appearance of terms depending on the specific network structure. In particular, an *annealed network* approximation has been successfully used in a number of models [36–42] in order to deal with these terms. This approximation involves replacing the original network by a complementary, weighted, fully connected network, whose weights are proportional to the probability of each pair of nodes being connected. In this way, highly precise analytical solutions can be obtained for all relevant dynamical variables.

A third group of analytical treatments has sought to establish an intermediate level of detail in its description of the dynamics, in-between the global-state and the node-state approaches. Similar to the global-state method, though significantly improving on it, a reduced set of variables is chosen to describe the state of the system. In particular, information is aggregated for nodes of a given degree. In this manner, some essential information about the network structure is kept—different equations for nodes with different degrees—while the tractability of the problem is greatly improved. As with the global-state approach, in order to write a master equation for this reduced set of variables, some further approximation is required to translate the individual particle transition rates that define the model into some effective transition rates depending only on the chosen set of variables. The number of these variables, as well as their nature, define different levels of approximation, which can be ordered from more to less detailed as: approximate master equation [16, 17], pair approximation (PA) [43–49] and heterogeneous mean-field [25, 37, 50, 51]. Many works based on these approaches have further neglected dynamical correlations, fluctuations and finite-size effects, which corresponds to a mean-field type of analysis. While there have been attempts to relax these restrictions and deal with finite-size effects, these efforts have been usually based on an adiabatic elimination of variables [37, 43, 48, 49, 51–53], whose range of validity is limited.

In this paper we consider this intermediate level of description focusing on the PA that chooses as a reduced set of relevant variables the number of nodes with a given degree in a given state and the number of links connecting nodes in different states. We develop a full stochastic treatment of the PA scheme in which we avoid mean-field analysis or adiabatic elimination of variables. In particular, after writing a master equation for the reduced set of variables, we obtain effective transition rates by introducing the PA as described in [43] and elaborating on the assumptions used in this approach as well as on their quantitative implications. Furthermore, after writing general equations for the moments and correlations, we propose two different strategies for their closure and solution. The first one, that we term **S1PA**, is based on a *van Kampen* type system size expansion [29, 54] around the deterministic solution of the dynamics, that is, around the lowest order term of the expansion which corresponds to the thermodynamic limit of infinite system size. This approach allows us to derive linear equations for the correlation matrix and first-order corrections to the average values. The second strategy, that we term **S2PA**, is based on a system size expansion around a dynamical attractor, and it allows us to obtain expressions not only for the stationary value of all relevant quantities but also for their dynamics.

Although our methodology is very general for any set of transition rates, only for the first closure strategy—**S1PA**—one can apply straightforwardly our general method to any model under study, to end up with the desired equations. The second one—**S2PA**—requires specific rates for the analysis to proceed. Thus, for the sake of concreteness, we focus on the noisy voter model as an explicit example for which we carry out a full analytical treatment. The noisy voter model is a variant of the original voter model [4, 55] which, apart from pairwise interactions in which a node copies the state of a randomly selected neighbor, it also includes random changes of state. It is a paradigmatic example of a stochastic binary-state model, with applications in the study of non-equilibrium systems in a wide range of fields, and it has been studied by, at least, five mutually independent strands of research, largely unaware of each other. Namely, random processes in genetics [56], percolation processes in strongly correlated systems [57], heterogeneous catalytic chemical reactions [58, 59], herding behavior in financial markets [60], and probability theory [61]. The behavior of the noisy voter model is characterized by the competition between two mechanisms: on the one hand, the pairwise copy interactions tend to order the system, driving it towards a homogeneous configuration—all spins in the same state, whether up or down—. On the other hand, the random changes of state tend to disorder the system, pulling it away from the homogeneous configurations. These homogeneous configurations are the absorbing states of the ordinary voter model, which lose their absorbing character due to the random changes of state. Depending on the values

of the model parameters, which set the relative strength of one mechanism over the other, the system can be found in a mostly ordered regime—dominated by pairwise interactions—or a mostly disordered regime—dominated by noise—. Furthermore, a noise-induced, finite-size transition can be observed between these two behavioral regimes. The finite-size character of the transition is due to the fact that the critical point tends to zero in the thermodynamic limit of large system sizes. Although most of the initial literature about the noisy voter model focused only on regular lattices [57, 61] and a fully connected network [60, 62], recent studies have addressed more complex topologies, both from an effective-field perspective [26, 63, 64] and using an annealed network approximation [42]. In particular, while the effective-field approach was only able to broadly capture the effect of the network size and mean degree on the results of the model for highly homogeneous and connected networks, the annealed approximation [42] was, in addition, able to reproduce the impact of the degree heterogeneity—variance of the underlying degree distribution—on the critical point of the transition and the temporal correlations with a high level of accuracy, as well as the main effects on the local order parameter, though with significantly less accuracy. Further studies of the noisy voter model include the effects of zealots [65], nonlinear group interactions [53] and aging [66].

When applied to the noisy voter model, the first of the approaches proposed in this paper—**S1PA**—is able to improve on the accuracy of previous methods only for the region above the critical point, while it fails in its proximity and in the region below it. This failure is due to the existence of a mathematical divergence as the noise intensity approaches zero in the analytical expressions obtained for the fluctuations and the local order parameter. The second approach proposed in this paper—**S2PA**—, on the contrary, leads to highly accurate results for all relevant variables and for all values of the parameters, whether below, within, or above the critical region, thus significantly improving on previous methods. Furthermore, this approach is not limited to stationary values, but rather it is able to provide highly accurate analytical expressions for their time evolution. Finally, this second approach allows us to show that finite-size effects in a complex network cannot be simply reduced to replacing the system size by an effective one, as most previous studies suggest [37, 43, 48, 49, 51].

Summing up, this paper contributes to the analysis of stochastic binary-state models in three main ways. First, it presents a general methodology with clearly identified and well-justified arguments and approximations, including an analytical assessment of their validity ranges. Second, our approach allows to find analytical expressions for the correlations, fluctuations and finite-size effects, as well as for the temporal evolution of all relevant variables, without needing to resort to crude adiabatic elimination of variables. Third, the results obtained for the noisy voter model represent a significant improvement in accuracy, being valid both above and below the critical point as well as in the critical region.

The paper is organized as follows: in section 2 we introduce the stochastic model and define the main quantities of interest, namely, the fluctuations of the density of nodes in a particular state and the density of active links. In section 3 we revisit some of the previous theoretical treatments introduced for the study of the stochastic binary-state dynamics in the two limiting cases of a global-state approach and a node-state approach. In the same section we explain the main approximations required to close and eventually solve the dynamical equations for the moments and correlations of the relevant dynamical variables. In section 4 we introduce our stochastic description based on a master equation for a reduced set of dynamical variables. In this section we use the PA to derive the hierarchy of equations for the moments and correlations of the dynamical variables. Two different closure schemes are presented and discussed in sections 5 and 6 based, respectively, on an expansion around the deterministic solution and around a stochastic dynamical attractor. The main results of both schemes as well as other approaches in the literature are compared against the results of numerical simulations in section 7 in the steady state. In section 8 we extend our method to determine the dynamical evolution of the correlation function and its comparison with numerical simulations. Finally, in section 9 we present the main conclusions of our study. The technical details of the solutions of some equations are presented in appendices A and C. In appendix B we present a simple example using linear equations in which one of our closure schemes can be carried out in full detail. In appendix D we explain the method of adiabatic elimination to obtain a partial solution of some of the variables under study.

## 2. Model

We consider an undirected network consisting of  $N$  nodes. Links between nodes are mapped into the usual (symmetric) adjacency matrix  $\mathbf{A} = (A_{ij})$  as  $A_{ij} = 1$  if nodes  $i$  and  $j$  are connected, and  $A_{ij} = 0$  otherwise. Connected nodes are ‘neighbors’ of each other. Each node  $i = 1, \dots, N$  holds a time-dependent binary variable  $n_i = 0, 1$ . We do not adopt in this paper any specific interpretation, but the state  $n_i = 1$  (resp. 0) will be labeled as ‘up’ (resp. ‘down’), and a link between nodes in different states will be defined as *active*. A stochastic dynamics is introduced by which the node variable  $n_i(t)$  can switch its state at time  $t$ , either  $0 \rightarrow 1$  with rate  $r_i^+$ , or  $1 \rightarrow 0$  with rate  $r_i^-$ . The rates  $r_i^\pm(\mathbf{n})$  depend, in general, on the whole set of node-states  $\mathbf{n} = (n_1, \dots, n_N)$  but are assumed to

not depend explicitly on time. Although our method is rather general, we will use throughout the paper the noisy voter (Kirman) model rates as an archetypal example

$$r_i^+ = a + \frac{h}{k_i} \sum_j A_{ij} n_j, \quad r_i^- = a + \frac{h}{k_i} \sum_j A_{ij} (1 - n_j), \quad (1)$$

being  $k_i = \sum_j A_{ij}$  the degree (number of connected nodes) of node  $i$ . The term proportional to  $h$  represents a ‘herding’ mechanism by which the transition rates are proportional to the fraction of neighbor nodes holding the opposite state, while the term proportional to  $a$  represents random jumps between states at a constant rate and is also known as the ‘noise term’.

We introduce the probability  $P(\mathbf{n}; t)$  of finding node configuration  $\mathbf{n}$  at time  $t$ . This probability can be understood at different levels, as the detailed configuration  $\mathbf{n}(t)$  depends on (i) the initial condition, (ii) the realization of the stochastic process, and (iii) the network realization  $\mathbf{A}$ . For a given network configuration, let  $N_k = \#\{i | \sum_j A_{ij} = k\}$  be the number of nodes with degree  $k$ . By definition,  $N_k = 0$  for  $k < k_{\min}$  and for  $k > k_{\max}$ , the minimum and maximum values of the degree for that given network. We introduce

$$P_k = \frac{N_k}{N}, \quad \mu_m = \sum_{k=k_{\min}}^{k_{\max}} P_k k^m, \quad (2)$$

as the probability  $P_k$  that a randomly chosen node has degree  $k$ , and the  $m$ -moment  $\mu_m$  of the degree distribution, and for brevity  $\mu \equiv \mu_1$ . The variance of the degree distribution is  $\sigma_k^2 = \mu_2 - \mu^2$  and we introduce the degree heterogeneity  $\kappa = \sigma_k^2 / \mu^2$ .

We will focus on the time evolution of the first moment  $\langle m(t) \rangle$  and variance  $\sigma^2[m(t)] = \langle m^2(t) \rangle - \langle m(t) \rangle^2$  of the rescaled density of nodes in the up state, as well as on the average  $\langle \rho(t) \rangle$  of the density of active links. The precise definitions of these quantities are

$$m = \frac{2}{N} \sum_i n_i - 1, \quad (3)$$

$$\rho = \frac{\frac{1}{2} \sum_{ij} A_{ij} (n_i (1 - n_j) + (1 - n_i) n_j)}{\frac{1}{2} \sum_{ij} A_{ij}} \equiv \frac{L}{\frac{1}{2} \mu N}, \quad (4)$$

where  $L$  is the number of active links, and the denominator  $\frac{1}{2} N \mu = \frac{1}{2} \sum_{ij} A_{ij}$  is the total number of links. While the ‘magnetization’  $m(t) \in [-1, 1]$  describes the global-state of the system,  $\rho(t) \in [0, 1]$  can be considered as a measure of disorder, corresponding the value  $\rho = 1/2$  to a random distribution of states, and  $\rho = 0$  to full order or consensus where all nodes are in the same state, either  $n_i = 0$  or  $n_i = 1$ . All averages  $\langle F[\mathbf{n}(t)] \rangle$  are understood with respect to the probability  $P(\mathbf{n}; t)$

$$\langle F[\mathbf{n}(t)] \rangle = \sum_{n_1=0,1;\dots;n_N=0,1} F(\mathbf{n}) P(\mathbf{n}; t). \quad (5)$$

Results can be later averaged over the ensemble of networks with a given degree distribution, always after performing the average in equation (5).

We also introduce the correlation functions

$$C_{ij}(t) = \langle \Delta_i(t) \Delta_j(t) \rangle = \langle n_i(t) n_j(t) \rangle - \langle n_i(t) \rangle \langle n_j(t) \rangle, \quad (6)$$

$$C_{ijk}(t) = \langle \Delta_i(t) \Delta_j(t) \Delta_k(t) \rangle, \quad (7)$$

...

where  $\Delta_i = n_i - \langle n_i \rangle$ . As  $n_i^2 = n_i$ , note that  $C_{ii} = \langle n_i \rangle - \langle n_i \rangle^2$  and that

$$\sigma^2[m(t)] = \frac{4}{N^2} \sum_{i,j} C_{ij}, \quad (8)$$

$$\langle \rho(t) \rangle = \frac{\sum_{ij} A_{ij} (\langle n_i \rangle (1 - \langle n_j \rangle) - C_{ij})}{\frac{1}{2} \mu N}. \quad (9)$$

### 3. Node-state and global-state approaches

#### 3.1. The node-state approach

For general rates  $r_i^\pm$ , the probability function satisfies the master equation [29, 67]

$$\frac{\partial P(\mathbf{n}; t)}{\partial t} = \sum_{i=1}^N \{ (E_i^+ - 1)[n_i r_i^- P(\mathbf{n}; t)] + (E_i^- - 1)[(1 - n_i) r_i^+ P(\mathbf{n}; t)] \}, \quad (10)$$

being  $E_i$  the step operator  $E_i^\pm f(n_1, \dots, n_i, \dots, n_N) = f(n_1, \dots, n_i \pm 1, \dots, n_N)$ . From the master equation it is possible to get general evolution equations for the moments as<sup>4</sup>

$$\frac{d\langle n_i(t) \rangle}{dt} = \langle r_i^+ \rangle - \langle (r_i^+ + r_i^-) n_i \rangle, \quad (11)$$

$$\frac{dC_{ij}(t)}{dt} = \langle (r_i^- + r_i^+) n_i \Delta_j \rangle + \langle (r_j^- + r_j^+) n_j \Delta_i \rangle + \langle r_i^+ \Delta_j \rangle + \langle r_j^+ \Delta_i \rangle, \quad i \neq j. \quad (12)$$

To proceed forward, one would need to solve these equations, introducing explicit expressions for the transition rates  $r_i^\pm$  and finding solutions for  $\langle n_i(t) \rangle$  and  $C_{ij}(t)$ , which can then be introduced in equations (3), (8), and (9) in order to obtain expressions for the desired quantities  $\langle m(t) \rangle$ ,  $\sigma^2[m(t)]$ , and  $\langle \rho(t) \rangle$ . When introducing explicit expressions for  $r_i^\pm$  for any given model, however, one might find two different cases depending on whether the resulting system of equations is closed or not, which will determine the need for different types of approximations:

- *Non-closed system of equations:* with the remarkable exception of the noisy voter model, the set of equations (11)–(12) is, in general, not closed. In particular, when the explicit expressions for  $r_i^\pm$  are introduced, the right-hand side of the equations will contain higher-order correlation functions,  $C_{i_1 i_2 \dots i_m}$  with  $m \geq 3$ . While it is also possible to obtain evolution equations for these higher-order correlation functions, the resulting hierarchy of equations is, in principle, infinite. Thus, some approximations are needed in order to close this hierarchy. In general, there are two basic schemes for the closure of equations (11) and (12) (see [35] for alternative techniques):

*Mean-field approximation:* in the mean-field approximation, all moments are replaced by products of individual averages, e.g.

$$\begin{aligned} \langle n_i(t) n_j(t) \rangle &\approx \langle n_i(t) \rangle \langle n_j(t) \rangle, \\ \langle n_i(t) n_j(t) n_k(t) \rangle &\approx \langle n_i(t) \rangle \langle n_j(t) \rangle \langle n_k(t) \rangle, \\ &\dots \end{aligned} \quad (13)$$

within this mean-field approximation, equation (11) is closed and, as  $C_{ij} = 0$ , equation (12) is not needed.

*System size scaling hypothesis of the correlation functions:* the usual van Kampen system size expansion [29] splits a relevant dynamical variable as  $x(t) = \langle x(t) \rangle + \Omega^{1/2} \xi(t)$ , being  $\Omega$  a large parameter (typically the system size or volume). This assumes that the dynamical variable scales as  $x(t) \sim \Omega$ . When trying to apply this approach to the master equation (10) one faces the fact the dynamical variables  $n_i(t) = 0, 1$  do not scale with system size  $N$ . A more sophisticated expansion was developed in [30] with the main ansatz that the correlation functions scale as  $C_{i_1 i_2 \dots i_m} = \langle \Delta_{i_1} \Delta_{i_2} \dots \Delta_{i_m} \rangle = O(\Omega^{-m/2})$  for all  $i_1, i_2, \dots, i_m$  different. Once this ansatz is assumed, it turns out that one can still use the mean-field results (13) in the evolution equations for the mean values (11), thus finding for  $\langle n_i(t) \rangle$  the same result as the mean-field approximation. Regarding the correlation  $C_{ij}$ , however, this approach leads to non-trivial evolution equations, allowing one to improve on the mean-field approach.

