System-size expansion of the moments of a master equation

A. F. Peralta, and R. Toral

Citation: Chaos **28**, 106303 (2018); doi: 10.1063/1.5039817
View online: https://doi.org/10.1063/1.5039817
View Table of Contents: http://aip.scitation.org/toc/cha/28/10
Published by the American Institute of Physics

System-size expansion of the moments of a master equation

A. F. Peralta[a] and R. Toral
IFISC (Instituto de Física Interdisciplinar y Sistemas Complejos), Universitat de les Illes Balears-CSIC, 07122 Palma de Mallorca, Spain

(Received 11 May 2018; accepted 2 July 2018; published online 4 October 2018)

We study an expansion method of the general multidimensional master equation, based on a system-size expansion of the time evolution equations of the moments. The method turns out to be more accurate than the traditional van Kampen expansion for the first and second moments, with an error that scales with system-size similar to an alternative expansion, also applied to the equations of the moments, called Gaussian approximation, with the advantage that it has less systematic errors. Besides, we analyze a procedure to find the solution of the expansion method and we show different cases where it greatly simplifies. This includes the analytical solution of the average value and fluctuations in the number of infected nodes of the SIS epidemic model in complex networks, under the degree-based approximation. *Published by AIP Publishing.* https://doi.org/10.1063/1.5039817

In all areas of science, there are situations in which the description of the system under study requires the consideration of probability theory. In these cases, a deterministic analysis of the occurring processes may be inadequate or lacking of the relevant information sought. Master equations are an extensively used mathematical tool to deal with such situations. We consider here and compare different methods to find approximate solutions to a general master equation. The presented techniques are applied to different cases, including models of gene transcription and epidemics, for which one is able to obtain highly accurate results.

I. INTRODUCTION

Rarely the repetition of an experiment in Biology will yield exactly the same result. That does not mean that a mathematical description of biological experiments is not possible, but that an element of randomness has to be introduced in the theory from the very beginning. Despite the general belief that physical theories are always deterministic and predictive, the truth is that Statistical and Nonlinear Physics is a well established discipline that incorporates randomness in the description of a physical system at a very fundamental level. Since the pioneering works of Ludwig Boltzmann, Albert Einstein, Paul Langevin, and many others, it has become clear that macroscopic laws can emerge from a probabilistic framework that takes into account the unavoidable elements of randomness, arising from our lack of knowledge of the exact microscopic description, including all forces between particles and their initial conditions. After more than a century and a half of expertise, we have learnt that beyond the predictions of the average values of the outcome of an experiment, the fluctuations around the mean values and the probability distribution of the possible outcomes bear much information about the microscopic mechanisms that underlie the process under scrutiny. For example, if the distribution of, say, the number of a given type of particles (mRNA molecules, enzymes, etc.) produced inside a cell follows a Poisson distribution, or alternatively, if the mean value is equal to the variance, then it is likely that the particles are produced and degraded both at constant rates and independently of each other.[1]

The stochastic description in biology is used at many different levels:[2] ranging from the production of complex molecules inside the cell to the population dynamics of competing species, as in the predator-prey environment, through the enzymatic reactions,[3] the excitation of neurons,[4,5] the neurochemical synaptic connection and neuronal signal transduction,[6,7] the regulation of the sleep-wake cycle,[8–10] the effect of drugs in the treatment of illnesses, random fluctuations in genetic networks,[11] intercellular calcium spiking,[12] gene transcription,[13–16] the spread of epidemics,[17] and so on.[18] In all these cases, an adequate description begins with a "model" that takes into account as much of the phenomenon as one wants to describe, but trying not to include too many details, so avoiding a level of mathematical complexity that could make the model unsolvable. For example, when studying the spread of an epidemic, we might assume that the contagion between two individuals is a stochastic process that occurs at some constant rate, but at the same time disregard the dependence of the contagion rate with the age, sex, or other features of the individuals; we might assume that all individuals have the same number of possible contacts potentially leading to a contagion or that some individuals have more risks than others; we can assume a constant mobility of people or that some move more and farther than others, etc.

The selection of the features one wants to include constitutes the modeling part that, given the stochastic ingredients already mentioned, requires a mathematical treatment in terms of *master equations*, or equations for the probability of the different outcomes of the process, for example, the number of infected people at a given time. Once the master equation has been derived, the more technical aspect of finding its solution begins. Although a complete analytical solution is usually very difficult and one must resource to numerical methods,[19] there are some analytical techniques to extract information from the master equation concerning either the probability

[a]Electronic mail: afperalta@ifisc.uib-csic.es

function itself or, at a not so detailed level, the moments and correlations of the stochastic variables.

Amongst these analytical techniques, one of the most powerful ones is the so-called system-size expansion initiated by van Kampen in his seminal work.[20] It is based on a singular expansion around a small parameter, usually the inverse of the volume or the total number of particles. The expansion is in most cases carried out only at first order, the so-called linear noise approximation. Given its technical subtleties, the van Kampen expansion might not be very intuitive to apply in a particular problem. There have been recent approaches in which the main results of the expansion are derived in alternative ways, using, for example, a stochastic differential equation approach[21,22] or a direct Gaussian approximation scheme.[23,24] It is the intention of this paper to show how the van Kampen expansion can be formulated at the level of the exact, but not yet closed, equations for the moments and correlations. Besides reproducing the main results of the van Kampen expansion, the proposed method yields, at the same order of the expansion, a better convergence for the moments.

The outline of the paper is as follows: in Sec. II, we explain the main lines of the van Kampen original method as expanded for multidimensional systems in Ref. 25. In Sec. III, we perform the system-size expansion directly as a method to close the equations for the moments, compared to the previous van Kampen method and also the Gaussian approximation. In Sec. IV, we show how to find the general solution of the equations for the correlations and corrections to the moments, together with relevant simplifications and analytical solutions for specific cases. In Sec. V, we apply and compare the different proposed methods to two low dimensional examples, the autocatalytic reaction and a model of gene transcription, while in Sec. V C, we explain how the proposed method applies to the SIS epidemic model in complex networks. We end with a summary and conclusions in Sec. VI.

II. THE EXPANSION OF THE MASTER EQUATION

Our starting point is a general stochastic process in which a set of M integer random variables $\mathbf{x} \equiv (x_1, \dots, x_M)$ can undergo a set of K random processes defined by the vectors $\boldsymbol{\ell}^{(\nu)} \equiv (\ell_1^{(\nu)}, \dots, \ell_M^{(\nu)})$, with $\nu = 1, \dots, K$ being the index of the process. In process ν, the variables change according to $x_i \to x_i + \ell_i^{(\nu)}$ with a rate $W^{(\nu)}(\mathbf{x})$. In Sec. V, we will be considering specific examples of stochastic processes, but for now we keep a very general notation. The probability $P(\mathbf{x}; t)$ that the state of the system is $\mathbf{x}$ at time t obeys the master equation:[19,20]

$$\frac{\partial P(\mathbf{x};t)}{\partial t} = \sum_{\nu=1}^{K} \left(\prod_{i=1}^{M} E_i^{-\ell_i^{(\nu)}} - 1 \right) \left[W^{(\nu)}(\mathbf{x}) P(\mathbf{x};t) \right]. \quad (1)$$

Here, E_i is the step operator acting on any function $f[\mathbf{x}]$ of the variables x_i as

$$E_i^{\ell} \left[f(x_1, \dots, x_i, \dots, x_M) \right] = f(x_1, \dots, x_i + \ell, \dots, x_M). \quad (2)$$

The step operator can also be written formally as $E_i = e^{\partial_{x_i}}$, where $\partial_{x_i} f = \frac{\partial f}{\partial x_i}$, as it can be readily seen from the Taylor expansion

$$f(x+\ell) = \sum_{k=0}^{\infty} \frac{\partial^k f}{\partial x^k} \frac{\ell^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \left[\ell \frac{\partial}{\partial x} \right]^k f = \left[e^{\ell \partial_x} \right] f. \quad (3)$$

The celebrated van Kampen's Ω-expansion[20] is one of the most powerful methods to find an approximate solution to the complicated partial differential master equation (1). Although the method can handle more complicated cases, in its simplest form, it can be used when the rates can be written as $W^{(\nu)}(\mathbf{x}) = \Omega w^{(\nu)}\left(\frac{\mathbf{x}}{\Omega}\right)$ in terms of a large parameter Ω and $w^{(\nu)}$ functions of the intensive variables $\frac{\mathbf{x}}{\Omega}$ (see Appendix A for the necessary modification for a more complex functional form of the rates). In most cases, Ω is proportional to the total volume V or the number of particles N, hence the name *system-size expansion* with which this technique is also known. It rests upon the expectation that the dynamical variables $\mathbf{x}(t)$ will follow a deterministic trajectory whose contribution scales as Ω plus some random deviations from this trajectory that scale as $\Omega^{1/2}$. This ansatz is based on the law of large numbers, but its validity has to be checked consistently at the end of the calculation. For example, the ansatz does not work near a critical point where anomalous fluctuations exist and a modified hypothesis is then needed.

Using this ansatz, one makes the following change of variables $\mathbf{x} = \Omega \boldsymbol{\phi} + \Omega^{1/2} \boldsymbol{\xi}$ or $x_i = \Omega \phi_i + \Omega^{1/2} \xi_i$ in coordinates. Here, $\boldsymbol{\phi} = (\phi_1, \dots, \phi_M)$ is a set of time-dependent functions to be determined self-consistently, and $\boldsymbol{\xi} = (\xi_1, \dots, \xi_M)$ are the new set of random variables. All the statistical properties of $\mathbf{x}$ as determined by the probability function $P(\mathbf{x}; t)$ are to be obtained from the statistical properties of $\boldsymbol{\xi}$ determined by the probability function $\Pi(\boldsymbol{\xi}; t)$ related to $P(\mathbf{x}; t)$ by $\Pi(\boldsymbol{\xi}; t) = \Omega^{M/2} P(\mathbf{x}; t)$, as $\Omega^{M/2}$ is the Jacobian of the change of variables $\mathbf{x} \to \boldsymbol{\xi}$.

