Binary-state dynamics on complex networks: Stochastic pair approximation and beyond

A. F. Peralta* and R. Toral

IFISC (Instituto de Física Interdisciplinar y Sistemas Complejos), Universitat de les Illes Balears-CSIC, 07122-Palma de Mallorca, Spain

(Received 14 April 2020; accepted 11 November 2020; published 14 December 2020)

Theoretical approaches to binary-state models on complex networks are generally restricted to infinite size systems, where a set of nonlinear deterministic equations is assumed to characterize its dynamical and stationary properties. We develop in this work the stochastic formalism of the different compartmental approaches, these are the approximate master equation (AME), pair approximation (PA), and heterogeneous mean-field (HMF), in descending order of accuracy. The stochastic formalism allows us to enlarge the range of validity and applicability of compartmental approaches. This includes (i) the possibility of studying the role of the size of the system in the different phenomena reproduced by the models together with a network structure, (ii) obtaining the finite-size scaling functions and critical exponents of the macroscopic quantities, and (iii) the extension of the rate description to a more general class of models.

DOI: 10.1103/PhysRevResearch.2.043370

I. INTRODUCTION

Binary-state models on complex network are a very general theoretical framework to study the effect of interactions in the dynamics of a population of individuals. They are composed by a set of nodes that are connected between them through a particular random network, where each node holds a binary ("spinlike," two values) variable that evolves in time by some given transition rates. Typical problems that can be mapped in this scheme include models of epidemic spreading [1–4], language competition [5–9], social interaction [10–15], financial markets [16–21], among many others.

Recently, there has been a lot of effort in the development of highly accurate mathematical descriptions of the dynamics of these models. Typically, we can distinguish between two types of approaches depending on the variables that one chooses to describe the system: (i) individual based-approaches [22–26], where the "spin" or state of each node of the network is considered as an independent variable, (ii) compartmental approaches [11,27–32], where nodes sharing the same topological property such as, for example, the number of neighbors in the network, are aggregated in a single variable, being this an integer (occupation) number. Depending on the level of description, i.e., the number of variables and its nature, one distinguishes between different compartmental approaches: approximate master equation (AME) [30,31,33,34], pair approximation (PA) [35–40], and heterogeneous mean field (HMF) [27–29]. Only the individual node and AME approaches can be considered as a complete description of the models, while the PA and HMF introduce constraints between variables which may or may not be fulfilled, thus they are generally a worse approximation.

Except in a few cases where stochastic effects are taken into account at some extent [41,42] and in its completeness [25,26], the approaches are usually followed by a deterministic type of description [23,27,29,43], where the stochastic nature of the models defined by the individual transitions rates is neglected. The deterministic approach enables one to obtain some important quantities of the models such as the critical point (e.g., the epidemic threshold), or the time evolution of the global state of the system (e.g., the density of infected individuals in the population). The accuracy and suitability of the different approaches have been widely discussed in the literature. For example, in the determination of the epidemic threshold, it has been shown that the two approaches, individual and compartmental, may give contradictory results [44,45] and a general recipe for choosing one or another was given in Ref. [46].

Although the deterministic approach gives us relevant information in all situations, it is an accurate description only in the strict infinite system size limit. Depending on the model, the variables chosen, the values of the parameters and the network, the difference between the deterministic approach and the numerical results may be very important on finite networks [47]. Stochastic effects may become relevant even for extremely large system sizes, specially if the system is close to a critical point, or the network has high degree heterogeneity. Besides, there are some models where the deterministic approach does not provide the relevant information sought.

For example, the noisy voter (Kirman) model [48–51] is an opinion model that considers neighbor imitation and random switching of opinion as basic ingredients. Different versions of the model have been applied in many different contexts, the most important in our perspective being the study price fluctuation in financial markets [16,17] and vote share distributions in electoral data [52–54]. In this context, the global opinion

*afperalta@ifisc.uib-csic.es

does not take a fixed deterministic value but shows heavy fluctuations around the mean instead. The statistics of these fluctuations is the most important feature of study, as it shows deviations from the Gaussian behavior (the global variable distributes as a Beta distribution [16,50]), with very similar properties to financial series and vote share fluctuations. For this reason a full stochastic treatment is necessary to study the model, as carried out in Ref. [26] within the individual approach and in Ref. [32] for the pair approximation. The model has a finite-size critical point that vanishes in the thermodynamic limit and thus a stochastic approach is mandatory in order to achieve the correct characterization [19,32]. Additionally, the noisy voter model is of major importance because of its simplicity and the possibility of obtaining analytical results, which are helpful to fully understand its properties. Recent generalizations of the model include: the effect of nonlinear copying mechanisms [55–57], non-Markovian memory effects [58–60], zealots [61], contrarians [62], more than two states [53,63,64], the role of different noise and copying mechanisms in the nature of the transition (continuous or discontinuous) [65,66], etc.

A second example, which we will study in detail here, is the contact process [67] epidemic model. In this case, a randomly selected individual (if infected) transmits the disease with a certain rate to a neighbor (also chosen at random). Curiously, as pointed out in Ref. [47], the model on scale free networks shows heavy fluctuations and finite-size effects that never die out in practice (even for extremely large system sizes). Thus, we have to resort to the finite-size scaling functions to obtain a useful theoretical description, as it is done with particular attention in Refs. [41,42] within a heterogeneous mean-field scheme, and in Ref. [40] for the pair approximation.

A third model for which stochastic effects are of special interest and that we will also study here is the threshold model [31,68]. This models has a particular type of degenerate dynamics, where the stochastic effects are not just weak fluctuations around a fixed deterministic value. Instead of that, even a very weak fluctuation may eventually drive the system to a completely different final state after some time steps. For this reason, the model shows heavy finite size effects even for large system sizes and a description of the stochastic dynamics is of crucial important to understand it.

The main aim of this work is to give a general theoretical approach to binary-state models on complex networks that takes into account stochastic effects, going beyond incomplete deterministic approaches. In a previous reference by some of us [32], this was done but only for the pair approximation, what we called stochastic pair approximation (SPA), and focusing mainly on the noisy voter model. In this work we will consider the stochastic version of all compartmental approaches (SAME, SPA, SHMF), in a more general formulation and applying them to several models. We start finding the general master equation of individual and compartmental approaches. The master equation corresponds to a full characterization of any Markovian process and one can derive easily the deterministic equations from it [69,70]. The formulation in terms of the master equation enables us to include a more general individual rate form. In the original work [31], individual rates only depend, for simplicity, on the number of neighbors in one of the two states and on the total number of neighbors. We will relax this assumption and include a more general class of models in the description.

In order to approximately solve the master equation, we will apply different expansion techniques. The first one is a van Kampen-like system-size expansion [69,71–74], where the variables are split between its deterministic value plus finite-size corrections. This approach is accurate only far from criticality with an increasing accuracy when the system size increases [71]. For completeness we will also derive the corresponding (continuous) Fokker-Planck and Langevin stochastic differential equations from the original master equation applying the Kramers-Moyal expansion [69], which is accurate, in principle, in the whole parameter region.

Close to a critical point, the models are well described by the finite-size scaling functions, see for example [56], which can not be derived, *a priori*, from the deterministic description nor the traditional van Kampen expansion. The method that we will use to obtain the theoretical scaling functions is a similar system-size expansion of the master equation, but with an anomalous scaling with system size [75,76]. These techniques will be applied to several models of interest on different network types, comparing altogether the different levels of approximation and accuracy of compartmental approaches.

The paper is organized as follows. In Sec. II, we introduce the general definitions and notation of binary state models and the main characteristics of the networks. In Sec. III, we define which are the characteristics of the stochastic effects that we will study. In Sec. IV, we construct the master equation for the individual and compartmental approaches. In Sec. IV, we apply the van Kampen expansion to the general master equation. We re-derive the deterministic nonlinear equations [31], together with a set of linear equations for the correlations and average values of the stochastic corrections. In Sec. V B, we compare the results of numerical simulations for small systems in the stationary state with the theoretical results of the van Kampen approach for different models and networks. In Sec. VI, we apply the expansion of the master equation close to a critical point, obtaining the finite-size scaling functions, and comparing with numerical simulations. In Sec. VII, we study the time dependence of the models. We end with a summary and conclusions in Section VIII. In Appendix A, we show the expressions of the matrices involved in the van Kampen and Kramers-Moyal expansions, while Appendix B contains the details of the expansion around the critical point.

II. GENERAL ASPECTS, MODELS, AND NETWORKS

A binary-state model is composed of a population of size N, where each member of the population can be in two states 1, "adopter," or 0, "nonadopter." Depending on the model and the context, the states may represent different properties of the individuals, for example magnetic spin, opinion on a topic, spoken language, infection state, etc. This is naturally described by a set of time-dependent binary variables $\mathbf{n}(t) \equiv \{n_i(t) = 0, 1\}_{i=1,\ldots,N}$. Each individual $i = 1, \ldots, N$ is regarded as a node of a, single-connected, undirected network, which can be mapped into the usual (symmetric) adjacency matrix $\mathbf{A} = \{A_{ij}\}$, with coefficients $A_{ij} = 1$ if nodes i and j are connected and $A_{ij} = 0$ otherwise, where self-loops are

avoided $A_{ii} = 0$. The degree of node i is defined as the total number of nodes connected to it $k_i = \sum_{j=1}^{N} A_{ij}$ (number of neighbors), taking values in between $k \in [k_{\min}, k_{\max}]$. The degree k can be heterogeneous within the population and one defines the number N_k of nodes with degree k, and the associated fraction $P_k = N_k/N$ called degree distribution. It is also useful to define the moments of degree m as $\mu_m = \sum_{k\in[k_{\min},k_{\max}]} P_k k^m$, with short notation $\mu_1 = \mu$, which corresponds to the average degree. Networks are assumed to be generated by the configuration model [77], with fixed degree distribution P_k, which produces uncorrelated networks if $k_{\max} \leqslant \sqrt{\mu N}$ (no degree-degree correlations and no triangles).

The dynamical model under study is defined by the individual rates $r_i^{\pm}$, which determine the time evolution of the state variables $n_i(t)$. They are defined as the probability per unit time that the transition $n_i = 0 \to 1$ occurs, with rate r_i^+, and $n_i = 1 \to 0$, with rate r_i^-. The rates may depend, in general, on the full set of states $r_i^{\pm}(\mathbf{n})$, however, most common models assume a dependence only through the number of neighbors in state 1, $q_i = \sum_{j=1}^{N} A_{ij} n_j$, in addition to the total number of neighbors k_i. For this reason, we will term the individual rates as $r_i^{\pm}(\mathbf{n}) \equiv R_{k_i,q_i}^{\pm}$, depending only on k_i, q_i, although we will relax this restriction to any rate dependence later in the next sections.

In our study, we will focus on global quantities, such as the total density of nodes in state 1, $m \equiv \frac{1}{N}\sum_{i=1}^{N} n_i \in (0, 1)$. For symmetrical models $R_{k,q}^{+} = R_{k,k-q}^{-}$, it is more natural to define the *magnetization* as $m_S \equiv 2m - 1 \in (-1, 1)$ and we will use one quantity or another depending on the symmetries. The density of active links ρ, i.e., links connecting nodes in different states, is computed as

$$\rho \equiv \frac{\sum_{i,j=1}^{N} A_{ij}(n_i(1-n_j) + (1-n_i)n_j)}{\sum_{i,j=1}^{N} A_{ij}}. \tag{1}$$

One of the interesting properties of ρ for binary-state models is that it can be used as an alternative to m or m_S to measure the level of order or agreement on one of the options, a situation in which ρ approaches zero, independently of the option.

III. STOCHASTIC EFFECTS

The stochastic dynamics produces variability across realizations/trajectories of the stochastic process. For this reason, one typically performs an average over realizations of the macroscopic quantities $\langle m(t)\rangle$, $\langle \rho(t)\rangle$ to characterize the global state of the system. A way to measure fluctuations and variability across realizations is by calculating the variance of the magnetization:

$$\chi \equiv N(\langle m^2\rangle - \langle m\rangle^2), \tag{2}$$

which is also traditionally called magnetic susceptibility in spin models, as it also quantifies how the system responds to an external perturbation such as a magnetic field. Usually, one observes that in the thermodynamic limit $N \to \infty$ the susceptibility converges to a given fixed value. This is the case when deviations of the magnetization (fluctuations) scale as $\sigma[m] \sim N^{-1/2}$, we will explore this ansatz throughout Sec. V. Some models have a special parameter value (critical point) where the susceptibility is not well defined. In this case, we must consider other, anomalous, N dependencies, which can be captured using the finite-size scaling functions.

Assume that the rates of the model depend on a parameter say T. If the model has a critical/bifurcation point $T = T_c$ where the dynamical properties change abruptly (e.g., the susceptibility is not well defined), the statistics over realizations can be described using the finite-size scaling functions. For a moment of order k of the magnetization (in the stationary state $t \to \infty$) that is

$$\langle m^k\rangle_{\rm st} = N^{-k\beta/\bar{\nu}} \widetilde{m}_k[N^{1/\bar{\nu}}(T - T_c)], \tag{3}$$

where β and $\bar{\nu}$ are critical exponents to be determined and $\widetilde{m}_k[x]$ are the scaling functions. The scaling functions are very useful to study the critical behavior of the system and its N dependence. If we know β, $\bar{\nu}$ and T_c, the results for different values of N collapse on a same curve when we plot $N^{k\beta/\bar{\nu}}\langle m^k\rangle_{\rm st}$ vs $N^{1/\bar{\nu}}(T - T_c)$. In the thermodynamic limit $N \to \infty$ we should recover the deterministic result in Eq. (3), that is $\lim_{x\to\infty} \widetilde{m}_k[x] \propto x^{k\beta}$. The scaling functions will be discussed in more detail, theoretically and numerically, in Sec. VI.

Note that after the average over the ensemble of realizations/trajectories is produced, one usually performs additional averages over the ensemble of networks generated with the configuration model with the same degree distribution P_k. This is because we consider the degree distribution as the only relevant characteristic of the network.

IV. THE MASTER EQUATION

The most detailed characterization of models whose dynamics comes defined by stochastic rules is achieved by the knowledge of the probability $P(\mathbf{x}, t)$ of finding the system in state $\mathbf{x}$ at time t. The time-evolution of this probability is governed by a *master equation*. In order to construct a general master equation, we consider: (i) a set of integer variables $\mathbf{x} \equiv (x_1, \ldots, x_M)$, and (ii) a set of processes $\nu = 1, \ldots, K$ characterized by the changes in the variables $x_j \to x_j + \ell_j^{(\nu)}$, $j = 1, \ldots, M$, with rates $W^{(\nu)}(\mathbf{x})$. Once we have these ingredients the general master equation reads [71,73]:

$$\frac{\partial P(\mathbf{x};t)}{\partial t} = \sum_{\nu=1}^{K} \left(\prod_{j=1}^{M} E_j^{-\ell_j^{(\nu)}} - 1 \right) [W^{(\nu)}(\mathbf{x}) P(\mathbf{x};t)], \tag{4}$$

where E_j is the step operator acting on any function $f(\mathbf{x})$ of the variable x_j as $E_j^{\ell}[f(x_1, \ldots, x_j, \ldots, x_M)] = f(x_1, \ldots, x_j + \ell, \ldots, x_M)$. For example, if we choose to include in our description the full set of node-state variables $\mathbf{x} = \mathbf{n}$, we have the following $K = 2N$ processes: $\nu = (i, +)$ where $n_i = 0 \to 1$, and $\nu = (i, -)$ where $n_i = 1 \to 0$, for $i = 1, \ldots, N$. The changes in the variables are $\ell_j^{(i,\pm)} = \pm\delta_{i,j}$ and the respective rates $W^{(i,+)} = \delta_{n_i,0} r_i^+$ and $W^{(i,-)} = \delta_{n_i,1} r_i^-$ ($\delta_{n_i,0} = 1 - n_i$, $\delta_{n_i,1} = n_i$).

When the individual rates $r_i^{\pm}$ depend only on the number k_i of neighbors and the number q_i of those in state 1, i.e., $r_i^{\pm} = R_{k_i,q_i}^{\pm}$, an alternative to the description based on the full set of node-state variables is to consider a compartmental approach

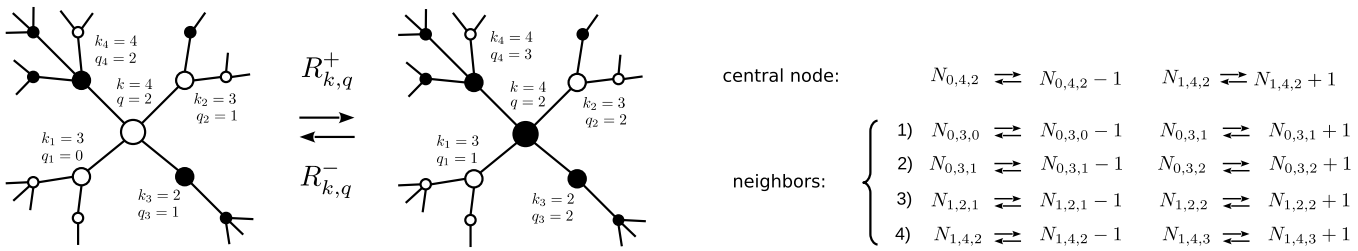


FIG. 1. Diagram of changes in the description variables $\{N_{n,k,q}\}$ when a node with $(k, q) = (4, 2)$ changes state in a particular given network configuration, with neighborhood $(k_1, k_2, k_3, k_4) = (3, 3, 2, 4)$ and $(q_1, q_2, q_3, q_4) = (0, 1, 1, 2) \rightleftharpoons (1, 2, 2, 3)$. Ten total changes are produced in the variables, two for the variables associated to the central node k, q and two additional changes for each one associated to the neighbors $\{k_i, q_i\}_{i=1,\ldots,k}$.

also known as AME.[1] This mesoscopic description in terms of the number of nodes with the same transition rate, was studied in detail in Refs. [30,31] and generalizations of this approach have been developed for multi-state models [78,79] and weighted networks [80]. The *occupation numbers* are defined as the number of nodes $\mathbf{x} \equiv \{N_{n,k,q}\}$ that are in state $n = 0, 1$ and have degree $k = k_{\min}, k_{\min}+1, \ldots, k_{\max}$ among which $q = 0, 1, \ldots, k$ are adopter neighbor nodes (nodes in state 1). The level of description consists of $M = \sum_{k,q} 2 = (1 + k_{\max} - k_{\min})(2 + k_{\max} + k_{\min})$ variables, which are not all independent. The total number of nodes that have degree k is fixed by the network, i.e., $N_k = \sum_{n,q} N_{n,k,q}$, which constitutes a total of $k_{\max} - k_{\min} + 1$ constraints between variables. Another more subtle constraint is that in an undirected network the number of 0-1 links is equal to the number of 1-0, i.e., $\sum_{k,q} qN_{0,k,q} = \sum_{k,q}(k-q)N_{1,k,q}$. Interestingly, in the limit of uncorrelated networks $k_{\max} \propto \sqrt{N}$ it is $M \propto N$, which indicates that the number of variables is of a similar magnitude compared to the node-state approach. Consequently, the occupation number approach will correspond to a significant decrease in the number of variables only when the degree distribution extends over a limited range of degree values $k_{\max} \ll \sqrt{N}$. The global variables of interest, used to portray the macroscopic state of the system, are the total number of adopter nodes $N_1 = \sum_{k,q} N_{1,k,q}$ and the number of active links (connecting nodes in state 0 to 1 or vice versa) $L = \sum_{k,q} qN_{0,k,q}$, and their respective densities $m = N_1/N$, $\rho = 2L/(\mu N)$ defined in Sec. II.