- *Closed system of equations:* in the particular case of the noisy voter model, and due to its linear transition rates, equations (11) and (12) are already closed, and thus there is no need for any closure approximation. Of course, closing the equations is just the first step towards solving them. Typically, after closing the equations one finds factors of the form  $\sum_j A_{ij} n_j$ , whose complexity for a general lattice structure makes further analytical progress very difficult. These terms can be dealt with by assuming the:

*Annealed approximation:* this approach replaces the original adjacency matrix  $A_{ij}$  by a complementary, weighted, fully connected adjacency matrix  $\tilde{A}_{ij}$ , whose weights are proportional to the probability of each pair of nodes being connected [38–41]. For uncorrelated networks of the configuration ensemble this can be written as

<sup>4</sup> For a step-by-step derivation, see the supplementary information of [42].



$$\tilde{A}_{ij} \approx \frac{k_i k_j}{N\mu}, \quad (14)$$

where the normalization of the weights is chosen so that both the degree sequence and the total number of links remain unchanged [37, 68–70]. This approximation reduces the aforementioned sums to

$$\sum_{j=1}^N A_{ij} f_j \approx \sum_{j=1}^N \tilde{A}_{ij} f_j = \frac{k_i}{N\mu} \sum_{j=1}^N k_j f_j, \quad (15)$$

being  $f_j[n_j, k_j]$  any function that depends on the degree  $k_j$  of the node  $j$  and/or on the node variable  $n_j$ . In the particular case that  $f_j[n_j, k_j] = f[k_j]$ , a further simplification is possible in terms of the degree distribution  $P_k = N_k/N$  as

$$\frac{1}{N} \sum_{j=1}^N k_j f_j = \sum_k P_k k f(k). \quad (16)$$

There is one case in which the annealed approximation is exact. This is when the adjacency matrix is  $A_{ij} = 1, \forall i, j$ , meaning that every node is connected to every other node, a situation known as all-to-all coupling or fully connected network. The results of the treatment in terms of the full set of node-states and the annealed approximation have been described in [42].

The main results for the noisy voter model using the mean-field and the annealed approximations will be compared in section 7 with our new treatment using a stochastic PA scheme. Before that, let us explain in the next subsection yet another, very simple, approximation which is capable of capturing the main phenomenological features of the noisy voter model.

### 3.2. The global-state approach

Instead of the full state  $\mathbf{n} = (n_1, \dots, n_N)$ , this approach proposes a much coarser description, in which the only relevant variable is the number of nodes in the up state  $N_1 = \#\{i | n_i = 1\}$ . This is clearly an important reduction with respect to the  $N$  variables needed in the  $\mathbf{n}$  representation, leading to a much simpler analytical treatment but also to less accurate results in general network topologies and exact results only for fully connected networks. As far as the global variable  $N_1$  is concerned, there are only two possible outcomes of the stochastic process:  $N_1 \rightarrow N_1 + 1$  and  $N_1 \rightarrow N_1 - 1$ . Let  $P(N_1; t)$  be the probability that the number of nodes in the up state takes the value  $N_1$  at time  $t$ . It obeys the master equation

$$\frac{\partial P(N_1; t)}{\partial t} = (E_{N_1}^+ - 1)[W^- P(N_1; t)] + (E_{N_1}^- - 1)[W^+ P(N_1; t)], \quad (17)$$

where  $W^+$  and  $W^-$  are effective rates, depending only on the global variable  $N_1$ , and the step operator now acts as  $E_{N_1}^\pm F(N_1) = F(N_1 \pm 1)$ . From here we derive equations for the average value and the second moment of the global variable as

$$\frac{d\langle N_1 \rangle}{dt} = \langle W^+ - W^- \rangle, \quad (18)$$

$$\frac{d\langle N_1^2 \rangle}{dt} = \langle (1 - 2N_1)W^- + (1 + 2N_1)W^+ \rangle. \quad (19)$$

Some approximation is needed to find the effective rates  $W^+$  and  $W^-$  in such a way that they depend only on the  $N_1$  variable and in a closed form. The simplest approximation is to assume that all nodes have the same rate of switching to the other state, such that we can replace the local density of nodes in the up state by the global density of nodes in the up state

$$\frac{\sum_j A_{ij} n_j}{k_i} \approx \frac{N_1}{N}, \quad (20)$$

leading to

$$W^+ = (N - N_1) \left( a + h \frac{N_1}{N} \right), \quad (21)$$

$$W^- = N_1 \left( a + h \frac{N - N_1}{N} \right). \quad (22)$$

Note that in this effective-field approximation, the details of the network (any dependence on the connectivity matrix) are completely lost. In the case of a fully connected network, where each node is connected to every other node, the effective-field approximation is exact.

Once the effective-field approximation has been assumed, no further approximations are needed, for after replacing equations (21)–(22) in (18)–(19), the evolution equations for the mean value  $\langle m(t) \rangle$  and the variance  $\sigma^2[m(t)]$  of  $m = 2N_1/N - 1$  satisfy

$$\frac{d\langle m \rangle}{dt} = -2a\langle m \rangle, \quad (23)$$

$$\frac{d\sigma^2[m]}{dt} = -2\left(2a + \frac{h}{N}\right)\sigma^2[m] - \frac{2h}{N}\langle m \rangle^2 + \frac{4a + 2h}{N}. \quad (24)$$

The density of active links  $\rho$  can be calculated using the all-to-all ansatz as the probability of finding a node in the up state times the probability of finding one of its neighbors in the down state or vice versa

$\rho = 2\frac{N_1 N - N_1}{N(N-1)} = \frac{N}{N-1}\frac{1-m^2}{2}$ . In the steady state we have

$$\langle m \rangle_{\text{st}} = 0, \quad (25)$$

$$\sigma_{\text{st}}^2[m] = \frac{2a + h}{2aN + h}, \quad (26)$$

$$\langle \rho \rangle_{\text{st}} = \frac{aN}{2aN + h}. \quad (27)$$

Further analysis shows that the noisy voter model has a transition separating regions in which the steady state probability  $P_{\text{st}}(N_1)$  has a single peak at  $N_1 = \frac{N}{2}$  and regions in which it has two peaks at  $N_1 = 0, N$ . The critical condition of the transition is thus that the probability distribution becomes flat  $P_{\text{st}}(N_1) = \frac{1}{N+1}$  for  $N_1 = 0, \dots, N$ . This condition leads to a variance  $\sigma_{\text{st}}^2[m] = \frac{N+2}{3N}$ , from which we can obtain the critical value  $a_c = h/N$ , such that  $P_{\text{st}}(N_1)$  is double peaked for  $a < a_c$  and single peaked for  $a > a_c$  (see [60, 62] for a different method of finding this critical value). These predictions for the transition point and the detailed dependence of  $\sigma_{\text{st}}^2[m]$  and  $\langle \rho \rangle_{\text{st}}$  on the system parameters will be analyzed in the next sections.

As the critical value  $a_c = h/N$  depends on  $N$ , the single to double peaked transition observed for  $P_{\text{st}}(N_1)$  is said to be a finite-size-transition [71]. Certainly, the transition disappears in the thermodynamic limit  $N \rightarrow \infty$  where, for all values of  $a$ , the distribution  $P_{\text{st}}(N_1)$  is always single peaked. Therefore, from the perspective of a strict statistical mechanics formulation, the transition does not exist. Nevertheless, the behavior near  $a = 0$  bears many similarities with a true phase transition, as there are valid scaling laws, divergence of fluctuations with characteristic exponents, critical slowing down, etc, as we will show in the next sections. For this reason, and being clearly an abuse of language, it is customary to refer to the small  $a \sim h/N$  values as the ‘critical region’. It is also possible to identify the critical dimension  $d_c = 2$  above which the ‘critical’ exponents are independent of dimension [61].

As shown in equation (26), which, we recall, is exact in the all-to-all coupling scenario, when  $a/h = O(N^{-1})$ , the fluctuations scale as  $\sigma[m] \sim O(N^0)$ , while far away from the critical point,  $a/h = O(1)$ , they scale as  $\sigma[m] \sim O(N^{-1/2})$ . As the critical dimension of this model is  $d_c = 2$ , this scaling with system size is the correct one in the all-to-all coupling. This simple global approach captures one of the main difficulties we will encounter when closing the equations for the moments in more complicated setups, namely, the fact that the fluctuations scale differently near the critical point and far away from it. A standard van Kampen approach [29] consisting in splitting  $m(t)$  in deterministic  $\phi(t)$  and stochastic  $\xi(t)$  contributions  $m(t) = \phi(t) + N^{-1/2}\xi(t)$ , will lead invariably to fluctuations of the order  $\sigma[m] = O(N^{-1/2})$ , and it is hence not appropriate near the critical point.

## 4. Stochastic pair approximation

In this section, we introduce the stochastic PA, starting from a master equation description of the dynamics and using an intermediate level of detail in our characterization, in-between the node-state and the global-state approaches. Our method is a full stochastic treatment of the PA considered by Vazquez and Eguíluz [43] in their study of the (noiseless) voter model. We elaborate further on this approach and we apply it to the case of the noisy voter model. We derive in this section the evolution equations for the moments and correlations of the dynamical variables and leave for the next two sections the closure of the hierarchy and the solution of the equations.

As explained, the starting point is not the full node-state configuration  $\mathbf{n} = (n_1, \dots, n_N)$ , but a reduced set of variables  $\{\mathbf{N}_l, L\}$ , where  $L$  is the number of active links and  $\mathbf{N}_l = (N_{l,k_{\min}}, N_{l,k_{\min}+1}, \dots, N_{l,k_{\max}})$  with  $N_{l,k}$  the number of nodes in the up state with degree  $k$ . The total number of variables in this description is thus  $k_{\max} - k_{\min} + 2$ , still much smaller than the  $N$  variables needed in the  $\mathbf{n}$  representation.

In the elementary process  $n_i = 0 \rightarrow n_i = 1$ , we have  $N_{l,k_i} \rightarrow N_{l,k_i} + 1$ , being  $k_i$  the degree of node  $i$ . At the same time, the number of active links  $L$  can vary in different amounts  $L \rightarrow L + k_i - 2q_i$  depending on the



number  $q_i \in (0, k_i)$  of neighbors of node  $i$  in the up state. Similarly, in the elementary process  $n_i = 1 \rightarrow n_i = 0$  it is  $N_{1,k_i} \rightarrow N_{1,k_i} - 1$  and  $L$  can vary in different amounts  $L \rightarrow L - k_i + 2q_i$ . Introducing  $P(\mathbf{N}_1, L; t)$  as the probability of the state  $\{\mathbf{N}_1, L\}$  at time  $t$ , we derive the master equation

$$\begin{aligned} \frac{\partial P(\mathbf{N}_1, L; t)}{\partial t} = & \sum_k \sum_{q=0}^k \{ (E_{N_{1,k}}^+ E_L^{k-2q} - 1) [W^{(-,k,q)} P(\mathbf{N}_1, L; t)] \\ & + (E_{N_{1,k}}^- E_L^{-k+2q} - 1) [W^{(+,k,q)} P(\mathbf{N}_1, L; t)] \}, \end{aligned} \quad (28)$$

with  $E_{N_{1,k}}$  and  $E_L$  the step operators acting on  $N_{1,k}$  and  $L$ , respectively. From this master equation and following standard techniques, one can derive the equations for the moments and correlations of the rescaled variables  $m_k = 2 \frac{N_{1,k}}{N_k} - 1$  and  $\rho = \frac{L}{\frac{1}{2}\mu N}$  as

$$\frac{d\langle m_k \rangle}{dt} = \langle F_k \rangle, \quad (29)$$

$$\frac{d\langle \rho \rangle}{dt} = \langle F_\rho \rangle, \quad (30)$$

$$\frac{d\langle m_k m_{k'} \rangle}{dt} = \langle m_{k'} F_k \rangle + \langle m_k F_{k'} \rangle + \delta_{k,k'} \frac{\langle G_k \rangle}{N}, \quad (31)$$

$$\frac{d\langle m_k \rho \rangle}{dt} = \langle \rho F_k \rangle + \langle m_k F_\rho \rangle + \frac{\langle G_{k,\rho} \rangle}{N}, \quad (32)$$

$$\frac{d\langle \rho^2 \rangle}{dt} = 2\langle \rho F_\rho \rangle + \frac{\langle G_\rho \rangle}{N}, \quad (33)$$

where

$$F_k = \frac{2}{N_k} \sum_{q=0}^k [W^{(+,k,q)} - W^{(-,k,q)}], \quad (34)$$

$$F_\rho = \frac{2}{\mu N} \sum_k \sum_{q=0}^k (k - 2q) [W^{(+,k,q)} - W^{(-,k,q)}], \quad (35)$$

$$G_k = \frac{4}{P_k N_k} \sum_{q=0}^k [W^{(+,k,q)} + W^{(-,k,q)}], \quad (36)$$

$$G_{k,\rho} = \frac{4}{\mu N_k} \sum_{q=0}^k (k - 2q) [W^{(+,k,q)} + W^{(-,k,q)}], \quad (37)$$

$$G_\rho = \frac{4}{\mu^2 N} \sum_k \sum_{q=0}^k (k - 2q)^2 [W^{(+,k,q)} + W^{(-,k,q)}]. \quad (38)$$

As in the global-state approach, some approximation is needed in order to find the effective rates  $W^{(\pm,k,q)}$  that appear in the master equation. Writing  $q_i$  as  $q$  and  $k_i$  as  $k$ , the rates of the elementary processes  $n_i = 0 \rightarrow n_i = 1$  and  $n_i = 1 \rightarrow n_i = 0$  are, respectively

$$R_{k,q}^+ = a + h \frac{q}{k}, \quad (39)$$

$$R_{k,q}^- = a + h \frac{k - q}{k}. \quad (40)$$

Therefore, the total rate of the process  $N_{1,k} \rightarrow N_{1,k} + 1, L \rightarrow L + k - 2q$  is calculated as the product of the elementary rate  $R_{k,q}^+$ , the number of nodes in the down state among the population with degree  $k$ , and the fraction (probability) of the population of nodes in the down state with degree  $k$  that have  $q$  neighbors in the up state  $P_0(k, q)$ , i.e.

$$W^{(+,k,q)} = (N_k - N_{1,k}) \cdot P_0(k, q) \cdot R_{k,q}^+. \quad (41)$$

A similar reasoning leads to the total rate of the process  $N_{1,k} \rightarrow N_{1,k} - 1, L \rightarrow L - k + 2q$ ,

$$W^{(-,k,q)} = N_{1,k} \cdot P_1(k, q) \cdot R_{k,q}^-. \quad (42)$$

In general, those probabilities depend on the detailed configuration  $\mathbf{n}$  and adjacency matrix  $\mathbf{A}$ , but we need to write them in terms only of the description variables  $\{\mathbf{N}_1, L\}$ . We then make the PA [43]

$$P_0(k, q) \approx \binom{k}{q} c_0^q (1 - c_0)^{k-q}, \quad (43)$$

$$P_1(k, q) \approx \binom{k}{q} c_1^{k-q} (1 - c_1)^q, \quad (44)$$

where

$$c_n = \frac{L}{\sum_{k=k_{\min}}^{k_{\max}} k N_{n,k}} = \frac{\rho}{\mu N \sum_{k=k_{\min}}^{k_{\max}} k N_{n,k}}, \quad n = 0, 1, \quad (45)$$

is the ratio between the total number of active links and the total number of links connected to nodes in state  $n = 0, 1$ . This  $c_{0/1}$  can be interpreted as the single event probability that a node in the state 0/1 has a neighbor in the opposite state 1/0. The PA assumes that the probability of having  $q$  nodes in the up state among the  $k$  nodes connected to a node in state 0/1 is a sequence of  $q$  independent processes and thus binomial. The PA is known to be accurate for uncorrelated networks [17, 43, 48].

For the noisy voter model, encoded in the transition rates (39) and (40) using the known moments of the binomial distribution  $\sum_{q=0}^k q P_{0/1}(k, q) = k c_{0/1}$  and  $\sum_{q=0}^k q^2 P_{0/1}(k, q) = k c_{0/1} (1 + (k - 1) c_{0/1})$ , after lengthy but straightforward algebra we arrive at

$$F_k = -2a m_k + \frac{2h\rho}{1 - m_L^2} (m_L - m_k), \quad (46)$$

$$F_\rho = 2a + 2(h - 2a)\rho - \frac{4h}{1 - m_L^2} \rho^2 + \frac{1}{\mu} \frac{4h\rho}{1 - m_L^2} \left[ \frac{1 + m_L^2 - 2m m_L}{1 - m_L^2} \rho - 1 + m m_L \right], \quad (47)$$

$$G_k = \frac{4}{P_k} \left( a + \frac{h\rho}{1 - m_L^2} (1 - m_k m_L) \right), \quad (48)$$

$$G_{k,\rho} = \frac{4}{\mu} \left( (h(k - 2) - 2ak) \frac{m_L - m_k}{1 - m_L^2} \rho - 2h(k - 1) \frac{m_k - 2m_L + m_k m_L^2}{(1 - m_L^2)^2} \rho^2 \right), \quad (49)$$

with the link (or ‘degree-weighted’) magnetization  $m_L$  defined as

$$m_L = \frac{1}{\mu} \sum_{k=k_{\min}}^{k_{\max}} P_k k m_k. \quad (50)$$

The complete expression for  $G_\rho$  being too long, only a simplified version valid in the stationary state will be displayed later (see equation (70)).

Equations (29)–(33) together with equations (46)–(50) are the basis of our subsequent analysis. They are not yet closed. Note, however, that  $m_L(t)$  satisfies the exact equation

$$\frac{d\langle m_L \rangle}{dt} = -2a \langle m_L \rangle, \quad (51)$$

whose solution  $\langle m_L(t) \rangle = \langle m_L(0) \rangle e^{-2at}$  indicates that  $\langle m_L \rangle_{st} = 0$ . Remarkably,  $\langle m_L \rangle$  is self-governed and fulfills an equation equivalent to  $\langle m \rangle$  in the effective-field approximation (see equation (23)). This will not be the only similarity between these two approximations and variables, as we will show in section 6.2.