To find the partial differential master equation satisfied by $\Pi(\boldsymbol{\xi}; t)$, one makes a systematic expansion of the step operator and the rates in powers of $\Omega^{-1/2}$. For the step operator, one uses that $\partial_{x_i} = \Omega^{-1/2} \partial_{\xi_i}$, which leads to

$$\begin{aligned} \prod_{i=1}^{M} E_i^{-\ell_i^{(\nu)}} &= \prod_{i=1}^{M} e^{-\ell_i^{(\nu)} \Omega^{-1/2} \partial_{\xi_i}} = e^{-\Omega^{-1/2} \boldsymbol{\ell}^{(\nu)} \cdot \vec{\nabla}_{\boldsymbol{\xi}}} \\ &= 1 - \Omega^{-1/2} \boldsymbol{\ell}^{(\nu)} \cdot \vec{\nabla}_{\boldsymbol{\xi}} + \frac{1}{2} \Omega^{-1} (\boldsymbol{\ell}^{(\nu)} \cdot \vec{\nabla}_{\boldsymbol{\xi}})^2 + O(\Omega^{-3/2}), \end{aligned} \quad (4)$$

where $\boldsymbol{\ell}^{(\nu)} \cdot \vec{\nabla}_{\boldsymbol{\xi}} = \sum_{i=1}^{M} \ell_i^{(\nu)} \partial_{\xi_i}$. The rates are expanded as

$$\begin{aligned} W^{(\nu)}(\mathbf{x}) &= \Omega w^{(\nu)}(\mathbf{x}/\Omega) = \Omega w^{(\nu)}(\boldsymbol{\phi} + \Omega^{-1/2} \boldsymbol{\xi}) \\ &= \Omega w^{(\nu)}(\boldsymbol{\phi}) + \Omega^{1/2} \sum_{j=1}^{M} \partial_{\phi_j} w^{(\nu)}(\boldsymbol{\phi}) \xi_j + O(\Omega^0) \\ &= \Omega w^{(\nu)}(\boldsymbol{\phi}) + \Omega^{1/2} \boldsymbol{\xi} \cdot \vec{\nabla}_{\boldsymbol{\xi}} w^{(\nu)}(\boldsymbol{\phi}) + O(\Omega^0), \end{aligned} \quad (5)$$

where $\partial_{\phi_j} w^{(\nu)}(\boldsymbol{\phi})$ is the partial derivative $\dfrac{\partial w^{(\nu)}(\boldsymbol{\phi})}{\partial \phi_j}$. For the time derivative of $P(\mathbf{x};t)$, we obtain

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial t}\left[\Omega^{-M/2}\Pi(\boldsymbol{\xi})\big|_{\boldsymbol{\xi}=\Omega^{-1/2}\mathbf{x}-\Omega^{1/2}\boldsymbol{\phi}}\right] = \Omega^{-M/2}\left[\sum_i \frac{\partial \Pi}{\partial \xi_i}\left(-\Omega^{1/2}\frac{d\phi_i(t)}{dt}\right) + \frac{\partial \Pi}{\partial t}\right]. \quad (6)$$

Replacing Eqs. (4)–(6) in Eq. (1) and equating powers of Ω, we have

$$\frac{d\phi_i}{dt} = \Phi_i, \quad (7)$$

and the Fokker-Planck equation

$$\frac{\partial \Pi(\boldsymbol{\xi};t)}{\partial t} = \sum_{i,j}\frac{\partial}{\partial \xi_i}\left[B_{ij}\xi_j\Pi + \frac{1}{2}G_{ij}\frac{\partial \Pi}{\partial \xi_j}\right], \quad (8)$$

with

$$\Phi_i(\boldsymbol{\phi}) = \sum_\nu \ell_i^{(\nu)} w^{(\nu)}(\boldsymbol{\phi}), \quad (9)$$

$$B_{ij}(\boldsymbol{\phi}) = -\sum_\nu \ell_i^{(\nu)}\partial_{\phi_j} w^{(\nu)}(\boldsymbol{\phi}) = -\partial_{\phi_j}\Phi_i, \quad (10)$$

$$G_{ij}(\boldsymbol{\phi}) = \sum_\nu \ell_i^{(\nu)}\ell_j^{(\nu)} w^{(\nu)}(\boldsymbol{\phi}). \quad (11)$$

Here, Φ_i is the drift term that determines the time evolution of the deterministic part ϕ_i, $\mathbf{B} = \{B_{ij}\}_{i,j=1,\dots,M}$ is the Jacobian matrix of $\boldsymbol{\Phi} = (\Phi_1,\dots,\Phi_M)$ (defined with a minus sign), and $\mathbf{G} = \{G_{ij}\}_{i,j=1,\dots,M}$ is an additional matrix that can be calculated using $\boldsymbol{\ell}^{(\nu)}$ and $w^{(\nu)}(\boldsymbol{\phi})$. Note that Eq. (8) is linear in $\Pi(\boldsymbol{\xi};t)$, which allows us to obtain closed time evolution equations for the moments $\langle \xi_k\rangle$ and correlations

$$C_{ij} \equiv \langle \xi_i\xi_j\rangle - \langle \xi_i\rangle\langle \xi_j\rangle = \Omega^{-1}\left[\langle x_i x_j\rangle - \langle x_i\rangle\langle x_j\rangle\right] \equiv \Omega^{-1}\sigma_{ij}, \quad (12)$$

where $\sigma_{ij} = \langle x_i x_j\rangle - \langle x_i\rangle\langle x_j\rangle$ is the correlation matrix of the original x_i variables. Multiplying both sides of Eq. (8) by the desired quantity, i.e., ξ_i or $\xi_i\xi_j$, and integrating over all variables $\int d\boldsymbol{\xi} = \int\left[\prod_{j=1}^M d\xi_j\right]$, we obtain

$$\frac{d\langle \xi_i\rangle}{dt} = -\sum_k B_{ik}\langle \xi_k\rangle, \quad (13)$$

$$\frac{dC_{ij}}{dt} = -\sum_k\left[C_{jk}B_{ik} + C_{ik}B_{jk}\right] + G_{ij}. \quad (14)$$

In matrix notation $\mathbf{C} = \{C_{ij}\}_{i,j=1,\dots,M}$, we have

$$\frac{d\langle \boldsymbol{\xi}\rangle}{dt} = -\mathbf{B}\langle \boldsymbol{\xi}\rangle, \quad (15)$$

$$\frac{d\mathbf{C}}{dt} = -\mathbf{C}\mathbf{B}^\intercal - \mathbf{B}\mathbf{C} + \mathbf{G}. \quad (16)$$

If we assume the existence of a steady-state of Eq. (7), obtained by setting $\boldsymbol{\Phi} = 0$, and $\langle \boldsymbol{\xi}\rangle_{\rm st} = 0$, then the correlation matrix $\mathbf{C}_{\rm st}$ in the steady-state satisfies

$$\mathbf{C}_{\rm st}\mathbf{B}_{\rm st}^\intercal + \mathbf{B}_{\rm st}\mathbf{C}_{\rm st} = \mathbf{G}_{\rm st}, \quad (17)$$

in terms of the steady-state matrices $\mathbf{B}_{\rm st}, \mathbf{G}_{\rm st}$. As shown in Ref. 20, the steady-state probability distribution has a Gaussian shape

$$\Pi_{\rm st}(\boldsymbol{\xi}) = \sqrt{\frac{|\mathbf{C}_{\rm st}|}{(2\pi)^M}}e^{-\frac{1}{2}\boldsymbol{\xi}^\intercal\cdot\mathbf{C}_{\rm st}^{-1}\cdot\boldsymbol{\xi}}. \quad (18)$$

In the time-dependent case, the probability distribution obeys also a Gaussian expression

$$\Pi(\boldsymbol{\xi},t) = \sqrt{\frac{|\mathbf{C}|}{(2\pi)^M}}e^{-\frac{1}{2}(\boldsymbol{\xi}-\langle\boldsymbol{\xi}\rangle)^\intercal\cdot\mathbf{C}^{-1}\cdot(\boldsymbol{\xi}-\langle\boldsymbol{\xi}\rangle)}, \quad (19)$$

provided the initial condition is also a Gaussian distribution (this includes as a limiting case the Dirac-delta or dispersion-free initial conditions).

This ends our brief summary of the van Kampen expansion method. In Sec. IV, we derive following Ref. 25 a general solution of Eq. (17) for the steady-state correlations $\mathbf{C}_{\rm st}$ obtained by diagonalizing matrix $\mathbf{B}_{\rm st}$. In the same section, we consider some simplifications and the explicit solution in some cases of interest. In those cases in which $\mathbf{B}_{\rm st}$ cannot be diagonalized, we propose in Appendix B a very efficient recursive numerical algorithm to obtain $\mathbf{C}_{\rm st}$. Before that, we explain in Sec. II alternative approaches to the van Kampen expansion, whose starting point is the derivation of equations for the first and second moments obtained directly from the master equation, without the need to use the Fokker-Planck approximation.

III. THE EXPANSION OF THE EQUATIONS FOR THE MOMENTS

Let us focus now on the time evolution of the first $\langle x_i\rangle$ and second $\langle x_i x_j\rangle$ moments, defined as

$$\langle x_i(t)\rangle = \int d\mathbf{x}\, x_i P(\mathbf{x};t), \quad (20)$$

$$\langle x_i(t)x_j(t)\rangle = \int d\mathbf{x}\, x_i x_j P(\mathbf{x};t), \quad (21)$$

with $\int d\mathbf{x} = \int\left[\prod_{j=1}^M dx_j\right]$. Multiplying both sides of the master equation (1) by x_i or $x_i x_j$ and integrating over $\int d\mathbf{x}$, the quantities $\langle x_i\rangle$ and $\langle x_i x_j\rangle$ are shown to satisfy the following evolution equations:

$$\frac{d\langle x_i\rangle}{dt} = \sum_{\nu=1}^K \ell_i^{(\nu)}\langle W^{(\nu)}\rangle, \quad (22)$$

$$\frac{d\langle x_i x_j\rangle}{dt} = \sum_{\nu=1}^K\left[\ell_i^{(\nu)}\ell_j^{(\nu)}\langle W^{(\nu)}\rangle + \ell_i^{(\nu)}\langle x_j W^{(\nu)}\rangle + \ell_j^{(\nu)}\langle x_i W^{(\nu)}\rangle\right]. \quad (23)$$

Except for a very few functional dependence forms of the rates $W^{(\nu)}(\mathbf{x})$, these equations are not closed as the right-hand-side contains higher order moments $\langle x_{i_1}x_{i_2}x_{i_3}\dots\rangle$ for which one must write down evolution equations leading to a,

usually infinite, hierarchy of equations which, with very few exceptions, cannot be solved.

In order to close this infinite hierarchy of equations, we will describe two alternative methods: (i) the Gaussian approximation and (ii) a system-size type of expansion, inspired by the van Kampen method, but performed directly at the level of the equations for the moments Eqs. (22) and (23), instead of the master equation (1).

A. The Gaussian approximation

As the method has been recently exposed by one of the authors, we just highlight here its main ingredients and refer to Ref. 23 for further details. The Gaussian approximation consists in writing any higher order moment $\langle x_{i_1}x_{i_2}x_{i_3}\ldots\rangle$ that might appear in Eqs. (22) and (23) as a function of the two first moments $\langle x_i\rangle$ and $\langle x_ix_j\rangle$, assuming that $\mathbf{x}$ distributes as a Gaussian random variable, a fact that is based on the van Kampen expansion as explained in Sec. II. For a single variable x_i, the Gaussian distribution comes fully specified if we know its first and second moments. Higher order moments are derived from these two lower order moments. For example, as needed in the examples of Sec. V, we have $\langle x_i^3\rangle = 3\langle x_i^2\rangle\langle x_i\rangle - 2\langle x_i\rangle^3$. For multi-dimensional distributions, any average $\langle x_{i_1}x_{i_2}x_{i_3}\ldots\rangle$ can be written in terms of the two first moments $\langle x_i\rangle$ and $\langle x_ix_j\rangle$ (or the correlations $\sigma_{ij} = \langle x_ix_j\rangle - \langle x_i\rangle\langle x_j\rangle$) using the Isserlis[26] theorem, also known as Wick's theorem.[27] For example, $\langle x_i^2x_j\rangle = \langle x_i^2\rangle\langle x_j\rangle + 2\langle x_i\rangle\langle x_ix_j\rangle - 2\langle x_i\rangle^2\langle x_j\rangle$.

As shown in the detailed study of Ref. 23, this method can be more precise than the traditional van Kampen expansion. For example, while the error of the van Kampen expansion in the variables $\left(\langle x_i\rangle, \langle x_ix_j\rangle, \sigma_{ij}\right)$ is of order $(\Omega^0, \Omega^1, \Omega^{1/2})$, respectively, the Gaussian approximation provides errors that scale at most as $(\Omega^{-1/2}, \Omega^{1/2}, \Omega^{1/2})$. This is a substantial improvement for $\left(\langle x_i\rangle, \langle x_ix_j\rangle\right)$ and the same scaling for σ_{ij}, although the numerical results of Ref. 23 indicate that the system-size expansion of the master equation gives smaller errors for σ_{ij} than the Gaussian approximation. One problem of the Gaussian approximation is that for very low Ω, the method may not converge. This occurs because the variables $\mathbf{x}$ are considered to be Gaussian, and thus unbounded, when actually they are restricted to be positive $\mathbf{x} > 0$.