In this occupation number approach $\{N_{n,k,q}\}$, however, the construction of the master equation is more cumbersome as we need to identify the possible processes ν and the associated rates $W^{(\nu)}$, and this will be our concern in the remainder of this section. Still, the elementary process of the dynamic is the state transition of a node i, compatible with the numbers $N_{n,k,q}$, changing from $n_i = 0$ to $n_i = 1$ or vice versa, but all processes that lead to the same change of the occupation number variables are grouped under the same label ν. In an elementary process, $2(k+1)$ changes of the set of description variables $\{N_{n,k,q}\}$ are produced, two for the variables associated to the central node and two for each one of its neighbors, see Fig. 1 as a schematic example. The variables that change during this process depend on the values k, q of the chosen node i, and additionally on the set $\{k_j, q_j\}_{j=1,\ldots,k}$ of the k neighbors of i. We adopt to order the list of neighbors such that $\{k_j, q_j\}_{j=1,\ldots,k-q}$ correspond to the neighbors in state 0, and $\{k_j, q_j\}_{j=k-q+1,\ldots,k}$ to the neighbors in state 1. Therefore the characterization of a process ν requires of the knowledge of the full set of variables, i.e., $\nu = (n, k, q, \{k_j, q_j\}_{j=1,\ldots,k})$. The problem now is that, in principle, one is not able to know from the variables $\{N_{n,k,q}\}$ the set $\{k_j, q_j\}_{j=1,\ldots,k}$, and consequently we need some approximation to attain a closed treatment of the dynamics. We make the ansatz that the rate of each process is calculated as the total change rate of the central node $N_{n,k,q}R^{\pm}_{k,q}$ times the probability of having a particular configuration of the neighborhood, this is

$$W^{(0,k,q,\{k_j,q_j\}_{j=1,\ldots,k})}(\{N_{n,k,q}\}) = N_{0,k,q}R^+_{k,q}\prod_{i=1}^{k-q} P_0(0,k_i,q_i)\prod_{j=k-q+1}^{k} P_0(1,k_j,q_j), \quad (5)$$

$$W^{(1,k,q,\{k_j,q_j\}_{j=1,\ldots,k})}(\{N_{n,k,q}\}) = N_{1,k,q}R^-_{k,q}\prod_{i=1}^{k-q} P_1(0,k_i,q_i)\prod_{j=k-q+1}^{k} P_1(1,k_j,q_j). \quad (6)$$

Here, we introduced $P_0(1, k_j, q_j)$, defined as the probability that an edge leaving a node in state 0 connects to a node in state 1 with k_j, q_j, and equivalently for $P_0(0, k_i, q_i)$, $P_1(0, k_i, q_i)$ and $P_1(1, k_j, q_j)$. These probabilities can be calculated using the description variables $N_{n,k,q}$ as

$$P_0(0,k,q) = \frac{(k-q)N_{0,k,q}}{\sum_{k,q}(k-q)N_{0,k,q}}, \quad (7)$$

$$P_0(1,k,q) = \frac{(k-q)N_{1,k,q}}{\sum_{k,q}(k-q)N_{1,k,q}}, \quad (8)$$

$$P_1(0,k,q) = \frac{qN_{0,k,q}}{\sum_{k,q} qN_{0,k,q}}, \quad (9)$$

$$P_1(1,k,q) = \frac{qN_{1,k,q}}{\sum_{k,q} qN_{1,k,q}}. \quad (10)$$

For example, $P_0(1, k, q)$ is the fraction of edges coming out of nodes in state 1 with k, q that go to nodes in state 0, divided by the total number of 0-1 edges and similarly for the other expressions. Note that the approximation in this method is that we assume the probability of the neighborhood configuration

[1]Note that in the acronym AME, or approximate master equation, the word "master" has a different meaning than the one used in this paper to describe an equation for the probability distribution.

of a node to be a product of independent single event probabilities, which is of general validity for uncorrelated networks.

It is also important to understand that the rate formulation, Eqs. (5) and (6), allows to consider a more general dependence of the individual rates as $R^{\pm}_{k,q,\{k_j,q_j\}_{j=1,\dots,k}}$. This is a further advantage of the stochastic formulation that the original deterministic approach did not consider. A relevant example that this new formulation includes, and that we will explore in Sec. VI, is the contact process, with rates: $R^{+}_{k,q,\{k_j\}^{k}_{j=k-q+1}} = \lambda \sum_{j=k-q+1}^{k} k_j^{-1}$; $R^{-}_{k,q} = \mu$ (the infection rate depends on the degrees of the infected neighbors). This epidemic model can be defined with two processes: (i) with rate λ a node transmits the disease to a randomly selected neighbor, and (ii) with rate μ the node recovers.

We now define $\ell^{(\nu)}_{n,k,q}$ as the change of the variable $N_{n,k,q} \to N_{n,k,q} + \ell^{(\nu)}_{n,k,q}$ in the process ν, which are computed as (see Fig. 1 for a guide):

$$\ell^{(0,k',q',\{k_j,q_j\}_{j=1,\dots,k})}_{0,k,q} = -\delta_{k,k'}\delta_{q,q'} + \sum_{j=1}^{k'-q'} \left(-\delta_{k,k_j}\delta_{q,q_j} + \delta_{k,k_j}\delta_{q,q_j+1}\right), \tag{11}$$

$$\ell^{(0,k',q',\{k_j,q_j\}_{j=1,\dots,k})}_{1,k,q} = \delta_{k,k'}\delta_{q,q'} + \sum_{j=k'-q'+1}^{k'} \left(-\delta_{k,k_j}\delta_{q,q_j} + \delta_{k,k_j}\delta_{q,q_j+1}\right), \tag{12}$$

$$\ell^{(1,k',q',\{k_j,q_j\}_{j=1,\dots,k})}_{0,k,q} = \delta_{k,k'}\delta_{q,q'} + \sum_{j=1}^{k'-q'} \left(-\delta_{k,k_j}\delta_{q,q_j} + \delta_{k,k_j}\delta_{q,q_j-1}\right), \tag{13}$$

$$\ell^{(1,k',q',\{k_j,q_j\}_{j=1,\dots,k})}_{1,k,q} = -\delta_{k,k'}\delta_{q,q'} + \sum_{j=k'-q'+1}^{k'} \left(-\delta_{k,k_j}\delta_{q,q_j} + \delta_{k,k_j}\delta_{q,q_j-1}\right). \tag{14}$$

Once the processes ν, rates $W^{(\nu)}$ and changes in the variables $\ell^{(\nu)}_{n,k,q}$ are defined, we can draw on the general theory of stochastic processes [69–71,81] in terms of the master equation (4).

Coarser levels of description are also possible. Let $N_{n,k} = \sum_q N_{n,k,q}$ be the number of nodes in state n with degree k, and $L_{n,k} = \sum_q qN_{n,k,q}$ the number of links that connect nodes of degree k and state n with nodes in state 1 (adopter nodes). The next level of description is the pair approximation (PA) that considers the set $\mathbf{x} = \{N_{1,k}, L_{0,k}, L_{1,k}\}$, with $N_{0,k} = N_k - N_{1,k}$, and $k \in [k_{\min}, k_{\max}]$. The pair approximation reduces the number of variables to $M = 3(k_{\max} - k_{\min} + 1)$ with the conservation of the total number of 0-1 links $\sum_k L_{0,k} = \sum_k (N_{1,k}k - L_{1,k})$ as the only constraint. The master equation requires to write the rates $W^{(\nu)}$ as a function only of the description variables. To achieve this, one introduces an approximation based on the ansatz that the variable $N_{n,k,q}$ appearing in the rates Eqs. (5) and (6) can be expressed as

$$N_{n,k,q} = N_{n,k}\mathrm{Bin}_{k,q}\left[\frac{L_{n,k}}{kN_{n,k}}\right], \tag{15}$$

where $\mathrm{Bin}_{k,q}[p] = \binom{k}{q}p^q(1-p)^{k-q}$ is the binomial distribution. In this paper, we restrict our study to this version of the pair approximation, but other variants exist in the literature, such as the so-called heterogeneous pair approximation [38], where one includes in the description the number of active links $L_{k,k'}$ that join nodes of degree k and k' that are in different states, or the original version [37] (also called homogeneous pair approximation) that takes into consideration just the global number L of active links.

An even cruder level of description is the heterogeneous mean field (HMF), which considers the set of variables $\mathbf{x} = \{N_{1,k}\}$, with $N_{0,k} = N_k - N_{1,k}$, reducing the number of variables to $M = k_{\max} - k_{\min} + 1$ with no constraints. The closure of the rates $W^{(\nu)}$ in terms of this set of variables is achieved by a similar binomial ansatz but with a simpler single event probability:

$$N_{n,k,q} = N_{n,k}\mathrm{Bin}_{k,q}\left[\frac{\sum_k kN_{1,k}}{\mu N}\right]. \tag{16}$$

The coarsest possible description is the mean field (MF) in which a single description variable $\mathbf{x} = N_1$ is used with closure ansatz $N_{1,k,q} = N_1\delta_{q,kN_1/N}$, $N_{0,k,q} = (N - N_1)\delta_{q,kN_1/N}$.

When we use the master equation with a compartmental approach (AME, PA, HMF) to obtain results beyond the deterministic description, we will include a S in the abbreviation (SAME, SPA, SHMF) to empathize that we are taking into account the stochastic effects. We will use this terminology in the figures of the next sections.

Note that the formulation of the node-state approach does not need any approximations, in principle, while the different occupation number approaches, whether AME, PA, HME or MF, use approximations in the calculation of the rates that limit the validity of their predictions. In particular, as discussed previously, we expect the AME to be accurate only for uncorrelated networks. The fact that the master equation for the node-state approach is free of approximations does not mean in general that we are able to solve such equation, and different approximations are then needed to obtain a solution [25,26]. The advantage of the occupation number approach is that it has some particularities that enables us to apply accurate methods to solve the master equation. These different techniques are explained and explored in the next sections.

V. APPROXIMATE SOLUTION OF THE MASTER EQUATION

A. Formulation

The main reason of the convenience of the occupation number approach is that the description variables are extensive. This means that for a fixed degree distribution P_k, if we increase the system size $N \to \lambda N$, the variables scale in the same way $N_{n,k,q} \to \lambda N_{n,k,q}$ and similarly for $N_{n,k}$, $L_{n,k}$, $N_{1,k}$, and N_1. This property is useful because it allows us to apply the well known system-size expansions of the master equation. Note that the rates Eqs. (5) and (6) are extensive functions $W^{(\nu)}(\mathbf{x}) = N w^{(\nu)}(\frac{\mathbf{x}}{N})$, where $N = \sum_{n,k,q} N_{n,k,q}$ is the total number of nodes and $w^{(\nu)}$ are the set of intensive rate functions. In this case, following [32,71], we can use a van Kampen type of system-size expansion, that we now explain in detail.

In the case of the AME, the expansion splits the variables as $\mathbf{x} = N\boldsymbol{\phi} + N^{1/2}\boldsymbol{a} + N^0\boldsymbol{b}$, in components $N_{n,k,q} = N\phi_{n,k,q} + N^{1/2}a_{n,k,q} + N^0 b_{n,k,q}$, where $\phi_{n,k,q}$ are a set of deterministic variables, while $a_{n,k,q}$ and $b_{n,k,q}$ are random variables. This is an expansion which is assumed to be of general validity in the thermodynamic limit $N \to \infty$ and which yields the first stochastic correction terms to the deterministic approach [30,31]. The deterministic evolution of the system fulfills a set of nonlinear differential equations

$$\frac{d\phi_{n,k,q}}{dt} = \Phi_{n,k,q}, \tag{17}$$

characterized by the drift term defined as $\Phi_{n,k,q}(\boldsymbol{\phi}) = \sum_\nu \ell^{(\nu)}_{n,k,q} w^{(\nu)}(\boldsymbol{\phi})$ which leads, after some algebra using Eqs. (5)–(14), to

$$\begin{aligned}\Phi_{0,k,q} = &-\phi_{0,k,q}R^+_{k,q} + \phi_{1,k,q}R^-_{k,q} - \phi_{0,k,q}(k-q)\beta^s \\ &+ \phi_{0,k,q-1}(k-q+1)\beta^s - \phi_{0,k,q}q\gamma^s \\ &+ \phi_{0,k,q+1}(q+1)\gamma^s,\end{aligned} \tag{18}$$

$$\begin{aligned}\Phi_{1,k,q} = &\phi_{0,k,q}R^+_{k,q} - \phi_{1,k,q}R^-_{k,q} - \phi_{1,k,q}(k-q)\beta^i \\ &+ \phi_{1,k,q-1}(k-q+1)\beta^i - \phi_{1,k,q}q\gamma^i \\ &+ \phi_{1,k,q+1}(q+1)\gamma^i.\end{aligned} \tag{19}$$

Here, β^s, γ^s, β^i, and γ^i are the individual rates $R^\pm_{k,q}$ at which a neighbor of a central node changes state averaged with the probabilities (7)–(10), where the symbol β, γ reflects the state of the neighbor node 0, 1, while the super index s, i reflects the state of the central node 0, 1, namely,

$$\beta^s \equiv \sum_{k,q} P_0(0,k,q)R^+_{k,q} = \frac{\sum_{k,q}(k-q)\,\phi_{0,k,q}R^+_{k,q}}{\sum_{k,q}(k-q)\,\phi_{0,k,q}}, \tag{20}$$

$$\gamma^s \equiv \sum_{k,q} P_0(1,k,q)R^-_{k,q} = \frac{\sum_{k,q}(k-q)\,\phi_{1,k,q}R^-_{k,q}}{\sum_{k,q}(k-q)\,\phi_{1,k,q}}, \tag{21}$$

$$\beta^i \equiv \sum_{k,q} P_1(0,k,q)R^+_{k,q} = \frac{\sum_{k,q} q\,\phi_{0,k,q}R^+_{k,q}}{\sum_{k,q} q\,\phi_{0,k,q}}, \tag{22}$$

$$\gamma^i \equiv \sum_{k,q} P_1(1,k,q)R^-_{k,q} = \frac{\sum_{k,q} q\,\phi_{1,k,q}R^-_{k,q}}{\sum_{k,q} q\,\phi_{1,k,q}}. \tag{23}$$

Note that, at the deterministic level, the set of differential equations (17)–(23) coincides with the original work of Gleeson [30,31], as it is naturally expected. The advantage of the stochastic formalism presented here Eqs. (4)–(14) is that, in addition, we will be able to obtain results for the average deviations $\langle a_{n,k,q}\rangle$, $\langle b_{n,k,q}\rangle$ from the deterministic solution, and also for the fluctuations/correlations $C_{n,k,q;n',k',q'} = \langle a_{n,k,q}a_{n',k',q'}\rangle - \langle a_{n,k,q}\rangle\langle a_{n',k',q'}\rangle$. In the van Kampen expansion, the set of differential equations for these quantities are linear and read in vector notation [71]:

$$\frac{d\langle \boldsymbol{a}\rangle}{dt} = -\mathbf{B}\langle \boldsymbol{a}\rangle, \tag{24}$$

$$\frac{d\langle \boldsymbol{b}\rangle}{dt} = -\mathbf{B}\langle \boldsymbol{b}\rangle + \boldsymbol{\Gamma}, \tag{25}$$

$$\frac{d\mathbf{C}}{dt} = -\mathbf{B}\mathbf{C} - \mathbf{C}\mathbf{B} + \mathbf{G}, \tag{26}$$

where $\mathbf{B}$ is the Jacobian matrix $B_{ij}(\boldsymbol{\phi}) = -\partial_{\phi_j}\Phi_i$; the noise $\mathbf{G}$ matrix is calculated as $G_{ij}(\boldsymbol{\phi}) = \sum_\nu \ell^{(\nu)}_i \ell^{(\nu)}_j w^{(\nu)}(\boldsymbol{\phi})$ and $\Gamma_i = \frac{1}{2}\sum_{j,k}\langle a_j a_k\rangle \partial^2_{\phi_j,\phi_k}\Phi_i$ is related to the Hessian matrices of $\boldsymbol{\Phi}$. For reason of space, the explicit expressions of these matrices are written down in Appendix A.

Note that Eqs. (18) and (19) are valid as long as the individual rates can be written in the form $R^\pm_{k,q}$. For a more general dependence $R^\pm_{k,q,\{k_j,q_j\}_{j=1,\dots,k}}$, the form of the deterministic drift functions $\Phi_{n,k,q}$ changes. For the contact process with rates $R^+_{k,q,\{k_j\}^k_{j=k-q+1}}$, similar equations apply after replacing: $R^+_{k,q} \to \widetilde{R}^+_{k,q} = \sum_{\{k_j,q_j\}} R^+_{k,q,\{k_j\}}\prod_j P_0(1,k_j,q_j)$, and $\beta^i \to \beta^i_k = \sum_{k',q'} P_1(0,k',q')\widetilde{R}^+_{k',q',k}$ with $\widetilde{R}^+_{k',q',k} = \sum_{\{k_j,q_j\}} R^+_{k',q',k,\{k_j\}}\prod_j P_0(1,k_j,q_j)$. Here $R^+_{k',q',k,\{k_j\}}$ is just the rate $R^+_{k',q',\{k_j\}}$ where one of the degrees (no matter which) is fixed to be k.

In the case of the pair approximation and proceeding with the general theory, we split the variables like $N_{1,k} = N\phi_k + N^{1/2}a_k + N^0 b_k$ and $L_{n,k} = N\phi_{n,k} + N^{1/2}a_{n,k} + N^0 b_{n,k}$. The evolution equations at the deterministic level are

$$\frac{d\phi_k}{dt} = \Phi_k, \tag{27}$$

$$\frac{d\phi_{n,k}}{dt} = \Phi_{n,k}. \tag{28}$$

In order to obtain the deterministic drift functions $\boldsymbol{\Phi}$, we have to perform sums in Eqs. (18) and (19) as $\Phi_k = \sum_q \Phi_{1,k,q}$ and $\Phi_{n,k} = \sum_q q\Phi_{n,k,q}$, which leads to

$$\Phi_k = \sum_q [\phi_{0,k,q}R^+_{k,q} - \phi_{1,k,q}R^-_{k,q}], \tag{29}$$

$$\begin{aligned}\Phi_{0,k} = &\sum_q [-q\phi_{0,k,q}R^+_{k,q} + q\phi_{1,k,q}R^-_{k,q}] \\ &+ \beta^s(kP_k - k\phi_k - \phi_{0,k}) - \gamma^s\phi_{0,k},\end{aligned} \tag{30}$$

$$\begin{aligned}\Phi_{1,k} = &\sum_q [q\phi_{0,k,q}R^+_{k,q} - q\phi_{1,k,q}R^-_{k,q}] \\ &+ \beta^i(k\phi_k - \phi_{1,k}) - \gamma^i\phi_{1,k},\end{aligned} \tag{31}$$

where one must replace $\phi_{n,k,q}$ by the binomial ansatz $\phi_{0,k,q} = (P_k - \phi_k)\mathrm{Bin}_{k,q}[p_{0,k}]$, $\phi_{1,k,q} = \phi_k \mathrm{Bin}_{k,q}[p_{1,k}]$ and $p_{0,k} = \phi_{0,k}/(k(P_k - \phi_k))$, $p_{1,k} = \phi_{1,k}/(k\phi_k)$. The corresponding Jacobian $\mathbf{B}$ and $\mathbf{G}$ matrices of this Pair Approximation can be found in Appendix A.

In the case of the heterogeneous mean field, the variable splitting is $N_{1,k} = N\phi_k + N^{1/2}a_k + N^0 b_k$, and the deterministic equation $\frac{d\phi_k}{dt} = \Phi_k$, where the drift functions $\mathbf{\Phi}$ are obtained by summing Eqs. (18) and (19) like $\Phi_k = \sum_q \Phi_{1,k,q}$, which leads also to

$$\Phi_k = \sum_q [\phi_{0,k,q} R^+_{k,q} - \phi_{1,k,q} R^-_{k,q}], \tag{32}$$

but now $\phi_{n,k,q}$ are given by $\phi_{0,k,q} = (P_k - \phi_k)\mathrm{Bin}_{k,q}[p]$, $\phi_{1,k,q} = \phi_k \mathrm{Bin}_{k,q}[p]$ with $p = \sum_k k\phi_k/\mu$ (independent of k). Again, the corresponding Jacobian $\mathbf{B}$ and $\mathbf{G}$ matrices can be found in Appendix A.

In a previous work [71], we explained how to solve equations (24)–(26) and we developed a very stable and fast convergent implicit Euler method to find the numerical solution of the correlation matrix $\mathbf{C}$. It is worth mentioning that a general result of the van Kampen expansion is that the stationary probability distribution $\Pi_{\mathrm{st}}(\boldsymbol{a})$ of the first stochastic correction $\boldsymbol{a}$ is Gaussian [69,71] with zero mean $\langle \boldsymbol{a} \rangle_{\mathrm{st}} = 0$:

$$\Pi_{\mathrm{st}}(\boldsymbol{a}) = \sqrt{\frac{|\mathbf{C}_{\mathrm{st}}|}{(2\pi)^M}} e^{-\frac{1}{2}\boldsymbol{a}^\intercal \cdot \mathbf{C}^{-1}_{\mathrm{st}} \cdot \boldsymbol{a}}. \tag{33}$$

Besides, if the initial condition $\Pi(\boldsymbol{a}, t = 0)$ is Gaussian, then the time-dependent $\Pi(\boldsymbol{a}, t)$ is also a Gaussian (33) replacing the stationary correlation matrix $\mathbf{C}_{\mathrm{st}} \to \mathbf{C}(t)$ by the time-dependent one.