## 5. Expansion around the deterministic solution

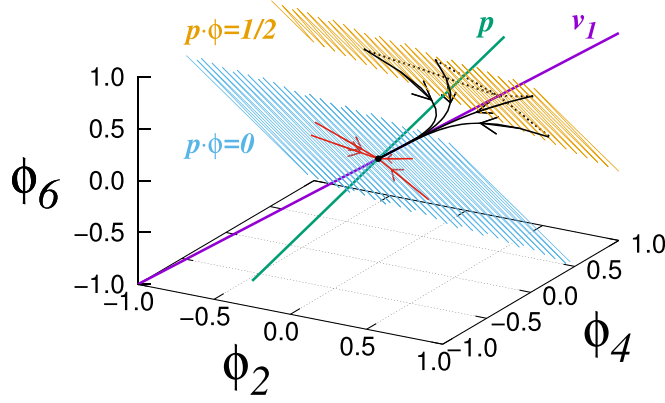
### 5.1. The deterministic solution

In order to close and eventually solve equations (29)–(33), we first study the deterministic solution using a mean-field assumption, i.e., neglecting all correlations. Introducing the notation  $\phi_k(t) = \langle m_k(t) \rangle$ ,  $\phi_\rho(t) = \langle \rho(t) \rangle$ ,  $\phi_L(t) = \langle m_L(t) \rangle = \phi_L(0) e^{-2at}$ , and  $\phi(t) = \langle m(t) \rangle = \sum_k P_k \phi_k(t)$ , and substituting equations (46) and (47) into (29) and (30), we find

$$\dot{\phi}_k = -2a\phi_k + 2h \frac{\phi_\rho}{1 - \phi_L^2} (\phi_L - \phi_k), \quad (52)$$

$$\dot{\phi}_\rho = 2a + 2(h - 2a)\phi_\rho - \frac{4h}{1 - \phi_L^2} \phi_\rho^2 + \frac{1}{\mu} \frac{4h\phi_\rho}{1 - \phi_L^2} \left( \frac{1 + \phi_L^2 - 2\phi\phi_L}{1 - \phi_L^2} \phi_\rho - 1 + \phi\phi_L \right), \quad (53)$$

where the dot indicates a time derivative. The equation for  $\phi$  would be identical to equation (52) with the exception of changing  $\phi_k \rightarrow \phi$ . This system of equations has only one (stable) fixed point  $\phi_k = 0$ ,  $\phi_\rho = \xi$ , where  $\xi$  is the positive solution of the quadratic equation



**Figure 1.** Simple example of different deterministic trajectories in the linear approximation with  $h = 1$ ,  $a = 0.01$ ,  $\mu = 4$ , degree sequence  $k = 2, 4, 6$  and normal vector  $\mathbf{p} = (1/8, 1/2, 3/8)$ .

$$2h\frac{\mu-1}{\mu}\xi^2 - \left(h\frac{\mu-2}{\mu} - 2a\right)\xi - a = 0, \quad (54)$$

namely

$$\xi = \frac{h(\mu-2) - 2\mu a}{4h(\mu-1)} + \sqrt{\left(\frac{h(\mu-2) - 2\mu a}{4h(\mu-1)}\right)^2 + \frac{\mu a}{2h(\mu-1)}}. \quad (55)$$

For  $a = 0$  this is  $\xi = \frac{\mu-2}{2(\mu-1)}$ . The other solution (unstable) of equation (54) is negative and thus non-physical, except in the case  $a = 0$ , which is  $\phi_\rho = 0$ . The stability of the fixed point  $\phi_k = 0$  and  $\phi_\rho = \xi$  can be assessed using the elements of the Jacobian matrix  $\mathbf{J}$ , defined as  $J_{a,b} = -\frac{\partial \dot{\phi}_a}{\partial \phi_b} \Big|_{\phi_k=0, \phi_\rho=\xi}$ , namely

$$J_{k,k'} = 2(a + h\xi)\delta_{k,k'} - 2h\xi\frac{P_{k'}k'}{\mu}, \quad (56)$$

$$J_{k,\rho} = J_{\rho,k} = 0, \quad (57)$$

$$J_{\rho,\rho} = 2(2a - h) + 8h\xi + \frac{4h}{\mu}(1 - 2\xi). \quad (58)$$

For future reference, let us also mention here the coefficients of the Hessian matrices  $\mathbf{H}$ , evaluated at the fixed point,  $H_{b,c}^a = \frac{\partial^2 \dot{\phi}_a}{\partial \phi_b \partial \phi_c} \Big|_{\phi_k=0, \phi_\rho=\xi}$ , namely

$$H_{k',\rho}^k = 2h\left(\frac{P_{k'}k'}{\mu} - \delta_{k,k'}\right), \quad (59)$$

$$H_{k,k'}^\rho = \frac{4h\xi}{\mu^2}P_kP_{k'}\left[(1 - 2\xi)(k + k') + \left(-\xi + \frac{3\xi - 1}{\mu}\right)2kk'\right], \quad (60)$$

$$H_{\rho,\rho}^\rho = -8h\frac{\mu-1}{\mu}, \quad (61)$$

$$H_{k',k''}^k = H_{\rho,\rho}^k = H_{k,\rho}^\rho = 0. \quad (62)$$

Note that the cross-terms  $J_{k,\rho}$  and  $J_{\rho,k}$  are zero, which means that the deterministic dynamics of  $\phi_k$  and  $\phi_\rho$  are uncoupled—at least in the linear regime—. The eigenvalues of the  $\mathbf{J}$  matrix (obviating the trivial  $\phi_\rho$  part) are (i)  $2a$  (slow) and (ii)  $2(a + h\xi)$  (fast). Both of them are positive, which means that the fixed point is a stable node (note that the Jacobian matrix has been defined with a minus sign). The first eigenvalue has multiplicity one while the second one has multiplicity  $d - 1$ , where  $d = k_{\max} - k_{\min} + 1$  is the dimensionality of the space  $\phi = \{\phi_k\}$ . The corresponding eigenvectors are (i)  $\mathbf{v}_1 = (1, \dots, 1)$  and (ii) the vectors  $\mathbf{v}_2$  contained in the plane  $\mathbf{p} \cdot \mathbf{v}_2 = 0$ , where  $\mathbf{p} = (k_{\min}P_{k_{\min}}, \dots, k_{\max}P_{k_{\max}})/\mu$  is the normal vector to the plane. The linear deterministic dynamics is then given by  $\phi(t) = \phi_L(0)\mathbf{v}_1e^{-2at} + (\phi(0) - \phi_L(0)\mathbf{v}_1)e^{-2(a+h\xi)t}$ , where  $\phi_L(0) = \mathbf{p} \cdot \phi(0)$ , which has a clear geometric interpretation. As shown in figure 1, in the early stages of the dynamics, the fast part dominates and trajectories evolve

close to the parallel plane  $\mathbf{p} \cdot \phi(t) = \phi_L(0)$ . Later, the slow part plays the important role and trajectories bend following the direction  $\mathbf{v}_1$ , an effect which is more pronounced the larger the separation of time scales is.

## 5.2. Expansion around the deterministic solution

The first approach that we will consider to close and solve equations (29)–(33) is a system size expansion around the deterministic trajectory. We term this approach as **SIPA**. In the spirit of van Kampen's expansion [29], we propose here the expansion

$$m_k(t) = \phi_k(t) + N^{-1/2}\gamma_k(t) + N^{-1}\eta_k(t) + \dots, \quad (63)$$

$$\rho(t) = \phi_\rho(t) + N^{-1/2}\gamma_\rho(t) + N^{-1}\eta_\rho(t) + \dots, \quad (64)$$

where  $\gamma_k, \gamma_\rho, \eta_k, \eta_\rho$  are stochastic processes, while  $\phi_k, \phi_\rho$  are the deterministic terms as given by the solution of equations (52) and (53).

If we introduce the ansatz (63) and (64) into (29)–(33) and equate powers of  $N$ , we are able to close the system of equations for the moments<sup>5</sup>. From now on, we restrict ourselves to the stationary state and set all time derivatives to zero. For the stochastic variables  $\gamma_k$  and  $\gamma_\rho$  we find  $\langle \gamma_k \rangle_{\text{st}} = 0$  and  $\langle \gamma_\rho \rangle_{\text{st}} = 0$  for the average values, and the following set of equations for the correlation matrix  $C_{a,b} \equiv \langle \gamma_a \gamma_b \rangle_{\text{st}}$

$$\sum_{k''} [C_{k',k''} J_{k,k''} + C_{k,k''} J_{k',k''}] + C_{k',\rho} J_{k,\rho} + C_{k,\rho} J_{k',\rho} = \delta_{k,k'} G_k^{\text{st}}, \quad (65)$$

$$\sum_{k''} [C_{\rho,k''} J_{k,k''} + C_{k,k''} J_{\rho,k''}] + C_{\rho,\rho} J_{k,\rho} + C_{k,\rho} J_{\rho,\rho} = G_{k,\rho}^{\text{st}}, \quad (66)$$

$$2 \sum_{k''} C_{\rho,k''} J_{\rho,k''} + 2 C_{\rho,\rho} J_{\rho,\rho} = G_\rho^{\text{st}}, \quad (67)$$

with  $G_k, G_{k,\rho}$  and  $G_\rho$  evaluated at the steady state

$$G_k^{\text{st}} = \frac{4(a + h\xi)}{P_k}, \quad (68)$$

$$G_{k,\rho}^{\text{st}} = 0, \quad (69)$$

$$G_\rho^{\text{st}} = -\frac{8\xi}{\mu^2} (a[\mu_2 - 2\mu + 2\xi(\mu - \mu_2)] + h[\mu - 2 + (\mu_2 - 7\mu + 6)\xi - 2(\mu_2 - 3\mu + 2)\xi^2]). \quad (70)$$

Although the system of linear equations (65)–(67) might be difficult to solve in general, the decoupling  $J_{k,\rho} = J_{\rho,k} = G_{k,\rho}^{\text{st}} = 0$ , which can be shown to be a general consequence of the up–down symmetry in the rates  $R_{k,q}^- = R_{k,k-q}^+$ , in addition to the simplicity of the terms  $J_{k,k'}$  for the noisy voter model, simplifies significantly the problem. As shown in appendix A, the solution is

$$C_{\rho,\rho} = G_\rho^{\text{st}}/2J_{\rho,\rho}, \quad (71)$$

$$C_{k,\rho} = 0, \quad (72)$$

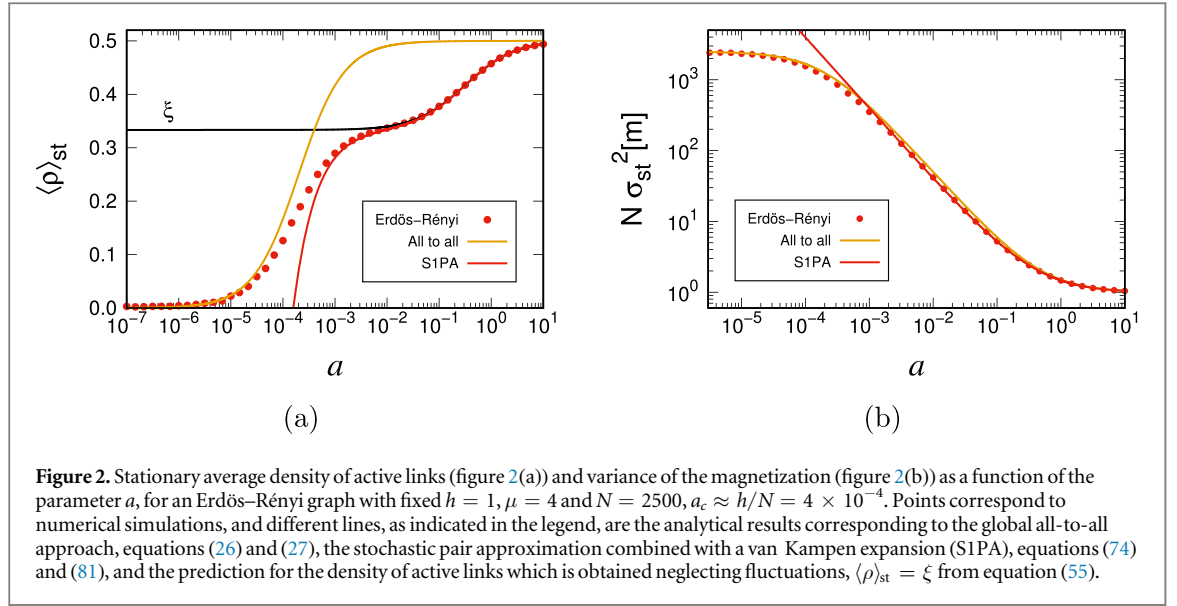
$$C_{k,k'} = \frac{h\xi}{2a + h\xi} \left( \frac{h\xi}{a} \frac{\mu_2}{\mu^2} + \frac{k + k'}{\mu} \right) + \frac{1}{P_k} \delta_{k,k'}. \quad (73)$$

The variance of the global magnetization  $m$  can now be obtained from equation (8), expressed in the form  $\sigma^2[m] = \sum_{k,k'} P_k P_{k'} C_{k,k'}$ ,

$$N\sigma_{\text{st}}^2[m] = \frac{h^2\xi^2}{2a + h\xi} \frac{N}{aN_{\text{eff}}} + \frac{2h\xi}{2a + h\xi} + 1, \quad (74)$$

where we have defined an effective system size  $N_{\text{eff}} \equiv N\mu^2/\mu_2$ . For a given finite network,  $N_{\text{eff}}$  adopts a finite value. In the large  $N$  limit, however,  $N_{\text{eff}}$  can be characterized by a non-trivial dependence on  $N$ , as the moments of the distribution can diverge. For instance, in a Barabási–Albert network it is  $N_{\text{eff}} \sim N/\log N$ . From equation (74) we obtain that, in the limit  $a \rightarrow 0$ , the fluctuations diverge as  $N\sigma_{\text{st}}^2[m] \sim \frac{h\xi}{a} \frac{\mu_2}{\mu^2}$ . In the same limit, the correlations behave as  $C_{k,k'} \sim N\sigma_{\text{st}}^2[m]$ , independent of  $k, k'$ , which is strongly related to the geometric picture of the linear deterministic dynamics that we described earlier, where variables fluctuate along the slow direction  $\mathbf{v}_1$  with an amplitude  $N^{1/2}\sigma_{\text{st}}[m]$ . In a homogeneous network, where  $\mu_2 = \mu^2$ , we find  $N\sigma_{\text{st}}^2[m] = \frac{a+h\xi}{a}$ , which in the limit  $N \rightarrow \infty$  coincides with the effective-field result in equation (26) whenever  $\xi = 1/2$ , which happens, for instance, in the limit of large connectivity  $\mu \rightarrow \infty$ . On the other hand, below or around the critical point, when  $a = O(N^{-1})$ , the

<sup>5</sup> The traditional van Kampen's expansion does not include the second order terms  $\eta_{k,\rho}$  and it is applied directly to the master equation (28). We will, however, do it at the level of the equations for the moments (29)–(33), and it can be proven [54] that only when we include these second order terms, the expansion leads to the correct results.



**Figure 2.** Stationary average density of active links (figure 2(a)) and variance of the magnetization (figure 2(b)) as a function of the parameter  $a$ , for an Erdős-Rényi graph with fixed  $h = 1$ ,  $\mu = 4$  and  $N = 2500$ ,  $a_c \approx h/N = 4 \times 10^{-4}$ . Points correspond to numerical simulations, and different lines, as indicated in the legend, are the analytical results corresponding to the global all-to-all approach, equations (26) and (27), the stochastic pair approximation combined with a van Kampen expansion (S1PA), equations (74) and (81), and the prediction for the density of active links which is obtained neglecting fluctuations,  $\langle \rho \rangle_{\text{st}} = \xi$  from equation (55).

van Kampen expansion fails to reproduce the all-to-all result and, hence, it provides the wrong scaling dependence with system size near the critical region (as can be seen in figure 2(b)). The critical point can be obtained from the condition of a flat probability distribution of the magnetization,  $\sigma_{\text{st}}^2[m] = \frac{N+2}{3N}$ , which leads to

$$a_c = \frac{3h}{2N_{\text{eff}}} \frac{\mu - 2}{\mu - 1} + O(N^{-2}). \quad (75)$$

It is also possible to calculate the  $O(N^{-1})$  stochastic corrections to the average value of the description variables defined in equations (63) and (64), which are led by  $\langle \eta_k \rangle_{\text{st}}$  and  $\langle \eta_\rho \rangle_{\text{st}}$ . In the stationary state, the equations for  $\langle \eta_k \rangle_{\text{st}}$  and  $\langle \eta_\rho \rangle_{\text{st}}$  read

$$\sum_{k'} J_{k,k'} \langle \eta_{k'} \rangle_{\text{st}} + J_{k,\rho} \langle \eta_\rho \rangle_{\text{st}} = \frac{1}{2} \sum_{k',k''} C_{k',k''} H_{k',k''}^k + \sum_{k'} C_{k',\rho} H_{k',\rho}^k + \frac{1}{2} C_{\rho,\rho} H_{\rho,\rho}^k, \quad (76)$$

$$\sum_{k'} J_{\rho,k'} \langle \eta_{k'} \rangle_{\text{st}} + J_{\rho,\rho} \langle \eta_\rho \rangle_{\text{st}} = \frac{1}{2} \sum_{k',k''} C_{k',k''} H_{k',k''}^\rho + \sum_{k'} C_{k',\rho} H_{k',\rho}^\rho + \frac{1}{2} C_{\rho,\rho} H_{\rho,\rho}^\rho. \quad (77)$$

Again, due to the particularities of the model, these equations greatly simplify, leading to

$$\sum_{k'} J_{k,k'} \langle \eta_{k'} \rangle_{\text{st}} = 0, \quad (78)$$

$$J_{\rho,\rho} \langle \eta_\rho \rangle_{\text{st}} = \frac{1}{2} \sum_{k',k''} C_{k',k''} H_{k',k''}^\rho + \frac{1}{2} C_{\rho,\rho} H_{\rho,\rho}^\rho, \quad (79)$$

from where we obtain  $\langle \eta_k \rangle_{\text{st}} = 0$ , a result arising again from the up-down symmetry of the model, and

$$\begin{aligned} \langle \eta_\rho \rangle_{\text{st}} = (J_{\rho,\rho})^{-1} & \left[ \frac{4h\xi(1-2\xi)}{\mu(2a+h\xi)} [h\xi D + 2(a+h\xi)] \right. \\ & \left. + 4 \left( -h\xi^2 + \frac{h\xi}{\mu} (3\xi - 1) \right) D + \frac{C_{\rho,\rho}}{2} H_{\rho,\rho}^\rho \right], \end{aligned} \quad (80)$$

where  $D$  is given by equation (A.5). Finally, the density of active links in the stationary state is

$$\langle \rho \rangle_{\text{st}} = \xi + N^{-1} \langle \eta_\rho \rangle_{\text{st}}. \quad (81)$$

In the all-to-all limit  $\mu_2 \rightarrow \mu^2 \rightarrow \infty$ , we have  $\langle \eta_\rho \rangle_{\text{st}} = -\frac{h}{4a}$  and  $\langle \rho \rangle_{\text{st}}$  coincides with the effective-field result in equation (27) to the order  $O(N^{-1})$ . On the other hand, in the limit of  $a \rightarrow 0$  we have

$$\lim_{a \rightarrow 0} \langle \eta_\rho \rangle_{\text{st}} = -\frac{h\xi^2}{a} \frac{\mu_2}{\mu^2}, \quad (82)$$

which leads to unphysical results for the density of active links, as, by definition, it should be  $\rho \geq 0$ .