B. System-size expansion

Our intention here is to close the hierarchy of Eqs. (22) and (23) using a system-size expansion similar to the one used in Sec. II to reduce the master equation to a Fokker-Planck equation. To this end, we propose an expansion of the stochastic term as $\boldsymbol{\xi} = \boldsymbol{a} + \Omega^{-1/2}\boldsymbol{b} + O(\Omega^{-1})$, where $\boldsymbol{a}$ and $\boldsymbol{b}$ are stochastic variables. We can determine the order of this approximation for the average value and correlations using $\mathbf{x} = \Omega\boldsymbol{\phi} + \Omega^{1/2}\boldsymbol{a} + \Omega^0\boldsymbol{b} + O(\Omega^{-1/2})$:

$$\langle x_i\rangle = \Omega\phi_i + \Omega^{1/2}\langle a_i\rangle + \Omega^0\langle b_i\rangle + O(\Omega^{-1/2}), \tag{24}$$

$$\langle x_ix_j\rangle = \Omega^2\phi_i\phi_j + \Omega^{3/2}\left(\phi_i\langle a_j\rangle + \phi_j\langle a_i\rangle\right) + \Omega^1\left(\langle a_ia_j\rangle + \phi_i\langle b_j\rangle + \phi_j\langle b_i\rangle\right) + O(\Omega^{1/2}), \tag{25}$$

$$\sigma_{ij} = \Omega\tilde{C}_{ij} + O(\Omega^{1/2}), \tag{26}$$

where we define $\tilde{C}_{ij} \equiv \langle a_ia_j\rangle - \langle a_i\rangle\langle a_j\rangle$. If we introduce Eqs. (24) and (25) in Eqs. (22) and (23) and equate powers of Ω, we obtain

$$\frac{d\phi_i}{dt} = \Phi_i, \tag{27}$$

$$\frac{d\langle a_i\rangle}{dt} = -\sum_j B_{ij}\langle a_j\rangle, \tag{28}$$

$$\frac{d\langle b_i\rangle}{dt} = -\sum_j B_{ij}\langle b_j\rangle + \frac{1}{2}\sum_{j,k}\left[\tilde{C}_{jk} + \langle a_j\rangle\langle a_k\rangle\right]\partial^2_{\phi_j,\phi_k}\Phi_i, \tag{29}$$

$$\frac{d\tilde{C}_{ij}}{dt} = -\sum_k\left[\tilde{C}_{jk}B_{ik} + \tilde{C}_{ik}B_{jk}\right] + G_{ij}. \tag{30}$$

Note that Eq. (30) for the correlations $\tilde{C}_{ij}$ coincides with Eq. (14) satisfied by C_{ij}, and hence, both correlations are identical. However, and this is the key point, had we not included the term $\Omega^0 b_i$ in the expansion, then Eq. (30) would have contained an additional incorrect term $\frac{1}{2}\sum_{k,q}\langle a_ka_q\rangle\left[\phi_i\partial^2_{\phi_k,\phi_q}\Phi_j + \phi_j\partial^2_{\phi_k,\phi_q}\Phi_i\right]$. Additionally, the equations for ϕ_i and $\langle a_i\rangle$ coincide with Eqs. (7) and (13), but, and this is the improvement of this method, we obtain an additional equation (29) for the average value of the higher order correction. According to Eqs. (24)–(26), the accuracy of this method is $(\Omega^{-1/2}, \Omega^{1/2}, \Omega^{1/2})$ for $\left(\langle x_i\rangle, \langle x_ix_j\rangle, \sigma_{ij}\right)$, respectively, the same scaling of the errors one finds in the Gaussian approximation discussed before, being therefore more accurate for $\left(\langle x_i\rangle, \langle x_ix_j\rangle\right)$ than the system-size expansion of the master equation. Beyond the scaling with system size of the errors, we will compute numerically in the examples of the next sections the prefactors of these scaling laws, and we will show that the actual values of the errors are substantially reduced using the expansion method based on the equations for the moments developed in this section.

1. Higher order expansions

It is of course possible to include higher order terms in the expansion $\mathbf{x} = \Omega\boldsymbol{\phi} + \Omega^{1/2}\boldsymbol{a} + \Omega^0\boldsymbol{b} + \Omega^{-1/2}\boldsymbol{c} + O(\Omega^{-1})$. One finds then the following expressions for $\langle x_i\rangle, \langle x_ix_j\rangle, \sigma_{ij}$:

$$\langle x_i\rangle = \Omega\phi_i + \Omega^{1/2}\langle a_i\rangle + \Omega^0\langle b_i\rangle + \Omega^{-1/2}\langle c_i\rangle + O(\Omega^{-1}), \tag{31}$$

$$\langle x_ix_j\rangle = \Omega^2\phi_i\phi_j + \Omega^{3/2}(\phi_i\langle a_j\rangle + \phi_j\langle a_i\rangle) + \Omega^1(\langle a_ia_j\rangle + \phi_i\langle b_j\rangle + \phi_j\langle b_i\rangle) + \Omega^{1/2}\left(\phi_i\langle c_j\rangle + \phi_j\langle c_i\rangle + \langle a_ib_j\rangle + \langle a_jb_i\rangle\right) + O(\Omega^0), \tag{32}$$

$$\sigma_{ij} = \Omega(\langle a_ia_j\rangle - \langle a_i\rangle\langle a_j\rangle) + \Omega^{1/2}\left(\langle a_ib_j\rangle + \langle a_jb_i\rangle - \langle a_i\rangle\langle b_j\rangle - \langle a_j\rangle\langle b_i\rangle\right) + O(\Omega^0). \tag{33}$$

If we introduce the new ansatz in the equations for the moments Eqs. (22) and (23), we obtain the same

Eqs. (27)–(30) and additional time evolution equations for $\langle c_i \rangle$ and $\langle a_i b_j \rangle$:

$$\frac{d\langle c_i \rangle}{dt} = -\sum_j B_{ij}\langle c_j \rangle + \sum_{j,k} \langle a_j b_k \rangle \partial^2_{\phi_j,\phi_k} \Phi_i + \frac{1}{6}\sum_{j,k,q} \langle a_j a_k a_q \rangle \partial^3_{\phi_j,\phi_k,\phi_q} \Phi_i, \tag{34}$$

$$\frac{d\langle a_i b_j \rangle}{dt} + \frac{d\langle a_j b_i \rangle}{dt} = -\sum_k B_{jk}\left(\langle b_i a_k \rangle + \langle a_i b_k \rangle\right) - \sum_k B_{ik}\left(\langle b_j a_k \rangle + \langle a_j b_k \rangle\right) + \frac{1}{2}\sum_{k,q} \langle a_i a_k a_q \rangle \partial^2_{\phi_k,\phi_q} \Phi_j + \frac{1}{2}\sum_{k,q} \langle a_j a_k a_q \rangle \partial^2_{\phi_k,\phi_q} \Phi_i + \sum_k \partial_{\phi_k} G_{ij} \langle a_k \rangle, \tag{35}$$

with the particularity that they depend also on the third moment $\langle a_i a_j a_k \rangle$. In order to find the time evolution equation for $\langle a_i a_j a_k \rangle$, we need to relate it with the third moment $\langle x_i x_j x_k \rangle$

$$\langle x_i x_j x_k \rangle = \mathcal{P}_{ijk}\left[\Omega^3 \phi_i \phi_j \phi_k + \Omega^{5/2} \phi_i \phi_k \langle a_j \rangle + \Omega^2\left(\phi_i \phi_j \langle b_k \rangle + \phi_i \langle a_j a_k \rangle\right) + \Omega^{3/2}\left(\phi_i \phi_j \langle c_k \rangle + \phi_i \langle a_j b_k \rangle + \langle a_i a_j a_k \rangle\right)\right] + O(\Omega^1), \tag{36}$$

where the permutation operator $\mathcal{P}_{ijk}$ adds to the expression inside the brackets any permutation of the indices i, j, k, which leads to a different result (e.g., $\mathcal{P}_{ijk}\left[\phi_i \phi_k \langle a_j \rangle\right] = \phi_i \phi_k \langle a_j \rangle + \phi_i \phi_j \langle a_k \rangle + \phi_j \phi_k \langle a_i \rangle$). Using now the exact evolution equation that follows from the master equation

$$\frac{d\langle x_i x_j x_k \rangle}{dt} = \sum_{\nu=1}^{K} \mathcal{P}_{ijk}\left[\ell_i^{(\nu)} \ell_j^{(\nu)} \ell_k^{(\nu)} \langle W^{(\nu)} \rangle + \ell_i^{(\nu)} \langle x_j x_k W^{(\nu)} \rangle + \ell_i^{(\nu)} \ell_j^{(\nu)} \langle x_k W^{(\nu)} \rangle\right], \tag{37}$$

we obtain

$$\frac{d\langle a_i a_j a_k \rangle}{dt} = \mathcal{P}_{ijk}\left[-\sum_q B_{kq} \langle a_i a_j a_q \rangle + G_{ij} \langle a_k \rangle\right]. \tag{38}$$

According to Eqs. (31)–(33), the method is of order $(\Omega^{-1}, \Omega^0, \Omega^0)$ for the variables $(\langle x_i \rangle, \langle x_i x_j \rangle, \sigma_{ij})$. In the steady state, it is $\langle c_i \rangle_{\rm st} = \langle a_i b_j \rangle_{\rm st} = \langle a_i a_j a_k \rangle_{\rm st} = 0$, which means that these higher order corrections are zero and the accuracy of the method of Sec. III B is higher than expected.

Higher order expansions applied directly to the master equation (not in the equations for the moments) are discussed in Ref. 20. In that case, the van Kampen ansatz $\mathbf{x} = \Omega \boldsymbol{\phi} + \Omega^{1/2} \boldsymbol{\xi}$ is kept as a change of variables, but we take into account the further orders of Eqs. (4) and (5) such that the probability is split as $\Pi(\boldsymbol{\xi}; t) = \Pi_0(\boldsymbol{\xi}; t) + \Omega^{-1/2}\Pi_1(\boldsymbol{\xi}; t) + O(\Omega^{-1})$ and Eq. (8) is only valid for $\Pi_0(\boldsymbol{\xi}; t)$. This higher order van Kampen expansion has the advantage of giving an equation for the probability distribution $\Pi(\boldsymbol{\xi}; t)$, but it is again less accurate $(\Omega^{-1/2}, \Omega^{1/2}, \Omega^0)$ for $\left(\langle x_i \rangle, \langle x_i x_j \rangle, \sigma_{ij}\right)$ compared to the expansion of the moments, for $\langle x_i \rangle$, $\langle x_i x_j \rangle$. In general, one can conclude that any higher order system-size expansion will be one order superior for $\langle x_i \rangle$, $\langle x_i x_j \rangle$ when performed in the equations for the moments and the same order for σ_{ij}, compared to the equivalent expansion of the master equation.

IV. SOLUTION OF THE EQUATION FOR CORRELATIONS

We now describe a general procedure as developed in Ref. 25 to solve Eq. (17) and find $\mathbf{C}_{\rm st}$ that, however, requires the diagonalization of matrix $\mathbf{B}_{\rm st}$. In those cases in which the diagonalization is not possible or it shows difficulties, we present in Appendix B an alternative recursive numerical scheme that can provide $\mathbf{C}_{\rm st}$ very efficiently and with high precision.