The van Kampen expansion will be accurate in the thermodynamic limit $N \to \infty$, for example in the determination of the magnetic susceptibility, this is

$$\chi = \sum_{k,q,k',q'} C_{1,k,q;1,k',q'}. \tag{34}$$

According to the van Kampen approach, the susceptibility defined as Eq. (2) does not depend on system size N, which is obviously not true for a finite system N. What we are obtaining in this approach is the thermodynamic limit $\lim_{N\to\infty} \chi_N$. With respect to the average values of the macroscopic quantities $\langle m(t) \rangle$, $\langle \rho(t) \rangle$ they are computed as

$$\langle m(t) \rangle = \sum_{k,q} \phi_{1,k,q} + \frac{1}{N} \sum_{k,q} \langle b_{1,k,q} \rangle, \tag{35}$$

$$\langle \rho(t) \rangle = \sum_{k,q} q\phi_{0,k,q} + \frac{1}{N} \sum_{k,q} q \langle b_{0,k,q} \rangle. \tag{36}$$

This is nothing but the deterministic solution plus a correcting factor of order $O(N^{-1})$ [note that $\langle \boldsymbol{a} \rangle = 0$, Eq. (24)].

An alternative less restrictive system size expansion is the Kramers-Moyal expansion, which transforms the master equation (4) into a continuous PDE for the intensive variables $\boldsymbol{\varphi}$. If we define the densities $\boldsymbol{\varphi} = \mathbf{x}/N$, the Kramers-Moyal expansion [69] leads to the Fokker-Planck equation [82] for the probability density $\Pi(\boldsymbol{\varphi}; t)$ of the intensive variables:

$$\frac{\partial \Pi(\boldsymbol{\varphi}; t)}{\partial t} = \sum_{i=1}^{K} \frac{\partial}{\partial \varphi_i} \left[-\Phi_i(\boldsymbol{\varphi})\Pi + \frac{1}{2N} \sum_{j=1}^{K} \frac{\partial}{\partial \varphi_j} [G_{ij}(\boldsymbol{\varphi})\Pi] \right], \tag{37}$$

where Φ_i and G_{ij} are the same functions defined previously. We can also derive the equivalent Langevin stochastic differential equations for the trajectories $\boldsymbol{\varphi}(t)$, in the Itô convention, it is

$$\frac{d\varphi_i}{dt} = \Phi_i(\boldsymbol{\varphi}) + N^{-1/2} \sum_{j=1}^{K} g_{ij}(\boldsymbol{\varphi}) \xi_j(t), \tag{38}$$

with $G_{ij} = \sum_{k=1}^{K} g_{ik} g_{jk}$, in matrix form $\mathbf{G} = \boldsymbol{g} \cdot \boldsymbol{g}^T$, and $\xi_j(t)$ are white noise variables with zero mean and correlations $\langle \xi_i(t) \xi_j(t') \rangle = \delta_{ij}\delta(t - t')$. A simple way of obtaining the matrix $\mathbf{g}$ from $\mathbf{G}$ is by diagonalizing both sides of equation $\mathbf{G} = \boldsymbol{g} \cdot \boldsymbol{g}^T$, and then it is straightforward to solve.

The problem with the Kramers-Moyal expansion is that, in most occasions, it is as complicated to solve as the original master equation (4), while the van Kampen expansion corresponds to a linearization of Eq. (37) (this is why it is also called linear noise approximation) where we assume $\boldsymbol{\varphi}$ to weakly fluctuate around the deterministic value $\boldsymbol{\phi}$, this is $\boldsymbol{\varphi} \approx \boldsymbol{\phi} + N^{-1/2}\boldsymbol{a}$. In the next subsection we apply the van Kampen expansion method to several models of interest and check its accuracy and validity.

B. Comparison with numerical simulations

We will now compare the results of the theory explained in the previous section to the numerical simulations. We will focus on stationary quantities in order to study how the results change with the parameters of the models.

1. Susceptible infected susceptible

The first model that we consider is the SIS (susceptible-infected-susceptible) epidemic model [1,83] on a scale-free network with rates $R^+_{k,q} = \varepsilon + \lambda q$ and $R^-_{k,q} = \mu$. Here λ is the transmission rate, μ is the recovery rate, and ε is the rate at which an outbreak appears in the system. Note that we incorporate the parameter ε in order for the system to have a properly defined stationary result, in principle one recovers the traditional SIS model by letting $\varepsilon \to 0$. In Fig. 2, we see that the van Kampen approach predicts accurately the stationary susceptibility for a small system size $N = 100$ with increasing accuracy as it increases to $N = 400$. The finite-size corrections to the average value $\langle \rho \rangle_{\mathrm{st}}$ are plotted in Fig. 2, with an improvement in the deterministic solution. If we focus on the comparison between the different approximations we observe, as expected, an increase in accuracy as AME > PA > HMF. Although the difference between the AME and the PA is small with respect to the deterministic $\boldsymbol{\phi}$ and fluctuations χ, the finite size corrections of the average values $N^{-1}\langle \boldsymbol{b} \rangle$ are well captured only by the AME. This indicates that the PA is a good approximation at the deterministic and linear (Jacobian)

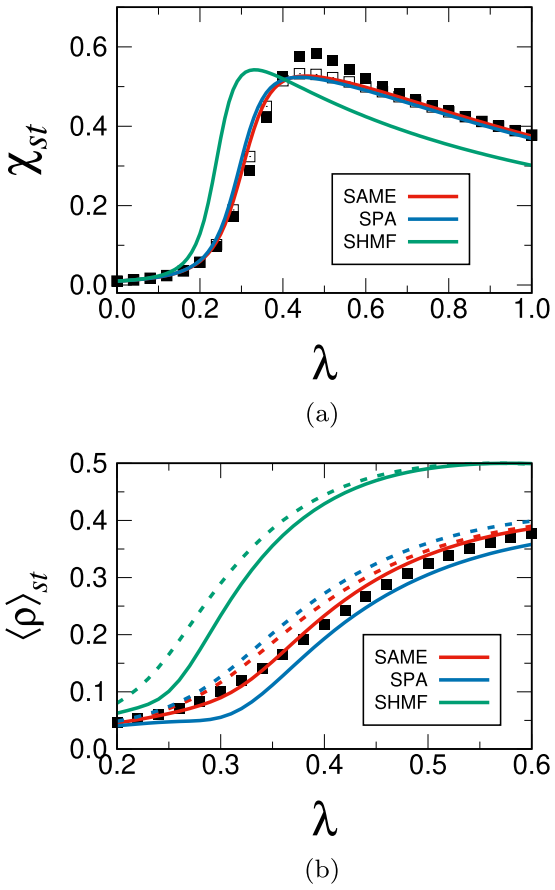


FIG. 2. (a) Stationary susceptibility $\chi_{\rm st}$ and (b) average density of active links $\langle\rho\rangle_{\rm st}$ as a function of the transmission rate λ for the SIS model. We choose as parameters $\varepsilon = 10^{-2}$ and $\mu = 1$ on a scale free network $P_k \sim k^{-2.5}$ with $k_{\rm min} = 2$ and $k_{\rm max} = 10$. Points correspond to numerical simulations of the model with $N = 100$ (solid squares) and $N = 400$ (empty squares) averaged over an ensemble of 100 networks. Lines of different colors are the theoretical prediction of the different approximations. In the top panel (a), the solid lines are the van Kampen result (34), while in the bottom panel (b), the dashed lines are the deterministic approaches and the solid lines the corrected average values (35).

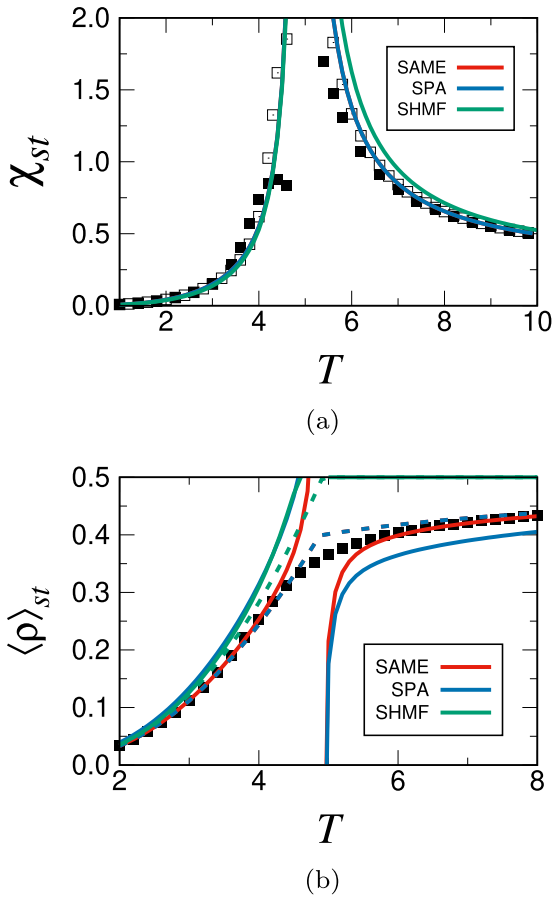


FIG. 3. (a) Stationary susceptibility $\chi_{\rm st}$ and (b) average density of active links $\langle\rho\rangle_{\rm st}$ as a function of the temperature T for the Ising model with Glauber rates. We choose as parameter $J = 1$ on an Erdős-Rényi network with average degree $\mu = 5$. Points correspond to numerical simulations of the model with $N = 100$ (solid squares) and 400 (empty squares) averaged over an ensemble of 100 networks. Lines of different colors are the theoretical prediction of the different approximations (the curves that do not appear are superposed). In the top panel (a), the solid lines are the van Kampen result Eq. (34), while in the bottom panel (b), the dashed lines are the deterministic approaches and the solid lines the corrected average values Eq. (35).

level Eqs. (24) and (26) but not for the second-order correction (Hessian) level (25).

2. Ising Glauber

The second model to which we apply our theory is the Ising model defined on an Erdős-Rényi network with Glauber rates [84] $R^+_{k,q} = (1 + e^{\frac{2J}{T}(k-2q)})^{-1}$ and $R^-_{k,q} = R^+_{k,k-q}$, where J is the coupling strength and T the temperature. Note that it is a symmetric model and thus, as discussed in Sec. II, we choose $m_S = 2m - 1$ to calculate the susceptibility. It is well known that the Ising model has a critical point T_c such that $\lim_{N\to\infty}\langle|m_S|\rangle_{\rm st} = 0$ if $T > T_c$ and $\lim_{N\to\infty}\langle|m_S|\rangle_{\rm st} \propto (T_c - T)^\beta$ if $T < T_c$, while $\langle m_S\rangle_{\rm st} = 0$ always for symmetry. For this reason, following the standard procedure, we compute the susceptibility numerically as $\chi_{\rm st} = N(\langle m_S^2\rangle_{\rm st} - \langle|m_S|\rangle^2_{\rm st})$ if $T < T_c$, and $\chi_{\rm st} = N\langle m_S^2\rangle_{\rm st}$ if $T > T_c$. In principle, we can not know the position of the critical point numerically for a single finite system size, we thus plot both quantities $\chi_{\rm st} = N(\langle m_S^2\rangle_{\rm st} - \langle|m_S|\rangle^2_{\rm st})$ and $\chi_{\rm st} = N\langle m_S^2\rangle_{\rm st}$ and eliminate those points to the right of the peak of the first expression. In Fig. 3, we see a good prediction of the stationary susceptibility for the small system $N = 100$ with increasing accuracy for $N = 400$. Note that the theory predicts the divergence of the susceptibility at the critical point T_c, as $\chi_{\rm st} \propto |T - T_c|^{-\gamma}$, which can be strictly true only in the thermodynamic limit $N \to \infty$. For a finite system it can be shown, see Section VI, that if we approach the critical point as an inverse power of the system size $|T - T_c| \propto N^{-r}$, the susceptibility scales as a positive power $\chi_{\rm st} \propto N^{2\upsilon-1}$, with appropriate exponents $r > 0$ and $\upsilon > 1/2$ determined in Sec. VI. This implies that the van Kampen expansion presents discrepancies with the numerical results that are important, for a finite system N, in a region of the critical point whose width decreases with system size. Similarly, we see in Fig. 3 that the stochastic correction to the

density of active links $\langle\rho\rangle_{\rm st}$ is only accurate outside the critical region, while it diverges at the critical point, at odds with the numerical result. In the comparison between the different approximations we observe again an increase in accuracy as AME > PA > HMF. As proven in Ref. [31], the deterministic part of the PA and AME are completely equivalent for all models fulfilling the microscopic reversibility condition $R^+_{k,q}/R^-_{k,q} = c^q R^+_{k,0}/R^-_{k,0}$ where c is a constant, for the Ising Glauber this is the case with $c = e^{4J/T}$. We also observe that the AME and the PA offer results for the susceptibility which are indistinguishable at the resolution of the figure. Although the AME and PA agree at the deterministic and fluctuation level, the finite size corrections to the average values $N^{-1}\langle \boldsymbol{b}\rangle$ are only accurate for the AME, confirming the results obtained for the SIS model.

3. Majority vote

The third model is the majority-vote model [85] on a z-regular network with rates $R^+_{k,q} = Q$ if $q < k/2$, $R^+_{k,q} = 1/2$ if $q = k/2$, $R^+_{k,q} = 1 - Q$ if $q > k/2$, and $R^-_{k,q} = R^+_{k,k-q}$, where Q is the rate of spontaneous opinion switching. It is also a symmetric model and it has similar phenomenology to the Glauber model with a critical point Q_c, see Fig. 4. The most notorious difference is that in this case the AME and PA results are very different even at the deterministic level, and thus for this model only the AME gives more accurate results. The reason for this difference is that the rates do not fulfill the microscopic reversibility condition, see Ref. [31]. Note also in Fig. 4 that the AME and PA predict similar critical points Q_c, but the scaling $\langle|m_S|\rangle_{\rm st} \propto (Q_c - Q)^\beta$ is $\beta = 1/2$ for the AME and $\beta = 1/4$ for the PA. In fact, according to our discussion in the next section, the scaling behavior of the magnetization and susceptibility around a critical point depends on the normal form of the bifurcation. For example, for a typical continuous phase transition as in the Ising model we have $\beta = 1/2$ and $\gamma = 1$ which corresponds to mean-field exponents. This justifies the common knowledge that critical exponents in complex networks coincide with those of mean-field theory, see [86] where the critical exponent of the heat capacity is determined to be $\alpha = 0$ (discontinuous heat capacity, which is the mean-field result) for the Ising model in a small-world network.

Note that this is true as long as the deterministic solution depends on degree moments μ_m that are well defined in the thermodynamic limit $N \to \infty$. This may not be the case on scale free networks [31], where the deterministic solution may depend on the $\mu_{2,3,4,\ldots}$ degree moments that diverge, depending on the value of the exponent of the power law degree distribution, as $N \to \infty$. In this case, this may imply that the critical exponents depend on the details of the degree distribution [87]. As explained in the next section, one may redefine the finite-size scaling functions and critical exponents to take into account the N dependence of the degree moments.

In the next section, we propose a different method for solving the master equation close to a critical region, where the van Kampen expansion fails. We also show how to determine the exponents and scaling properties of the models close to a critical point, for a finite-system and also in the thermodynamic limit $N \to \infty$.

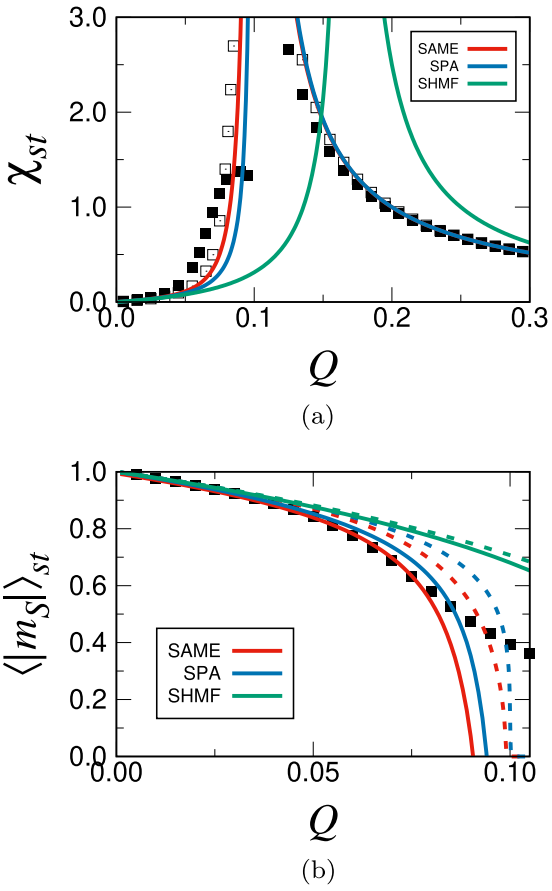


FIG. 4. (a) Stationary susceptibility $\chi_{\rm st}$ and (b) average magnetization $\langle|m_S|\rangle_{\rm st}$ as a function of Q on a three-regular random network for the majority-vote model. Points correspond to numerical simulations of the model with $N = 100$ (solid squares) and 400 (empty squares) averaged over an ensemble of 100 networks. Lines of different colors are the theoretical prediction of the different approximations (the curves that do not appear are superposed). In the top panel (a), the solid lines are the van Kampen result (34), while in the bottom panel (b), the dashed lines are the deterministic approach and the solid lines the corrected average values (35).

VI. THE EXPANSION AROUND A CRITICAL POINT

A. Formulation

Usually, the rates of the model depend on a set of parameters. Take, for example, a single parameter T for simplicity. It may happen that at a determined value $T = T_c$, one of the eigenvalues of the linearized deterministic dynamics becomes equal to zero $D_1 = 0$, this is called critical or bifurcation point. The proposed system size expansion $\mathbf{x} = N\boldsymbol{\phi} + N^{1/2}\boldsymbol{a} + N^0\boldsymbol{b}$ in this case leads to singular, divergent, results for the correlations and average value corrections [71]. The mathematical divergence of the correlations near the critical point is an accurate description only in the strict thermodynamic limit $N \to \infty$. When N is finite, near the critical point, we have an anomalous scaling with system size, which implies that we have to consider a different ansatz for the system size expansion [32].

In order to deal with such situations, we start by finding the linear transformation that diagonalizes the Jacobian

matrix $\mathbf{B}_{\rm st} = \mathbf{PDP}^{-1}$ being $\mathbf{D}$ the diagonal matrix composed by the eigenvalues and $\mathbf{P}$ the matrix of change of basis whose columns are the corresponding eigenvectors, all evaluated at the critical point $T = T_c$. We define the transformed variables in the eigenvector basis $\mathbf{u} = \mathbf{P}^{-1}\boldsymbol{\phi}$, such that the deterministic dynamics of the new variables is

$$\frac{d\mathbf{u}}{dt} = \mathbf{U} \equiv \mathbf{P}^{-1}\mathbf{\Phi}(\mathbf{Pu}). \tag{39}$$

At the critical point we have $U_i(T_c, \mathbf{u}_{\rm st}) = 0$ and $\partial_{u_j} U_i(T_c, \mathbf{u}_{\rm st}) = -D_i\delta_{ij}$ with $D_1 = 0$. The center manifold theory [88–90] states that, in this case, there exists a special trajectory or center manifold $u_i = h_i(T, u_1)$ for $i \neq 1$ with $u_i^{\rm st} = h_i(T, u_1^{\rm st})$ and $\partial_{u_1} h_i(T_c, u_1^{\rm st}) = 0$, that describes locally the dynamics of $\mathbf{u}$ close to the critical point T_c and near the fixed point $\mathbf{u}_{\rm st}$. This implies that the time dependence of the fast variables $u_{i>1}(t)$ is enslaved to the slow variable $u_1(t)$. We can write the dependence of $h_i(T, u_1)$ as a series expansion

$$\begin{aligned} h_i(T, u_1) = u_i^{\rm st} + \alpha_i^{(10)}(T - T_c) + \alpha_i^{(02)}\left(u_1 - u_1^{\rm st}\right)^2 \\ + \alpha_i^{(11)}(T - T_c)\left(u_1 - u_1^{\rm st}\right) + \dots, \end{aligned} \tag{40}$$

where the other terms of the expansion are neglected, for example, $(T - T_c)^2$, $(T - T_c)(u_1 - u_1^{\rm st})^2$, etc. The coefficients $\alpha_i^{(10)}$, $\alpha_i^{(11)}$, $\alpha_i^{(02)}$ can be determined expanding the dynamical equation $\dot{u}_i = \partial_{u_1} h_i \cdot \dot{u}_1 = U_i(T, u_1, h_2, h_3, \dots)$, the expressions are displayed in Appendix B.