In figure 2 we test the validity of these analytical results for a wide range of values of  $a$  for an Erdős-Rényi network with  $N = 2500$  nodes and average connectivity  $\mu = 4$  (a more detailed comparison for other network types is postponed to the next section). The results in this figure suggest that the expansion S1PA around the deterministic solution works perfectly for  $a \gg a_c = hN^{-1}$ , while it fails below the critical point  $a \ll a_c$ , where it is not able to reproduce even qualitatively the observed dependence. We also plot the prediction for the density

of active links which is obtained neglecting fluctuations,  $\langle \rho \rangle_{\text{st}} = \xi$ , which performs very poorly, specially for small  $a$ . In the same panel we show the results of the simple global approach, equations (25) and (26), which is surprisingly accurate for the variance of the magnetization but not so good for the density of active links. Nevertheless, we note that the global approach, equivalent to the all-to-all approximation, completely disregards the details of the network connectivity and, as shown in the next sections, is not capable of explaining the differences observed between different networks. The reason for the failure near the critical region of the stochastic PA combined with a van Kampen expansion lies in the ansatz stated in equations (63) and (64), which requires  $\sigma^2[m]$  to scale as  $\sim N^{-1}$  and also to be small enough for the expansion to be accurate, which is only true in the limit  $a \gg a_c$  (see equation (26)). This is strongly related to the associated deterministic dynamics and to the slow eigendirection  $\mathbf{v}_1$  having an infinite time scale as  $a \rightarrow 0$ , which eventually leads to large fluctuations following this direction. If we want an expansion to be also valid in the critical region, we have to explore the full set of nonlinear deterministic dynamical equations (52)–(53) and find a solution (in this case a privileged slow trajectory) around which the dynamics has a finite time scale in the entire parameter region. Then, we can propose an expansion similar to equations (63) and (64) around this solution. This is the method that we are going to apply in the next section and whose procedure is made clear in appendix B with a simple example of a set of linear equations.

## 6. Expansion around a stochastic dynamical attractor

### 6.1. The dynamical attractor

As explained before, the second method we propose to tackle equations (29)–(33) starts by finding a particular deterministic solution around which we can include stochastic corrections. The solution of equations (52) and (53) depends on the initial conditions  $\phi_k(0), \phi_\rho(0)$  besides the time  $t$ . We are interested, however, in a particular solution that we will name  $\phi_k^*(t), \phi_\rho^*(t)$  and that depends on time only through  $\phi_L(t)$ , namely  $\phi_k^*(\phi_L(t)), \phi_\rho^*(\phi_L(t))$ , i.e., in a particular solution such that the system reacts instantly to the self-governed variable  $\phi_L(t) = \phi_L(0)e^{-2at}$ . This solution has to fulfill the constraints  $\phi^* = \pm 1$  and  $\phi_\rho^* = 0$  when  $\phi_L = \pm 1$ , which arise from the definition of the variables when all the nodes are up or down. By inspection, we note that  $\phi_k^* = \phi_L$  is an exact solution of equation (52) that fulfills these constraints, while the dependence  $\phi_\rho^*(\phi_L)$  is not so trivial.

Introducing  $\phi = \phi_L$  in equation (53) and changing the time derivative  $\frac{d}{dt} = -2a\phi_L \frac{d}{d\phi_L}$ , we get

$$-2a\phi_L \frac{d\phi_\rho^*}{d\phi_L} = 2a + 2(h - 2a)\phi_\rho^* - \frac{4h}{1 - \phi_L^2}\phi_\rho^{*2} - \frac{4h}{\mu}\phi_\rho^* \left(1 - \frac{\phi_\rho^*}{1 - \phi_L^2}\right), \quad (83)$$

a Riccati type of equation whose exact solution, carried out in appendix C, depends on a combination of Gauss hypergeometric functions. A simpler form, more amenable to analytical treatment, is also derived in appendix C as

$$\phi_\rho^* = \xi(1 - \phi_L^2) + O(N^{-1}). \quad (84)$$

In order to show that this simpler  $\phi_\rho^*$  is a stable attractor of the deterministic dynamics, we linearize equations (52) and (53) around the particular trajectory  $\phi_k = \phi_k^* + \delta_k, \phi_\rho = \phi_\rho^* + \delta_\rho$ , finding

$$\dot{\delta}_k = -2(a + h\xi)\delta_k, \quad (85)$$

$$\dot{\delta}_\rho = -J_{\rho,\rho}\delta_\rho + \frac{4h}{\mu}\xi(1 - 2\xi)\phi_L\delta + 2a(1 - 2\xi)\phi_L^2, \quad (86)$$

with  $\delta(t) = \sum_k P_k \delta_k(t)$  and  $J_{\rho,\rho}$  given by equation (58). The solution of these linear equations is

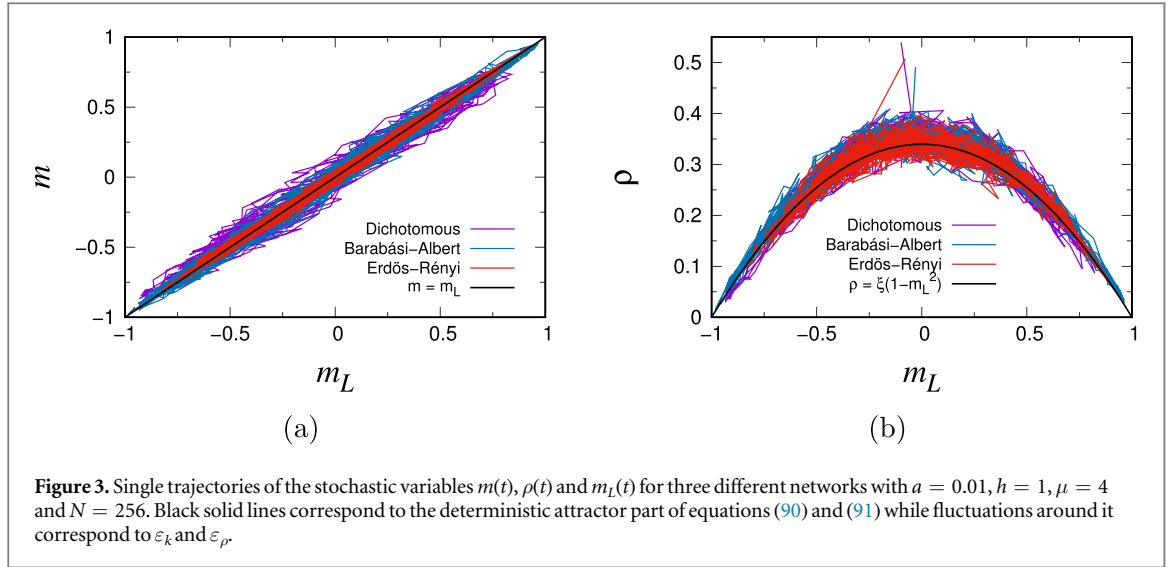
$$\delta_k(t) = \delta_k(0)e^{-2(a+h\xi)t}, \quad (87)$$

$$\delta(t) = \delta(0)e^{-2(a+h\xi)t}, \quad (88)$$

$$\delta_\rho(t) = ce^{-J_{\rho,\rho}t} + \frac{4h}{\mu} \frac{\xi(1 - 2\xi)}{J_{\rho,\rho} - 4a - 2h\xi} \phi_L(t)\delta(t) + \frac{2a(1 - 2\xi)}{J_{\rho,\rho} - 4a} \phi_L(t)^2, \quad (89)$$

where  $c$  is a constant determined by the initial conditions. The time scale of  $\delta_k$  is fast (compared to the time scale of  $\phi_L(t)$ ) and corresponds to the second eigenvalue calculated in section 5.1, while  $\delta_\rho$  has three terms: the first two are fast and the last one is slow (for  $a \ll 1$ ). It is important to note that this last slow term appears because we used the approximate expression of the trajectory in equation (84), instead of the exact one. According to the discussion in appendix C, this term can be neglected, since it is of order  $O(N^{-1})$  as long as the approximation is accurate. We thereby conclude that this trajectory is an attractor of the dynamics and that the attractor is reached in a time scale, which is shorter than the time scale within the attractor. This approach captures a critical feature





that the expansion around the deterministic solution overlooked. While  $\phi_k(t)$  is exactly the same for both approaches,  $\phi_\rho(t)$  depends on  $\phi_L(t)$  following equation (84), and the dynamics are not uncoupled as in the previous case.

## 6.2. Expansion around the dynamical attractor

In figures 3(a) and (b) we plot, respectively,  $m(t)$  and  $\rho(t)$  versus  $m_L(t)$  for different realizations of the stochastic process in different network structures. It is apparent from this figure that  $m(t)$  and  $\rho(t)$  fluctuate around the dynamical attractor  $m(t) = m_L(t)$ ,  $\rho(t) = \phi_\rho^*(m_L(t))$ . Based upon this picture of the phase space, our second approach to solve equations (29)–(33), which we name **S2PA**, consists in splitting the  $m_k(t)$  and  $\rho(t)$  variables into the dynamical attractor contribution plus additional fluctuations as

$$m_k(t) = m_L(t) + \varepsilon_k(t), \quad (90)$$

$$\rho(t) = \phi_\rho^*(m_L(t)) + \varepsilon_\rho(t), \quad (91)$$

where  $m_L$ ,  $\varepsilon_k$  and  $\varepsilon_\rho$  are all stochastic variables. According to the discussion above, section 6.1,  $\varepsilon_k$  and  $\varepsilon_\rho$  are fast variables and we assume that they can be approximated by a van Kampen type expansion

$$\varepsilon_k(t) = \delta_k(t) + N^{-1/2}\lambda_k(t) + N^{-1}\nu_k(t) + \dots, \quad (92)$$

$$\varepsilon_\rho(t) = \delta_\rho(t) + N^{-1/2}\lambda_\rho(t) + \dots, \quad (93)$$

$\delta_k$  and  $\delta_\rho$  being the deterministic equations (87) and (89), while  $\lambda_k, \lambda_\rho, \nu_k$  are stochastic. The order  $O(N^{-1})$  of  $\varepsilon_\rho$  will not be necessary for the expansion procedure in this case.

We now proceed to find the statistical properties of  $\varepsilon_k$  and  $\varepsilon_\rho$ . Introducing equations (90)–(93) in the equations for the moments (29)–(33) and equating powers of  $N$ , we find

$$\frac{d\langle\lambda_k\rangle}{dt} = -2(a + h\xi)\langle\lambda_k\rangle, \quad (94)$$

$$\frac{d\langle\lambda_\rho\rangle}{dt} = -J_{\rho,\rho}\langle\lambda_\rho\rangle + \frac{4h}{\mu}\xi(1 - 2\xi)\sum_k P_k\langle\lambda_k m_L\rangle, \quad (95)$$

$$\frac{d\langle\nu_k\rangle}{dt} = -2(a + h\xi)\langle\nu_k\rangle - 2h\left\langle\frac{\lambda_k\lambda_\rho}{1 - m_L^2}\right\rangle, \quad (96)$$

for the average values and

$$\frac{d\langle\lambda_k m_L\rangle}{dt} = -2(2a + h\xi)\langle\lambda_k m_L\rangle, \quad (97)$$

$$\begin{aligned}
\frac{d\langle\lambda_k\lambda_{k'}\rangle}{dt} + \frac{d\langle\nu_k m_L\rangle}{dt} + \frac{d\langle\nu_{k'} m_L\rangle}{dt} + N\frac{d\langle m_L^2\rangle}{dt} = & -2(2a + h\xi)\langle\nu_k m_L\rangle - 2(2a + h\xi)\langle\nu_{k'} m_L\rangle \\
& - 4aN\langle m_L^2\rangle - 4(a + h\xi)\langle\lambda_k\lambda_{k'}\rangle + \frac{4}{P_k}(a + h\xi(1 - \langle m_L^2\rangle))\delta_{k,k'} \\
& - 2h\left\langle\frac{(\lambda_k + \lambda_{k'})\lambda_\rho m_L}{1 - m_L^2}\right\rangle,
\end{aligned} \tag{98}$$

for the correlations, where, for simplicity and brevity of the expressions, we have assumed that the deterministic contributions  $\delta_k(t)$ ,  $\delta_\rho(t)$  are well captured by the linear system of equations (85) and (86) and that  $\delta_k(0) = \delta_\rho(0) = 0$ . An important difference with respect to the expansion around the deterministic solution, equations (63)–(64), is that the equations for the moments are not closed, as indicated by the presence of the last term in equation (98). Nevertheless, we will argue later on in this same section that this term can be neglected at this level of approximation.

From these equations, it follows directly that  $\langle\lambda_k\rangle_{st} = \langle\lambda_\rho\rangle_{st} = \langle\lambda_k m_L\rangle_{st} = 0$ . Furthermore, due to the up–down symmetry of the model, we know that  $\langle\nu_k\rangle_{st} = 0$ , and thus it follows that  $\left\langle\frac{\lambda_k\lambda_\rho}{1 - m_L^2}\right\rangle_{st} = 0$ . Performing in equation (98) the double sum  $\sum_{k,k'} kk' P_k P_{k'}$  and taking into account that, by definition,  $\sum_k k P_k \varepsilon_k(t) = 0$  (the same identity holds for  $\delta_k(t)$ ,  $\lambda_k(t)$  and  $\nu_k(t)$  separately), we find

$$\frac{d\langle m_L^2\rangle}{dt} = -4\left(a + \frac{\mu_2 h\xi}{\mu^2 N}\right)\langle m_L^2\rangle + \frac{\mu_2 4(a + h\xi)}{\mu^2 N}, \tag{99}$$

which, together with equation (51), proves that  $m_L(t)$  behaves as an autonomous stochastic variable. Furthermore, its first and second moments fulfill equivalent equations to those of  $m(t)$  in the all-to-all scenario (see equations (23) and (24)) if we change  $h \rightarrow 2h\xi$  and  $N \rightarrow N_{\text{eff}}$ . The stationary solution reads

$$\langle m_L^2\rangle_{st} = \frac{a + h\xi}{aN_{\text{eff}} + h\xi}, \tag{100}$$

and the transient regime

$$\langle m_L(t)^2\rangle = \langle m_L^2\rangle_{st} + [\langle m_L(0)^2\rangle - \langle m_L^2\rangle_{st}]e^{-4\left(a + \frac{h\xi}{N_{\text{eff}}}\right)t}. \tag{101}$$

Note that  $aN\langle m_L^2\rangle_{st} = O(1)$ , implying that the order of equation (98) is consistently  $O(1)$  and also supporting the reasoning of appendix C. As shown in appendix D, it is also possible to obtain these statistical properties starting from a closed master equation for the link magnetization  $m_L(t)$ .