Let us assume, then, that it is possible to write $\mathbf{B}_{\rm st} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, with $\mathbf{D}$ being a diagonal matrix and $\mathbf{P}$ being the matrix of change of basis, whose columns are the corresponding eigenvectors. Replacing $\mathbf{Q} = \mathbf{P}^{-1}\mathbf{C}_{\rm st}\left(\mathbf{P}^{\intercal}\right)^{-1}$ and $\mathbf{F} = \mathbf{P}^{-1}\mathbf{G}_{\rm st}\left(\mathbf{P}^{\intercal}\right)^{-1}$ (note that $\mathbf{Q}^{\intercal} = \mathbf{Q}$ and $\mathbf{F}^{\intercal} = \mathbf{F}$) in Eq. (17), we arrive at

$$\mathbf{Q}\mathbf{D} + \mathbf{D}\mathbf{Q} = \mathbf{F}. \tag{39}$$

In index notation, $Q_{ij}D_j + D_i Q_{ij} = F_{ij}$, leading to

$$Q_{ij} = \frac{F_{ij}}{D_i + D_j}. \tag{40}$$

Once we have $\mathbf{Q}$, we obtain $\mathbf{C}_{\rm st} = \mathbf{P}\mathbf{Q}\mathbf{P}^{\intercal}$. It is also possible to write the solution in index notation as

$$C^{\rm st}_{ij} = \sum_{k,q} G^{\rm st}_{kq} \sum_{n,m} \frac{P_{in}P^{-1}_{nk}P_{jm}P^{-1}_{mq}}{D_n + D_m}. \tag{41}$$

The corrections to the mean value $\langle \mathbf{b} \rangle_{\rm st}$ can also be determined. Applying the linear transformation that diagonalizes the Jacobian matrix, we define $\mathbf{g} = \mathbf{P}^{-1}\mathbf{b}$, $\mathbf{u} = \mathbf{P}^{-1}\boldsymbol{\phi}$, and $\mathbf{U} \equiv \mathbf{P}^{-1}\boldsymbol{\Phi}$. Multiplying both sides of Eq. (29) by $\mathbf{P}^{-1}(\cdots)$, taking into account that $\partial_{\phi_i} = \sum_k P^{-1}_{ki}\partial_{u_k}$, and setting the time derivative equal to zero in the steady-state, we obtain

$$\langle \mathrm{g}_i \rangle_{\rm st} = \frac{1}{2D_i}\sum_{k,q} Q_{kq}\partial^2_{u_k,u_q} U_i. \tag{42}$$

One can now undo the transformation $\langle \mathbf{b} \rangle_{\rm st} = \mathbf{P}\langle \mathbf{g} \rangle_{\rm st}$.

Alternatively, a more useful equation for the average value $\langle b_i \rangle_{\rm st}$ can also be obtained if we determine first $\Gamma_i \equiv \frac{1}{2}\sum_{k,q} C^{\rm st}_{kq}\partial^2_{\phi_k,\phi_q}\Phi_i$, and then Eq. (42) can be rewritten as $\langle \mathrm{g}_i \rangle_{\rm st} = \frac{1}{D_i}\sum_n P^{-1}_{in}\Gamma_n$, which leads to

$$\langle \mathrm{b}_i \rangle_{\rm st} = \sum_{n,m} \frac{P_{in}P^{-1}_{nm}\Gamma_m}{D_n}. \tag{43}$$

The above expressions, besides an explicit solution of Eqs. (17) and (29) that can be used both analytically or numerically, provide a close relation between the solution of the linearization of the deterministic dynamics equation (7) and the correlations between variables. Basically, if one solves the linear dynamics, finding the steady-state correlations requires much less effort than approaching the problem Eq. (17)

directly as a system of equations. In some cases of interest, the solution can be simplified or directly found with this method. Next, we show an interesting collection of these cases.

A. Degenerate eigenvalues

The *star node* dynamics correspond to a situation where all the eigenvalues are equal $D_i = D$, $\forall i$, such that any arbitrary vector $\mathbf{v}$ is an eigenvector of $\mathbf{B}_{\rm st}$ with the same eigenvalue, i.e., $\mathbf{B}_{\rm st}\mathbf{v} = D\mathbf{v}$. In this case, $\mathbf{Q} = \frac{1}{2D}\mathbf{F} = \frac{1}{2D}\mathbf{P}^{-1}\mathbf{G}_{\rm st}\,(\mathbf{P}^{\intercal})^{-1}$, and the expression $\mathbf{C}_{\rm st} = \mathbf{P}\mathbf{Q}\mathbf{P}^{\intercal}$ becomes

$$\mathbf{C}_{\rm st} = \mathbf{P}\left[\frac{1}{2D}\mathbf{P}^{-1}\mathbf{G}_{\rm st}\,(\mathbf{P}^{\intercal})^{-1}\right]\mathbf{P}^{\intercal} = \frac{1}{2D}\mathbf{G}_{\rm st}. \tag{44}$$

This is, the correlation matrix is equal to the G-matrix rescaled with $1/2D$.

A more complex case is when there is a non-degenerate eigenvalue D_1 with associated eigenvector $\mathbf{v}_1^{\intercal} = (u_1, \ldots, u_M)$ and, additionally, the rest of the eigenvectors $\mathbf{v}_2$ form a star node, $\mathbf{B}_{\rm st}\mathbf{v}_2 = D_2\mathbf{v}_2$ with an eigenvalue D_2 of degeneracy $M-1$, in the plane $\mathbf{n}^{\intercal}\cdot\mathbf{v}_2 = 0$, where $\mathbf{n}^{\intercal} = (n_1, n_2, \ldots)$ is a normal vector to the plane. The matrix $\mathbf{P}$ fulfills then $P_{i1} = u_i$, $\forall i$, $\sum_i n_i P_{ij} = \delta_{1j}\mathbf{n}^{\intercal}\cdot\mathbf{v}_1$, and we can determine the first row of the inverse matrix as $P^{-1}_{1j} = \frac{n_j}{\mathbf{n}^{\intercal}\cdot\mathbf{v}_1}$. This information is enough to simplify Eq. (41). If we consider separately the terms $n = 1$, $m = 1$, we find

$$\begin{aligned} C^{\rm st}_{ij} &= \sum_{k,q} G^{\rm st}_{kq}\left(\frac{u_i u_j P^{-1}_{1k}P^{-1}_{1q}}{2D_1}\right. \\ &\quad + \frac{u_j P^{-1}_{1q}(\delta_{ik} - u_i P^{-1}_{1k}) + u_i P^{-1}_{1k}(\delta_{jq} - u_j P^{-1}_{1q})}{D_1 + D_2} \\ &\quad \left. + \frac{(\delta_{ik} - u_i P^{-1}_{1k})(\delta_{jq} - u_j P^{-1}_{1q})}{2D_2}\right) \\ &= u_i u_j \frac{F_{11}}{2D_1} + \frac{\alpha_i u_j + \alpha_j u_i - 2F_{11}u_i u_j}{D_1 + D_2} \\ &\quad + \frac{G^{\rm st}_{ij} - \alpha_i u_j - \alpha_j u_i + F_{11}u_i u_j}{2D_2}, \end{aligned} \tag{45}$$

with $F_{11} = \sum_{k,q} P^{-1}_{1k}P^{-1}_{1q}G^{\rm st}_{kq}$ and $\alpha_i \equiv \sum_k P^{-1}_{1k}G^{\rm st}_{ik}$. In general, if $\mathbf{B}_{\rm st}$ has degenerate eigenvalues, it is natural to split the sum of Eq. (41) for the indices n, m that have common eigenvalues $D_n = D$, $D_m = D'$ and the analysis is greatly simplified.

Using this same trick in Eq. (43), the corrections to the average value $\langle b_i\rangle_{\rm st}$ are

$$\langle \mathrm{b}_i\rangle_{\rm st} = \frac{\Gamma_i}{D_2} + \frac{D_2 - D_1}{D_1 D_2}u_i\sum_j \Gamma_j P^{-1}_{1j}. \tag{46}$$

B. Time scale separation

If the linearization of the dynamics reveals an eigenvalue D_1 which is very small compared to the rest $D_i \gg D_1$, an approximate solution is directly found. This corresponds to a situation where the trajectories bend following the slow eigendirection $\mathbf{v}_1$. In this case, we have that $Q_{11} = \frac{F_{11}}{2D_1} \gg Q_{ij}$ and then

$$C^{\rm st}_{ij} = \sum_{k,q} P_{ik}Q_{kq}P_{jq} \approx P_{i1}Q_{11}P_{j1}. \tag{47}$$

We only have to compute the first column P_{i1} and the first row P^{-1}_{1i}, and we avoid tedious matrix multiplications. Note that the correlations are the product of the coordinates of the eigenvector $\mathbf{v}_1$ times the global value Q_{11}. In this same limit, Eq. (42) reduces to

$$\langle \mathrm{g}_i\rangle_{\rm st} = \frac{Q_{11}}{2D_i}\partial^2_{u_1}U_i, \tag{48}$$

and the leading correction to the average value is $\langle \mathrm{b}_i\rangle_{\rm st} = P_{i1}\langle \mathrm{g}_1\rangle_{\rm st}$.

C. Conservation laws

It may happen that not all the variables are independent, i.e., the variables are related by what we will assume a set $i = 1, \ldots, M_c$ of linear relations $\sum_{j=1}^{M} H_{ij}x_j = \mathcal{C}_i$, where $\mathcal{C}_i$ are a given set of constants. For example, this could be the stoichiometric relations of the different species involved in a chemical reaction.

If the master equation respects these relations, we expect the constants not to change during any process $x_j \to x_j + \ell_j^{(\nu)}$ and then $\sum_{j=1}^{M} H_{ij}\ell_j^{(\nu)} = 0$, or equivalently multiplying by $w^{(\nu)}(\boldsymbol{\phi})$ and summing over all ν we have $\sum_{j=1}^{M} H_{ij}\Phi_j = 0$. Consequently, the matrix $\mathbf{B}_{\rm st}$ fulfills the relations $\sum_{j=1}^{M} H_{ij}B^{\rm st}_{jk} = 0$, for any k. Comparing this to the equations for the eigenvalues $\mathbf{P}^{-1}\mathbf{B}_{\rm st} = \mathbf{D}\mathbf{P}^{-1}$, in index notation $\sum_{j=1}^{M} P^{-1}_{ij}B^{\rm st}_{jk} = D_i P^{-1}_{ik}$, we can associate nil eigenvalues $D_i = 0$, and also $H_{ij} = P^{-1}_{ij}$, for $i = 1, \ldots, M_c$.

For two values $i = 1, \ldots, M_c$ and $j = 1, \ldots, M_c$, the denominator of Eq. (40) vanishes but also does the numerator:

$$\begin{aligned} F_{ij} &= \sum_{k,l} P^{-1}_{ik}G^{\rm st}_{kl}P^{-1}_{jl} \\ &= \sum_{\nu}\left(\sum_k P^{-1}_{ik}\ell^{(\nu)}_k\right)\left(\sum_l P^{-1}_{jl}\ell^{(\nu)}_l\right)w^{(\nu)} = 0, \end{aligned} \tag{49}$$

and then we can set $Q_{ij} = 0$. Thus, the matrix $\mathbf{Q}$ has a block of zeros that corresponds to the set of conserved quantities by the master equation.

V. EXAMPLES

In order to test the accuracy and differences of the methods, we will apply them to some examples: the autocatalytic reaction, already considered in Ref. 20, a model of gene transcription,[13] and the susceptible-infected-susceptible (SIS) model of epidemic spreading. For each example, we will consider the van Kampen system-size expansion of the master equation (vKE) as given by Eqs. (7), (15), and (16), the Gaussian approximation as explained in Ref. 23, and the system-size expansion carried out at the level of the equations for the moments (SSM) given by Eqs. (27)–(30).