The dynamics of u_1 inside the center manifold is $\dot{u}_1 = U_1(T, u_1, h_2, h_3, \dots)$ whose series expansion reads

$$\begin{aligned} \dot{u}_1 = \beta^{(0m)}\left(u_1 - u_1^{\rm st}\right)^m \\ + \beta^{(1n)}(T - T_c)\left(u_1 - u_1^{\rm st}\right)^n + \dots, \end{aligned} \tag{41}$$

where $\beta^{(0m)}$, $m \geqslant 2$, and $\beta^{(1n)}$, $n \geqslant 0$, are the lowest nonzero terms in the expansion in powers of $(u_1 - u_1^{\rm st})$ and higher-order terms are neglected. The expressions of the first coefficients $\beta^{(10)}$, $\beta^{(11)}$, $\beta^{(02)}$, $\beta^{(03)}$ are given in Appendix B.

Equation (41) is called the normal form of the bifurcation [90] and depending on the value of the coefficients it characterizes three types of critical points/bifurcations. If $\beta^{(10)} \neq 0$ the bifurcation is a *saddle node*; while if $\beta^{(10)} = 0$, but $\beta^{(11)} \neq 0$, the bifurcation is said to be *transcritical* for m even, or *pitchfork* for m odd.

From the normal form one can determine the critical exponent β. Setting the time derivative of Eq. (41) equal to zero and keeping in mind that $u_1^{\rm st}$ refers to the fixed point at the critical point $u_1^{\rm st}(T_c)$. For the saddle node, with normal form $0 = \beta^{(0m)}(u_1^{\rm st} - u_1^{\rm st})^m + \beta^{(10)}(T - T_c)$, we obtain $u_1^{\rm st}(T) - u_1^{\rm st}(T_c) \propto |T - T_c|^{1/m}$, while for the transcritical and pitchfork bifurcations, with normal form $0 = \beta^{(0m)}(u_1^{\rm st} - u_1^{\rm st})^m + \beta^{(11)}(u_1 - u_1^{\rm st})(T - T_c)$, we have $u_1^{\rm st}(T) - u_1^{\rm st}(T_c) \propto |T - T_c|^{1/(m-1)}$ (m even for the transcritical and odd for the pitchfork).

Thus $\beta = 1/m$ for the saddle and $\beta = 1/(m-1)$ for the transcritical and pitchfork bifurcations. Note that in Eq. (41), we only keep the two most important terms of the expansion to study the behavior of the stable fixed point close to the transition, and the others can be neglected. This can be checked introducing the first-order result $u_1^{\rm st}(T) - u_1^{\rm st}(T_c) \propto |T - T_c|^{\beta}$ in the expansion Eq. (41) and evaluating the order of each term.

Once this is understood, we propose the following system-size expansion based on the results of Refs. [32,75,76]. If we approach the critical point as $(T - T_c) \sim N^{-r}$, $0 < r < 1$ and we define the transformed y_i variables as $y_i = \sum_j P_{ij}^{-1} x_j$, then y_i follows the center manifold with small deviations of order $N^{1/2}$, while the stochastic part of y_1 has an anomalous scaling N^{υ}, $1/2 < \upsilon < 1$, namely,

$$T = T_c + N^{-r}\xi_0, \tag{42}$$

$$y_1 = N u_1^{\rm st} + N^{\upsilon}\xi_1, \tag{43}$$

$$y_i = N h_i\left(T, \frac{y_1}{N}\right) + N^{1/2}\xi_i. \tag{44}$$

Note that r and υ are parameters to be determined and that fluctuations inside the slow center manifold are assumed to scale differently than fluctuations outside it. Using this change of variables $(T, y_1, y_{i>1}) \to (\xi_0, \xi_1, \xi_{i>1})$, we can expand the master equation (4) in powers of N, this is done in detail in Appendix B. During the expansion we determine that for a saddle node, $\beta^{(10)} \neq 0$, it is $r = \upsilon = \frac{m}{m+1}$, while for the transcritical or pitchfork bifurcations, $\beta^{(10)} = 0$, it is $\upsilon = \frac{m}{m+1}$, $r = \frac{m-1}{m+1}$. After the expansion of the master equation we obtain a Fokker-Planck equation for the probability $\Pi(\xi_1; t)$ of the slow variable ξ_1, which for the transcritical and pitchfork bifurcation reads

$$\frac{\partial \Pi(\xi_1; t)}{\partial t} = N^{-\frac{m-1}{m+1}} \frac{\partial}{\partial \xi_1}\left[-\left(\beta^{(11)}\xi_0\xi_1 + \beta^{(0m)}\xi_1^m\right)\Pi + \frac{1}{2}F_{11}\frac{\partial \Pi}{\partial \xi_1}\right], \tag{45}$$

with a noise intensity $F_{11} = \sum_{i,j} P_{1i}^{-1} P_{1j}^{-1} G_{ij}$. For the saddle node, we obtain the same equation but replacing $\beta^{(11)}\xi_0\xi_1$ by $\beta^{(10)}\xi_0$. Note that the equation evolves at a slow time scale $\tau = N^{(m-1)/(m+1)}$, this is known in the literature as critical slowing down. In the stationary state for the transcritical and pitchfork bifurcations, we have

$$\Pi_{\rm st}(\xi_1) \propto \exp\left(\frac{\beta^{(11)}\xi_0}{F_{11}}\xi_1^2 + \frac{2\beta^{(0m)}}{(m+1)F_{11}}\xi_1^{m+1}\right). \tag{46}$$

This corresponds to a Gaussian distribution with a saturation term that the van Kampen approach does not take into account. For the saddle node one should replace $\beta^{(11)}\xi_0\xi_1^2$ by $2\beta^{(10)}\xi_0\xi_1$ and the distribution is no longer Gaussian. Note that if m is even we can not integrate the probability (46) in the entire range of ξ_1 and we have to restrict it to the "stable" zone, where fluctuations are not big enough to drive the dynamics to a zone where the deterministic dynamics is unstable and evolves towards infinity.

Any moment of the y_1 variable can be computed integrating the distribution Eq. (46), for example, the variance:

$$\sigma^2[y_1] \equiv N^{-1}\left(\langle y_1^2\rangle - \langle y_1\rangle^2\right) = N^{2\upsilon-1}\cdot\widetilde{\sigma}^2[N^r(T-T_c)], \tag{47}$$

where $\widetilde{\sigma}^2[\xi_0]$ is the variance of the ξ_1 variable. Now, the average values and correlations of the $\mathbf{x}$ variables can be related to the transformed variables $\mathbf{y}$ as

$$\frac{\langle x_i\rangle}{N} = \phi_i^{\rm st} + N^{\upsilon-1}P_{i1}\langle\xi_1\rangle, \tag{48}$$

$$N^{-1}[\langle x_i x_j\rangle - \langle x_i\rangle\langle x_j\rangle] = P_{i1}P_{j1}\sigma^2[y_1], \tag{49}$$

and from this it is straightforward to determine $\langle m\rangle_{\rm st}$, $\langle\rho\rangle_{\rm st}$, and $\chi_{\rm st}$ with the definitions given in Sec. II.

In the thermodynamic limit $N\to\infty$, one can show that the van Kampen result is recovered naturally. Take, for example, a pitchfork bifurcation with $\langle m_S\rangle_{\rm st}=0$. In this case, according to Eqs. (48) and (49) the scaling properties of $\langle|m_S|\rangle_{\rm st}$ and $\chi_{\rm st}$ with N are

$$\langle|m_S|\rangle_{\rm st} = N^{\upsilon-1}\widetilde{m}[N^r(T-T_c)], \tag{50}$$

$$\chi_{\rm st} = N^{2\upsilon-1}\widetilde{\chi}[N^r(T-T_c)], \tag{51}$$

where $\widetilde{m}(\xi_0)$ and $\widetilde{\chi}(\xi_0)$ are the respective scaling functions determined from Eqs. (46)–(49). In the limit $N\to\infty$ of Eqs. (50) and (51), the argument $\xi_0 = N^r(T-T_c)\to\infty$ (for $T\neq T_c$) and if we assume the scaling relations $\widetilde{m}\sim\xi_0^{\beta}$ and $\widetilde{\chi}\sim\xi_0^{-\gamma}$, with appropriate exponents β and γ such that $\langle|m_S|\rangle_{\rm st}$ and $\chi_{\rm st}$ are N-independent, we obtain consistently $\beta=\frac{1-\upsilon}{r}=\frac{1}{m-1}$ and $\gamma=\frac{2\upsilon-1}{r}=1$. Another quantity that is of great interest and that we will use in the next section is the Binder cumulant, defined as the ratio of moments $U_4\equiv 1-\frac{\langle m_S^4\rangle}{3\langle m_S^2\rangle^2}$. It is easy to show that the scaling of this function is given by

$$U_4 = \widetilde{u}[N^r(T-T_c)], \tag{52}$$

where $\widetilde{u}(\xi_0) = 1-\frac{\langle\xi_1^4\rangle_{\rm st}}{3\langle\xi_1^2\rangle_{\rm st}^2}$ is nothing but the Binder cumulant of the ξ_1 variable, that can be determined using the probability (46) (note that this is independent of the eigenvector coefficients P_{i1}, since they cancel out when computing the ratio of moments).

The scaling functions $\widetilde{m}$ and $\widetilde{\chi}$ generally depend on the degree moments μ_m which may scale with system-size N in a nontrivial way for certain types of highly heterogeneous networks, such as scale-free. It is possible, depending on the model, to reabsorb this N dependence by redefining the scaling functions (50) and (51), see Ref. [56], and this may imply network dependent critical exponents β, γ, see Ref. [87].

In the next section, we apply this method to the models studied in Sec. V B and we check if it corrects the problems of the van Kampen expansion in the critical zone.

1. Epidemic models (contact process)

Most epidemic models cannot be studied simply applying the previous techniques and some extra considerations have to be taken into account. The basic problem is that the models have an absorbing state for $\varepsilon=0$, and the noise intensity becomes equal to zero in this case, $F_{11}=0$ (evaluated at the absorbing state and at the critical point $\varepsilon=0$, $\lambda=\lambda_c$) and thus, the stationary probability Eq. (46) is ill-defined. In order to solve this problem let us consider that the deterministic analysis reveals a transcritical transition with $\beta^{(01)}=0$ and $\beta^{(02)}\neq 0$. Additionally, due to the existence of the absorbing state we have $F_{11}=0$ and we must consider a second-order term. In this case, Eq. (45) transforms into (see Appendix B):

$$\frac{\partial\Pi(\xi_1;t)}{\partial t} = N^{-1/2}\frac{\partial}{\partial\xi_1}\left[-\left(\beta^{(11)}\xi_0\xi_1+\beta^{(02)}\xi_1^2\right)\Pi + \frac{\gamma}{2}\frac{\partial}{\partial\xi_1}[\xi_1\Pi]\right], \tag{53}$$

with $\gamma=\frac{\partial F_{11}}{\partial u_1}$ (evaluated at the absorbing state and at the critical point $\varepsilon=0$, $\lambda=\lambda_c$). The stationary solution of Eq. (53) is

$$\Pi_{\rm st}(\xi_1)\propto\xi_1^{-1}\exp\left[\frac{2}{\gamma}\left(\beta^{(11)}\xi_0\xi_1+\frac{\beta^{(02)}}{2}\xi_1^2\right)\right], \tag{54}$$

which is singular around $\xi_1=0$, i.e., the absorbing state. This is a problem in the simulation too, when averaging over realizations of the stochastic process, and different solutions to avoid the absorbing state can be found in Ref. [4]. In this case, the scaling function follows the relation:

$$\langle m\rangle_{\rm st} = N^{-1/2}\widetilde{m}[N^{1/2}(\lambda-\lambda_c)]. \tag{55}$$

In order to check if the theory is correct we will compare it with the numerical results in Ref. [47] for the contact process on scale free networks.

The rates of the contact process are $R^+_{k,q,\{k_j\}_{j=k-q+1}^k} = \lambda\sum_{j=k-q+1}^k k_j^{-1}$ and $R^-_{k,q}=1$. The averaged rate then reads $\widetilde{R}^+_{k,q}=\lambda q\frac{\sum_k\phi_k}{\sum_k\phi_k k}$ and the deterministic equations in the HMF description are

$$\frac{d\phi_k}{dt} = \lambda\frac{P_k k}{\mu}\phi-\phi_k-\lambda\frac{\phi_k k}{\mu}\phi, \tag{56}$$

with $\phi_k^{\rm st}=0$ as the trivial (absorbing state) fixed point. Linearizing around the absorbing state we find that the critical point (in the HMF) is $\lambda_c=1$. The eigenvalues are: $D_1=\lambda_c-\lambda$ with associated eigenvector $\mathbf{v}_1=(P_{k_{\rm min}}k_{\rm min},\ldots,P_{k_{\rm max}}k_{\rm max})/\mu$ and $D_2=1$ with any vector $\mathbf{v}_2$ in the plane $\mathbf{n}^{\intercal}\cdot\mathbf{v}_2=0$ with $\mathbf{n}^{\intercal}=(1,\ldots,1)$.

Trajectories in the center manifold fulfill $\phi_k=\frac{P_k k}{\mu}\phi+O((\lambda-1)\phi)+O(\phi^2)$, thus we can obtain a closed equation for the global (slow) variable ϕ:

$$\frac{d\phi}{dt} = (\lambda-1)\phi-\frac{\mu_2}{\mu^2}\phi^2+O((\lambda-1)\phi^2)+O(\phi^3). \tag{57}$$

From this we conclude that the bifurcation is transcritical with coefficients $\beta^{(11)}=1$ and $\beta^{(02)}=-\mu_2/\mu^2$. The

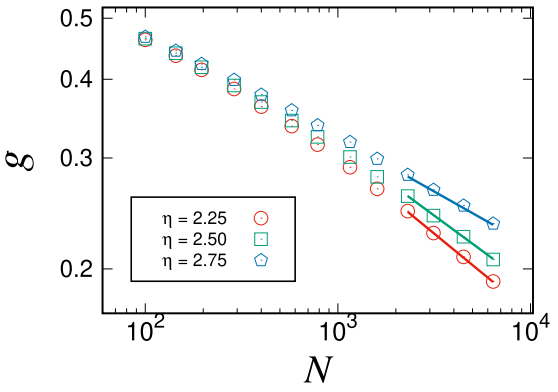


FIG. 5. Prefactor $g \equiv \sum_i P_{i1}$ for the contact process at the AME compartmental approach description level as a function of the system size N with a power law degree distribution $P_k \sim k^{-\eta}$, $k_{\min} = 2$, $k_{\max} = \sqrt{N}$. The solid lines are linear regressions of the last four points.

prefactor from Eq. (48) to obtain the scaling function of m is $g \equiv \sum_i P_{i1} = \sum_k \frac{P_k k}{\mu} = 1$.

Now assume that the degree distribution is power law $P_k \sim k^{-\eta}$ and that we have the scaling for the moments as $\mu_2/\mu^2 \sim N^{\frac{3-\eta}{2}}$ [28] when $N \to \infty$. If we replace the average value with the maximum of the probability Eq. (54) in order to avoid the absorbing state problem, we have $\langle \xi_1 \rangle_{\rm st} \approx -\frac{2\beta^{(11)}\xi_0}{\beta^{(02)}}$. We can rewrite Eq. (55) reabsorbing the N dependence of the coefficient $\beta^{(02)}$ as

$$\langle m \rangle_{\rm st} = N^{-\beta/\bar{\nu}} \widetilde{\widetilde{m}}[N^{1/\bar{\nu}}(\lambda - \lambda_c)], \tag{58}$$

with a new scaling function $\widetilde{\widetilde{m}}$ and exponents $\beta/\bar{\nu} = 1/2$ and $\bar{\nu} = 2/(\eta - 2)$ and $\beta = 1/(\eta - 2)$, which coincide with the theoretical results obtained in Ref. [47].

The advantage of the method presented here is that it can be applied to all compartmental approaches. In the same reference [47], the exponents obtained from numerical simulation show deviation from the heterogeneous mean-field prediction. For example, $\beta/\bar{\nu}$ does depend on the exponent of the power law η as: $\beta/\bar{\nu} = 0.76(5)$ for $\eta = 2.25$, $\beta/\bar{\nu} = 0.70(3)$ for $\eta = 2.50$ and $\beta/\bar{\nu} = 0.63(4)$ for $\eta = 2.75$. These deviations are well captured by the stochastic AME approach. In Fig. 5, we showed that the prefactor $g \equiv \sum_i P_{i1}$ of the scaling function depends on system size as $g \sim N^{1/2-\beta/\bar{\nu}}$ at the AME description level, we obtain $\beta/\bar{\nu} = 0.73$ for $\eta = 2.25$, $\beta/\bar{\nu} = 0.71$ for $\eta = 2.50$, and $\beta/\bar{\nu} = 0.67$ for $\eta = 2.75$, in perfect agreement with the numerical results in Ref. [47].

B. Comparison with numerical simulations

1. Ising Glauber

We will start with the Ising model with Glauber rates for the network and parameter specifications in the caption of Fig. 3. The critical point predicted by the AME and PA approximations is $T_c(\text{AME/PA}) = 4.93\ldots$, while for the HMF, it is $T_c(\text{HMF}) = 4.96\ldots$ In order to determine the critical point numerically from the Monte Carlo (MC) simulations, we use a standard technique of statistical mechanics [91], which consists in computing the Binder cumulant defined as

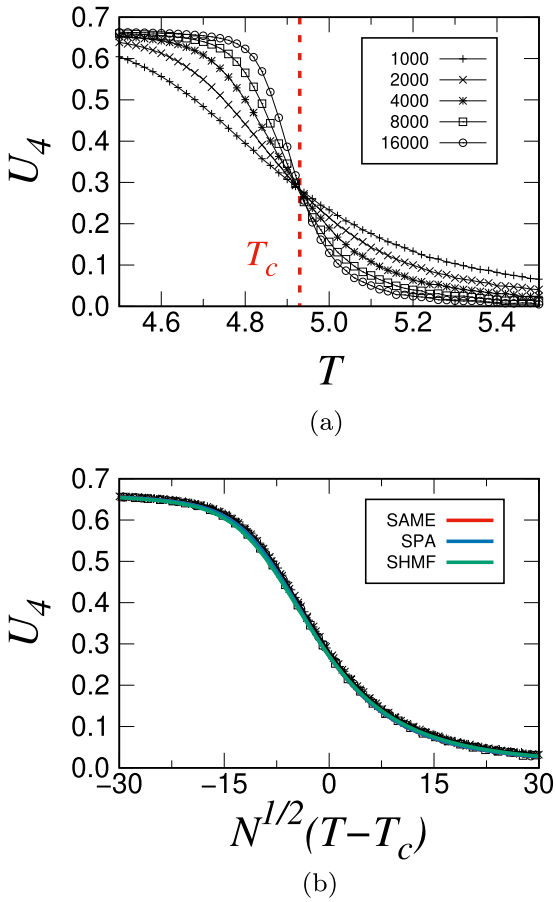


FIG. 6. Binder cumulant as a function of the temperature T (a) and as a function of the rescaled temperature $N^{1/2}(T - T_c)$ (b), for the Ising Glauber model with different system sizes N, specified in the figure. The parameters are $J = 1$ on an Erdős-Rényi network with average degree $\mu = 5$ and the results were averaged over an ensemble of 100 networks. Points correspond to numerical simulations of the model with different system sizes N specified in the legend, while lines are the theoretical scaling functions determined from Eqs. (46) and (52).