Performing in equation (98) the partial sum  $\sum_{k'} k' P_{k'}$  we obtain an expression for  $\langle\nu_k m_L\rangle$  that depends on  $\langle m_L^2\rangle$ ,

$$\begin{aligned}
\frac{d\langle\nu_k m_L\rangle}{dt} + N\frac{d\langle m_L^2\rangle}{dt} = & -4aN\langle m_L^2\rangle - 2(2a + h\xi)\langle\nu_k m_L\rangle \\
& + \frac{4k}{\mu}(a + h\xi(1 - \langle m_L^2\rangle)) - 2h\left\langle\frac{\lambda_k\lambda_\rho m_L}{1 - m_L^2}\right\rangle.
\end{aligned} \tag{102}$$

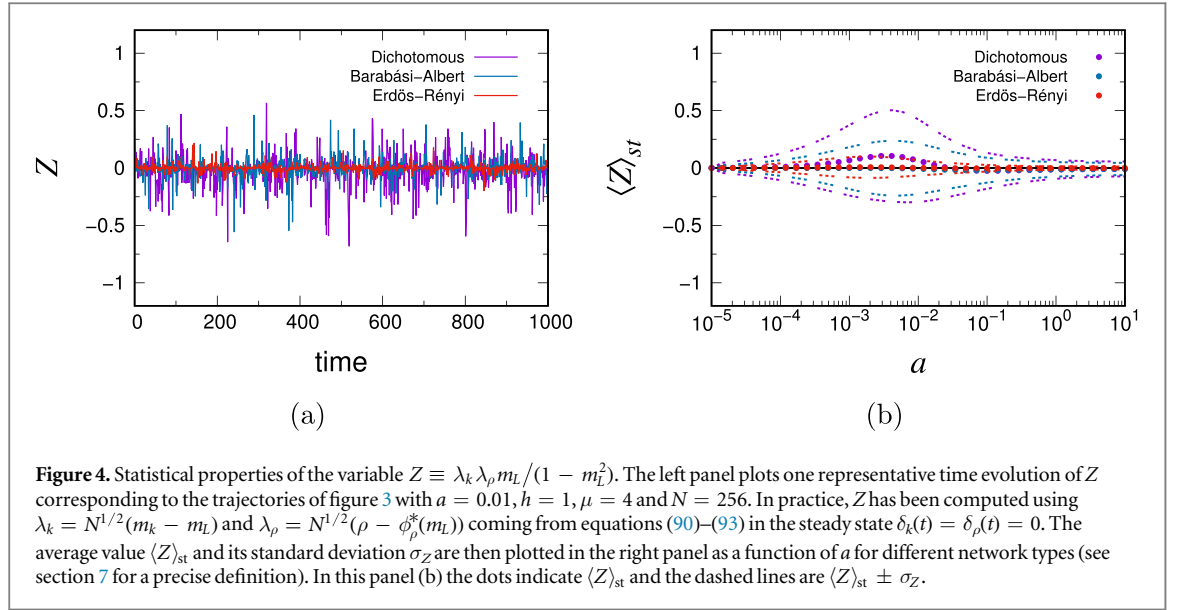
We now set

$$\left\langle\frac{\lambda_k\lambda_\rho m_L}{1 - m_L^2}\right\rangle_{st} \approx 0. \tag{103}$$

A justification of this assumption relies on the use of the van Kampen expansion equations (63), (64), that determines that at order  $O(N^{-1/2})$  the average of equation (103) is proportional to  $\langle\gamma_k\gamma_{k'}\gamma_\rho\rangle$ , which is zero as the stochastic variables  $\gamma_a$  are Gaussian [54]. This analysis is strictly valid only in the regime of large  $a \gg a_c$ , but in figure 4 we provide numerical evidence of its validity in the whole range of parameters of  $a$ . We observe that the assumption works very well for the Erdős–Rényi and Barabási–Albert networks, while for the dichotomous network there is a narrow parameter region around the critical point where the average equation (103) is non-zero but still negligible.

Once this term has been neglected, the solution of equation (102) in the stationary state is

$$\langle\nu_k m_L\rangle_{st} = -\frac{2aN + 2h\xi\frac{k}{\mu}}{2a + h\xi}\langle m_L^2\rangle_{st} + \frac{2k}{\mu}\frac{a + h\xi}{2a + h\xi}. \tag{104}$$



**Figure 4.** Statistical properties of the variable  $Z \equiv \lambda_k \lambda_{\rho} m_L / (1 - m_L^2)$ . The left panel plots one representative time evolution of  $Z$  corresponding to the trajectories of figure 3 with  $a = 0.01$ ,  $h = 1$ ,  $\mu = 4$  and  $N = 256$ . In practice,  $Z$  has been computed using  $\lambda_k = N^{1/2}(m_k - m_L)$  and  $\lambda_{\rho} = N^{1/2}(\rho - \phi_{\rho}^*(m_L))$  coming from equations (90)–(93) in the steady state  $\delta_k(t) = \delta_{\rho}(t) = 0$ . The average value  $\langle Z \rangle_{st}$  and its standard deviation  $\sigma_Z$  are then plotted in the right panel as a function of  $a$  for different network types (see section 7 for a precise definition). In this panel (b) the dots indicate  $\langle Z \rangle_{st}$  and the dashed lines are  $\langle Z \rangle_{st} \pm \sigma_Z$ .

Replacing this result in equation (98), we obtain the following correlations in the stationary state

$$\langle \lambda_k \lambda_{k'} \rangle_{st} = -\frac{k + k'}{\mu} + \frac{aN\mu + h\xi(k + k')}{\mu(a + h\xi)} \langle m_L^2 \rangle_{st} + \frac{\delta_{k,k'}}{P_k} \left( 1 - \frac{h\xi}{a + h\xi} \langle m_L^2 \rangle_{st} \right). \quad (105)$$

We have now all necessary ingredients to calculate the first and second moments of the global magnetization  $m = \sum_k P_k m_k$ , both in the stationary and transient regimes. For example, the stationary variance is

$$\sigma_{st}^2[m] = \langle m_L^2 \rangle_{st} + N^{-1} \left( 2 \sum_k P_k \langle \nu_k m_L \rangle_{st} + \sum_{k,k'} P_k P_{k'} \langle \lambda_k \lambda_{k'} \rangle_{st} \right), \quad (106)$$

which leads to

$$N\sigma_{st}^2[m] = \frac{h^2\xi^2}{2a + h\xi} \frac{N}{aN_{\text{eff}} + h\xi} + \frac{2h\xi}{2a + h\xi} + \frac{aN_{\text{eff}}}{aN_{\text{eff}} + h\xi}. \quad (107)$$

The critical point is found by solving  $\sigma_{st}^2[m] = \frac{N+2}{3N}$ , taking into account that  $\xi$  depends on  $a$ ,

$$a_c = \frac{h}{N_{\text{eff}}} \frac{\mu - 2}{\mu - 1} + O(N^{-2}). \quad (108)$$

The average value of  $\rho$  can be determined in the stationary regime as

$$\langle \rho \rangle_{st} = \xi(1 - \langle m_L^2 \rangle_{st}) + O(N^{-1}). \quad (109)$$

At variance with the results of the expansion around the deterministic solution (see equations (81) and (82)), the above expression is not divergent for  $a \rightarrow 0$ .

It is important to stress the similarities and differences of our stochastic PA method with the one developed in [43] for the noiseless voter model. In that paper the authors use a master equation approach for the link magnetization  $m_L$  which can readily be extended to the noisy case (see appendix D) yielding for  $m_L$  the same statistical properties as the ones obtained by our more sophisticated approach. However, the authors of [43] make, implicitly, the extra assumption that the link magnetization  $m_L$  and the node magnetization  $m$  share the same statistical properties. To check the validity of this assumption, and using  $\langle mm_L \rangle_{st} = \langle m_L^2 \rangle_{st} + N^{-1} \sum_k P_k \langle \nu_k m_L \rangle_{st}$ , we obtain

$$\langle (m - m_L)^2 \rangle_{st} = \left( 1 - \frac{N_{\text{eff}}}{N} \right) \frac{a}{a + h\xi} \langle m_L^2 \rangle_{st}, \quad (110)$$

a function of the degree heterogeneity and the parameters of the model (the ratio of time scales). From this expression we conclude that it is true that in the noiseless voter model, with  $a = 0$ ,  $\langle (m - m_L)^2 \rangle_{st}$  becomes zero (rather, of order  $N^{-1}$ , which is the accuracy of our expansion), validating the approach of [43] in that case. Remarkably, in the limit of a homogeneous network, with  $N_{\text{eff}} = N$ , the statistical properties of  $m$  and  $m_L$  also coincide, consistently with the fact that  $m$  and  $m_L$  are the same quantity for those networks. Furthermore, in the critical zone  $a = O(N^{-1})$  and for a heterogeneous network we obtain that  $\langle (m - m_L)^2 \rangle_{st} = O(N^{-1})$ , meaning that when there is time scale separation  $m$  does not differ significantly from  $m_L$ . Our analysis shows that in the

noisy voter model the assumption that the statistical properties of the magnetization are those of the link magnetization cannot be held for heterogeneous networks far away from the critical region, while for the voter model without noise ( $a = 0$ ) the statistical differences between both quantities turn out to be negligible. The relevance of this difference for other stochastic processes defined in heterogeneous networks is under study.

## 7. Comparison with numerical simulations

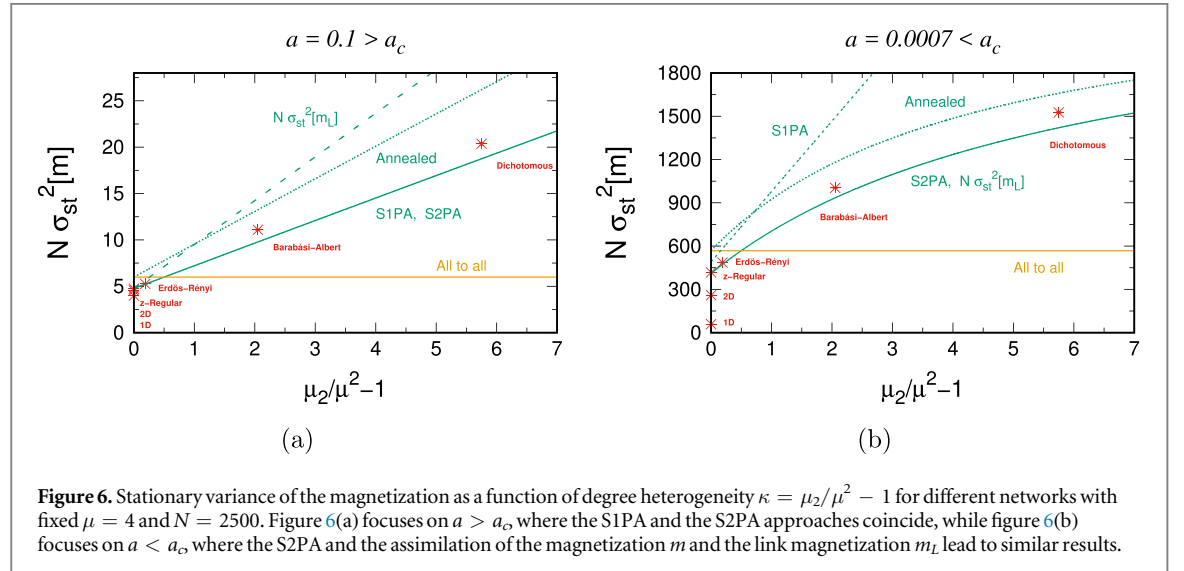
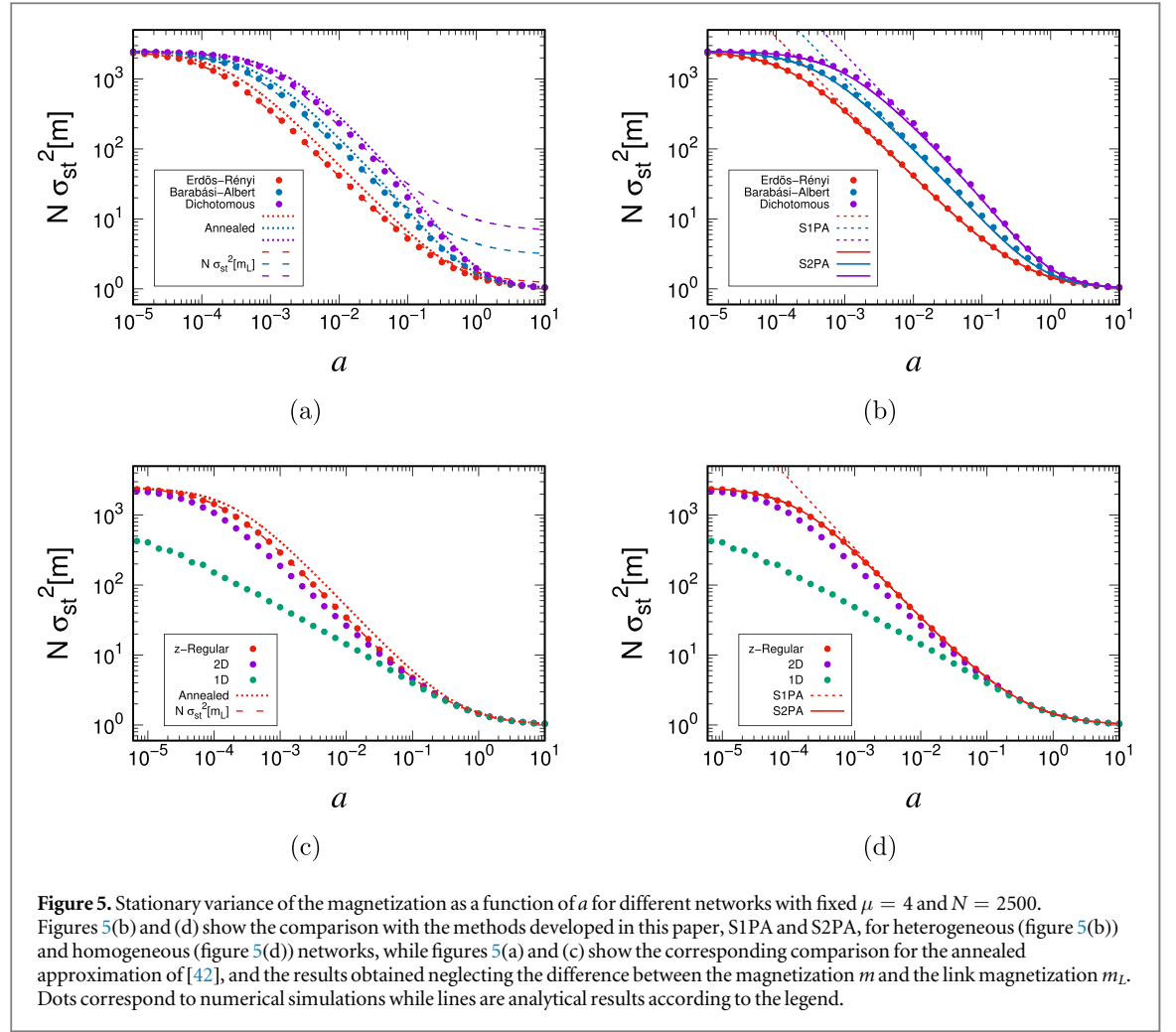
We now compare the different theoretical approaches based on the stochastic PA and others with the results of numerical simulations. We focus on the steady state values  $\sigma_{\text{st}}^2[m]$  and  $\langle \rho \rangle_{\text{st}}$  as functions of the noise parameter  $a$  and the degree heterogeneity  $\kappa \equiv \mu_2/\mu^2 - 1$ , for six different network structures keeping a constant value for the average connectivity  $\mu = 4$  and the total number of nodes  $N = 2500$ , namely:

- (1) an Erdős–Rényi random network with a fixed number of links  $\mu N/2$  that are connected randomly between pairs of nodes, thus leading to a degree heterogeneity  $\kappa = 0.19$ ;
- (2) a Barabási–Albert preferential-attachment network, with a power-law degree distribution  $P_k \sim k^{-3}$  and a degree heterogeneity  $\kappa = 3.05$ ;
- (3) a dichotomous network where a fraction  $23/24$  of nodes has  $k = 2$  neighbors and the remaining fraction  $1/24$  has  $k = 50$  neighbors, leading to a degree heterogeneity  $\kappa = 5.75$ ;
- (4) a  $z$ -regular network where each node has exactly  $k = 4$  neighbors but the connections amongst them are chosen at random;
- (5) a one-dimensional (1D) linear lattice where a node is connected to its closest 4 neighbors (2 to the right and 2 to the left);
- (6) and a two-dimensional square lattice (2D) where each node is connected to its 4 nearest neighbors.