A. Autocatalytic reaction

Consider the following chemical reaction:

$$A \xrightarrow{\alpha} X, \quad 2X \xrightarrow{\gamma} Y. \tag{50}$$

With rate α molecules of substance X are created out of a reservoir of size Ω; with rate γ, when two molecules meet, they react to give a molecule of substance Y. We only consider one variable, the number n of molecules of X. This variable can change by any of the two processes outlined in (50). In the notation of Sec. II, we have two processes $\nu = 1, 2$, creation and reaction. The changes in the variable n and the rate of the processes are $\ell^{(1)} = 1$ (one particle is created from the reservoir), $W^{(1)} = \Omega\alpha$ and $\ell^{(2)} = -2$ (two particles disappear when they react), $W^{(2)} = \dfrac{\gamma}{\Omega} n(n-1)$, where we assume that the production rate is proportional to Ω, and the probability of two X-particles to meet is inversely proportional to Ω. Using Eqs. (22) and (23), we find equations for the first and second moments $\langle n\rangle$, $\langle n^2\rangle$:

$$\frac{d\langle n\rangle}{dt} = \Omega\alpha - \frac{2\gamma}{\Omega}(\langle n^2\rangle - \langle n\rangle), \tag{51}$$

$$\frac{d\langle n^2\rangle}{dt} = \Omega\alpha(1 + 2\langle n\rangle) + \frac{4\gamma}{\Omega}\left(-\langle n^3\rangle + 2\langle n^2\rangle - \langle n\rangle\right), \tag{52}$$

which are not closed due to the presence of $\langle n^3\rangle$ on the right-hand-side of Eq. (52). In the Gaussian approximation, we use the replacement $\langle n^3\rangle = 3\langle n^2\rangle\langle n\rangle - 2\langle n\rangle^3$ and Eqs. (51) and (52) become closed. When using the system-size expansion, either at the level of the master equation or the equations for the moments, we need to determine Φ and G given by Eqs. (9) and (11) in the one-variable case:

$$\Phi = \alpha - 2\gamma\phi^2, \tag{53}$$

$$G = \alpha + 4\gamma\phi^2. \tag{54}$$

The equations for the correlations [Eq. (30)] involve the first derivative $\partial_\phi\Phi = -4\gamma\phi$ and the corrections to the average value Eq. (29) the second derivative $\partial^2_\phi\Phi = -4\gamma$. The rate $W^{(2)}$ does not strictly follow the scaling form $W^{(2)} = \Omega w^{(2)}(n/\Omega)$ as we have instead $W^{(2)} = \Omega\gamma \cdot \left[\left(\frac{n}{\Omega}\right)^2 - \frac{n}{\Omega^2}\right]$, one needs to redo the expansions taking into account the correct presence of the different Ω terms. In Appendix A, we show that the only effect is to add an extra term $\Delta = 2\gamma\phi$ to the right-hand side of Eq. (29), and the original van Kampen expansion is not modified.

In Fig. 1, we plot the results of the time evolution of $\langle n\rangle$, $\langle n^2\rangle$, σ^2 for the three different proposed expansion methods, compared to the numerical results of the Gillespie algorithm applied to Eq. (50). For the average value $\langle n\rangle$ [Fig. 1(a)], we have SSM > Gaussian > vKE in order of accuracy; for the second moment $\langle n^2\rangle$ [Fig. 1(b)], Gaussian >SSM > vKE; and for the variance σ^2 [Fig. 1(c)], SSM = vKE > Gaussian. This confirms our expectations that the expansion of the moments and the Gaussian approximation work better than the standard van Kampen expansion for the first two moments. In Fig. 2, we plot the errors of $\langle n\rangle$, $\langle n^2\rangle$, σ^2 as a function of the volume Ω, which are compatible with the scaling relations derived in Sec. III. For the second moment $\langle n^2\rangle$ [Fig. 2(b)], there is the particularity that the error of the Gaussian approximation scales as Ω^{-1}, better than expected, which is a particular feature of the autocatalytic reaction due to some cancellations. Although the error of the Gaussian approximation and SSM scale equivalently, one can appreciate in Figs. 2(a) and 2(c) that SSM presents smaller systematic errors. We also observe the predicted problems of the Gaussian approximation for extremely low system sizes, where the errors soar or the method does not even converge.

B. Gene transcription

We study now the model of gene transcription proposed in Ref. 13. In that model, the system is described by two variables n_r and n_p, respectively, the number of molecules of *mRNA* and proteins (p). The model is constructed as follows: (1) with rate k_r, a fraction of DNA (D) is copied into a *mRNA* molecule; (2) with rate k_p, a molecule of *mRNA* is translated into a protein; (3) *mRNA* degrades with rate γ_r; and (4) protein degrades with rate γ_p. Schematically in the reaction notation this can be written as

$$D \xrightarrow{k_r} mRNA, \quad mRNA \xrightarrow{k_p} p,$$
$$mRNA \xrightarrow{\gamma_r} \emptyset, \quad p \xrightarrow{\gamma_p} \emptyset. \tag{55}$$

We have then four processes $\nu = 1, 2, 3, 4$, transcription, translation, degradation of *mRNA*, and degradation of protein, with changes in the variables and rates: $\ell^{(1)}_r = +1$, $W^{(1)} = \Omega k_r$; $\ell^{(2)}_r = -1$, $\ell^{(2)}_p = +1$, $W^{(2)} = n_r k_p$; $\ell^{(3)}_r = -1$,

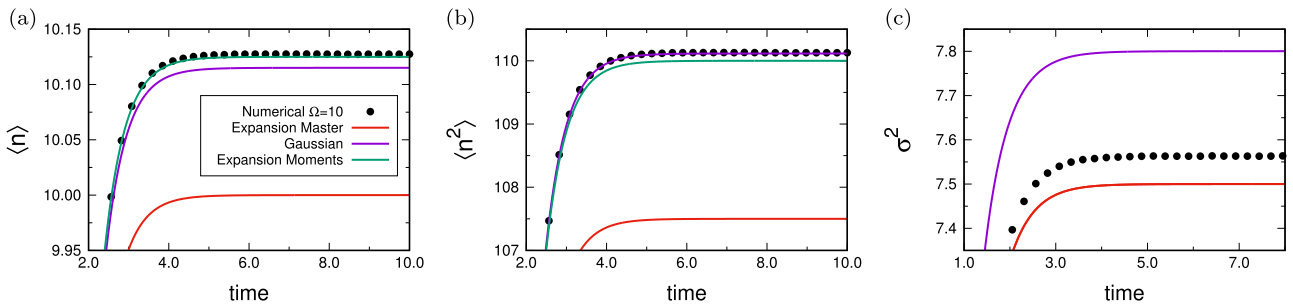

FIG. 1. Time evolution of $\langle n\rangle$, $\langle n^2\rangle$, and σ^2 shown in panels (a), (b), and (c), respectively, for the autocatalytic reaction (50) with parameters $\alpha = 1$, $\gamma = 1/2$, $\Omega = 10$. The black dots are the numerical results of the Gillespie algorithm, averaged over 10^9 trajectories, compared to the results of the standard van Kampen expansion (red lines), the Gaussian approximation (purple), and the system-size expansion of the moments (green, coincident with the red lines in the case of the variance).

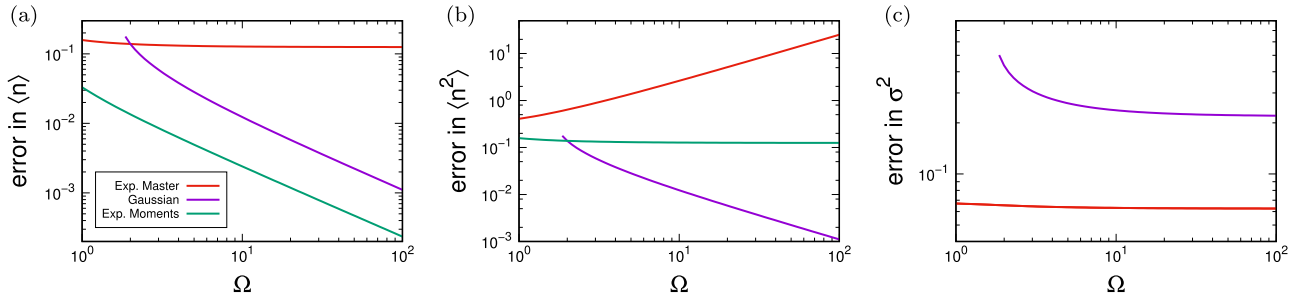

FIG. 2. Error of the methods in the steady-state state of variables $\langle n\rangle_{\rm st}$, $\langle n^2\rangle_{\rm st}$, and $\sigma^2_{\rm st}$ shown in panels (a), (b), and (c) respectively, for the autocatalytic reaction (50) with parameters $\alpha = 1$, $\gamma = 1/2$, as a function of the volume Ω. The standard van Kampen expansion (red lines) displays errors that scale as $(\Omega^0, \Omega^1, \Omega^0)$; the Gaussian approximation (purple) as $(\Omega^{-1}, \Omega^{-1}, \Omega^0)$; and the system-size expansion of the moments (green) as $(\Omega^{-1}, \Omega^0, \Omega^0)$. The exact solution with which the error has been computed can be found in Ref. 20.

$W^{(3)} = n_r\gamma_r$; $\ell_p^{(4)} = -1$, $W^{(4)} = n_p\gamma_p$. The model considers the auto-regulation of proteins by assuming that the transcription rate is a nonlinear decreasing function of the concentration of proteins, such as $k_r(n_p) = \dfrac{k_r^{\max}}{1 + [n_p/(\Omega k_d)]^m}$, where $k_r^{\max}$, k_d, and m are parameters to be determined by fitting to experimental results. Using Eqs. (22) and (23), we find the time evolution equations of the first and second moments and cross-correlations of n_r, n_p:

$$\frac{d\langle n_r\rangle}{dt} = \Omega\langle k_r(n_p)\rangle - \left(\gamma_r + k_p\right)\langle n_r\rangle, \tag{56}$$

$$\begin{aligned}
\frac{d\langle n_p\rangle}{dt} &= k_p\langle n_r\rangle - \gamma_p\langle n_p\rangle,\\
\frac{d\langle n_r^2\rangle}{dt} &= \Omega\langle(2n_r+1)k_r(n_p)\rangle + (\gamma_r + k_p)\left(\langle n_r\rangle - 2\langle n_r^2\rangle\right),\\
\frac{d\langle n_p^2\rangle}{dt} &= k_p(\langle n_r\rangle + 2\langle n_r n_p\rangle) + \gamma_p\left(\langle n_p\rangle - 2\langle n_p^2\rangle\right),\\
\frac{d\langle n_r n_p\rangle}{dt} &= \Omega\langle n_p k_r(n_p)\rangle - (\gamma_r + \gamma_p + k_p)\langle n_r n_p\rangle\\
&\quad + k_p\left(\langle n_r^2\rangle - \langle n_r\rangle\right).
\end{aligned} \tag{57}$$

Applying the Gaussian approximation in this case would be cumbersome because of the appearance of non-trivial moments such as $\langle n_r k_r(n_p)\rangle$, given the functional form of $k_r(n_p)$. For the system-size expansion, however, we only have to construct $\mathbf{\Phi}$ and $\mathbf{G}$ using Eqs. (9) and (11)

$$\Phi_r = k_r(\phi_p) - \left(\gamma_r + k_p\right)\phi_r, \quad \Phi_p = k_p\phi_r - \gamma_p\phi_p, \tag{58}$$

$$G_{rr} = k_r(\phi_p) + \left(\gamma_r + k_p\right)\phi_r, \quad G_{pp} = k_p\phi_r + \gamma_p\phi_p, \tag{59}$$

$$G_{rp} = G_{pr} = -k_p\phi_r. \tag{60}$$

The deterministic steady-state is determined by the equation $k_r(\phi_p^{\rm st}) = \gamma_p\left(1 + \frac{\gamma_r}{k_p}\right)\phi_p^{\rm st}$ with $\phi_r^{\rm st} = \frac{\gamma_p}{k_p}\phi_p^{\rm st}$, which has to be solved numerically in general. The solution of the procedure Eqs. (27) and (30) for this model is plotted in Fig. 3. We appreciate good agreement in comparison to numerical simulations, with an increasing accuracy as Ω increases. The corrections to the average values $\langle b_r\rangle$, $\langle b_p\rangle$, coming from the system-size expansion of the moments Figs. 3(a) and 3(b), reduce the discrepancies of the bare deterministic solution of the standard van Kampen expansion.