$U_4 \equiv 1 - \frac{\langle m_S^4 \rangle}{3\langle m_S^2 \rangle^2}$ for different system sizes N, such that all the different curves cross at the critical T_c, see Fig. 6. We obtain in this case $T_c(\text{MC}) = 4.93 \pm 0.01$ in perfect accordance to the AME/PA results. After computing the coefficients of the normal form of the bifurcation Eq. (41), we obtain for all three approaches AME/PA/HMF that $\beta^{(10)} = 0$, $\beta^{(11)} < 0$, $\beta^{(02)} = 0$ and $\beta^{(03)} < 0$ which indicates that, according to our discussion in Sec. VI, we have a pitchfork bifurcation with $m = 3$, $r = 1/2$, and $\upsilon = 3/4$. If the theory is correct, and the scaling properties Eqs. (50)–(52) are valid, if we rescale $\langle |m_S| \rangle_{\rm st}$, $\chi_{\rm st}$, U_4 by $N^{1/4}$, $N^{1/2}$, and N^0, respectively, and the temperature by $N^{1/2}$, all the curves should collapse on a single universal one $\widetilde{m}(\xi_0)$, $\widetilde{\chi}(\xi_0)$ and $\widetilde{u}(\xi_0)$. In Figs. 6 and 7, we compute this numerically and compare it with the theoretical scaling functions derived from Eq. (46). The matching between numerical and theory is very good which proves the validity of the method. Note also that the scaling functions of the AME and PA coincide, while the HMF shows some deviations. This indicates that there is a strong relation

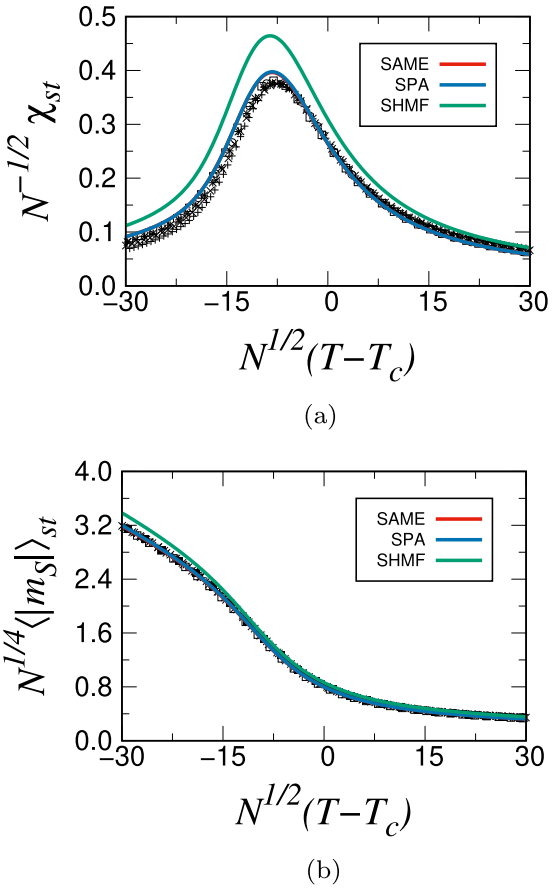


FIG. 7. (a) Rescaled stationary susceptibility $N^{-1/2}\chi_{\rm st} = N^{1/2}(\langle m_S^2\rangle_{\rm st} - \langle|m_S|\rangle_{\rm st}^2)$ and (b) average magnetization $N^{1/4}\langle|m_S|\rangle_{\rm st}$ as a function of $N^{1/2}(T - T_c)$ on an Erdős-Rényi network with average degree $\mu = 5$ for the Glauber Ising model with $J = 1$. Points correspond to numerical simulations of the model with different system sizes N specified in Fig. 6, while lines are the theoretical scaling functions determined from Eqs. (46), (50), and (51).

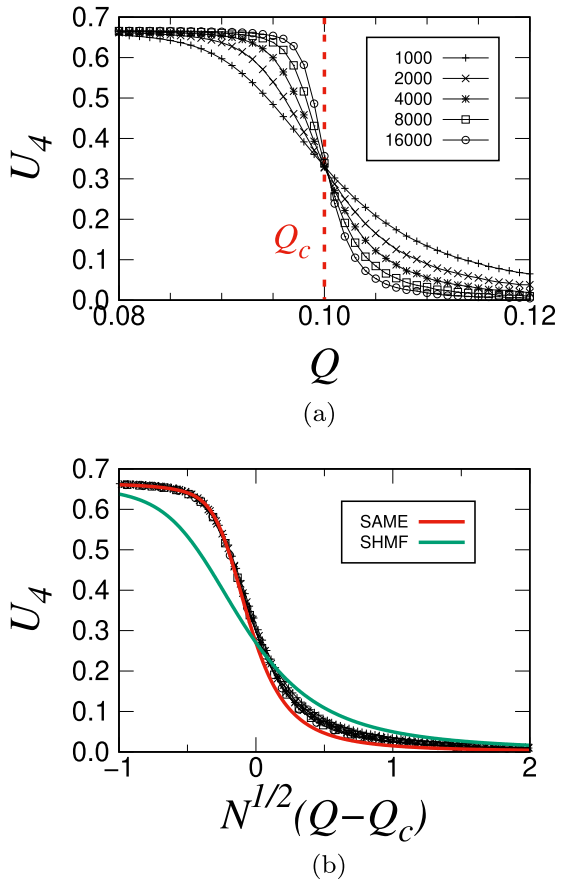


FIG. 8. Binder cumulant as a function of Q (a) and as a function of the rescaled $N^{1/2}(Q - Q_c)$ (b), for the majority-vote model with different system sizes N specified in the legend on a three-regular random network, and results were averaged over an ensemble of 100 networks. Points correspond to numerical simulations of the model with different system sizes N specified in the legend, while lines are the theoretical scaling functions determined from Eqs. (46) and (52). The finite-size scaling for the PA result is not displayed as it predicts incorrect scaling properties.

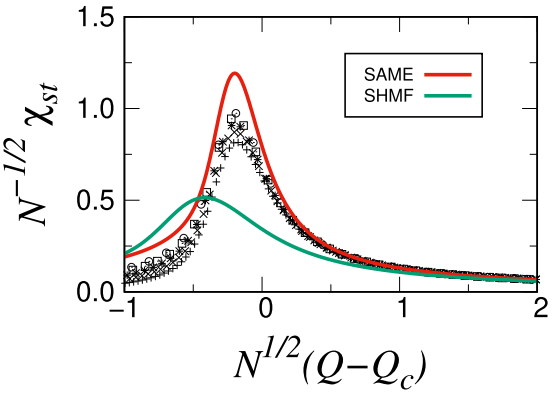
between the validity of the deterministic solution and the scaling functions.

2. Majority vote

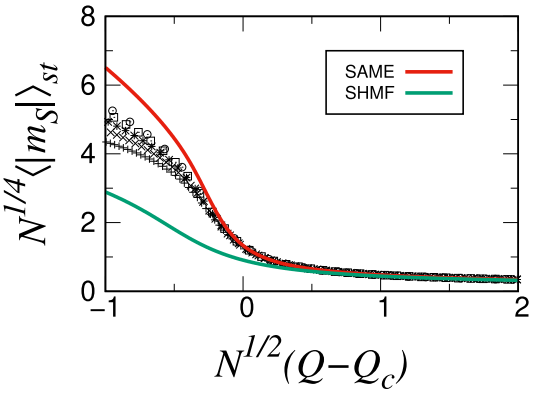
The next model that we study is the majority-vote model with the same specifications of Fig. 4. The critical point predicted by the AME is $Q_c(\text{AME}) = 0.099\ldots$, for the PA it is $Q_c(\text{PA}) = 0.100\ldots$ and the HMF is $Q_c(\text{HMF}) = 0.167\ldots$. The numerical critical point obtained from the Binder cumulant in Fig. 8 is $Q_c(\text{MC}) = 0.10 \pm 0.01$, compatible with the results of the AME and PA but not with the HMF. When we compute the coefficients of the normal form of the bifurcation Eq. (41) we obtain $\beta^{(10)} = 0$, $\beta^{(11)} < 0$, $\beta^{(02)} = 0$, and $\beta^{(03)} < 0$ for the AME and HMF, which corresponds again to a pitchfork bifurcation with $m = 3$, $r = 1/2$, and $\upsilon = 3/4$. Surprisingly, for the PA, we obtain instead $\beta^{(03)} = 0$ which suggests a different type of pitchfork with $m = 5$, $r = 2/3$, and $\upsilon = 5/6$. This could be already seen in Fig. 4, as $\langle|m_S|\rangle_{\rm st} \propto (Q_c - Q)^{1/2}$ for the AME but for the PA is more abrupt $\langle|m_S|\rangle_{\rm st} \propto (Q_c - Q)^{1/4}$. As a consequence, we conclude that the PA is not able to capture correctly the scaling properties in this case. In Figs. 8 and 9, we compared the theoretical scaling functions with the numerical simulations, where we see that the theoretical scaling of the AME offers a reasonable agreement. Note, however, how the convergence to the theoretical scaling is very slow for $Q < Q_c$. This is because for the AME $\beta^{(03)}$ is small, and the other higher-order terms of the normal form Eq. (41) may be important, unless N is extremely large. This also explains the failure of the PA that actually predicts $\beta^{(03)} = 0$.

VII. TIME DEPENDENCE

In the previous sections, we have focused on stationary averages. The methods, however, are straightforwardly generalized for time dependent results. For the van Kampen approach, we have to solve the deterministic dynamics $\frac{d\boldsymbol{\phi}}{dt} = \boldsymbol{\Phi}$ and, at the same time, the dynamics of the average values and correlations Eqs. (24)–(26). On the other hand, if we are close to a critical point in the parameter space, we assume that dynamics evolve following the center manifold and we

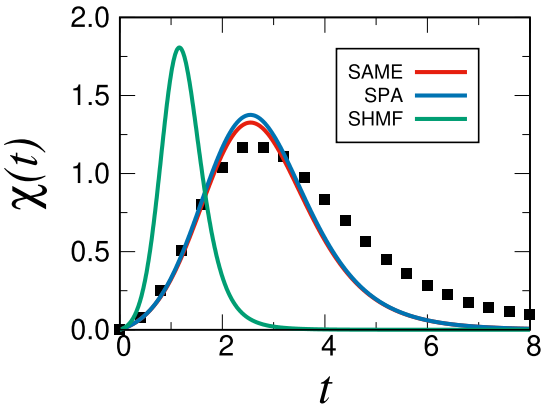


(a)

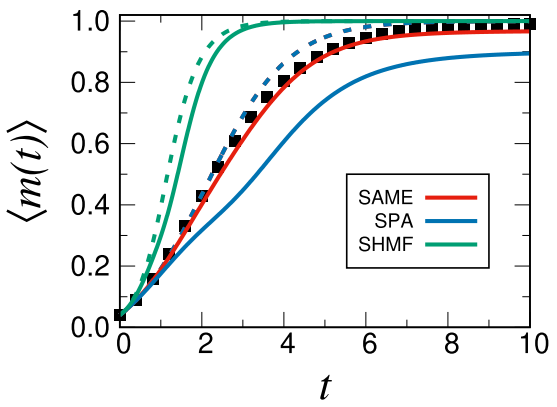


(b)

FIG. 9. (a) Rescaled stationary susceptibility $N^{-1/2}\chi_{\rm st} = N^{1/2}(\langle m_S^2\rangle_{\rm st} - \langle|m_S|\rangle_{\rm st}^2)$ and (b) average magnetization $N^{1/4}\langle|m_S|\rangle_{\rm st}$ as a function of $N^{1/2}(Q - Q_c)$ on a three-regular random network for the majority-vote model. Points correspond to numerical simulations of the model with different system sizes N specified in Fig. 8, while lines are the theoretical scaling functions determined from Eqs. (46), (50), and (51). The finite-size scaling for the PA result is not displayed as it predicts incorrect scaling properties.



(a)



(b)

FIG. 10. (a) Density of active nodes $\langle m(t)\rangle$ and (b) susceptibility $\chi(t)$ as a function of time t, for the SI epidemic dynamics with $\lambda = 1$ on a scale free network with $P_k \sim k^{-2.5}$, $k_{\rm min} = 2$ and $k_{\rm max} = 5$, and $N = 25$. Dots are numerical simulations averaged over 100 trajectories and 100 networks, while lines are (a) in the top panel it is the result of solving the dynamical Eq. (26) for the different approaches, (b) in the bottom panel the dashed lines are the deterministic results $\boldsymbol{\phi}(t)$, while the solid lines are corrected by the second-order term $\boldsymbol{\phi}(t) + \langle \boldsymbol{b}(t)\rangle/N$, Eq. (25).

have to solve Eq. (45), obtaining $\Pi(\xi_1; t)$. This corresponds to a separation of time scales, which implies that the dynamics outside the manifold is very fast compared to the dynamics inside and thus negligible. If neither of these two methods capture correctly the stochastic dynamics, we can still resort to the Langevin or Fokker-Planck Eqs. (37) and (38).

We will apply the methods to two different models of interest, not considered in the previous sections, the SI (susceptible-infected) epidemic model and the Threshold model. We chose these models because their dynamics are more interesting than the stationary properties.

A. Susceptible infected

We start with the SI epidemic model with rates $R^+_{k,q} = \lambda q$ and $R^-_{k,q} = 0$. This model does not have a critical point, and it shows weak finite-size effects. Thus, we expect the van Kampen expansion to work accurately. In Fig. 10, we compare the results of numerical simulations for a very small system of $N = 25$ nodes, with the results of integrating the dynamics of the van Kampen expansion Eqs. (17), (24), (25), and (26). For the susceptibility $\chi(t)$, the SAME and SPA give a good approximation with slight differences between both approaches, while the SHMF shows important discrepancies. For the average value $\langle m(t)\rangle$, the deterministic AME and PA give the same results, while HMF again shows discrepancies. We observe that, similarly to what happens in the stationary state Fig. 2, although the deterministic part of the AME and PA part is equal, the stochastic corrections happen to be only accurate for the SAME approach.

B. Threshold model

Next, we explore the threshold model [68,92], with rates $R^+_{k,q} = 1$ if $q \geqslant M_k$ and $R^+_{k,q} = 0$ if $q < M_k$ (where M_k are a set of integer parameters), and $R^-_{k,q} = 0$. The results plotted in Figs. 11 and 12 show some peculiarities. The model has an absorbing state for $m(t) = 1$ as the recovery rate is zero $R^-_{k,q} = 0$. In Fig. 11, we show that the deterministic AME predicts

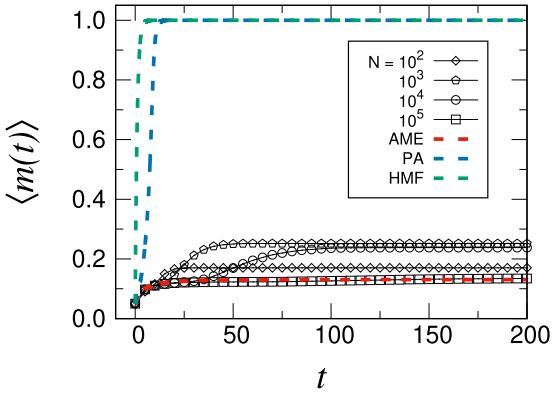


FIG. 11. Average density of active nodes $\langle m(t)\rangle$ as a function of time t for the threshold model with $M_k = 2$, $\forall k$ on a five-regular random network. Dots ($N = 10^2$ rhombuses, $N = 10^3$ pentagons, $N = 10^4$ circles, and $N = 10^5$ squares) are numerical simulations averaged over 100 trajectories and 100 networks, while dashed lines are the deterministic result $\phi(t)$ for the different approaches.

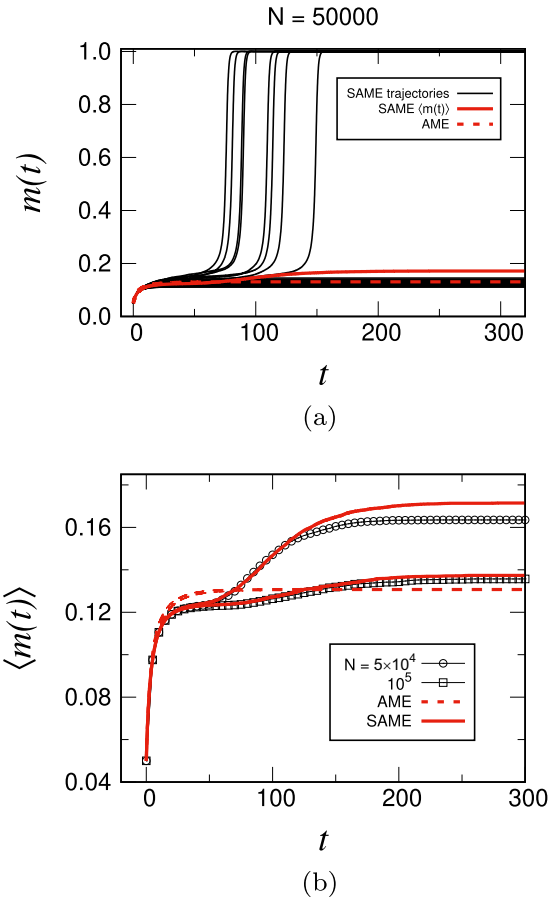


FIG. 12. Trajectories (a) of the density of active nodes $m(t)$ obtained by numerical integration of the Langevin Eq. (38) for the AME approach, with the same model specifications of Fig. 11. Solid red lines are the average value $\langle m(t)\rangle$ of these trajectories for $N = 5 \times 10^4$ (a) and (b) and $N = 10^5$ (b), while the dashed red line is the deterministic AME approach (a) and (b). In the bottom panel (b) dots correspond to the average value obtained from numerical simulations ($N = 5 \times 10^4$ circles and $N = 10^5$ squares).

that the system reaches a frozen staten with $m(t) < 1$ while the PA and HMF approach the absorbing state $m(t) \to 1$. In the same figure, we observe that numerical results for finite N are closer to the AME prediction but show large deviations from the deterministic value, even for large system sizes.

Curiously, finite-size corrections increase with system size until a certain value around $N = 10^3$–10^4 and then start decreasing and approaching the deterministic value $\langle m(t)\rangle \to \phi(t)$. It can be proven that neither the van Kampen expansion nor the expansion around a critical point work properly in this case, and they do not capture correctly finite-size effects. The van Kampen expansion of Sec. V does not work because the deterministic dynamics have eigenvalues with zero real part. The expansion around the critical point of Sec. VI neither does, because there are degenerate eigenvalues whose multiplicity is greater than the dimension of the space of eigenvectors of the corresponding eigenvalue, in other words the Jacobian matrix is not diagonalizable.

In order to elucidate what is the issue with finite-size effects in the Threshold model, we integrated directly the Langevin Eq. (38) using an Euler-Maruyama method [70] for the AME approach. We see that most trajectories evolve towards a frozen state with $m(t) < 1$, close to the deterministic AME prediction, with some variability over realizations. There are some trajectories, however, that eventually deviate to the absorbing state $m(t) = 1$ and this is the source of error of the deterministic description. We check that the stochastic description, in terms of a Langevin equation, is able to reproduce perfectly these deviations. In the bottom panel (b) of Fig. 12, we compare the average density of active nodes $\langle m(t)\rangle$ with simulation of the model with a good agreement as compared to the theoretical results of the stochastic AME. The PA and HMF results are omitted as they predict all trajectories to evolve towards the absorbing state.

VIII. SUMMARY AND CONCLUSIONS

In this paper, we have introduced theoretical tools to study stochastic effects in binary-state models on complex networks. First, we constructed the general master equation of the different compartmental approaches: approximate master equation (AME), pair approximation (PA) and heterogeneous mean field (HMF). After that, we elaborated on the different approximate methods for solving the master equation, in particular we explored the van Kampen expansion, valid far from a critical point, and a critical expansion, accurate at the critical zone. From the van Kampen expansion, we were able to obtain equations for the correlation matrix of the set of variables and the corrections to their average values, while from the critical expansion, we got their finite-size scaling functions.

We applied these techniques to characterize the stationary properties of the SIS and contact process epidemic, Glauber Ising, and majority-vote models. When comparing the performance of the different compartmental approaches to numerical simulations we conclude that, if AME and PA have equal or similar results at the deterministic level, the same goes for the fluctuations of the van Kampen expansion and the scaling functions of the critical expansion, but that is not the case for the finite-size corrections to the average

values which are only accurate for the AME. This is what we observe for the Glauber model where PA and AME have equivalent deterministic, fluctuations and scaling functions but different finite-size corrections to the average values. This is an indication that, although the PA may work very well in the determination of certain quantities such as average values, the binomial restriction between variables is not necessarily fulfilled by the stochastic trajectories. For the majority-vote model, the AME and PA give different results at all levels, where the PA even predicts incorrectly the scaling coefficients (critical exponents). For the contact process, we were able to obtain the finite-size scaling functions and corresponding critical exponents using the expansion around a critical point. For the stochastic AME, it matches perfectly the numerical results reported in Ref. [47].

In general, we can highlight that the probabilistic description using the AME gives very accurate results for stationary and also time-dependent results (as it is shown in Sec. VII for the SI epidemic model) within the range of validity of the expansion methods. For some model like the threshold model the van Kampen expansion or the expansion around a critical point do not reproduce correctly the numerical results. In this case, we resorted to the Kramers-Moyal expansion, in terms of a Langevin stochastic differential equation. We showed that the theoretical description, in terms of the stochastic AME, predicts that some stochastic trajectories evolve towards a frozen state while others towards an absorbing (consensus) state. This is not the case of the deterministic AME, which only predicts the evolution towards the frozen state.