The last three of these networks are called ‘homogeneous’, as  $\kappa = 0$ . For each one of these networks we plot the theoretical predictions for the variance of the magnetization  $\sigma_{\text{st}}^2[m]$  and the average density of active links  $\langle \rho \rangle_{\text{st}}$  in the steady state as given by:

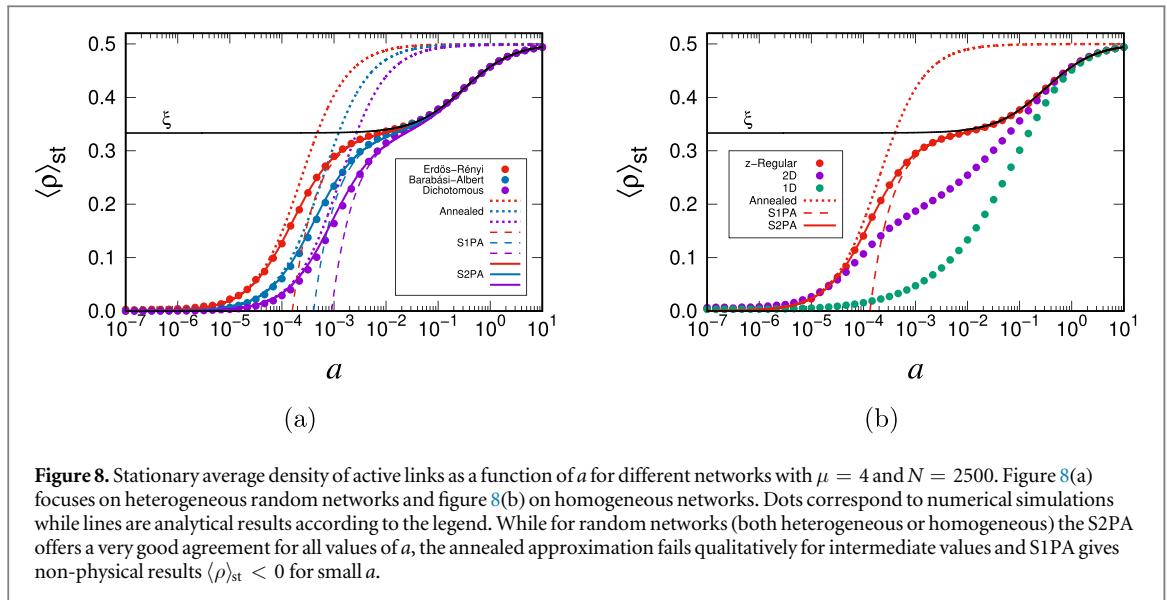
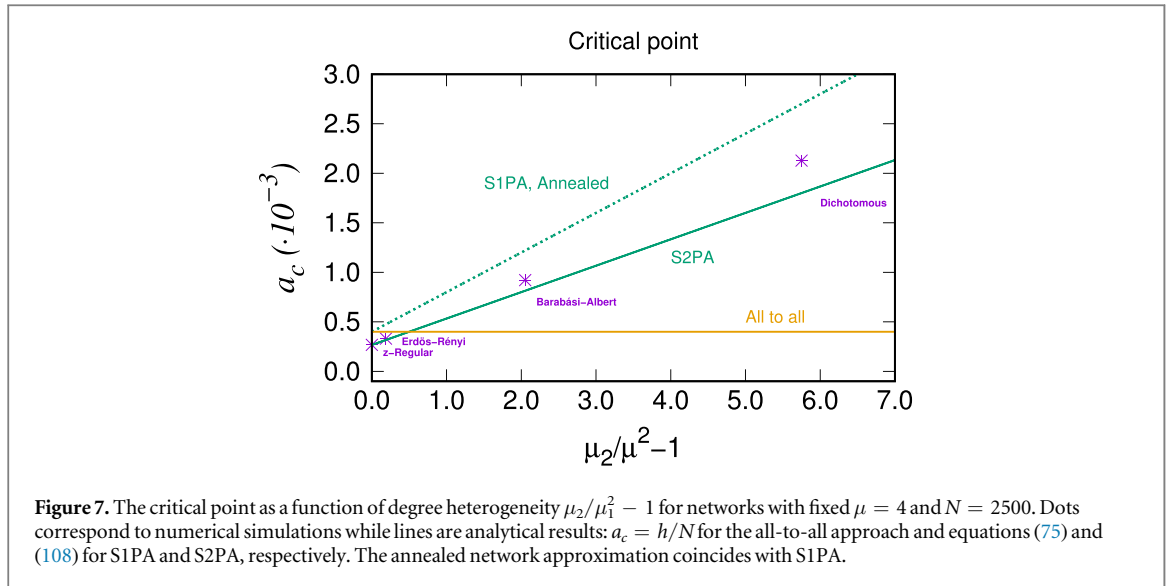
- (a) the global-state-approach as presented in equations (26) and (27), which is independent of all network details (the comparison has already been presented in figure 2);
- (b) the annealed approximation as developed in [42];
- (c) the expansion around the deterministic solution (S1PA) as given by equations (74) and (81);
- (d) the expansion around the dynamical attractor (S2PA) as given by equations (107) and (109);
- (e) and, for comparison, we have also included the variance  $\sigma_{\text{st}}^2[m_L]$ , which would have been the result of extending the theory developed at [43] to the noisy voter model, neglecting the difference between the magnetization  $m(t)$  and the link magnetization  $m_L(t)$ .

In figure 5 we plot  $N\sigma_{\text{st}}^2[m]$  as a function of the noise parameter  $a$ . For clarity, we have split the comparison in several panels. We can appreciate in figures 5(b) and (d) how S2PA reproduces very accurately the numerical results for the heterogeneous and the  $z$ -regular networks in the whole range of values of  $a$ , both above and below the critical point  $a_c$ . The results of S1PA are indistinguishable from those of S2PA above the critical point  $a > a_c$ , while below it  $a < a_c$  the S1PA prediction for  $N\sigma_{\text{st}}^2[m]$  diverges as  $a \rightarrow 0$ , at odds with S2PA and the numerical simulations. We expect a full agreement between these two approaches only in the thermodynamic limit, in which  $N \rightarrow \infty$  with  $a$  fixed, as equations (107) and (74) would then coincide. The data is also analyzed in figure 6 as a function of the degree heterogeneity  $\kappa$ . While for  $a > a_c$  (figure 6(a))  $\sigma_{\text{st}}^2[m]$  is proportional to the degree heterogeneity in accordance with equation (74), for  $a < a_c$  (figure 6(b)) the dependence is not linear but well captured by equation (107). On the other hand,  $N\sigma_{\text{st}}^2[m_L]$  coincides with  $N\sigma_{\text{st}}^2[m]$  only below the critical point  $a < a_c$  or for homogeneous networks  $\kappa \rightarrow 0$ , where the difference between  $m$  and  $m_L$  is irrelevant, as discussed in the previous section. The annealed approximation developed in [42] reproduces correctly both the dependence on  $a$  and  $\kappa$ , but it is less accurate in general than the PA (S2PA). The global-state approach equation (26) has the correct scaling, but it is accurate only for networks with low levels of heterogeneity and does not predict any dependence with  $\kappa$ . In the limit  $\mu_2 = \mu^2 \rightarrow \infty$ , and for all values of  $a$  below and above the critical point, the relative difference between equations (107) and (26) is of order  $N^{-1}$ . We also check the predicted



values for the critical point  $a_c$  in figure 7. Again, the S2PA approach offers the best fit to the numerical data, while the S1PA and the annealed approximations coincide for all values of the degree heterogeneity.

In figure 8 we plot the results of  $\langle \rho \rangle_{st}$  as a function of  $a$ . We have split the comparison for heterogeneous random (figure 8(a)) and homogeneous (figure 8(b)) networks. We can observe again a very high degree of accuracy of S2PA for random networks in the whole range of values of  $a$ . The results using the expansion around the deterministic solution S1PA are accurate above the critical point, but below it they show a divergence leading to non-physical results  $\rho < 0$ . Comparing with other treatments, such as the annealed and all-to-all



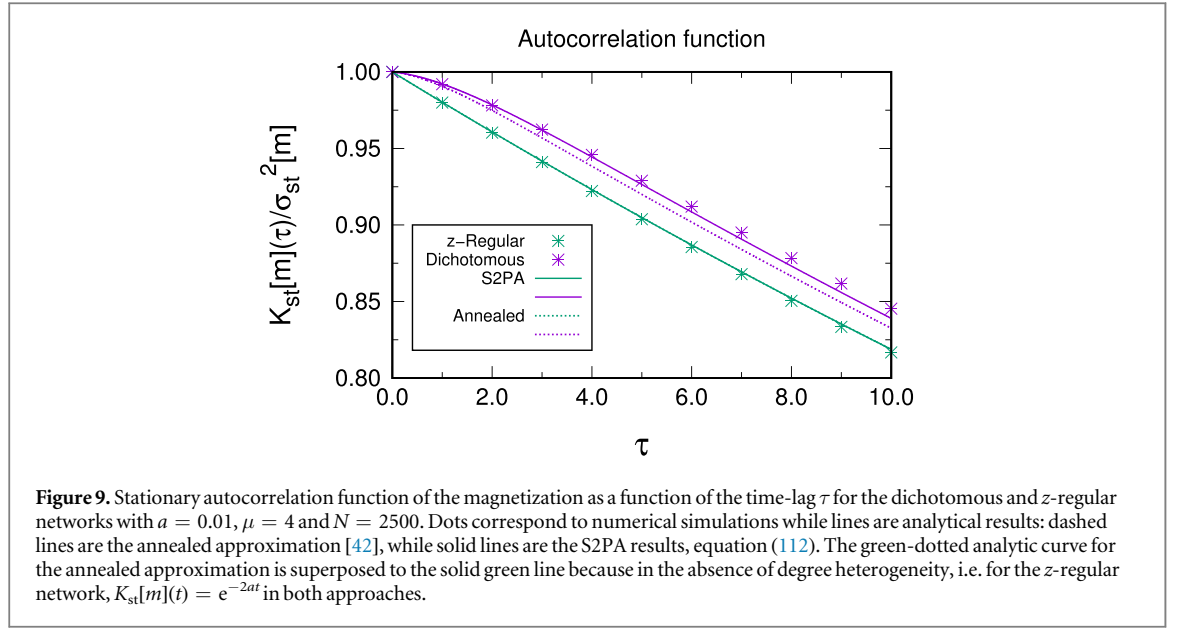
approximations, it is clear that the stochastic PA (S2PA) developed here perfectly captures the shape of the curve  $\langle \rho \rangle_{st}$  versus  $a$ , while the annealed approximation discussed in [42] fails even qualitatively. This failure of the annealed and all-to-all approximations is directly related to the fact that the average degree  $\mu$  is finite, as one can prove that in the limit  $\kappa \rightarrow 0$  equations (109) and (27) coincide only when  $\mu \rightarrow \infty$  and thus  $\xi \rightarrow 1/2$ . Note that the expression of S2PA for  $\langle \rho \rangle_{st}$  given at equation (109) is independent on whether the spin  $m$  and link  $m_L$  magnetizations are assumed to be equal or not, and thus this approximation leads to the same result as the extension of the theory developed at [43] for  $\langle \rho \rangle_{st}$ .

Both the variance of the magnetization and the average density of active links show significant discrepancies between the results of numerical simulations and the proposed analytical methods in the case of regular networks of low dimensionality  $d = 1, 2$ . This is specially notorious for the 1D case, where even the scaling  $\langle m^2 \rangle_{st} \sim O(N^0)$  if  $a \sim O(N^{-1})$  does not apply and, instead, it is  $\langle m^2 \rangle_{st} \sim O(N^{-1/2})$ . This is the reason why there is no finite-size critical point in figure 7, as defined previously, for the 1D and 2D cases. As the critical dimension of the noisy voter model is  $d = 2$ , for any dimension above this value the same scaling properties are expected, in accordance to the all-to-all result. A detailed analysis of the finite-size scaling for regular lattices will be published elsewhere [72].

## 8. Time dependence

The method S2PA, developed in the previous sections, allows us to determine the time evolution of  $\langle m(t) \rangle$  and  $\langle \rho(t) \rangle$ . From the construction of the variables we have  $\langle m_k(t) \rangle = \delta_k(t) + \langle m_L(t) \rangle$ , which leads to





$$\langle m(t) \rangle = m_L(0)e^{-2at} + [m(0) - m_L(0)]e^{-2(a+h\xi)t}, \quad (111)$$

assuming that we start with a fixed initial condition  $m(0)$  and  $m_L(0)$ . A preferable quantity to measure this time evolution, that does not depend on the initial conditions, is the stationary autocorrelation function

$$K_{st}[m](t) \equiv \langle m(t)m(0) \rangle_{st} = \langle mm_L \rangle_{st} e^{-2at} + [\langle m^2 \rangle_{st} - \langle mm_L \rangle_{st}] e^{-2(a+h\xi)t}. \quad (112)$$

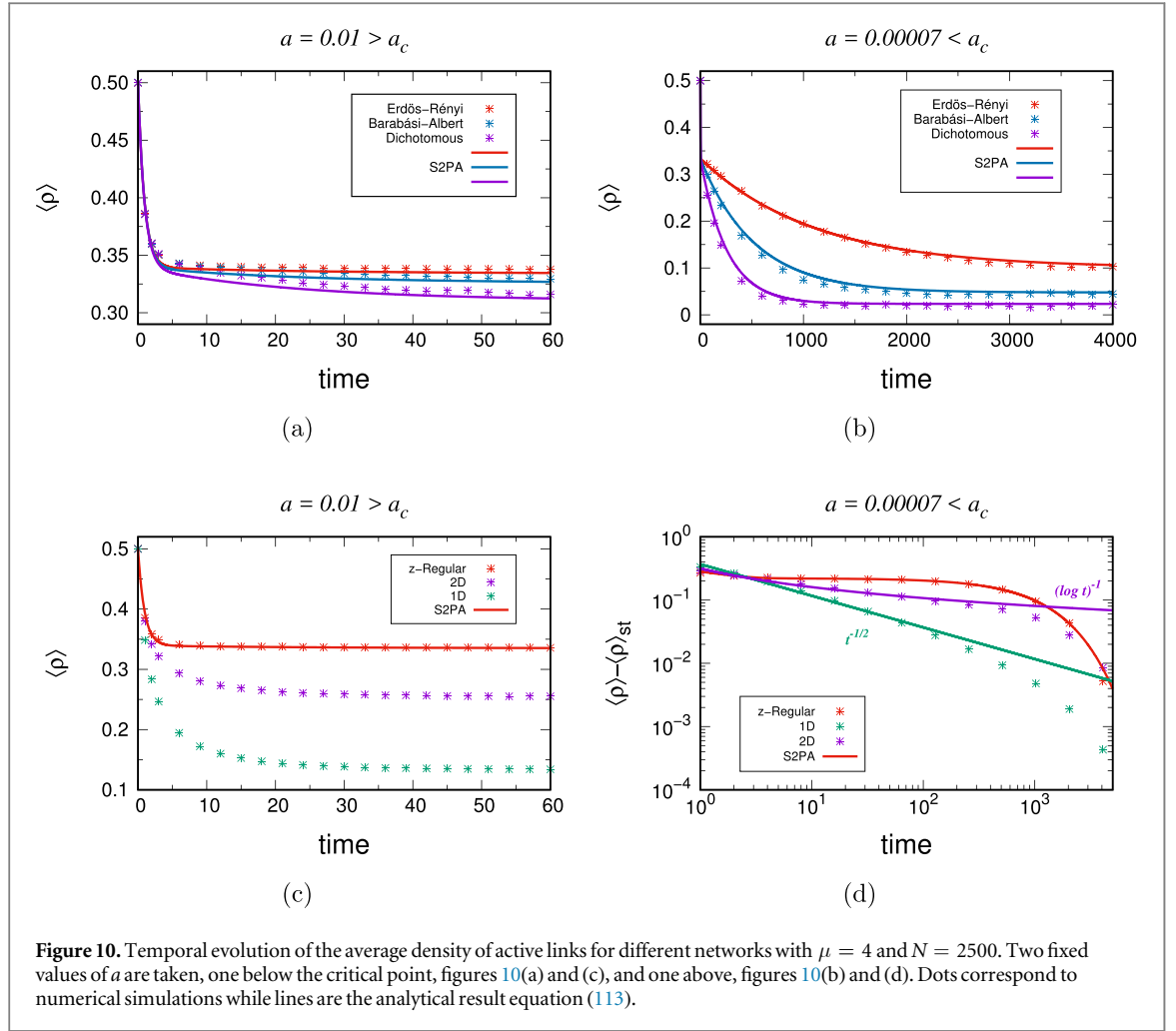
The existence of two exponentials in this expression is directly related to the degree heterogeneity, as already pointed out with the annealed approximation in [42]. In figure 9 we plot the time evolution of  $K_{st}[m](t)$  for the two extreme cases of a network with no degree heterogeneity (z-regular random network) and a highly heterogeneous degree distribution (dichotomous network). While for the z-regular network there is only one slow exponential, for the dichotomous network the early stages are dominated by the fast exponential and later the slow part dominates, in accordance to equation (112). Hence, by fitting a double exponential to the autocorrelation data we are then capable of obtaining information about the network heterogeneity.

While it has not been studied before in the context of the noisy version of the model, the time evolution of the interface density of active links  $\rho$  is commonly used to describe the dynamics of the (noiseless,  $a = 0$ ) voter model. In that case, the voter model is characterized by the existence of two absorbing states ( $m = -1$  and  $m = 1$ ), both of them corresponding to full order or consensus ( $\rho = 0$ ). Therefore, the focus is on how the system approaches consensus and if it does so in the thermodynamic limit. In particular, the ordering process is found to be determined by the dimensionality  $d$  of the underlying topology. For  $d \leq 2$ , and in the infinite size limit, the system orders by a coarsening process, i.e., by a continuous and unbounded growth of spatial domains of one of the opinions. The ensemble average interface density decays as  $\langle \rho \rangle \sim t^{-1/2}$  for  $d = 1$  and as  $\langle \rho \rangle \sim (\log t)^{-1}$  for  $d = 2$ . After an initial transient, infinite systems with  $d > 2$  are observed to fall into a metastable, highly disordered state ( $0 < \rho < 1/2$ ) where coarsening processes have stopped. In these cases, the average interface density behaves as  $\langle \rho \rangle \sim b - c \cdot t^{-d/2}$ , and thus complete order is never reached. Finite-size systems, on the contrary, are always led to complete order by finite-size fluctuations. As a consequence, finite systems follow the above described behaviors only for a finite time, after which an exponential decay to complete order is observed. Notably, it has been shown that this characteristic decay time is proportional to the effective system size  $N_{\text{eff}}$  (for  $d > 2$ ).

According to the theoretical description presented in the previous section 6, the time dependence of  $\langle \rho(t) \rangle$  can be determined as

$$\langle \rho(t) \rangle = \delta_\rho(0)e^{-J_{\rho,\rho}t} + \xi(1 - \langle m_L(t)^2 \rangle) + O(N^{-1}), \quad (113)$$

where  $\delta_\rho(0) = \langle \rho(0) \rangle - \xi$  and, for simplicity, we assume  $\langle m_k(0) \rangle = \langle m_L(0) \rangle = 0$  so that the term  $\delta_\rho(0)e^{-J_{\rho,\rho}t}$  is the only one that remains from equation (89). The form of this expression is in line with the qualitative description given in [64], where the average interface density is said to be characterized by a short initial transient after which a plateau is reached at a value  $\rho^* = \xi > 0$ . In a more quantitative description, equation (113) contains the initial transient  $\delta_\rho(0)e^{-J_{\rho,\rho}t}$ , the exponential decay due to finite-size effects after the plateau  $\xi \langle m_L^2 \rangle_{st} e^{-t/\tau}$  with  $\tau = \frac{N_{\text{eff}}}{4(aN_{\text{eff}} + h\xi)}$  and the final steady state described by equation (109). Note how the decay time of the initial transient is independent of the degree heterogeneity, unlike the one due to finite-size effects.