C. The SIS epidemic model

A relevant example, where one can directly apply the methods presented in Sec. IV A to obtain analytical results, is the susceptible-infected-susceptible (SIS) epidemic model.[28,31] In this model of epidemic spreading, we have a population of size $N \equiv \Omega$ such that each member of the population can be susceptible to the disease, S, or infected, I. With a rate ε, an outbreak is created in the system and some individuals get infected. Infected individuals can spread the disease to the neighbors with which they have contact, with rate λ. Eventually, infected individuals recover and become susceptible again with rate $\beta = 1$, which is taken as unit of time for simplicity. Each individual has contact with an integer number $k \in (k_{\min}, k_{\max})$ of people to whom the disease can be transmitted. This number k, called degree, is heterogeneous within the population N and one can define the number N_k of people with k contacts, and the associated fraction $p_k = N_k/N$. For what comes next, it is useful to also define the averages $\mu_m = \sum_{k\in(k_{\min},k_{\max})} p_k k^m$, with short notation $\mu_1 = \mu$. Using the reaction description, one can portray the model as

$$S_k \xrightarrow{\varepsilon} I_k, \quad S_k + I_{k'} \xrightarrow{\lambda} I_k + I_{k'}, \quad I_k \xrightarrow{1} S_k. \tag{61}$$

We define the description variables $\mathbf{x} = (x_{k_{\min}}, \ldots, x_{k_{\max}})$ as the number of infected individuals with $k = k_{\min}, \ldots, k_{\max}$ contacts. By lumping together the first two reactions of (61), there are two main processes: either a susceptible person (with k contacts) becomes infected or it recovers, depending on its present state. In our general notation, a process is identified by the double index $\nu = (s, k)$ with $s = 1$ (infection) or $s = 2$ (recovery) and $k = k_{\min}, \ldots, k_{\max}$, with associated changes in the variables $\ell_k^{(1,k')} = \delta_{kk'}$ and $\ell_k^{(2,k')} = -\delta_{kk'}$. The total rate at which these processes happen can be calculated as

$$W^{(1,k)} = (N_k - x_k)\left(\varepsilon + \lambda k\sum_{k'}\frac{x_{k'}k'}{\mu N}\right), \quad W^{(2,k)} = x_k. \tag{62}$$

Here, $\sum_{k'}\frac{x_{k'}k'}{\mu N}$ is the global fraction of contacts connected to the infected nodes such that the total infection rate of a k-individual is λ times its number of contacts, k, times the "probability" that a contact is infected, i.e., $\sum_{k'}\frac{x_{k'}k'}{\mu N}$.

If the network of contacts is quenched, i.e., the people with whom one interacts is fixed, Eq. (62) is only an approximation[29] of the real process Eq. (61), known as

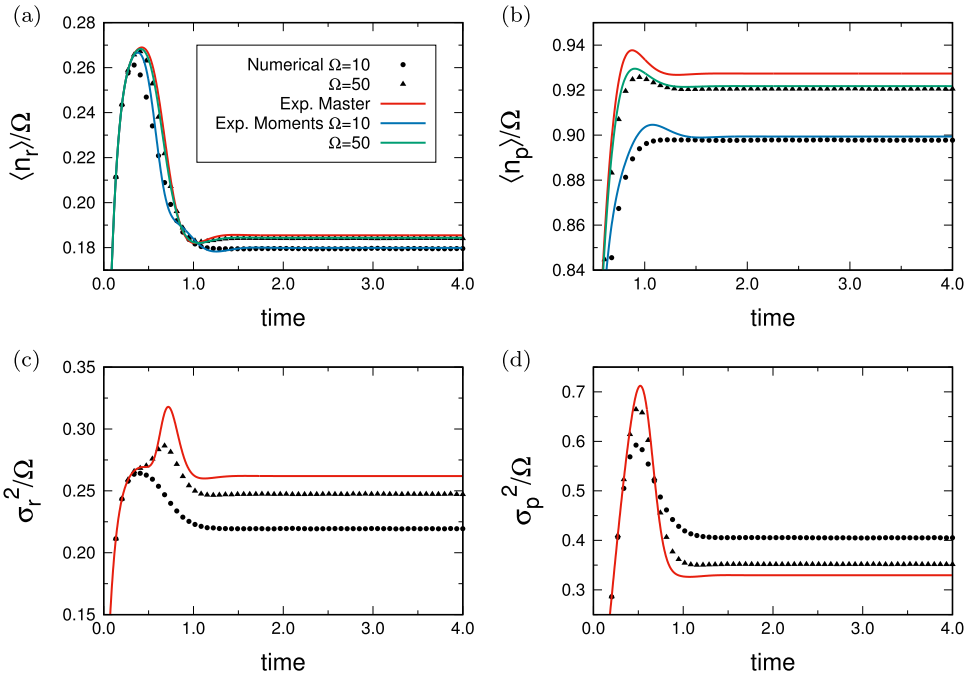

FIG. 3. Time evolution of $\langle n_r\rangle$, $\langle n_p\rangle$, σ_r^2, and σ_p^2 shown in panels (a), (b), (c), and (d) respectively, for the gene transcription model (55) with parameters $k_r^{\max}=3$, $k_d=1$, $k_p=10$, $\gamma_r=1$, $\gamma_p=2$, $m=10$. The black dots are the numerical results of the Gillespie algorithm for two different volumes $\Omega=10$ (circles) and $\Omega=50$ (triangles), averaged over 10^7 trajectories. The results of the system-size expansion ϕ_r, ϕ_p, C_{rr}, C_{pp} are plotted as red lines, while the blue ($\Omega=10$) and green ($\Omega=50$) lines include the corrections to the average value $\langle b_r\rangle/\Omega$, $\langle b_p\rangle/\Omega$.

degree-based approach or heterogeneous mean field.[30,31] The reason why this is an approximation lies already in the selection of the description variables x_k. The correlations of the (infected or recovered) state of any two people are expected to depend only on their degrees k, k' and not on the detailed structure of the network of contacts (whether these two people are in actual contact or not). If we consider instead an annealed network, i.e., at each elementary time step, the neighbors of each person are shuffled in the population, the neighborhood-dependence effect vanishes and the correlations depend only on the degree, thus making Eq. (62) exact, see Refs. 32 and 33.

Applying the regular procedure Eqs. (9) and (11), we find

$$\Phi_k = -\phi_k + (p_k - \phi_k)\,(\varepsilon + \lambda k\Theta)\,, \tag{63}$$

$$G_{kk'} = \delta_{kk'}\left[\phi_k + (p_k - \phi_k)\,(\varepsilon + \lambda k\Theta)\right], \tag{64}$$

with $\Theta \equiv \sum_k \frac{\phi_k k}{\mu}$.

The steady-state solution fulfills $\Phi_k = 0$, which leads to

$$\phi_k^{\rm st} = p_k \frac{\varepsilon + \lambda k\Theta^{\rm st}}{1+\varepsilon+\lambda k\Theta^{\rm st}}, \tag{65}$$

$$\Theta^{\rm st} = \sum_k \frac{p_k k}{\mu}\frac{\varepsilon + \lambda k\Theta^{\rm st}}{1+\varepsilon+\lambda k\Theta^{\rm st}}. \tag{66}$$

One can solve (numerically) the self-consistent equation (66) and then determine $\phi_k^{\rm st}$ by Eq. (65). In Fig. 4(a), we plot $\phi^{\rm st} \equiv \sum_k \phi_k^{\rm st}$, which corresponds to the deterministic solution of the global fraction of infected individuals. For $\varepsilon > 0$, there is a stable solution $\phi^{\rm st} > 0$ that monotonically increases with the infection rate λ. For $\varepsilon = 0$, one can identify a critical λ_c (epidemic threshold) that we will determine later, that separates the healthy phase $\phi^{\rm st} = 0$, $\lambda < \lambda_c$, from the endemic phase $\phi^{\rm st} > 0$, $\lambda > \lambda_c$. The Jacobian B-matrix, the G-matrix, and the second partial derivatives of Φ_k evaluated at this stable solution read as

$$B_{kk'}^{\rm st} = \delta_{kk'}\left(1+\varepsilon+\lambda k\Theta^{\rm st}\right) - \frac{\lambda p_k k}{1+\varepsilon+\lambda k\Theta^{\rm st}}\frac{k'}{\mu}, \tag{67}$$

$$G_{kk'}^{\rm st} = \delta_{kk'} 2p_k \frac{\varepsilon + \lambda k\Theta^{\rm st}}{1+\varepsilon+\lambda k\Theta^{\rm st}}, \tag{68}$$

$$\partial^2_{\phi_{k'},\phi_{k''}}\Phi_k = -\frac{\lambda k}{\mu}\left(\delta_{kk''}k' + \delta_{kk'}k''\right). \tag{69}$$

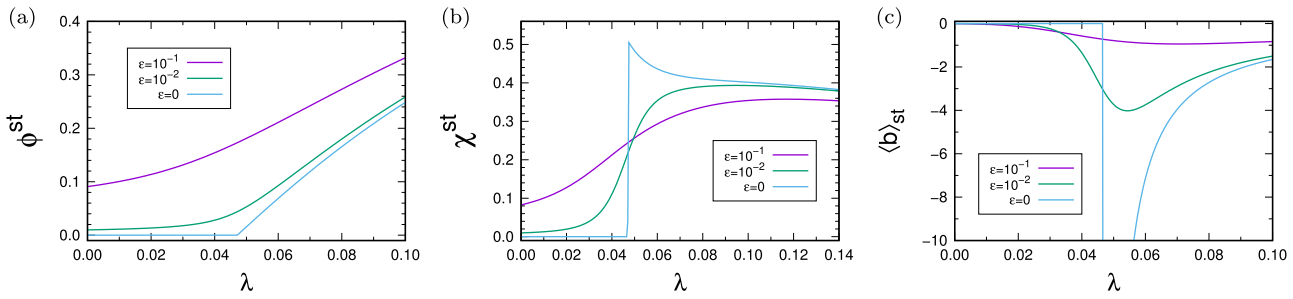

FIG. 4. For the SIS model (61), we plot $\phi^{\rm st} \equiv \sum_k \phi_k^{\rm st}$, $\chi^{\rm st} \equiv \sum_{kk'} C_{kk'}^{\rm st}$, and $\langle b\rangle_{\rm st} \equiv \sum_k \langle b_k\rangle_{\rm st}$ shown in panels (a), (b), and (c), respectively, as a function of the infection rate λ, obtained by solving Eqs. (27)–(30) taking a power-law distribution of contacts $p_k \sim k^{-2.5}$ with $k_{\min}=5$ and $k_{\max}=100$. Colors correspond to different values of $\varepsilon = 10^{-1}, 10^{-2}, 0$ as indicated in the legend.