The solutions of the equations for the average values and the fluctuations have been performed numerically using an efficient method developed in Ref. [71]. It is left for a future work to explore the possibility of obtaining analytical results of the models and the general conditions for the AME–PA equivalence at the stochastic level. A particularly interesting case that has not been considered in this work is the noisy-voter (Kirman) model for which the linearity of the rates allows one to close the equations for the moments and correlations without the need to resort to the van Kampen approximation. This was done for the homogeneous pair approximation in Ref. [32] and it would be interesting to extend these results to the more complicated compartmental models considered in this work. Initial results indicate that the AME corrects the lacks of the PA in specific cases. The differences being specially notorious for low dense—highly heterogeneous networks [93].

ACKNOWLEDGMENTS

Partial financial support has been received from the Agencia Estatal de Investigación (AEI, MCI, Spain) and Fondo Europeo de Desarrollo Regional (FEDER, UE), under Project PACSS (RTI2018-093732-B-C21/C22) and the Maria de Maeztu Program for units of Excellence in R&D (MDM-2017-0711). A.F.P. acknowledges support by the Formación de Profesorado Universitario (FPU14/00554) program of Ministerio de Educación, Cultura y Deportes (MECD) (Spain). A.F.P. thanks James Gleeson for the warm hospitality and useful discussions concerning the methods of Ref. [31] during a stay at the University of Limerick.

APPENDIX A: EXPRESSIONS OF B, G, AND HESSIAN MATRICES

In this section, we will calculate all the necessary ingredients involved in the equations of the correlations and corrections (24)–(26). This essentially includes the Jacobian **B**-matrix, **G**-matrix, and the set of Hessian matrices $\partial^2_{\phi_j,\phi_k}\Phi_i$ that can be later used to determine $\Gamma_i = \frac{1}{2}\sum_{j,k}\langle a_j a_k\rangle \partial^2_{\phi_j,\phi_k}\Phi_i$. We will proceed for the different levels of description explained in the main text. All the expressions developed here have been incorporated in a FORTRAN code that can be obtained from the authors upon request.

1. Approximate master equation

Starting with the variables $\{N_{n,k,q}\}$, the **B** matrix can be calculated by taking the derivatives $B_{n,k,q;n',k',q'} = -\partial_{\phi_{n',k',q'}}\Phi_{n,k,q}$, with definitions Eqs. (18) and (19), this is

$$B_{0,k,q;0,k',q'} = \delta_{k,k'}\delta_{q,q'}(R^+_{k,q} + (k-q)\beta^s + q\gamma^s) - \delta_{k,k'}\delta_{q-1,q'}(k-q+1)\beta^s - \delta_{k,k'}\delta_{q+1,q'}(q+1)\gamma^s + [(k-q)\phi_{0,k,q} - (k-q+1)\phi_{0,k,q-1}]\frac{\partial\beta^s}{\partial\phi_{0,k',q'}}, \tag{A1}$$

$$B_{0,k,q;1,k',q'} = -\delta_{k,k'}\delta_{q,q'}R^-_{k,q} + [q\phi_{0,k,q} - (q+1)\phi_{0,k,q+1}]\frac{\partial\gamma^s}{\partial\phi_{1,k',q'}}, \tag{A2}$$

$$B_{1,k,q;0,k',q'} = -\delta_{k,k'}\delta_{q,q'}R^+_{k,q} + [(k-q)\phi_{1,k,q} - (k-q+1)\phi_{1,k,q-1}]\frac{\partial\beta^i}{\partial\phi_{0,k',q'}}, \tag{A3}$$

$$B_{1,k,q;1,k',q'} = \delta_{k,k'}\delta_{q,q'}(R^-_{k,q} + (k-q)\beta^i + q\gamma^i) - \delta_{k,k'}\delta_{q-1,q'}(k-q+1)\beta^i - \delta_{k,k'}\delta_{q+1,q'}(q+1)\gamma^i + [q\phi_{1,k,q} - (q+1)\phi_{1,k,q+1}]\frac{\partial\gamma^i}{\partial\phi_{1,k',q'}}. \tag{A4}$$

Here, the derivatives of the rates (20)–(23) are

$$\frac{\partial \beta^s}{\partial \phi_{0,k,q}} = \frac{(k-q)(R^+_{k,q}-\beta^s)}{\sum_{k,q}(k-q)\phi_{0,k,q}}, \tag{A5}$$

$$\frac{\partial \gamma^s}{\partial \phi_{1,k,q}} = \frac{(k-q)(R^-_{k,q}-\gamma^s)}{\sum_{k,q}(k-q)\phi_{1,k,q}}, \tag{A6}$$

$$\frac{\partial \beta^i}{\partial \phi_{0,k,q}} = \frac{q(R^+_{k,q}-\beta^i)}{\sum_{k,q} q\phi_{0,k,q}}, \tag{A7}$$

$$\frac{\partial \gamma^i}{\partial \phi_{1,k,q}} = \frac{q(R^-_{k,q}-\gamma^i)}{\sum_{k,q} q\phi_{1,k,q}}. \tag{A8}$$

The **G**-matrix can be calculated as $G_{n,k,q;n',k',q'} = \sum_\nu \ell^{(\nu)}_{n,k,q}\ell^{(\nu)}_{n',k',q'}w^{(\nu)}(\boldsymbol{\phi})$, where $\ell^{(\nu)}_{n,k,q}$ are Eqs. (11)–(14) and $w^{(\nu)}(\boldsymbol{\phi})$ the intensive version of Eqs. (5) and (6) evaluated at $N_{n,k,q} \to N\phi_{n,k,q}$. This leads, after lengthy algebra, to

$$\begin{aligned}
&G_{0,k,q;0,k',q'}\\
&= \phi_{0,k,q}R^+_{k,q}\delta_{k,k'}\delta_{q,q'} - \phi_{0,k,q}R^+_{k,q}(k-q)(-P_0(0,k',q') + P_0(0,k',q'-1))\\
&\quad - \phi_{0,k',q'}R^+_{k',q'}(k'-q')(-P_0(0,k,q) + P_0(0,k,q-1)) + \beta^{nss}[P_0(0,k,q)P_0(0,k',q') - P_0(0,k,q-1)P_0(0,k',q')\\
&\quad - P_0(0,k,q)P_0(0,k',q'-1) + P_0(0,k,q-1)P_0(0,k',q'-1)] + \beta^{ns}[\delta_{k,k'}\delta_{q,q'}P_0(0,k,q) - \delta_{k,k'}\delta_{q-1,q'}P_0(0,k,q-1)\\
&\quad - \delta_{k,k'}\delta_{q+1,q'}P_0(0,k,q) + \delta_{k,k'}\delta_{q,q'}P_0(0,k,q-1)]\\
&\quad + \phi_{1,k,q}R^-_{k,q}\delta_{k,k'}\delta_{q,q'} + \phi_{1,k,q}R^-_{k,q}(k-q)(-P_1(0,k',q') + P_1(0,k',q'+1))\\
&\quad + \phi_{1,k',q'}R^-_{k',q'}(k'-q')(-P_1(0,k,q) + P_1(0,k,q+1)) + \gamma^{nss}[P_1(0,k,q)P_1(0,k',q') - P_1(0,k,q)P_1(0,k',q'+1)\\
&\quad - P_1(0,k,q+1)P_1(0,k',q') + P_1(0,k,q+1)P_1(0,k',q'+1)] + \gamma^{ns}[\delta_{k,k'}\delta_{q,q'}P_1(0,k,q) - \delta_{k,k'}\delta_{q-1,q'}P_1(0,k,q)\\
&\quad - \delta_{k,k'}\delta_{q+1,q'}P_1(0,k,q+1) + \delta_{k,k'}\delta_{q,q'}P_1(0,k,q+1)],
\end{aligned} \tag{A9}$$

$$\begin{aligned}
G_{0,k,q;1,k',q'} &= -\phi_{0,k,q}R^+_{k,q}\delta_{k,k'}\delta_{q,q'} - \phi_{0,k,q}R^+_{k,q}q(-P_0(1,k',q') + P_0(1,k',q'-1))\\
&\quad + \phi_{0,k',q'}R^+_{k',q'}(k'-q')(-P_0(0,k,q) + P_0(0,k,q-1)) + \beta^{nsi}[P_0(0,k,q)P_0(1,k',q') - P_0(0,k,q)P_0(1,k',q'-1)\\
&\quad - P_0(0,k,q-1)P_0(1,k',q') + P_0(0,k,q-1)P_0(1,k',q'-1)] - \phi_{1,k,q}R^-_{k,q}\delta_{k,k'}\delta_{q,q'}\\
&\quad + \phi_{1,k,q}R^-_{k,q}q(-P_1(1,k',q') + P_1(1,k',q'+1)) - \phi_{1,k',q'}R^-_{k',q'}(k'-q')(-P_1(0,k,q) + P_1(0,k,q+1))\\
&\quad + \gamma^{nsi}[P_1(0,k,q)P_1(1,k',q') - P_1(0,k,q)P_1(1,k',q'+1)\\
&\quad - P_1(0,k,q+1)P_1(1,k',q') + P_1(0,k,q+1)P_1(1,k',q'+1)],
\end{aligned} \tag{A10}$$

$$\begin{aligned}
&G_{1,k,q;1,k',q'}\\
&= \phi_{0,k,q}R^+_{k,q}\delta_{k,k'}\delta_{q,q'} + \phi_{0,k,q}R^+_{k,q}q(-P_0(1,k',q') + P_0(1,k',q'-1))\\
&\quad + \phi_{0,k',q'}R^+_{k',q'}q'(-P_0(1,k,q) + P_0(1,k,q-1)) + \beta^{nii}[P_0(1,k,q)P_0(1,k',q') - P_0(1,k,q-1)P_0(1,k',q')\\
&\quad - P_0(1,k,q)P_0(1,k',q'-1) + P_0(1,k,q-1)P_0(1,k',q'-1)] + \beta^{ni}[\delta_{k,k'}\delta_{q,q'}P_0(1,k,q) - \delta_{k,k'}\delta_{q-1,q'}P_0(1,k,q-1)\\
&\quad - \delta_{k,k'}\delta_{q+1,q'}P_0(1,k,q) + \delta_{k,k'}\delta_{q,q'}P_0(1,k,q-1)]\\
&\quad + \phi_{1,k,q}R^-_{k,q}\delta_{k,k'}\delta_{q,q'} - \phi_{1,k,q}R^-_{k,q}q(-P_1(1,k',q') + P_1(1,k',q'+1)) - \phi_{1,k',q'}R^-_{k',q'}q'(-P_1(1,k,q) + P_1(1,k,q+1))\\
&\quad + \gamma^{nii}[P_1(1,k,q)P_1(1,k',q') - P_1(1,k,q)P_1(1,k',q'+1) - P_1(1,k,q+1)P_1(1,k',q') + P_1(1,k,q+1)P_1(1,k',q'+1)]\\
&\quad + \gamma^{ni}[\delta_{k,k'}\delta_{q,q'}P_1(1,k,q) - \delta_{k,k'}\delta_{q-1,q'}P_1(1,k,q) - \delta_{k,k'}\delta_{q+1,q'}P_1(1,k,q+1) + \delta_{k,k'}\delta_{q,q'}P_1(1,k,q+1)].
\end{aligned} \tag{A11}$$

The **G**-matrix is symmetric, so that $G_{1,k,q;0,k',q'} = G_{0,k',q';1,k,q}$. The probabilities (7)–(10) must be understood again as evaluated at the deterministic $N_{n,k,q} \to N\phi_{n,k,q}$, and we have defined the new averaged rates

$$\beta^{nss} = \sum_{k,q}(k-q)(k-q-1)\phi_{0,k,q}R^+_{k,q}, \tag{A12}$$

$$\beta^{ns} = \sum_{k,q}(k-q)\phi_{0,k,q}R^+_{k,q}, \tag{A13}$$

$$\gamma^{nss} = \sum_{k,q}(k-q)(k-q-1)\phi_{1,k,q}R^{-}_{k,q}, \tag{A14}$$

$$\gamma^{ns} = \sum_{k,q}(k-q)\phi_{1,k,q}R^{-}_{k,q}, \tag{A15}$$

$$\beta^{nsi} = \sum_{k,q} q(k-q)\phi_{0,k,q}R^{+}_{k,q}, \tag{A16}$$

$$\gamma^{nsi} = \sum_{k,q} q(k-q)\phi_{1,k,q}R^{-}_{k,q}, \tag{A17}$$

$$\beta^{nii} = \sum_{k,q} q(q-1)\phi_{0,k,q}R^{+}_{k,q}, \tag{A18}$$

$$\beta^{ni} = \sum_{k,q} q\phi_{0,k,q}R^{+}_{k,q}, \tag{A19}$$

$$\gamma^{nii} = \sum_{k,q} q(q-1)\phi_{1,k,q}R^{-}_{k,q}, \tag{A20}$$

$$\gamma^{ni} = \sum_{k,q} q\phi_{1,k,q}R^{-}_{k,q}. \tag{A21}$$

Similarly as before, they can be interpreted as the total rate at which links, connecting the first and second neighbors of the central node, change from being 0-0, 0-1, or 1-1 when the neighbor node changes state. Here, the symbol β, γ reflects the state of the neighbor node, while the first super index s, i reflects the state of the central node and the second super index the sate of the second neighbor.

The Hessians $\frac{\partial^2 \Phi_{n,k,q}}{\partial\phi_{n',k',q'}\partial\phi_{n'',k'',q''}}$ can be calculated by taking the derivative of Eqs. (A1)–(A4) which leads to

$$\begin{aligned}\frac{\partial^2\Phi_{0,k,q}}{\partial\phi_{0,k',q'}\partial\phi_{0,k'',q''}} &= \delta_{k,k'}[-\delta_{q,q'}(k-q)+\delta_{q-1,q'}(k-q+1)]\frac{\partial\beta^s}{\partial\phi_{0,k'',q''}}\\ &+\delta_{k,k''}[-\delta_{q,q''}(k-q)+\delta_{q-1,q''}(k-q+1)]\frac{\partial\beta^s}{\partial\phi_{0,k',q'}}\\ &+[-(k-q)\phi_{0,k,q}+(k-q+1)\phi_{0,k,q-1}]\frac{\partial^2\beta^s}{\partial\phi_{0,k',q'}\partial\phi_{0,k'',q''}},\end{aligned} \tag{A22}$$

$$\frac{\partial^2\Phi_{0,k,q}}{\partial\phi_{0,k',q'}\partial\phi_{1,k'',q''}} = \delta_{k,k'}[-\delta_{q,q'}q+\delta_{q+1,q'}(q+1)]\frac{\partial\gamma^s}{\partial\phi_{1,k'',q''}}, \tag{A23}$$

$$\frac{\partial^2\Phi_{0,k,q}}{\partial\phi_{1,k',q'}\partial\phi_{1,k'',q''}} = [-q\phi_{0,k,q}+(q+1)\phi_{0,k,q+1}]\frac{\partial^2\gamma^s}{\partial\phi_{1,k',q'}\partial\phi_{1,k'',q''}}, \tag{A24}$$

$$\frac{\partial^2\Phi_{1,k,q}}{\partial\phi_{0,k',q'}\partial\phi_{0,k'',q''}} = [-(k-q)\phi_{1,k,q}+(k-q+1)\phi_{1,k,q-1}]\frac{\partial^2\beta^i}{\partial\phi_{0,k',q'}\partial\phi_{0,k'',q''}}, \tag{A25}$$

$$\frac{\partial^2\Phi_{1,k,q}}{\partial\phi_{0,k',q'}\partial\phi_{1,k'',q''}} = \delta_{k,k''}[-\delta_{q,q''}(k-q)+\delta_{q-1,q''}(k-q+1)]\frac{\partial\beta^i}{\partial\phi_{0,k',q'}}, \tag{A26}$$

$$\begin{aligned}\frac{\partial^2\Phi_{1,k,q}}{\partial\phi_{1,k',q'}\partial\phi_{1,k'',q''}} &= \delta_{k,k'}[-\delta_{q,q'}q+\delta_{q+1,q'}(q+1)]\frac{\partial\gamma^i}{\partial\phi_{1,k'',q''}}\\ &+\delta_{k,k''}[-\delta_{q,q''}q+\delta_{q+1,q''}(q+1)]\frac{\partial\gamma^i}{\partial\phi_{1,k',q'}}\\ &+[-q\phi_{1,k,q}+(q+1)\phi_{1,k,q+1}]\frac{\partial^2\gamma^i}{\partial\phi_{1,k',q'}\partial\phi_{1,k'',q''}}.\end{aligned} \tag{A27}$$

The second derivatives of the rates Eqs. (20)–(23) are

$$\frac{\partial^2\beta^s}{\partial\phi_{0,k',q'}\partial\phi_{0,k'',q''}} = \frac{(k'-q')(k''-q'')(2\beta^s-R^{+}_{k',q'}-R^{+}_{k'',q''})}{\left(\sum_{k,q}(k-q)\phi_{0,k,q}\right)^2}, \tag{A28}$$

$$\frac{\partial^2\gamma^s}{\partial\phi_{1,k',q'}\partial\phi_{1,k'',q''}} = \frac{(k'-q')(k''-q'')(2\gamma^s-R^{-}_{k',q'}-R^{-}_{k'',q''})}{\left(\sum_{k,q}(k-q)\phi_{1,k,q}\right)^2}, \tag{A29}$$

$$\frac{\partial^2\beta^i}{\partial\phi_{0,k',q'}\partial\phi_{0,k'',q''}} = \frac{q'q''(2\beta^i - R^+_{k',q'} - R^+_{k'',q''})}{\left(\sum_{k,q}(k-q)\phi_{0,k,q}\right)^2}, \tag{A30}$$

$$\frac{\partial^2\gamma^i}{\partial\phi_{1,k',q'}\partial\phi_{1,k'',q''}} = \frac{q'q''(2\gamma^i - R^-_{k',q'} - R^-_{k'',q''})}{\left(\sum_{k,q}(k-q)\phi_{1,k,q}\right)^2}. \tag{A31}$$

2. Contact process

For the rates of the contact process, the previous matrices slightly change. In Eq. (A9), we must replace $R^+_{k,q} \to \widetilde{R}^+_{k,q} = \sum_{\{k_j,q_j\}} R^+_{k,q,\{k_j\}} \prod_j P_0(1,k_j,q_j)$. In Eq. (A10), in the first term of the sum $R^+_{k,q} \to \widetilde{R}^+_{k,q}$ and in the second term $R^+_{k,q} \to \widetilde{R}^+_{k,q,k'} = \sum_{\{k_j,q_j\}} R^+_{k,q,k',\{k_j\}} \prod_j P_0(1,k_j,q_j)$ with $j = k-q+2,\ldots,k$. In the third term, $R^+_{k',q'} \to \widetilde{R}^+_{k',q'}$. In the fourth term, $\beta^{nsi} \to \beta^{nsi}_{k'} = \sum_{k,q} q(k-q)\phi_{0,k,q}\widetilde{R}^+_{k,q,k'}$.

For Eq. (A11), in the first term of the sum, $R^+_{k,q} \to \widetilde{R}^+_{k,q}$, in the second term, $R^+_{k,q} \to \widetilde{R}^+_{k,q,k'}$, and in the third term, $R^+_{k',q'} \to \widetilde{R}^+_{k',q',k}$. The rates transform $\beta^{nii} \to \beta^{nii}_{k,k'} = \sum_{k'',q''} q''(q''-1)\phi_{0,k'',q''}\widetilde{R}^+_{k'',q'',k,k'}$ with $\widetilde{R}^+_{k'',q'',k,k'} = \sum_{\{k_j,q_j\}} R^+_{k'',q'',k,k',\{k_j\}} \prod_j P_0(1,k_j,q_j)$; $j = k-q+3,\ldots,k$, and $\beta^{ni} \to \beta^{ni}_k = \sum_{k',q'} q'\phi_{0,k',q'}\widetilde{R}^+_{k',q',k}$.