**Figure 10.** Temporal evolution of the average density of active links for different networks with  $\mu = 4$  and  $N = 2500$ . Two fixed values of  $a$  are taken, one below the critical point, figures 10(a) and (c), and one above, figures 10(b) and (d). Dots correspond to numerical simulations while lines are the analytical result equation (113).

For  $a < a_c$ ,  $\tau \sim N_{\text{eff}} \gg 1/J_{\rho,\rho}$ , whereas  $\tau$  and  $1/J_{\rho,\rho}$  are both of the order  $N^0$  for  $a > a_c$ . The difference between the plateau and the steady state  $\xi \langle m_L(t)^2 \rangle$  is of order  $1/N$  for  $a > a_c$  and of order  $aN$  for  $a < a_c$ .

Numerical results for the time evolution of the average order parameter are presented in figure 10 for the six different networks studied. Note that, due to their particular ordering dynamics and for comparison with the voter model, we have included both 1D and 2D regular networks. In this figure, we can clearly distinguish the two stages in the dynamics of the average interface density identified in [64] and described by equation (113): an initial ordering transient, with a significant decrease of the order parameter from its initial, close to complete disorder value  $\rho = 1/2$ , corresponding to a random initial distribution of states; and a final steady state, where the order parameter fluctuates around a characteristic level of disorder. As shown in the figure, for  $a < a_c$  the initial ordering behavior of the noisy voter model is remarkably similar to the ordering dynamics of the voter model described above: a power-law decay as  $\langle \rho \rangle \sim t^{-1/2}$  for  $d = 1$  (regular 1D), a logarithmic decay as  $\langle \rho \rangle \sim (\ln t)^{-1}$  for  $d = 2$  (regular 2D), and a fast decay to the plateau followed by a slow exponential decay  $\langle \rho \rangle \sim \rho^* e^{-t/\tau}$  for  $d > 2$  (z-regular, Erdős-Rényi, Barabási-Albert, dichotomous). Therefore, as in the voter model, the initial ordering of systems with dimension  $d \leq 2$  is characterized by a coarsening process with growth of spacial domains, while the ordering of systems with dimension  $d > 2$  occurs mainly by finite-size fluctuations after a period in a metastable, highly disordered state. We can also note that, apart from coarsening, finite-size fluctuations are present too for  $d \leq 2$ , being responsible for the exponential decay interrupting the coarsening process. It is important to note, however, that, while in the voter model any ordering process in a finite system is a transient towards complete order, in the noisy voter model it is a transient towards a non-zero steady state of the order parameter. Thus, the voter model-like initial ordering of the noisy voter model is interrupted whenever the average order parameter reaches its steady state value. In general, the behavior for  $a > a_c$  is quite different, as one can notice from the figure. This is because the influence of finite-size effects on  $\langle \rho \rangle$  becomes weaker as  $a$  is increased. Thus, the value at the plateau  $\rho = \rho^*$  is closer to the steady state solution  $\langle \rho \rangle_{st}$  and the time scale of the fast initial transient is not that different from the one of finite-size effects  $\tau$ , which mixes the two behaviors. Furthermore, for  $d \leq 2$  the previous voter-like time dependence is no longer true.

## 9. Conclusions

We have developed a full stochastic approach to the PA scheme to study binary-state dynamics on heterogeneous networks. Our starting point is the master equation for the set of variables  $\{L, N_{1,k}\}$ , being  $N_{1,k}$  the number of nodes with degree  $k$  in the state up and  $L$  the number of active links. The effective rates of the master equation are obtained using the PA as discussed in [43] in the context of the noiseless voter model. For the noisy voter model, however, we show that the spin magnetization and the link magnetization cannot be identified. Once the equations for the time evolution of the moments and correlations are formulated, we proceed to close the hierarchy by using two different closure schemes. The first one, S1PA, is based on a van Kampen type of expansion around the deterministic solution of the dynamics corresponding to the thermodynamic limit, while the second one, S2PA, is based on an expansion around a stochastic dynamical attractor.

In the first approach, we expand the solution around the deterministic fixed point. In this way we are able to derive a linear set of equations for the correlation matrix and first corrections to the average values. Although our methodology is very general for any set of rates, we have focused on the noisy voter model, for which we have been able to carry out a full analytical treatment. This first analysis, based on the ansatz that the fluctuations of the magnetization scale with system size as  $\sigma_{\text{st}}^2[m] \sim N^{-1}$ , predicts that the fluctuations depend linearly on the degree heterogeneity of the network. We also obtained an expression for the average density of active links  $\langle \rho \rangle_{\text{st}}$ . Both expressions have a mathematical divergence as the noise intensity approaches  $a \rightarrow 0$ . Compared to numerical simulations, these results turn out to be very accurate well above the critical point of the noisy voter model,  $a/h \gg N^{-1}$ , but fail below or close to it. The failure of this approach is ultimately linked to the ansatz of the scaling of the correlations with system size. As we approach the critical point  $a \rightarrow 0$ , it is not true that fluctuations diverge but there is a change in the scaling with  $N$ , in the noisy voter case becoming  $\sigma_{\text{st}}^2[m] \sim N^0$ .

Our second approach is based on a system size expansion around a stochastic dynamical attractor. Although this method turns out to be more precise than the expansion around the deterministic solution and fits much better the numerical data for the noisy voter model, it is much more involved and its applicability must be analyzed for the particular model under study. Exploring the phase space of deterministic solutions, we were able to identify a special trajectory  $m(t) = m_L(t)$  and  $\rho(t) \approx \xi(1 - m_L(t)^2)$  ( $m_L$  being the link magnetization) around which the dynamics is fast and the fluctuations do fulfill a van Kampen type of scaling, namely  $\sigma_{\text{st}}^2[m - m_L] \sim N^{-1}$ , for all values of  $a$ . Using this expansion we are able to obtain expressions for all the desired quantities, not limited to the stationary values but including the dynamics of  $\langle \rho(t) \rangle$  and  $\langle m(t) \rangle$ , which match perfectly compared to numerical simulations and work better than previous approaches to the same model, such as all-to-all or annealed network approximations. We highlight that the results of  $\sigma_{\text{st}}^2[m_L]$  are the same as the all-to-all network if we rescale the herding as  $h \rightarrow 2h\xi$  and the system size as  $N \rightarrow N_{\text{eff}} = N \frac{\mu^2}{\mu_2}$  in terms of the first  $\mu$  and second  $\mu_2$  moments of the degree distribution, a result that was previously reported for the noiseless voter model. However, due to the effect of noise  $a \neq 0$ , it is not true that  $\sigma_{\text{st}}^2[m] \approx \sigma_{\text{st}}^2[m_L]$ , which implies that finite-size effects in a complex network cannot be reduced to replacing the system size  $N$  by an effective one  $N_{\text{eff}}$ .

## Acknowledgments

We acknowledge financial support from the Agencia Estatal de Investigación (AEI, Spain) and Fondo Europeo de Desarrollo Regional under project ESOTECOS FIS2015-63628-C2-2-R (MINECO/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). AFP acknowledges support by the Formación de Profesorado Universitario (FPU14/0554) program of Ministerio de Educación, Cultura y Deportes (MECD) (Spain).

## Appendix A. Solution of equations (65)–(67)

Replacing the Jacobian equations (56)–(58) and the steady state value equations (68)–(49) into equations (65)–(67), these latter can be reduced to

$$4(a + h\xi)C_{k,k'} - 2h\xi \sum_{k''} \frac{P_{k''}k''}{\mu} (C_{k,k''} + C_{k',k''}) = \frac{4(a + h\xi)}{P_k} \delta_{k,k'}, \quad (\text{A.1})$$

$$(2a + 2h\xi + J_{\rho,\rho})C_{k,\rho} - 2h\xi \sum_{k''} \frac{P_{k''}k''}{\mu} C_{k'',\rho} = 0, \quad (\text{A.2})$$

$$2C_{\rho,\rho}J_{\rho,\rho} = G_\rho. \quad (\text{A.3})$$

From equation (A.3) we get  $C_{\rho,\rho} = G_\rho/2J_{\rho,\rho}$ , while from equation (A.2), by inspection, we note that  $C_{k,\rho}$  is a constant that does not depend on  $k$ . Reinserting this independence of  $k$  back into the equation, we notice that the

only possibility is the trivial solution  $C_{k,\rho} = 0$ . Equation (A.1) is more involved. Let us define  $D_k = \sum_{k'} \frac{P_k k'}{\mu} C_{k,k'}$  and  $D = \sum_k \frac{P_k k}{\mu} D_k$ , so the equation becomes

$$4(a + h\xi) C_{k,k'} - 2h\xi(D_k + D_{k'}) = \frac{4(a + h\xi)}{P_k} \delta_{k,k'}. \quad (\text{A.4})$$

We now perform the double sum  $\sum_k \frac{P_k k}{\mu} \sum_{k'} \frac{P_{k'} k'}{\mu}$ , obtaining in this way a closed relation for  $D$

$$D = \frac{a + h\xi}{a} \frac{\mu_2}{\mu^2}. \quad (\text{A.5})$$

If we perform instead the single sum  $\sum_{k'} \frac{P_{k'} k'}{\mu}$ , we obtain an expression for  $D_k$  as a function of  $D$

$$D_k = \frac{h\xi}{2a + h\xi} D + \frac{k}{\mu} \frac{2(a + h\xi)}{2a + h\xi}, \quad (\text{A.6})$$

and we then solve  $C_{k,k'}$

$$C_{k,k'} = \frac{h\xi}{2(a + h\xi)} (D_k + D_{k'}) + \frac{1}{P_k} \delta_{k,k'}. \quad (\text{A.7})$$

Finally, replacing equation (A.5) into (A.6) and this latter into (A.7), we obtain equation (73).

## Appendix B. A linear noise example

Consider the simple example of two coupled linear stochastic equations for variables  $x(t)$  and  $y(t)$

$$\dot{x} = -\frac{1}{\tau_x} x + N^{-1/2} \xi_x(t), \quad (\text{B.1})$$

$$\dot{y} = -\frac{1}{\tau_y} y + \frac{1}{\tau_y} (x - y) + N^{-1/2} \xi_y(t), \quad (\text{B.2})$$

where  $\xi_{x,y}(t)$  are two uncorrelated Gaussian white noise variables  $\langle \xi_a(t) \xi_b(t') \rangle = \delta_{a,b} \delta(t - t')$ . The  $x$  variable is independent of  $y$  and has a stationary variance  $\langle x^2 \rangle_{\text{st}} = \tau_x / (2N)$ . This variance fulfills the van Kampen ansatz as long as  $\tau_x = O(1)$ , while when the time scale diverges  $\tau_x = O(N^\alpha)$  with  $0 \leq \alpha \leq 1$ , the variance has an anomalous scaling  $\langle x^2 \rangle_{\text{st}} = O(N^{\alpha-1})$ . The deterministic part of the equation has the solution  $y(t) = x(t)$  which is an attractor of trajectories. We then propose to split the  $y$  variable as  $y(t) = x(t) + \varepsilon(t)$ , with two stochastic parts: a slow one  $x(t)$ , based on the solution of the deterministic attractor, and a fast one  $\varepsilon(t)$ . Introducing this change of variables in equation (B.2), we find

$$\dot{\varepsilon} = -\left( \frac{1}{\tau_x} + \frac{1}{\tau_y} \right) \varepsilon + N^{-1/2} \xi_y(t) - N^{-1/2} \xi_x(t), \quad (\text{B.3})$$

whose variance is  $\langle \varepsilon^2 \rangle_{\text{st}} = N^{-1} \frac{\tau_x \tau_y}{\tau_x + \tau_y}$ . Hence, as long as  $\tau_y = O(1)$ , and regardless of the scaling of  $\tau_x$ , the variance of  $\varepsilon$  always scales as  $\langle \varepsilon^2 \rangle_{\text{st}} = O(N^{-1})$ , which justifies the use of a van Kampen type of expansion for this variable in the whole range of values of  $\tau_x$ .

## Appendix C. Solution of equation (83)

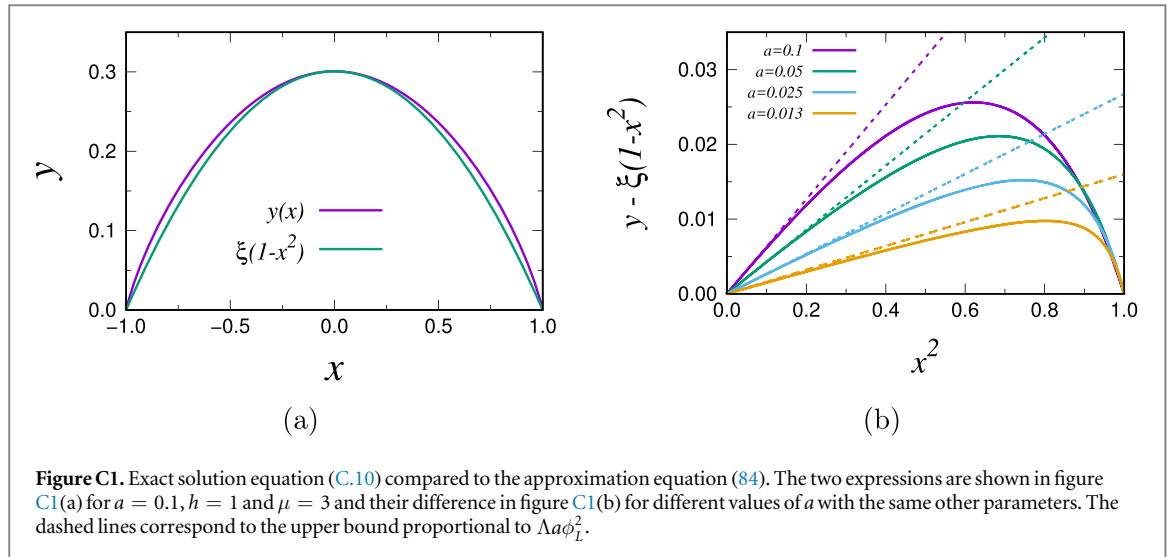
The equation is

$$-\phi_L \frac{d\phi_\rho^*}{d\phi_L} = 1 + \alpha \phi_\rho^* - \beta \frac{\phi_\rho^{*2}}{1 - \phi_L^2}, \quad (\text{C.1})$$

with  $\alpha = \frac{h}{a}(1 - 2/\mu) - 2$ ,  $\beta = \frac{2h}{a}(1 - 1/\mu)$  and boundary conditions  $\phi_\rho^*(\phi_L = \pm 1) = 0$ . Note that, the value  $\phi_\rho^*(\phi_L = 0) \equiv y_0$  is determined by the (positive) solution of  $1 + \alpha y_0 - \beta y_0^2 = 0$  or  $\phi_\rho^*(0) = \xi$ , where  $\xi$  is given by equation (54).

To find the general solution of the Ricatti-type equation (C.1) we follow a standard procedure. We first introduce the change of variables  $s = \phi_L^2$ , which reduces the equation to

$$-2s \frac{d\phi_\rho^*}{ds} = 1 + \alpha \phi_\rho^* - \beta \frac{\phi_\rho^{*2}}{1 - s}. \quad (\text{C.2})$$



**Figure C1.** Exact solution equation (C.10) compared to the approximation equation (84). The two expressions are shown in figure C1(a) for  $a = 0.1$ ,  $h = 1$  and  $\mu = 3$  and their difference in figure C1(b) for different values of  $a$  with the same other parameters. The dashed lines correspond to the upper bound proportional to  $\Lambda a \phi_L^2$ .