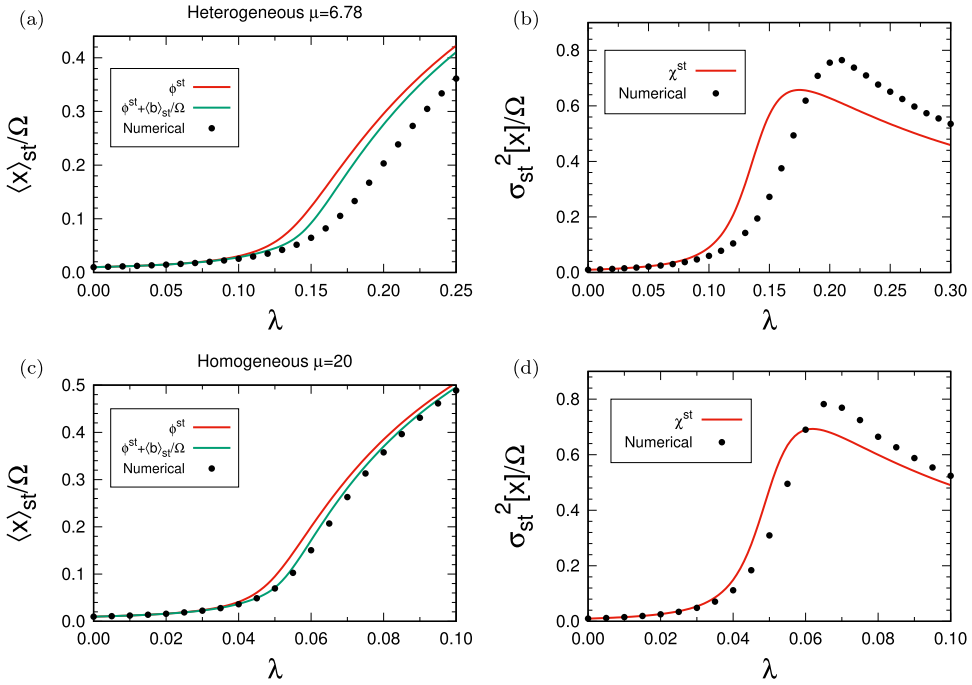

FIG. 5. For the SIS model (61), we plot the rescaled average value [panels (a) and (c)] and variance [panels (b) and (d)] of the number of infected individuals $x \equiv \sum_k x_k$ as a function of the infection rate λ. The system size is fixed $\Omega = 100$ and $\varepsilon = 10^{-2}$, while two different distribution of contacts are used: [panels (a) and (b)] a heterogeneous distribution $p_k \sim k^{-2.5}$ with $k_{\min} = 5$ and $k_{\max} = 10$ and [panels (c) and (d)] a homogeneous distribution $p_k = \delta_{k,20}$. Dots correspond to numerical simulations using the Gillespie algorithm for the SIS model on a single quenched network of contacts, while lines are the theoretical predictions obtained using Eqs. (27)–(30).

In Figs. 4(b) and 4(c), we show the results of solving Eqs. (29) and (30) with the proposed algorithm of Appendix B. The global fluctuations of the number of infected nodes is measured by $\chi^{st} \equiv \sum_{kk'} C^{st}_{kk'}$. For $\varepsilon > 0$, it increases with λ until it reaches a maximum and then decreases, whereas for $\varepsilon = 0$, we have $\chi^{st} = 0$ for $\lambda < \lambda_c$ and $\chi^{st} > 0$ for $\lambda > \lambda_c$ with a finite jump discontinuity at λ_c. Similar results are displayed for the correction to the average value $\langle b\rangle_{st} = \sum_k \langle b_k\rangle_{st}$, which is always negative $\langle b\rangle_{st} < 0$ and, in this case, it has an infinite jump discontinuity at $\varepsilon = 0$, $\lambda = \lambda_c$.

In Fig. 5, we plot the results of simulations of the SIS model Eq. (61) in two given quenched networks of contacts, constructed by taking a finite sample $\Omega = 100$ of the distribution $p_k \sim k^{-2.5}$ or $p_k = \delta_{k,20}$ and connecting individuals at random afterwards, as described in Ref. 34. The matching of the simulations with the theoretical results of the expansion methods is acceptable with some discrepancies. There are two sources of error that can explain this discrepancy, one is the approximation introduced when defining the rates Eq. (62), the degree-based approach, and the second one is the effect of finite Ω. The first source contributes the most for the low dense heterogeneous case $\mu = 6.78$, as one can appreciate how the correction $\langle b\rangle_{st}/\Omega$ in Fig. 5(a) slightly improves the result of the average value. For highly dense $\mu = 20$ homogeneous networks, however, one can see in Fig. 5(c) how the finite-size correction greatly improves the deterministic result ϕ^{st}.

1. The degenerate limit

There is an interesting limit for which one can find an analytical solution, namely, $\Theta^{st} \approx 0$ which requires $\varepsilon \ll 1$. If we expand Eq. (66) to third order in powers of Θ^{st} and explore the solution, we find

$$\Theta^{st} = \frac{\varepsilon}{1 - \frac{\lambda}{\lambda_c}}, \quad \lambda < \lambda_c, \tag{70}$$

$$\Theta^{st} = \frac{\varepsilon}{\frac{\lambda}{\lambda_c} - 1} + \frac{\mu_2^2}{\mu\mu_3}\left(\frac{\lambda}{\lambda_c} - 1\right), \quad \lambda > \lambda_c, \tag{71}$$

where $\lambda_c = \frac{\mu}{\mu_2}$. This corresponds to the lowest order in $O(\varepsilon)$ and it will be accurate only when $\Theta^{st} \approx 0$, that is to say $\varepsilon \ll \left|1 - \frac{\lambda}{\lambda_c}\right|$. In this limit, Eqs. (65), (67), and (68) simplify to

$$\phi^{st}_k = p_k\left(\varepsilon + \lambda k \Theta^{st}\right), \tag{72}$$

$$B^{st}_{kk'} = \delta_{kk'} - \lambda p_k k \frac{k'}{\mu}, \tag{73}$$

$$G^{st}_{kk'} = \delta_{kk'} 2p_k\left(\varepsilon + \lambda k \Theta^{st}\right). \tag{74}$$

Instead of directly diagonalizing the Jacobian, we will perform an equivalent calculation by solving the linearization of the dynamics $\phi_k = \phi^{st}_k + \delta_k$, $\Theta = \Theta^{st} + \theta$, that fulfill $\frac{d\delta_k}{dt} = -\sum_{k'} B^{st}_{kk'}\delta_{k'}$, this is

$$\frac{d\delta_k}{dt} = -\delta_k + \lambda p_k k \theta. \tag{75}$$

Summing $\sum_k \frac{k}{\mu}$

$$\frac{d\theta}{dt} = (-1 + \lambda/\lambda_c)\,\theta, \tag{76}$$

and then $D_1 = \left|1 - \frac{\lambda}{\lambda_c}\right|$ is an eigenvalue of the system[35] and $\theta(t) = \theta_0 e^{-D_1 t}$. Introducing this in Eq. (75), we have

$$\delta_k = C_k e^{-t} + \theta_0 \lambda_c p_k k e^{-D_1 t}, \tag{77}$$

and then the other eigenvalue is $D_2 = 1$. The scenario in this case is equivalent to what we explained in Sec. IV A, where now from Eq. (77) we have $\mathbf{v}_1^{\mathsf{T}} = (p_{k_{\min}}k_{\min}, \ldots, p_{k_{\max}}k_{\max})$ and the normal vector to the star node plane is $\mathbf{n}^{\mathsf{T}} =$

$(k_{\min},\dots,k_{\max})$, with $\mathbf{n}^{\intercal}\cdot\mathbf{v}_1=\mu_2$. Introducing $P_{k1}=p_k k$, $\sum_k kP_{kk'}=\mu_2\delta_{1k'}$, and $P^{-1}_{1k}=k/\mu_2$ in Eq. (45), we have

$$C^{\rm st}_{kk'}=p_kp_{k'}kk'\frac{F_{11}}{2D_1}+\frac{p_{k'}k'\alpha_k+p_k k\alpha_{k'}-2F_{11}p_kkp_{k'}k'}{D_1+D_2}+\frac{G^{\rm st}_{kk'}-p_{k'}k'\alpha_k-p_k k\alpha_{k'}+F_{11}p_kkp_{k'}k'}{2D_2}, \tag{78}$$

with

$$F_{11}=\sum_{kk'}\frac{k}{\mu_2}\frac{k'}{\mu_2}G^{\rm st}_{kk'}=\frac{2}{\mu_2}\left(\varepsilon+\lambda\frac{\mu_3}{\mu_2}\Theta^{\rm st}\right), \tag{79}$$

$$\alpha_k=\sum_{k'}\frac{k'}{\mu_2}G^{\rm st}_{kk'}=2\frac{p_kk}{\mu_2}\left(\varepsilon+\lambda k\Theta^{\rm st}\right). \tag{80}$$

From Eq. (78), it is straightforward to obtain a closed expression for the global $\chi^{\rm st}=\sum_{kk'}C^{\rm st}_{kk'}$ that we will not write down for reasons of brevity. For $\lambda=0$, it leads to $\chi^{\rm st}=\varepsilon$, the known variance of the Bernoulli process $S\underset{1}{\overset{\varepsilon}{\rightleftharpoons}}I$ for $\varepsilon\ll 1$. For $\varepsilon=0$ and $\lambda\to\lambda_c$, there is time-scale separation and only survives the first term of Eq. (78) leading to the value at the jump discontinuity $\chi^{\rm st}=\frac{\mu^2}{\mu_2}$.

The corrections to the average value $\langle b_k\rangle_{\rm st}$ can be calculated as described in Sec. IV. The inhomogeneous term Γ_k reads

$$\Gamma_k=\frac{1}{2}\sum_{k',k''}C^{\rm st}_{k'k''}\partial^2_{\phi_{k'},\phi_{k''}}\Phi_k=-\lambda k\sum_{k'}\frac{k'}{\mu}C^{\rm st}_{kk'}, \tag{81}$$

which can be obtained from Eq. (78). Using Eq. (46), we have

$$\langle b_k\rangle_{\rm st}=\frac{\Gamma_k}{D_2}+\frac{p_kk}{\mu_2}\frac{D_1-D_2}{D_1D_2}\sum_{k'}k'\Gamma_{k'}. \tag{82}$$

For $\lambda=0$, we have $\langle b_k\rangle_{\rm st}=0$, while for $\varepsilon=0$ and $\lambda\to\lambda_c$, the term $1/D_1$ dominates and $\langle b\rangle_{\rm st}=\sum_k\langle b_k\rangle_{\rm st}=-\frac{\mu\mu_3}{\mu_2^2}\frac{1}{\frac{\lambda}{\lambda_c}-1}$, which diverges at λ_c in accordance with the results shown in Fig. 4(c).

VI. SUMMARY AND CONCLUSIONS

In this paper, we have studied two different expansion methods of a general multidimensional master equation: (i) the traditional van Kampen's Ω expansion and (ii) the system-size expansion of the moments, and we have compared them altogether to an alternative expansion called Gaussian approximation. The traditional van Kampen expansion is based on the splitting of the description variables of the master equation as $\mathbf{x}=\Omega\boldsymbol{\phi}+\Omega^{1/2}\boldsymbol{\xi}$ and focusing on the probability distribution of $\boldsymbol{\xi}$. The Gaussian approximation, described in detail in Ref. 23, assumes, based on the results of the van Kampen expansion, that the variables $\mathbf{x}$ follow a Gaussian distribution and gives an explicit recipe for the closure of the equations of the moments. Finally, the system-size expansion that we propose here assumes that the variables can be split as $\mathbf{x}=\Omega\boldsymbol{\phi}+\Omega^{1/2}\boldsymbol{a}+\Omega^0\boldsymbol{b}+\cdots$ in terms of new stochastic variables $\boldsymbol{a},\boldsymbol{b},\dots$ and replace this expansion on the exact evolution equations for the moments and correlations $\langle x_i\rangle$, $\langle x_ix_j\rangle$, etc. We find that the system-size expansion of the moments that recovers the results of the van Kampen expansion for the correlations σ_{ij} is always one order more accurate in $\Omega^{1/2}$ for the first and second moments $\langle x_i\rangle$, $\langle x_ix_j\rangle$. At the lowest order, the errors of the van Kampen expansion scale at most as $(\Omega^0,\Omega^1,\Omega^{1/2})$ for $(\langle x_i\rangle,\ \langle x_ix_j\rangle,\ \sigma_{ij})$, while for the expansion of the moments scale as $(\Omega^{-1/2},\Omega^{1/2},\Omega^{1/2})$, in a similar way to the Gaussian approximation. Although the errors of the Gaussian approximation and the expansion of the moments scale equivalently, we show that the latter displays less systematic errors and can be applied without any singularities for small system sizes. This has been shown with numerical and theoretical results throughout the paper, with the application to some examples of relevance: the autocatalytic reaction, a model of gene transcription, and the SIS epidemic model on networks.