3. Pair approximation

The case of the pair approximation can be seen as a reduction and change of variables of the previous more complex case, where $\phi_{0,k,q} = (P_k - \phi_k)\mathrm{Bin}_{k,q}[p_{0,k}]$, $\phi_{1,k,q} = \phi_k \mathrm{Bin}_{k,q}[p_{1,k}]$ and $p_{0,k} = \phi_{0,k}/(k(P_k-\phi_k))$, $p_{1,k} = \phi_{1,k}/(k\phi_k)$. In this way, the Jacobian matrix of the new variables can be calculated using the chain rule as, e.g., $B_{0,k;0,k'} = \sum_{q,q''} qB_{0,k,q;0,k'',q''} \frac{\partial\phi_{0,k'',q''}}{\partial\phi_{0,k'}}$. This procedure leads to the different matrix elements:

$$B_{k;k'} = -\delta_{k,k'} \sum_q \left[\frac{\partial\phi_{0,k,q}}{\partial\phi_k} R^+_{k,q} - \frac{\partial\phi_{1,k,q}}{\partial\phi_k} R^-_{k,q} \right], \tag{A32}$$

$$B_{k;0,k'} = -\delta_{k,k'} \sum_q \frac{\partial\phi_{0,k,q}}{\partial\phi_{0,k}} R^+_{k,q}, \tag{A33}$$

$$B_{k;1,k'} = \delta_{k,k'} \sum_q \frac{\partial\phi_{1,k,q}}{\partial\phi_{1,k}} R^-_{k,q}, \tag{A34}$$

$$B_{0,k;k'} = \delta_{k,k'} \sum_q \left[\frac{\partial\phi_{0,k,q}}{\partial\phi_k} qR^+_{k,q} - \frac{\partial\phi_{1,k,q}}{\partial\phi_k} qR^-_{k,q} \right] - \frac{\partial\beta^s}{\partial\phi_{k'}}(kP_k - k\phi_k - \phi_{0,k}) + \frac{\partial\gamma^s}{\partial\phi_{k'}}\phi_{0,k} + \delta_{k,k'}k\beta^s, \tag{A35}$$

$$B_{0,k;0,k'} = \delta_{k,k'} \sum_q \frac{\partial\phi_{0,k,q}}{\partial\phi_{0,k}} qR^+_{k,q} - \frac{\partial\beta^s}{\partial\phi_{0,k'}}(kP_k - k\phi_k - \phi_{0,k}) + \delta_{k,k'}(\beta^s + \gamma^s), \tag{A36}$$

$$B_{0,k;1,k'} = -\delta_{k,k'} \sum_q \frac{\partial\phi_{1,k,q}}{\partial\phi_{1,k}} qR^-_{k,q} + \frac{\partial\gamma^s}{\partial\phi_{1,k'}}\phi_{0,k}, \tag{A37}$$

$$B_{1,k;k'} = -\delta_{k,k'} \sum_q \left[\frac{\partial\phi_{0,k,q}}{\partial\phi_k} qR^+_{k,q} - \frac{\partial\phi_{1,k,q}}{\partial\phi_k} qR^-_{k,q} \right] - \frac{\partial\beta^i}{\partial\phi_{k'}}(k\phi_k - \phi_{1,k}) + \frac{\partial\gamma^i}{\partial\phi_{k'}}\phi_{1,k} - \delta_{k,k'}k\beta^i, \tag{A38}$$

$$B_{1,k;0,k'} = -\delta_{k,k'} \sum_q \frac{\partial\phi_{0,k,q}}{\partial\phi_{0,k}} qR^+_{k,q} - \frac{\partial\beta^i}{\partial\phi_{0,k'}}(k\phi_k - \phi_{1,k}), \tag{A39}$$

$$B_{1,k;1,k'} = \delta_{k,k'} \sum_q \frac{\partial\phi_{1,k,q}}{\partial\phi_{1,k}} qR^-_{k,q} + \frac{\partial\gamma^i}{\partial\phi_{1,k'}}\phi_{1,k} + \delta_{k,k'}(\beta^i + \gamma^i), \tag{A40}$$

where

$$\frac{\partial\phi_{0,k,q}}{\partial\phi_k} = -\mathrm{Bin}_{k,q}[p_{0,k}] + \mathrm{Bin}'_{k,q}[p_{0,k}]p_{0,k}, \tag{A41}$$

$$\frac{\partial\phi_{1,k,q}}{\partial\phi_k} = \mathrm{Bin}_{k,q}[p_{1,k}] - \mathrm{Bin}'_{k,q}[p_{1,k}]p_{1,k}, \tag{A42}$$

$$\frac{\partial\phi_{0,k,q}}{\partial\phi_{0,k}} = \frac{1}{k}\mathrm{Bin}'_{k,q}[p_{0,k}], \tag{A43}$$

$$\frac{\partial\phi_{1,k,q}}{\partial\phi_{1,k}} = \frac{1}{k}\mathrm{Bin}'_{k,q}[p_{1,k}], \tag{A44}$$

and the derivative of the binomial distribution is $\mathrm{Bin}'_{k,q}[x] = \frac{q-xk}{x(1-x)}\mathrm{Bin}_{k,q}[x]$. The derivatives of the rates β^s, β^i, γ^s, γ^i can be calculated using the chain rule, e.g., $\partial\beta^s/\partial\phi_k = \sum_q(\partial\beta^s/\partial\phi_{0,k,q})(\partial\phi_{0,k,q}/\partial\phi_k)$.

Similarly, the **G**-matrix of the pair approximation can be obtained taking a partial sum of the previous one, e.g. $G_{k;0,k'} = \sum_{q,q'} q' G_{1,k,q;0,k',q'}$. This procedure reads for the different elements:

$$G_{k;k'} = \delta_{k,k'}\sum_q[\phi_{0,k,q}R^+_{k,q} + \phi_{1,k,q}R^-_{k,q}], \tag{A45}$$

$$\begin{aligned} G_{k;0,k'} = &-\delta_{k,k'}\sum_q[q\phi_{0,k,q}R^+_{k,q} + q\phi_{1,k,q}R^-_{k,q}] + \sum_q(k-q)\phi_{0,k,q}R^+_{k,q}\sum_{q'}P_0(0,k',q') \\ &+ \sum_q(k-q)\phi_{1,k,q}R^-_{k,q}\sum_{q'}P_1(0,k',q'), \end{aligned} \tag{A46}$$

$$\begin{aligned} G_{k;1,k'} = &\delta_{k,k'}\sum_q[q\phi_{0,k,q}R^+_{k,q} + q\phi_{1,k,q}R^-_{k,q}] + \sum_q q\phi_{0,k,q}R^+_{k,q}\sum_{q'}P_0(1,k',q') \\ &+ \sum_q q\phi_{1,k,q}R^-_{k,q}\sum_{q'}P_1(1,k',q'), \end{aligned} \tag{A47}$$

$$\begin{aligned} G_{0,k;0,k'} = &\delta_{k,k'}\sum_q[q^2\phi_{0,k,q}R^+_{k,q} + q^2\phi_{1,k,q}R^-_{k,q}] \\ &- \sum_q q(k-q)\phi_{0,k,q}R^+_{k,q}\sum_{q'}P_0(0,k',q') - \sum_{q'}q'(k'-q')\phi_{0,k',q'}R^+_{k',q'}\sum_q P_0(0,k,q) \\ &- \sum_q q(k-q)\phi_{1,k,q}R^-_{k,q}\sum_{q'}P_1(0,k',q') - \sum_{q'}q'(k'-q')\phi_{1,k',q'}R^-_{k',q'}\sum_q P_1(0,k,q) \\ &+ \beta^{nss}\sum_q P_0(0,k,q)\sum_{q'}P_0(0,k',q') + \gamma^{nss}\sum_q P_1(0,k,q)\sum_{q'}P_1(0,k',q') \\ &+ \delta_{k,k'}\left[\beta^{ns}\sum_q P_0(0,k,q) + \gamma^{ns}\sum_q P_1(0,k,q)\right], \end{aligned} \tag{A48}$$

$$\begin{aligned} G_{0,k;1,k'} = &-\delta_{k,k'}\sum_q[q^2\phi_{0,k,q}R^+_{k,q} + q^2\phi_{1,k,q}R^-_{k,q}] \\ &- \sum_q q^2\phi_{0,k,q}R^+_{k,q}\sum_{q'}P_0(1,k',q') + \sum_{q'}q'(k'-q')\phi_{0,k',q'}R^+_{k',q'}\sum_q P_0(0,k,q) \\ &- \sum_q q^2\phi_{1,k,q}R^-_{k,q}\sum_{q'}P_1(1,k',q') + \sum_{q'}q'(k'-q')\phi_{1,k',q'}R^-_{k',q'}\sum_q P_1(0,k,q) \\ &+ \beta^{nsi}\sum_q P_0(0,k,q)\sum_{q'}P_0(1,k',q') + \gamma^{nsi}\sum_q P_1(0,k,q)\sum_{q'}P_1(1,k',q'), \end{aligned} \tag{A49}$$

$$\begin{aligned} G_{1,k;1,k'} = &\delta_{k,k'}\sum_q[q^2\phi_{0,k,q}R^+_{k,q} + q^2\phi_{1,k,q}R^-_{k,q}] \\ &+ \sum_q q^2\phi_{0,k,q}R^+_{k,q}\sum_{q'}P_0(1,k',q') + \sum_{q'}(q')^2\phi_{0,k',q'}R^+_{k',q'}\sum_q P_0(1,k,q) \\ &+ \sum_q q^2\phi_{1,k,q}R^-_{k,q}\sum_{q'}P_1(1,k',q') + \sum_{q'}(q')^2\phi_{1,k',q'}R^-_{k',q'}\sum_q P_1(1,k,q) \\ &+ \beta^{nii}\sum_q P_0(1,k,q)\sum_{q'}P_0(1,k',q') + \gamma^{nii}\sum_q P_1(1,k,q)\sum_{q'}P_1(1,k',q') \\ &+ \delta_{k,k'}\left[\beta^{ni}\sum_q P_0(1,k,q) + \gamma^{ni}\sum_q P_1(1,k,q)\right]. \end{aligned} \tag{A50}$$

The Hessians can be obtained taking a second derivative of the Eqs. (A32)–(A40), which leads to

$$\frac{\partial^2\Phi_k}{\partial\phi_{k'}\partial\phi_{k''}} = \delta_{k,k'}\delta_{k,k''}\sum_q\left[\frac{\partial^2\phi_{0,k,q}}{\partial\phi_k^2}R^+_{k,q} - \frac{\partial^2\phi_{1,k,q}}{\partial\phi_k^2}R^-_{k,q}\right], \tag{A51}$$

$$\frac{\partial^2\Phi_k}{\partial\phi_{k'}\partial\phi_{0,k''}} = \delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{0,k,q}}{\partial\phi_k\partial\phi_{0,k}}R^+_{k,q}, \tag{A52}$$

$$\frac{\partial^2\Phi_k}{\partial\phi_{k'}\partial\phi_{1,k''}} = -\delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{1,k,q}}{\partial\phi_k\partial\phi_{1,k}}R^-_{k,q}, \tag{A53}$$

$$\frac{\partial^2\Phi_k}{\partial\phi_{0,k'}\partial\phi_{0,k''}} = \delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{0,k,q}}{\partial\phi_{0,k}^2}R^+_{k,q}, \tag{A54}$$

$$\frac{\partial^2\Phi_k}{\partial\phi_{0,k'}\partial\phi_{1,k''}} = 0, \tag{A55}$$

$$\frac{\partial^2\Phi_k}{\partial\phi_{1,k'}\partial\phi_{1,k''}} = -\delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{1,k,q}}{\partial\phi_{1,k}^2}R^-_{k,q}, \tag{A56}$$

$$\frac{\partial^2\Phi_{0,k}}{\partial\phi_{k'}\partial\phi_{k''}} = -\delta_{k,k'}\delta_{k,k''}\sum_q\left[\frac{\partial^2\phi_{0,k,q}}{\partial\phi_k^2}qR^+_{k,q} - \frac{\partial^2\phi_{1,k,q}}{\partial\phi_k^2}qR^-_{k,q}\right]$$
$$+\frac{\partial^2\beta^s}{\partial\phi_{k'}\phi_{k''}}(kP_k - k\phi_k - \phi_{0,k}) - \frac{\partial^2\gamma^s}{\partial\phi_{k'}\phi_{k''}}\phi_{0,k} - \delta_{k,k''}k\frac{\partial\beta^s}{\partial\phi_{k'}} - \delta_{k,k'}k\frac{\partial\beta^s}{\partial\phi_{k''}}, \tag{A57}$$

$$\frac{\partial^2\Phi_{0,k}}{\partial\phi_{k'}\partial\phi_{0,k''}} = -\delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{0,k,q}}{\partial\phi_k\partial\phi_{0,k}}qR^+_{k,q}$$
$$+\frac{\partial^2\beta^s}{\partial\phi_{k'}\phi_{0,k''}}(kP_kk - k\phi_k - \phi_{0,k}) - \delta_{k,k''}\left(\frac{\partial\gamma^s}{\partial\phi_{k'}} + \frac{\partial\beta^s}{\partial\phi_{k'}}\right) - \delta_{k,k'}k\frac{\partial\beta^s}{\partial\phi_{0,k''}}, \tag{A58}$$

$$\frac{\partial^2\Phi_{0,k}}{\partial\phi_{k'}\partial\phi_{1,k''}} = \delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{1,k,q}}{\partial\phi_k\partial\phi_{1,k}}qR^-_{k,q} - \frac{\partial^2\gamma^s}{\partial\phi_{k'}\phi_{1,k''}}\phi_{0,k}, \tag{A59}$$

$$\frac{\partial^2\Phi_{0,k}}{\partial\phi_{0,k'}\partial\phi_{0,k''}} = -\delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{0,k,q}}{\partial\phi_{0,k}^2}qR^+_{k,q}$$
$$+\frac{\partial^2\beta^s}{\partial\phi_{0,k'}\phi_{0,k''}}(kP_kk - k\phi_k - \phi_{0,k}) - \delta_{k,k''}\frac{\partial\beta^s}{\partial\phi_{0,k'}} - \delta_{k,k'}\frac{\partial\beta^s}{\partial\phi_{0,k''}}, \tag{A60}$$

$$\frac{\partial^2\Phi_{0,k}}{\partial\phi_{0,k'}\partial\phi_{1,k''}} = -\delta_{k,k'}\frac{\partial\gamma^s}{\partial\phi_{1,k''}}, \tag{A61}$$

$$\frac{\partial^2\Phi_{0,k}}{\partial\phi_{1,k'}\partial\phi_{1,k''}} = \delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{1,k,q}}{\partial\phi_{1,k}^2}qR^-_{k,q} - \frac{\partial^2\gamma^s}{\partial\phi_{1,k'}\phi_{1,k''}}\phi_{0,k}, \tag{A62}$$

$$\frac{\partial^2\Phi_{1,k}}{\partial\phi_{k'}\partial\phi_{k''}} = \delta_{k,k'}\delta_{k,k''}\sum_q\left[\frac{\partial^2\phi_{0,k,q}}{\partial\phi_k^2}qR^+_{k,q} - \frac{\partial^2\phi_{1,k,q}}{\partial\phi_k^2}qR^-_{k,q}\right]$$
$$+\frac{\partial^2\beta^i}{\partial\phi_{k'}\phi_{k''}}(\phi_kk - \phi_{1,k}) - \frac{\partial^2\gamma^i}{\partial\phi_{k'}\phi_{k''}}\phi_{1,k} + \delta_{k,k''}k\frac{\partial\beta^i}{\partial\phi_{k'}} + \delta_{k,k'}k\frac{\partial\beta^i}{\partial\phi_{k''}}, \tag{A63}$$

$$\frac{\partial^2\Phi_{1,k}}{\partial\phi_{k'}\partial\phi_{0,k''}} = \delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{0,k,q}}{\partial\phi_k\partial\phi_{0,k}}qR^+_{k,q} + \frac{\partial^2\beta^i}{\partial\phi_{k'}\phi_{0,k''}}(\phi_kk - \phi_{1,k}) + \delta_{k,k'}k\frac{\partial\beta^i}{\partial\phi_{0,k''}}, \tag{A64}$$

$$\frac{\partial^2\Phi_{1,k}}{\partial\phi_{k'}\partial\phi_{1,k''}} = -\delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{1,k,q}}{\partial\phi_k\partial\phi_{1,k}}qR^-_{k,q} - \delta_{k,k''}\left(\frac{\partial\beta^i}{\partial\phi_{k'}} + \frac{\partial\gamma^i}{\partial\phi_{k'}}\right) - \frac{\partial^2\gamma^i}{\partial\phi_{k'}\phi_{1,k''}}\phi_{1,k}, \tag{A65}$$

$$\frac{\partial^2\Phi_{1,k}}{\partial\phi_{0,k'}\partial\phi_{0,k''}} = \delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{0,k,q}}{\partial\phi_{0,k}^2}qR^+_{k,q} + \frac{\partial^2\beta^i}{\partial\phi_{0,k'}\phi_{0,k''}}(\phi_kk - \phi_{1,k}), \tag{A66}$$

$$\frac{\partial^2\Phi_{1,k}}{\partial\phi_{0,k'}\partial\phi_{1,k''}} = -\delta_{k,k''}\frac{\partial\beta^i}{\partial\phi_{0,k'}}, \tag{A67}$$

$$\frac{\partial^2\Phi_{1,k}}{\partial\phi_{1,k'}\partial\phi_{1,k''}} = -\delta_{k,k'}\delta_{k,k''}\sum_q \frac{\partial^2\phi_{1,k,q}}{\partial\phi_{1,k}^2}qR^-_{k,q} - \frac{\partial^2\gamma^i}{\partial\phi_{1,k'}\phi_{1,k''}}\phi_{1,k} - \delta_{k,k''}\frac{\partial\gamma^i}{\partial\phi_{1,k'}} - \delta_{k,k'}\frac{\partial\gamma^i}{\partial\phi_{1,k''}}, \tag{A68}$$

where

$$\frac{\partial^2\phi_{0,k,q}}{\partial\phi_k^2} = \mathrm{Bin}''_{k,q}[p_{0,k}]\frac{p_{0,k}^2}{P_k-\phi_k}, \tag{A69}$$

$$\frac{\partial^2\phi_{0,k,q}}{\partial\phi_k\partial\phi_{0,k}} = \mathrm{Bin}''_{k,q}[p_{0,k}]\frac{p_{0,k}}{k(P_k-\phi_k)}, \tag{A70}$$

$$\frac{\partial^2\phi_{0,k,q}}{\partial\phi_{0,k}^2} = \mathrm{Bin}''_{k,q}[p_{0,k}]\frac{1}{k^2(P_k-\phi_k)}, \tag{A71}$$

$$\frac{\partial^2\phi_{1,k,q}}{\partial\phi_k^2} = \mathrm{Bin}''_{k,q}[p_{1,k}]\frac{p_{1,k}^2}{\phi_k}, \tag{A72}$$

$$\frac{\partial^2\phi_{1,k,q}}{\partial\phi_k\partial\phi_{1,k}} = -\mathrm{Bin}''_{k,q}[p_{1,k}]\frac{p_{1,k}}{k\phi_k}, \tag{A73}$$

$$\frac{\partial^2\phi_{1,k,q}}{\partial\phi_{1,k}^2} = \mathrm{Bin}''_{k,q}[p_{1,k}]\frac{1}{k^2\phi_k}, \tag{A74}$$

and the second derivative of the binomial distribution is $\mathrm{Bin}''_{k,q}[x] = \frac{-kx(1-x)+(q-xk)(q-xk-1+2x)}{x^2(1-x)^2}\mathrm{Bin}_{k,q}[x]$. The second derivatives of the rates β^s, β^i, γ^s, γ^i can be calculated using the chain rule again, e.g.,

$$\frac{\partial^2\beta^s}{\partial\phi_k\partial\phi_{k'}} = \sum_{q,q'}\frac{\partial^2\beta^s}{\partial\phi_{0,k,q}\partial\phi_{0,k',q'}}\frac{\partial\phi_{0,k,q}}{\partial\phi_k}\frac{\partial\phi_{0,k',q'}}{\partial\phi_{k'}} + \delta_{k,k'}\sum_q\frac{\partial\beta^s}{\partial\phi_{0,k,q}}\frac{\partial^2\phi_{0,k,q}}{\partial\phi_k^2}. \tag{A75}$$

4. Heterogeneous mean field

This is the simplest of the approximations where we reduce variables following $\phi_{0,k,q} = (P_k-\phi_k)\mathrm{Bin}_{k,q}[p]$, $\phi_{1,k,q} = \phi_k\mathrm{Bin}_{k,q}[p]$ with $p = \sum_k \phi_k k/\mu$. The Jacobian in this case can be obtained taking the derivative of Eq. (32), this is

$$B_{k;k'} = -\sum_q\left[\frac{\partial\phi_{0,k,q}}{\partial\phi_{k'}}R^+_{k,q} - \frac{\partial\phi_{1,k,q}}{\partial\phi_{k'}}R^-_{k,q}\right], \tag{A76}$$

$$\frac{\partial\phi_{0,k,q}}{\partial\phi_{k'}} = -\delta_{k,k'}\mathrm{Bin}_{k,q}[p] + (P_k-\phi_k)\frac{k'}{\mu}\mathrm{Bin}'_{k,q}[p], \tag{A77}$$

$$\frac{\partial\phi_{1,k,q}}{\partial\phi_{k'}} = \delta_{k,k'}\mathrm{Bin}_{k,q}[p] + \phi_k\mathrm{Bin}'_{k,q}[p]\frac{k'}{\mu}. \tag{A78}$$