Then the change of variables

$$\phi_\rho^*(s) = -\frac{2s(1-s)}{\beta z(s)} \frac{dz(s)}{ds}, \quad (\text{C.3})$$

leads to the second order linear differential equation

$$\frac{d^2 z}{ds^2} + \frac{2 + \alpha - (4 + \alpha)s}{2s(1-s)} \frac{dz}{ds} - \frac{\beta}{4(1-s)s^2} z = 0, \quad (\text{C.4})$$

whose solution is

$$z(s) = Cs^{\lambda_1} F(\lambda_1, 1 + \lambda_2; 1 + \lambda_1 + \lambda_2; s) + s^{-\lambda_2} F(-\lambda_2, 1 - \lambda_1; 1 - \lambda_1 - \lambda_2; s), \quad (\text{C.5})$$

where  $C$  is an integration constant (another irrelevant global multiplicative constant has been set equal to one).  $F(a, b; c; s)$  is the Gauss hypergeometric function and we have introduced the notation

$$\lambda_1 = \frac{1}{4}(-\alpha + \sqrt{\alpha^2 + 4\beta}), \quad (\text{C.6})$$

$$\lambda_2 = \frac{1}{4}(\alpha + \sqrt{\alpha^2 + 4\beta}) = \frac{\beta}{2}\xi. \quad (\text{C.7})$$

We note that, independently of the value of the constant  $C$ , the solution verifies  $\phi_\rho^*(0) = \xi$ . The constant  $C$  is obtained by demanding that  $\phi_\rho^*(1) = 0$ . Using the known expansion [73] of the hypergeometric function

$$F(a, b; a + b; s) = -\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \log(1-s) + O((1-s)^1), \quad (\text{C.8})$$

which is valid for  $s \rightarrow 1$ , we arrive at

$$C = -\frac{\Gamma(1 - \lambda_1 - \lambda_2)\Gamma(\lambda_1)\Gamma(1 + \lambda_2)}{\Gamma(1 + \lambda_1 + \lambda_2)\Gamma(1 - \lambda_1)\Gamma(-\lambda_2)}. \quad (\text{C.9})$$

The solution is finally written as

$$\phi_\rho^*(\phi_L) = \xi(1 - \phi_L^2)G(\phi_L^2), \quad (\text{C.10})$$

where

$$G(s) = -\frac{sz'(s)}{\lambda_2 z(s)}. \quad (\text{C.11})$$

For a more explicit expression of  $\phi_\rho^*(\phi_L)$  in terms of hypergeometric functions we could perform the derivative  $z'(s) = \frac{dz(s)}{ds}$  using

$$\frac{dF(a, b; c; s)}{ds} = \frac{ab}{c} F(a+1, b+1; c+1; s). \quad (\text{C.12})$$

By expanding  $G(s)$  around  $s = 0$  it is possible to find the following bound

$$|\phi_\rho^*(\phi_L) - \xi(1 - \phi_L^2)| = \xi(1 - \phi_L^2)(G(\phi_L^2) - 1) < \Lambda a \phi_L^2, \quad (\text{C.13})$$

with

$$\Lambda = \frac{2\mu}{h(\mu - 2)^2 + 2a\mu^2 + \mu\sqrt{4(1 - \mu)h^2 + \mu^2(h + 2a)^2}} = \frac{1}{2a + h} \frac{1}{\mu} + O(\mu^{-2}). \quad (\text{C.14})$$

Note that the difference  $|\phi_\rho^*(\phi_L) - \xi(1 - \phi_L^2)|$  is a measure of the error introduced by using the approximate form  $\xi(1 - \phi_L^2)$  instead of the exact function  $\phi_\rho^*(\phi_L)$ , as done in equation (84), and thus the previous equation establishes an upper bound for this error. The validity of the approximation and its error bound are graphically proven in figure C1, where we plot the difference  $|\phi_\rho^*(\phi_L) - \xi(1 - \phi_L^2)|$  as a function of  $m_L^2$  and the straight line  $\Lambda am_L^2$ .

If we now assume that  $m_L$  fulfills the same scaling properties as  $m$  described earlier in equations (25)–(26), then  $\langle m_L^2 \rangle_{\text{st}} \sim O(N^0)$  if  $a \sim O(N^{-1})$  and  $\langle m_L^2 \rangle_{\text{st}} \sim O(N^{-1})$  if  $a \sim O(N^0)$ , and it turns out that the difference between the exact function  $\phi_\rho^*$  and the approximation equation (84) is of order  $N^{-1}$  and thus it can be neglected at the order of our approximation.

## Appendix D. Master equation for the link magnetization

We present in this appendix a simpler method, based on a master equation approach, to compute the statistical properties of the link magnetization, including fluctuations and finite-size effects [53]. This approach is very similar to the all-to-all approximation, and it can be derived following the same steps. We only consider one description variable  $m_L$  to be relevant. The possible processes are  $m_L \rightarrow m_L \pm \Delta_k$ , with  $\Delta_k = \frac{2k}{\mu N}$ , where the sign  $-$  (resp.  $+$ ) corresponds to the switching of a node of degree  $k$  from the up (resp. down) to the down (resp. up) state. The master equation then reads

$$\frac{\partial P(m_L; t)}{\partial t} = \sum_k (E_{m_L}^{\Delta_k} - 1)[W_k^- P] + (E_{m_L}^{-\Delta_k} - 1)[W_k^+ P], \quad (\text{D.1})$$

where  $W_k^\pm$  are the effective rates of the proposed processes, that only depend on the description variable  $m_L$ . We assume that all nodes with degree  $k$  have the same rate of switching to the other state, replacing the local density  $\frac{\sum_j A_{ij} n_j}{k_i}$  by the global density of up-nodes connected to nodes in state up or down, which coincides with the  $c_{0/1}$  calculated in equation (45) as  $\frac{\rho}{1 \pm m_L}$ . Additionally, and this is the key point, we assume that all the other variables follow strictly the deterministic attractor and do not deviate at all, i.e.,  $m_k = m_L$ ,  $\rho = \xi(1 - m_L^2)$ . We then have, for the noisy voter model

$$W_k^+ = \frac{N_k}{2}(1 - m_L)(a + h\xi(1 + m_L)), \quad (\text{D.2})$$

$$W_k^- = \frac{N_k}{2}(1 + m_L)(a + h\xi(1 - m_L)). \quad (\text{D.3})$$

Following [43], we take the continuum limit by expanding the master equation (D.1) in powers of  $\Delta_k$  to second order

$$\frac{\partial P(m_L; t)}{\partial t} = -\frac{\partial}{\partial m_L}[f_L(m_L)P] + \frac{1}{2}\frac{\partial^2}{\partial m_L^2}[g_L(m_L)P], \quad (\text{D.4})$$

where now  $P(m_L; t)$  must be understood as a probability density function and the functions  $f_L(m_L)$  and  $g_L(m_L)$  are defined as

$$f_L(m_L) = \sum_k \Delta_k [W_k^+ - W_k^-] = -2am_L, \quad (\text{D.5})$$

$$g_L(m_L) = \sum_k \Delta_k^2 [W_k^+ + W_k^-] = 4(a + h\xi(1 - m_L^2))/N_{\text{eff}}. \quad (\text{D.6})$$

The time evolution of the moments  $\langle m_L(t) \rangle$  and  $\langle m_L(t)^2 \rangle$  can be obtained by integrating both sides of the Fokker–Planck equation (D.4) respectively as  $\int m_L dm_L$  and  $\int m_L^2 dm_L$ , such that one reproduces the results obtained in equations (51) and (99).

If we take the same continuum limit in the master equation (17) of the all-to-all approach, we find, for the probability  $P(m; t)$ , an equation similar to equation (D.4), but dependent now on functions  $f(m)$  and  $g(m)$ , defined as

$$f(m) = \frac{2}{N}[W^+ - W^-] = -2am, \quad (\text{D.7})$$

$$g(m) = \frac{4}{N^2}[W^+ + W^-] = 2(2a + h(1 - m^2))/N. \quad (\text{D.8})$$



These functions are equivalent to  $f_L$  and  $g_L$  if we replace  $h \rightarrow 2h\xi$ ,  $N \rightarrow N_{\text{eff}}$  and  $m \rightarrow m_L$ , in agreement with the equivalence of equations (100) and (26) under the same replacement.

## ORCID iDs

A F Peralta  <https://orcid.org/0000-0002-5118-5725>

A Carro  <https://orcid.org/0000-0001-9838-3027>

M San Miguel  <https://orcid.org/0000-0001-6237-425X>

R Toral  <https://orcid.org/0000-0003-2046-7620>

## References

- [1] Gunton J D, San Miguel M and Sahni P S 1983 The dynamics of first order phase transitions *Phase Transit. Crit. Phenom.* **8** 269
- [2] Marro J and Dickman R 1999 *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge: Cambridge University Press)
- [3] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities *Proc. Natl Acad. Sci.* **79** 2554
- [4] Clifford P and Sudbury A 1973 A model for spatial conflict *Biometrika* **60** 581
- [5] Crawley M J and May R M 1987 Population dynamics and plant community structure: competition between annuals and perennials *J. Theor. Biol.* **125** 475
- [6] Anderson R M, May R M and Anderson B 1991 *Infectious Diseases of Humans: Dynamics and Control* (Oxford: Oxford University Press)
- [7] Pastor-Satorras R and Vespignani A 2001 Epidemic spreading in scale-free networks *Phys. Rev. Lett.* **86** 3200
- [8] Watts D J 2002 A simple model of global cascades on random networks *Proc. Natl Acad. Sci.* **99** 5766
- [9] Castellano C, Fortunato S and Loreto V 2009 Statistical physics of social dynamics *Rev. Mod. Phys.* **81** 591
- [10] Castellano C and Pastor-Satorras R 2012 Competing activation mechanisms in epidemics on networks *Sci. Rep.* **2** 371
- [11] Albert R and Barabási A-L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47
- [12] Newman M E J and Park J 2003 Why social networks are different from other types of networks *Phys. Rev. E* **68** 36122
- [13] Barrat A, Barthélemy M and Vespignani A 2008 *Dynamical Processes on Complex Networks* (Cambridge: Cambridge University Press)
- [14] Newman M E J 2010 *Networks: An Introduction* (Oxford: Oxford University Press)
- [15] Lambiotte R 2007 How does degree heterogeneity affect an order-disorder transition? *Europhys. Lett.* **78** 68002
- [16] Gleeson J P 2011 High-accuracy approximation of binary-state dynamics on networks *Phys. Rev. Lett.* **107** 68701
- [17] Gleeson J P 2013 Binary-state dynamics on complex networks: pair approximation and beyond *Phys. Rev. X* **3** 21004
- [18] Dorogovtsev S N, Goltsev A V and Mendes J F F 2002 Ising model on networks with an arbitrary distribution of connections *Phys. Rev. E* **66** 16104
- [19] Leone M, Vázquez A, Vespignani A and Zecchina R 2002 Ferromagnetic ordering in graphs with arbitrary degree distribution *Eur. Phys. J. B* **28** 191
- [20] Viana Lopes J, Pogorelov Y G, dos Santos J M B and Toral R 2004 Exact solution of ising model on a small-world network *Phys. Rev. E* **70** 26112
- [21] Boguñá M, Pastor-Satorras R and Vespignani A 2003 Absence of epidemic threshold in scale-free networks with degree correlations *Phys. Rev. Lett.* **90** 28701
- [22] Durrett R 2010 Some features of the spread of epidemics and information on a random graph *Proc. Natl Acad. Sci.* **107** 4491
- [23] Castellano C and Pastor-Satorras R 2010 Thresholds for epidemic spreading in networks *Phys. Rev. Lett.* **105** 218701
- [24] Parshani R, Carmi S and Havlin S 2010 Epidemic threshold for the susceptible-infectious-susceptible model on random networks *Phys. Rev. Lett.* **104** 258701
- [25] Pastor-Satorras R, Castellano C, Van Mieghem P and Vespignani A 2015 Epidemic processes in complex networks *Rev. Mod. Phys.* **87** 925
- [26] Alfaro S and Milaković M 2009 Network structure and N-dependence in agent-based herding models *J. Econ. Dyn. Control* **33** 78
- [27] Chakrabarti D, Wang Y, Wang C, Leskovec J and Faloutsos C 2008 Epidemic thresholds in real networks *ACM Trans. Inf. Syst. Secur.* **10** 1
- [28] Gómez S, Arenas A, Borge-Holthoefer J, Meloni S and Moreno Y 2010 Discrete-time markov chain approach to contact-based disease spreading in complex networks *Europhys. Lett.* **89** 38009
- [29] Van Kampen N G 1981 *Stochastic Processes in Physics and Chemistry* (Amsterdam: North-Holland)
- [30] Lafuerza L F and Toral R 2013 On the effect of heterogeneity in stochastic interacting-particle systems *Sci. Rep.* **3** 1189
- [31] Keeling M J 2000 Multiplicative moments and measures of persistence in ecology *J. Theor. Biol.* **205** 269
- [32] Rand D A 1999 Correlation equations and pair approximations for spatial ecologies *CWI Q.* **12** 329
- [33] Demirel G, Vazquez F, Böhme G A and Gross M 2014 Approximations for discrete adaptive networks *Physica D* **267** 68
- [34] Bauch C T 2005 The spread of infectious diseases in spatially structured populations: an invasy pair approximation *Math. Biosci.* **198** 217
- [35] Kuehn C 2016 Moment closure—a brief review *Control of Self-Organizing Nonlinear Systems. Understanding Complex Systems* (Berlin: Springer)
- [36] Sood V and Redner S 2005 Voter model on heterogeneous graphs *Phys. Rev. Lett.* **94** 178701
- [37] Sood V, Antal T and Redner S 2008 Voter models on heterogeneous networks *Phys. Rev. E* **77** 41121
- [38] Vilone D and Castellano C 2004 Solution of voter model dynamics on annealed small-world networks *Phys. Rev. E* **69** 16109
- [39] Dorogovtsev S N, Goltsev A V and Mendes J F F 2008 Critical phenomena in complex networks *Rev. Mod. Phys.* **80** 1275
- [40] Guerra B and Gómez-Gardeñes J 2010 Annealed and mean-field formulations of disease dynamics on static and adaptive networks *Phys. Rev. E* **82** 35101
- [41] Sonnenschein B and Schimansky-Geier L 2012 Onset of synchronization in complex networks of noisy oscillators *Phys. Rev. E* **85** 51116
- [42] Carro A, Toral R and San Miguel M 2016 The noisy voter model on complex networks *Sci. Rep.* **6** 24775
- [43] Vazquez F and Eguíluz V M 2008 Analytical solution of the voter model on uncorrelated networks *New J. Phys.* **10** 63011
- [44] Vazquez F, Eguíluz V M and San Miguel M 2008 Generic absorbing transition in coevolution dynamics *Phys. Rev. Lett.* **100** 108702
- [45] Vazquez F, Castelló X and San Miguel M 2010 Agent based models of language competition: macroscopic descriptions and order-disorder transitions *J. Stat. Mech.* **P04007**

- [46] De Oliveira M J, Mendes J F F and Santos M A 1993 Nonequilibrium spin models with ising universal behaviour *J. Phys. A: Math. Gen.* **26** 2317
- [47] Dickman R 1986 Kinetic phase transitions in a surface-reaction model: mean-field theory *Phys. Rev. A* **34** 4246
- [48] Pugliese E and Castellano C 2009 Heterogeneous pair approximation for voter models on networks *Europhys. Lett.* **88** 58004
- [49] Mata A S, Ferreira R S and Ferreira S C 2014 Heterogeneous pair-approximation for the contact process on complex networks *New J. Phys.* **16** 053006
- [50] Vespignani A 2011 Modelling dynamical processes in complex socio-technical systems *Nat. Phys.* **8** 32
- [51] Ferreira S C, Ferreira R S and Pastor-Satorras R 2011 Quasistationary analysis of the contact process on annealed scale-free networks *Phys. Rev. E* **83** 066113
- [52] Ferreira S C, Ferreira R S, Castellano C and Pastor-Satorras R 2011 Quasistationary simulations of the contact process on quenched networks *Phys. Rev. E* **84** 066102
- [53] Peralta A F, Carro A, San Miguel M and Toral R 2018 Analytical and numerical study of the non-linear noisy voter model on complex networks *Chaos* **28** 075516
- [54] Peralta A F and Toral R 2018 System-size expansion of the moments of a master equation *Chaos* **28** 106303
- [55] Holley R A and Liggett T M 1975 Ergodic theorems for weakly interacting infinite systems and the voter model *Ann. Probab.* **3** 643
- [56] Moran P A P 1958 Random processes in genetics *Math. Proc. Camb. Phil. Soc.* **54** 60
- [57] Lebowitz J L and Saleur H 1986 Percolation in strongly correlated systems *Physica. A* **38** 194
- [58] Fichthorn K, Gulari E and Ziff R 1989 Noise-induced bistability in a Monte Carlo surface-reaction model *Phys. Rev. Lett.* **63** 1527
- [59] Considine D, Redner S and Takayasu H 1989 Comment on 'Noise-induced bistability in a Monte Carlo surface-reaction model' *Phys. Rev. Lett.* **63** 2857
- [60] Kirman A 1993 Ants, rationality and recruitment *Q. J. Econ.* **108** 137
- [61] Granovsky B L and Madras N 1995 The noisy voter model *Stoch. Process. Appl.* **55** 23
- [62] Alfarano S, Lux T and Wagner F 2008 Time variation of higher moments in a financial market with heterogeneous agents: an analytical approach *J. Econ. Dyn. Control* **32** 101
- [63] Alfarano S, Milaković M and Raddant M 2013 A note on institutional hierarchy and volatility in financial markets *Eur. J. Finance* **19** 449
- [64] Diakonova M, Eguíluz V M and San Miguel M 2015 Noise in coevolving networks *Phys. Rev. E* **92** 32803
- [65] Khalil N, Toral R and San Miguel M 2018 Zealots in the mean-field noisy voter model *Phys. Rev. E* **97** 012310
- [66] Artime O, Peralta A F, Toral R, Ramasco J and San Miguel M 2018 Aging-induced continuous phase transition *Phys. Rev. E* **98** 032104
- [67] Toral R and Colet P 2014 *Stochastic Numerical Methods: An Introduction for Students and Scientists* (New York: Wiley)
- [68] Newman M E J 2003 The structure and function of complex networks *SIAM Rev.* **45** 167
- [69] Boguñá M, Pastor-Satorras R and Vespignani A 2004 Cut-offs and finite size effects in scale-free networks *Eur. Phys. J. B* **38** 205
- [70] Bianconi G 2009 Entropy of network ensembles *Phys. Rev. E* **79** 36114
- [71] Toral R and Tessone C J 2007 Finite size effects in the dynamics of opinion formation *Commun. Comput. Phys.* **2** 177
- [72] Peralta A F and Toral R 2018 Work in progress
- [73] Wolfram Research, Inc. 2018 Expansion of the Gauss hypergeometric function as a generalized power series around  $z = 1$  (<http://functions.wolfram.com/HypergeometricFunctions/Hypergeometric2F1/06/01/04/01/03/>) (Accessed: 1 March 2018)