Additionally, we have explored different algorithms to solve the general set of equations for the correlations and first corrections to the average values, coming out of the expansion method. In some cases of interest, it is possible to obtain analytical results of the correlation matrix and finite-size corrections, that is when the dynamics show degenerate eigenvalues. The SIS model is one of these cases under some limits of the parameters, for which we have been able to obtain closed expressions of the average number of infected individuals and fluctuations.

ACKNOWLEDGMENTS

It is a pleasure to dedicate this paper to Hans Braun, on the occasion of his 70th birthday. We acknowledge financial support from Agencia Estatal de Investigación (AEI, Spain) and Fondo Europeo de Desarrollo Regional under Project ESoTECoS Grant No. FIS2015-63628-C2-2-R (AEI/FEDER/UE), and the Spanish State Research Agency, through the María de Maeztu Program for Units of Excellence in R&D (MDM-2017-0711). A.F.P. acknowledges support by the Formación de Profesorado Universitario (FPU14/00554) program of Ministerio de Educación, Cultura y Deportes (MECD) (Spain).

APPENDIX A: THE EXPANSION FOR MORE GENERAL RATES

Let us consider a more general rate formulation

$$W^{(\nu)}(\mathbf{x})=\Omega w^{(\nu)}\left(\frac{\mathbf{x}}{\Omega}\right)+\Omega^0 w_1^{(\nu)}\left(\frac{\mathbf{x}}{\Omega}\right), \tag{A1}$$

with the addition of an extra term as in the example of Sec. V A. The van Kampen expansion of the master equation is not modified, as this extra term appears as $O(\Omega^0)$ in the expansion Eq. (5), which is neglected. In the expansion of the moments Eqs. (22) and (23), however, we have the new term $\Delta_i\equiv\sum_\nu \ell_i^{(\nu)}w_1^{(\nu)}(\boldsymbol{\phi})$, $O(\Omega^0)$ of Eq. (22) and, $\phi_j\Delta_i+\phi_i\Delta_j$, $O(\Omega^1)$ of Eq. (23). This modification leads to the same set of Eqs. (28) and (30) but different Eq. (29)

$$\frac{d\langle b_i\rangle}{dt}=-\sum_j B_{ij}\langle b_j\rangle+\frac{1}{2}\sum_{k,q}\langle a_ka_q\rangle\partial^2_{\phi_k,\phi_q}\Phi_i+\Delta_i. \tag{A2}$$

APPENDIX B: NUMERICAL SOLUTION OF EQ. (17)

We describe an implicit Euler iterative method that converges very fast to the exact steady-state solution of Eq. (17). Starting from some initial value $\mathbf{C}_0$ (we do not use the subindex "st" for brevity in the notation), we propose the implicit recursion

$$\mathbf{C}_{n+1} = \mathbf{C}_n - \Delta t\,(\mathbf{C}_{n+1}\mathbf{B}^\intercal + \mathbf{B}\mathbf{C}_{n+1} - \mathbf{G})\,. \tag{B1}$$

Adding $\mathbf{B}\mathbf{C}_{n+1}\mathbf{B}^\intercal(\Delta t)^2$ on the right-hand side of Eq. (B1) and $\mathbf{B}\mathbf{C}_n\mathbf{B}^\intercal(\Delta t)^2$ on the left-hand side, one can isolate $\mathbf{C}_{n+1}$:

$$\mathbf{C}_{n+1} = (1 + \Delta t\mathbf{B})^{-1}\,(\mathbf{C}_n + \Delta t\mathbf{G} + (\Delta t)^2\mathbf{B}\mathbf{C}_n\mathbf{B}^\intercal)\,(1 + \Delta t\mathbf{B}^\intercal)^{-1}. \tag{B2}$$

This method has the advantage of being very stable compared to other integration methods; in this way, we can set Δt large enough so that the method converges in few iterations. This is very useful for systems with a very large number of variables. Computationally, one may prefer the iterative method Eq. (B2) than Eq. (41) to prevent divisions by zero when the matrix $\mathbf{B}$ is singular and other complications. For example, when there are conserved quantities (Sec. IV C), if the initial value $\mathbf{C}_0$ respects the given set of constraints between variables, we expect, by the own structure of the algorithm, that $\mathbf{C}_n$ will also respect them, avoiding in this way the problem discussed in Eq. (49).

[1] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 4th ed. (Springer-Verlag, Berlin, 2009).
[2] T. M. Earnest, J. A. Cole, and Z. Luthey-Schulten, "Simulating biological processes: Stochastic physics from whole cells to colonies," Rep. Prog. Phys. **81**, 052601 (2018).
[3] B. P. English, W. Min, A. M. Van Oijen, T. L. Kang, G. Luo, H. Sun, B. J. Cherayil, S. C. Kou, and X. S. Xie, "Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited," Nat. Chem. Biol. **2**, 87–94 (2006).
[4] E. M. Izhikevich, "Which model to use for cortical spiking neurons?," IEEE Trans. Neural Netw. **15**, 1063–1070 (2004).
[5] M. Benayoun, J. D. Cowan, W. van Drongelen, and E. Wallace, "Avalanches in a stochastic model of spiking neurons," PLoS Comput. Biol. **6**, e1000846 (2010).
[6] H. A. Braun, M. T. Huber, K. S. Afer, and K. Voigt, "Computer simulations of neuronal signal transduction: The role of nonlinear dynamics and noise," Int. J. Bifurcat. Chaos **8**, 881–889 (1998).
[7] U. Feudel, A. Neiman, X. Pei, W. Wojtenek, H. Braun, M. Huber, and F. Moss, "Homoclinic bifurcation in a Hodgkin-Huxley model of thermally sensitive neurons," Chaos **10**, 231–239 (2000).
[8] M. Patriarca, S. Postnova, H. Braun, E. Hernández-García, and R. Toral, "Diversity and noise effects in a model of homeostatic regulation of the sleep-wake cycle," PLoS Comput. Biol. **8**, e1002650 (2012).
[9] M. Patriarca, E. Hernández-García, and R. Toral, "Constructive effects of diversity in a multi-neuron model of the homeostatic regulation of the sleep-wake cycle," Chaos Solitons Fractals **81**, 567–574 (2015).
[10] S. Postnova, K. Voigt, and H. A. Braun, "A mathematical model of homeostatic regulation of sleep-wake cycles by Hypocretin/Orexin," J. Biol. Rhythms **24**, 523–535 (2009).
[11] J. Paulsson, "Summing up the noise in gene networks," Nature **427**, 415–418 (2004).
[12] M. E. Gracheva, R. Toral, and J. D. Gunton, "Stochastic effects in intercellular calcium spiking in hepatocytes," J. Theor. Biol. **212**, 111–25 (2001).
[13] M. Thattai and A. van Oudenaarden, "Intrinsic noise in gene regulatory networks," Proc. Natl. Acad. Sci. **98**, 8614–8619 (2001).
[14] M. Kærn, T. C. Elston, W. J. Blake, and J. J. Collins, "Stochasticity in gene expression: From theories to phenotypes," Nat. Rev. Genet. **6**, 451–464 (2005).
[15] D. Fraser and M. Kærn, "A chance at survival: Gene expression noise and phenotypic diversification strategies," Mol. Microbiol. **71**, 1333–1340 (2009).
[16] P. C. Bressloff, "Stochastic switching in biology: From genotype to phenotype," J. Phys. A: Math. Theor. **50**, 133001 (2017).
[17] R. H. Wang, Z. Jin, Q. X. Liu, J. van de Koppel, and D. Alonso, "A simple stochastic model with environmental transmission explains multi-year periodicity in outbreaks of avian flu," PLoS ONE **7**, e28873 (2012).
[18] L. S. Tsimring, "Noise in biology," Rep. Prog. Phys. **77**, 026601 (2014).
[19] R. Toral and P. Colet, *Stochastic Numerical Methods: An Introduction for Students and Scientists* (Wiley, New York, 2014).
[20] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (North-Holland, Amsterdam, 2007).
[21] E. W. J. Wallace, "A simplified derivation of the linear noise approximation," e-print arXiV:1004.4280v4 (2012).
[22] E. W. J. Wallace, L. R. Petzold, D. T. Gillespie, and K. R. Sanft, "Linear noise approximation is valid over limited times for any chemical system that is sufficiently large," IET Syst. Biol. **6**, 102–115 (2012).
[23] L. F. Lafuerza and R. Toral, "On the Gaussian approximation for master equations," J. Stat. Phys. **140**, 917–933 (2010).
[24] R. Grima, "A study of the accuracy of moment-closure approximations for stochastic chemical kinetics," J. Chem. Phys. **136**, 154105 (2012).
[25] J. Elf and M. Ehrenberg, "Fast evaluation of fluctuations in biochemical networks with the linear noise approximation," Genome Res. **13**, 2475–2484 (2003).
[26] L. Isserlis, "On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables," Biometrika **12**, 134 (1918).
[27] D. J. Amit and V. M. Mayor, *Field Theory, the Renormalization Group and Critical Phenomena*, 3rd ed. (World Scientific Publishing Co. Pte. Ltd., Singapore, 2005).
[28] E. Vynnycky and R. G. White, *An Introduction to Infectious Disease Modelling* (Oxford University Press, New York, 2010).
[29] 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).
[30] A. Vespignani, "Modelling dynamical processes in complex socio-technical systems," Nat. Phys. **8**, 32–39 (2012).
[31] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," Rev. Mod. Phys. **87**, 925–979 (2015).
[32] A. Carro, R. Toral, and M. San Miguel, "The noisy voter model on complex networks," Sci. Rep. **6**, 24775 (2016).
[33] 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).
[34] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, "Generation of uncorrelated random scale-free networks," Phys. Rev. E **71**, 027103 (2005).

[35] We can determine a better estimation of D_1 if, instead of using the first order Jacobian Eq. (73), we take the full Eq. (67) to find $\frac{d\theta}{dt} = \left(-1 - \varepsilon + \frac{\lambda}{\mu}\sum_k \frac{p_k k^2}{1+\varepsilon+\lambda k\Theta^{\rm st}}\right)\theta - \lambda\Theta^{\rm st}\sum_k \frac{k^2}{\mu}\delta_k$, which is not closed for θ. If we, however, assume $\delta_k \propto p_k k\lambda_c\theta$ in accordance with Eq. (77), we get $D_1 = 1 + \varepsilon - \frac{\lambda}{\mu}\sum_k \frac{p_k k^2}{1+\varepsilon+\lambda k\Theta^{\rm st}} + \lambda_c\lambda\frac{\mu_3}{\mu}\Theta^{\rm st}$, and expanding for ε small, we have $D_1 = 1 - \lambda/\lambda_c + \varepsilon\,(1 + \lambda/\lambda_c) + \lambda\frac{\mu_3}{\mu}\Theta^{\rm st}\,(\lambda + \lambda_c)$.