The $\mathbf{G}$-matrix is simply $G_{k;k'} = \sum_{q,q'} G_{1,k,q;1,k',q'}$ which leads to

$$G_{k;k'} = \delta_{k,k'}\sum_q[\phi_{0,k,q}R^+_{k,q} + \phi_{1,k,q}R^-_{k,q}]. \tag{A79}$$

The second derivatives can be computed as

$$\frac{\partial^2\Phi_k}{\partial\phi_{k'}\partial\phi_{k''}} = \sum_q\left[\frac{\partial^2\phi_{0,k,q}}{\partial\phi_{k'}\partial\phi_{k''}}R^+_{k,q} - \frac{\partial^2\phi_{1,k,q}}{\partial\phi_{k'}\partial\phi_{k''}}R^-_{k,q}\right], \tag{A80}$$

$$\frac{\partial^2\phi_{0,k,q}}{\partial\phi_{k'}\partial\phi_{k''}} = -\mathrm{Bin}'_{k,q}[p]\left(\delta_{k,k'}\frac{k''}{\mu} + \delta_{k,k''}\frac{k'}{\mu}\right) + (P_k-\phi_k)\mathrm{Bin}''_{k,q}[p]\frac{k'}{\mu}\frac{k''}{\mu}, \tag{A81}$$

$$\frac{\partial^2\phi_{1,k,q}}{\partial\phi_{k'}\partial\phi_{k''}} = \mathrm{Bin}'_{k,q}[p]\left(\delta_{k,k'}\frac{k''}{\mu} + \delta_{k,k''}\frac{k'}{\mu}\right) + \phi_k\mathrm{Bin}''_{k,q}[p]\frac{k'}{\mu}\frac{k''}{\mu}. \tag{A82}$$

APPENDIX B: THE CENTER MANIFOLD

The coefficients $\alpha_i^{(10)}$, $\alpha_i^{(11)}$, $\alpha_i^{(02)}$ of the center manifold $u_i = h_i(T, u_1)$, Eq. (40), are

$$D_i\alpha_i^{(10)} = \partial_T U_i, \tag{B1}$$

$$D_i\alpha_i^{(02)} = \frac{1}{2}\partial^2_{u_1}U_i, \tag{B2}$$

$$D_i\alpha_i^{(11)} = -\alpha_i^{(02)}\partial_T U_1 + \partial^2_{Tu_1}U_i + \sum_{j\neq 1}\partial^2_{u_1u_j}U_i\alpha_j^{(10)}. \tag{B3}$$

The coefficients $\beta^{(10)}$, $\beta^{(11)}$, $\beta^{(02)}$, $\beta^{(03)}$ of the normal form of the bifurcation Eq. (41) are

$$\beta^{(10)} = \partial_T U_1, \tag{B4}$$

$$\beta^{(02)} = \frac{1}{2}\partial^2_{u_1} U_1, \tag{B5}$$

$$\beta^{(11)} = \partial^2_{Tu_1} U_1 + \sum_{j\neq 1} \partial^2_{u_1 u_j} U_1 \alpha_j^{(10)}, \tag{B6}$$

$$\beta^{(03)} = \frac{1}{3!}\partial^3_{u_1} U_1 + \sum_{j\neq 1} \partial^2_{u_1 u_j} U_1 \alpha_j^{(02)}. \tag{B7}$$

Using the change of variables Eqs. (42)–(44), we can expand the master equation Eq. (4) in powers of N and derive a Fokker-Planck equation for the probability of the new variable $\Pi(\xi_1;t)$. The step operator $\prod_{i=1}^{M} E_i^{-\ell_i^{(\nu)}} = e^{-\ell^{(\nu)}\cdot\nabla_{\mathbf{x}}}$ in the right hand side of Eq. (4) for the $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ variables with $\nabla_{\mathbf{x}} = \mathbf{P}^{-1}\nabla_{\mathbf{y}}$ transforms as $-\mathbf{P}^{-1}\ell^{(\nu)}\cdot\nabla_{\mathbf{y}} + \frac{1}{2}(\mathbf{P}^{-1}\boldsymbol{\ell}^{(\nu)}\cdot\nabla_{\mathbf{y}})^2 + \ldots$ Taking into account that the derivatives change like $\partial_{y_i} = N^{-1/2}\partial_{\xi_i}$ and $\partial_{y_1} = N^{-\upsilon}\partial_{\xi_1} - N^{-1/2}\sum_j (N^{-r}\alpha_j^{(11)}\xi_0 + N^{\upsilon-1}2\alpha_j^{(20)}\xi_1 + \ldots)\partial_{\xi_j}$, and integrating the full equation $\int [\prod_{i\neq 1} d\xi_i]$ the only terms that remain are, (i) associated to the first derivative $\partial_{\xi_1}[\ldots\Pi(\xi_1;t)]$ we have

$$-N^{-\upsilon}\sum_{\nu}\sum_{j} P_{1j}^{-1}\ell_j^{(\nu)}W^{(\nu)} = -N^{1-\upsilon-r}\beta^{(10)}\xi_0 - N^{-r}\beta^{(11)}\xi_0\xi_1 - N^{(m-1)(\upsilon-1)}\beta^{(0m)}\xi_1^m + \ldots \tag{B8}$$

and (ii) associated to the second derivative $\frac{1}{2}\partial^2_{\xi_1}[\ldots\Pi(\xi_1;t)]$ we have

$$N^{-2\upsilon}\sum_{\nu}\sum_{ij} P_{1i}^{-1}P_{1j}^{-1}\ell_i^{(\nu)}\ell_j^{(\nu)}W^{(\nu)} = N^{1-2\upsilon}F_{11} + \ldots, \tag{B9}$$

with $F_{11} = \sum_{i,j} P_{1i}^{-1}P_{1j}^{-1}G_{ij}$. Both terms must be of the same order of N, thus we must choose r and υ properly obtaining: $r = \upsilon = \frac{m}{m+1}$ if $\beta^{(10)} \neq 0$ and $\upsilon = \frac{m}{m+1}$, $r = \frac{m-1}{m+1}$ if $\beta^{(10)} = 0$. Putting both terms together we obtain finally the Fokker-Planck equation (45).

In the case of epidemic models with an absorbing state and $F_{11} = 0$, we must consider a further term in Eq. (B9), this is

$$N^{-2\upsilon}\sum_{\nu}\sum_{ij} P_{1i}^{-1}P_{1j}^{-1}\ell_i^{(\nu)}\ell_j^{(\nu)}W^{(\nu)} = N^{-\upsilon}\gamma\xi_1 + \ldots, \tag{B10}$$

with $\frac{\partial F_{11}}{\partial u_1}$ evaluated at the absorbing state and critical point. For a transcritical bifurcation with $\beta^{(01)} = 0$ and $\beta^{(02)} \neq 0$, equating both orders of N in Eqs. (B8) and (B10), we obtain $r = \upsilon = 1/2$.

[1] R. M. Anderson, B. Anderson, and R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, 1991).

[2] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*, 1st ed. (Princeton University Press, Princeton, NJ, 2007).

[3] F. Brauer and C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology*, Texts in Applied Mathematics (Springer, New York, NY, 2011).

[4] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, Rev. Mod. Phys. **87**, 925 (2015).

[5] D. M. Abrams and S. H. Strogatz, Modelling the dynamics of language death, Nature (London) **424**, 900 (2003).

[6] J. Mira and Á. Paredes, Interlinguistic similarity and language death dynamics, Europhys. Lett. **69**, 1031 (2005).

[7] X. Castelló, A. Baronchelli, and V. Loreto, Consensus and ordering in language dynamics, Eur. Phys. J. B **71**, 557 (2009).

[8] F. Vazquez, X. Castelló, and M. San Miguel, Agent based models of language competition: macroscopic descriptions and order–disorder transitions, J. Stat. Mech.: Theory Exp. (2010) P04007.

[9] M. Patriarca, X. Castelló, J. R. Uriarte, V. M. Eguíluz, and M. San Miguel, Modeling two-language competition dynamics, Adv. Complex Systems **15**, 1250048 (2012).

[10] S. Galam, Minority opinion spreading in random geometry, Eur. Phys. J. B **25**, 403 (2002).

[11] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).

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

[13] M. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).

[14] C. Castellano, Social influence and the dynamics of opinions: The approach of statistical physics, Manage. Decis. Econ. **33**, 311 (2012).

[15] P. Nyczka and K. Sznajd-Weron, Anticonformity or Independence?—Insights from Statistical Physics, J. Stat. Phys. **151**, 174 (2013).

[16] S. Alfarano, T. Lux, and F. Wagner, Estimation of agent-based models: The case of an asymmetric herding model, Comp. Economics **26**, 19 (2005).

[17] S. Alfarano, T. Lux, and F. Wagner, Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach, J. Econ. Dyn. Control **32**, 101 (2008).

[18] S. Alfarano and M. Milaković, Network structure and N-dependence in agent-based herding models, J. Econ. Dyn. Control **33**, 78 (2009).

[19] A. Carro, R. Toral, and M. San Miguel, Markets, herding and response to external information, PLOS ONE **10**, e0133287 (2015).

[20] A. L. M. Vilela, C. Wang, K. P. Nelson, and H. Eugene Stanley, Majority-vote model for financial markets, Physica A **515**, 762 (2019).

[21] A. Kononovicius and J. Ruseckas, Order book model with herd behavior exhibiting long-range memory, Physica A **525**, 171 (2019).

[22] H. W. Hethcote and J. Yorke, *Gonorrhea Transmission Dynamics and Control*, Lecture Notes in Biomathematics (Springer, Berlin, 1984).

[23] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, Epidemic spreading in real networks: An eigenvalue viewpoint, in *22nd International Symposium on Reliable Distributed Systems, 2003* (IEEE, 2003), pp. 25–34.

[24] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, Epidemic thresholds in real networks, ACM Trans. Inf. Syst. Secur. **10**, 1 (2008).

[25] L. F. Lafuerza and R. Toral, On the effect of heterogeneity in stochastic interacting-particle systems, Sci. Rep. **3**, 1189 (2013).

[26] A. Carro, R. Toral, and M. San Miguel, The noisy voter model on complex networks, Sci. Rep. **6**, 24775 (2016).

[27] R. Pastor-Satorras and A. Vespignani, Epidemic Spreading in Scale-Free Networks, Phys. Rev. Lett. **86**, 3200 (2001).

[28] V. Sood, T. Antal, and S. Redner, Voter models on heterogeneous networks, Phys. Rev. E **77**, 041121 (2008).

[29] A. Vespignani, Modelling dynamical processes in complex socio-technical systems, Nat. Phys. **8**, 32 (2012).

[30] J. P. Gleeson, High-Accuracy Approximation of Binary-State Dynamics on Networks, Phys. Rev. Lett. **107**, 068701 (2011).

[31] J. P. Gleeson, Binary-State Dynamics on Complex Networks: Pair Approximation and Beyond, Phys. Rev. X **3**, 021004 (2013).

[32] A. F. Peralta, A. Carro, M. San Miguel, and R. Toral, Stochastic pair approximation treatment of the noisy voter model, New J. Phys. **20**, 103045 (2018).

[33] V. Marceau, P.-A. Noël, L. Hébert-Dufresne, A. Allard, and L. J. Dubé, Adaptive networks: Coevolution of disease and topology, Phys. Rev. E **82**, 036116 (2010).

[34] J. Lindquist, Junling Ma, P. van den Driessche, and F. H. Willeboordse, Effective degree network disease models, J. Math. Biol. **62**, 143 (2011).

[35] R. Dickman, Kinetic phase transitions in a surface-reaction model: Mean-field theory, Phys. Rev. A **34**, 4246 (1986).

[36] M. J. de Oliveira, J. F. F. Mendes, and M. A. Santos, Nonequilibrium spin models with Ising universal behavior, J. Phys. A: Math. Gen. **26**, 2317 (1993).

[37] F. Vazquez and V. M. Eguíluz, Analytical solution of the voter model on uncorrelated networks, New J. Phys. **10**, 063011 (2008).

[38] E. Pugliese and C. Castellano, Heterogeneous pair approximation for voter models on networks, Europhys. Lett. **88**, 58004 (2009).

[39] F. Schweitzer and L. Behera, Nonlinear voter models: the transition from invasion to coexistence, Eur. Phys. J. B **67**, 301 (2009).

[40] A. S. Mata, R. S. Ferreira, and S. C. Ferreira, Heterogeneous pair-approximation for the contact process on complex networks, New J. Phys. **16**, 053006 (2014).

[41] S. C. Ferreira, R. S. Ferreira, C. Castellano, and R. Pastor-Satorras, Quasistationary simulations of the contact process on quenched networks, Phys. Rev. E **84**, 066102 (2011).

[42] S. C. Ferreira, R. S. Ferreira, and R. Pastor-Satorras, Quasistationary analysis of the contact process on annealed scale-free networks, Phys. Rev. E **83**, 066113 (2011).

[43] M. Boguñá and R. Pastor-Satorras, Epidemic spreading in correlated complex networks, Phys. Rev. E **66**, 047104 (2002).

[44] C. Castellano and R. Pastor-Satorras, Thresholds for Epidemic Spreading in Networks, Phys. Rev. Lett. **105**, 218701 (2010).

[45] M. Boguñá, C. Castellano, and R. Pastor-Satorras, Nature of the Epidemic Threshold for the Susceptible-Infected-Susceptible Dynamics in Networks, Phys. Rev. Lett. **111**, 068701 (2013).

[46] S. C. Ferreira, R. S. Sander, and R. Pastor-Satorras, Collective versus hub activation of epidemic phases on networks, Phys. Rev. E **93**, 032314 (2016).

[47] C. Castellano and R. Pastor-Satorras, Non-Mean-Field behavior of the Contact Process on Scale-Free Networks, Phys. Rev. Lett. **96**, 038701 (2006).

[48] K. Fichthorn, E. Gulari, and R. Ziff, Noise-Induced Bistability in a Monte Carlo Surface-Reaction Model, Phys. Rev. Lett. **63**, 1527 (1989).

[49] D. Considine, S. Redner, and H. Takayasu, Comment on "Noise-Induced Bistability in a Monte Carlo Surface-Reaction Model", Phys. Rev. Lett. **63**, 2857 (1989).

[50] A. Kirman, Ants, rationality, and recruitment, Q. J. Econ. **108**, 137 (1993).

[51] B. L. Granovsky and N. Madras, The noisy voter model, Stochastic Processes and their Applications **55**, 23 (1995).

[52] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguíluz, Is the Voter Model a Model for Voters? Phys. Rev. Lett. **112**, 158701 (2014).

[53] A. Kononovicius, Modeling of the parties' vote share distributions, Acta Physica Polonica A **133**, 1450 (2018).

[54] A. Kononovicius, Compartmental voter model, J. Stat. Mech.: Theory Exp. (2019) 103402.

[55] A. Jędrzejewski, Pair approximation for the q-voter model with independence on complex networks, Phys. Rev. E **95**, 012307 (2017).

[56] A. F. Peralta, A. Carro, M. San Miguel, and R. Toral, Analytical and numerical study of the non-linear noisy voter on complex networks, Chaos **28**, 075516 (2018).

[57] A. R. Vieira and C. Anteneodo, Threshold q-voter model, Phys. Rev. E **97**, 052106 (2018).

[58] O. Artime, A. Carro, A. F. Peralta, J. J. Ramasco, M. San Miguel, and R. Toral, Herding and idiosyncratic choices:

Nonlinearity and aging-induced transitions in the noisy voter model, C. R. Phys. **20**, 262 (2019).
[59] O. Artime, A. F. Peralta, R. Toral, J. J. Ramasco, and M. San Miguel, Aging-induced phase transition, Phys. Rev. E **98**, 032104 (2018).
[60] A. F. Peralta, N. Khalil, and R. Toral, Ordering dynamics in the voter model with aging, Physica A **552**, 122475 (2019).
[61] N. Khalil, M. San Miguel, and R. Toral, Zealots in the mean-field noisy voter model, Phys. Rev. E **97**, 012310 (2018).
[62] N. Khalil and R. Toral, The noisy voter model under the influence of contrarians, Physica A **515**, 81 (2019).
[63] F. Vazquez, E. S. Loscar, and G. Baglietto, Multistate voter model with imperfect copying, Phys. Rev. E **100**, 042301 (2019).
[64] F. Herrerías-Azcué and T. Galla, Consensus and diversity in multistate noisy voter models, Phys. Rev. E **100**, 022304 (2019).
[65] B. Nowak and K. Sznajd-Weron, Homogeneous symmetrical threshold model with nonconformity: Independence versus anticonformity, Complexity **2019**, 5150825 (2019).
[66] A. Abramiuk, J. Pawłowski, and K. Sznajd-Weron, Is independence necessary for a discontinuous phase transition within the q-voter model? Entropy **21**, 521 (2019).
[67] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models*, Collection Alea-Saclay: Monographs and Texts in Statistical Physics (Cambridge University Press, Cambridge, 1999).
[68] A. Montanari and A. Saberi, The spread of innovations in social networks, Proc. Natl. Acad. Sci. USA **107**, 20196 (2010).
[69] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd. ed. (North-Holland, Amsterdam, 2007).
[70] R. Toral and P. Colet, *Stochastic Numerical Methods: An Introduction for Students and Scientists* (Wiley, New York, NY, 2014).
[71] A. F. Peralta and R. Toral, System-size expansion of the moments of a master equation, Chaos **28**, 106303 (2018).
[72] M. Asslani, F. Di Patti, and D. Fanelli, Stochastic turing patterns on a network, Phys. Rev. E **86**, 046105 (2012).
[73] R. Grima, A study of the accuracy of moment-closure approximations for stochastic chemical kinetics, J. Chem. Phys. **136**, 154105 (2012).
[74] C. Cianci, D. Schnoerr, A. Piehler, and R. Grima, An alternative route to the system-size expansion, J. Phys. A: Math. Theor. **50**, 395003 (2017).
[75] T. Nakanishi and K. Yamamoto, System size effect in the critical fluctuation of a nonequilibrium stochastic system, Phys. Lett. A **147**, 257 (1990).
[76] T. Nakanishi, Y. Kobayashi, and K. Yamamoto, Expansion of the master equation in the vicinity of a critical point, Phys. Lett. A **273**, 53 (2000).
[77] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, Generation of uncorrelated random scale-free networks, Phys. Rev. E **71**, 027103 (2005).
[78] C. Kyriakopoulos, G. Grossmann, V. Wolf, and L. Bortolussi, Lumping of degree-based mean-field and pair-approximation equations for multistate contact processes, Phys. Rev. E **97**, 012301 (2018).
[79] P. G. Fennell and J. P. Gleeson, Multistate dynamical processes on networks: Analysis through degree-based approximation frameworks, SIAM Rev. **61**, 92 (2019).
[80] S. Unicomb, G. Iñiguez, and M. Karsai, Threshold driven contagion on weighted networks, Sci. Rep. **8**, 3094 (2018).
[81] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer Series in Synergetics, Springer, Berlin, 2009).
[82] M. San Miguel and R. Toral, Stochastic effects in physical systems, in *Instabilities and Nonequilibrium Structures VI*, edited by E. Tirapegui, J. Martínez, and R. Tiemann (Springer Netherlands, Dordrecht, 2000), pp. 35–127.
[83] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and its Applications*, Mathematics in Medicine Series (Griffin, London, 1975).
[84] R. J. Glauber, Time-dependent statistics of the Ising model, J. Math. Phys. **4**, 294 (1963).
[85] M. J. de Oliveira, Isotropic majority-vote model on a square lattice, J. Stat. Phys. **66**, 273 (1992).
[86] J. Viana Lopes, Y. G. Pogorelov, J. M. B. Lopes dos Santos, and R. Toral, Exact solution of Ising model on a small-world network, Phys. Rev. E **70**, 026112 (2004).
[87] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Critical phenomena in complex networks, Rev. Mod. Phys. **80**, 1275 (2008).
[88] A. Fernández, Center-manifold extension of the adiabatic-elimination method, Phys. Rev. A **32**, 3070 (1985).
[89] G. L. Oppo and A. Politi, Improved adiabatic elimination in laser equations, Europhys. Lett. **1**, 549 (1986).
[90] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Applied Mathematical Sciences (Springer, New York, NY, 2002).
[91] K. Binder, Finite size scaling analysis of Ising model block distribution functions, Z. Phys. B **43**, 119 (1981).
[92] S. Morris, Contagion, Rev. Economic Studies **67**, 57 (2000).
[93] A. F. Peralta, Stochastic binary-state dynamics on complex networks: A theoretical analysis, Ph.D. thesis, University of the Balearic Islands, Palma de Mallorca, Spain, 2020.