Stochastic population dynamics: The Poisson approximation

Hernán G. Solari*
Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón I, Ciudad Universitaria, 1428 Buenos Aires, Argentina

Mario A. Natiello†
Centre for Mathematical Sciences, Lund University, Box 118 S-221 00 Lund, Sweden

(Received 4 December 2002; published 26 March 2003)

We introduce an approximation to stochastic population dynamics based on almost independent Poisson processes whose parameters obey a set of coupled ordinary differential equations. The approximation applies to systems that evolve in terms of *events* such as death, birth, contagion, emission, absorption, etc., and we assume that the event-rates satisfy a generalized mass-action law. The dynamics of the populations is then the result of the projection from the space of events into the space of populations that determine the state of the system (phase space). The properties of the Poisson approximation are studied in detail. Especially, error bounds for the moment generating function and the generating function receive particular attention. The deterministic approximation for the population fractions and the Langevin-type approximation for the fluctuations around the mean value are recovered within the framework of the Poisson approximation as particular limit cases. However, the proposed framework allows to treat other limit cases and general situations with small populations that lie outside the scope of the standard approaches. The Poisson approximation can be viewed as a general (numerical) integration scheme for this family of problems in population dynamics.

DOI: 10.1103/PhysRevE.67.031918 PACS number(s): 87.23.Cc, 02.50.Ey, 02.50.Ga

I. INTRODUCTION

In this paper, we consider the time evolution of interacting populations, i.e., the time evolution of systems which are described by non-negative integers counting the members of the different species acting in the problem. Molecules [1–5], photons [6–8], predators, preys, infected individuals, etc., can be regarded as populations under a diversity of situations [9]. The different nature of the systems will be apparent by the characteristic interactions of each particular problem in consideration. Henceforth, we will consider a rather general class of interactions.

The time evolution of discrete populations is described with jumps in the population values that occur ''instantaneously'' (meaning that the time taken to dissociate a molecule, emit a photon, hatch an egg, cut the umbilical cord, etc., must be considerably shorter than the time between observations, and essentially unimportant for the purposes of the analysis). We shall refer to each of these jumps as an *event*.

We will consider that this time evolution responds to a Markov process, i.e., the probability of occurrence of any event in an infinitesimal time interval $(t,t+dt)$ will only depend on the number of individuals at the time t and on parameters that might depend on t. We shall further consider that the time between events is exponentially distributed with a characteristic frequency (transition rate) that only depends on the state of the system [16].

This setting has been successfully applied to a large and diverse class of systems and is usually modeled with Monte Carlo simulations (the name comes from solving an inverse problem: the approximation of solutions to differential equations by stochastic jumps obtained by the Monte Carlo method). Monte Carlo simulations are simple realizations of the Markov process. This approach is sometimes called *Feller process* [10] in the mathematical literature.

Considering the trasition rates, a widely used assumption which is reasonable for a large number of natural process is the ''mass-action'' law which formalizes the following intuitive idea: if a system is made twice as large duplicating each population as well as the environment, then the number of interactions per unit time will be roughly twice the original figure.

When the mass-action law holds and in addition the populations are large, the fractions of the total population represented by each species are the relevant variables. A Markov jump process complying with the mass-action law can be approximated by the combination of a deterministic differential equation and a stochastic correction describing the departure from the deterministic law in the form of a Langevin equation (Brownian process) [10–13]. The approximation is sound provided that the description is not applied to relatively small-time intervals hosting too few events.

Large-system limits where the description can be performed with a deterministic law or Langevin equations rely on two requirements: first, that we are interested in fractions of the total population and/or fluctuations of these fractions rather than in the actual population numbers and second, that the total populations are as large as needed (infinite population size) to make the approximation valid. In terms of *natural sciences* this limit needs to be reinterpreted. Normally, natural scientists are not allowed to change their problem, i.e., we cannot change city if our model is not good enough for the small city where the study of, for example, an epi-

*Electronic address: solari@df.uba.ar
†Electronic address: Mario.Natiello@math.lth.se

demic outbreak is required. In this work, we shall deal with this difficulty indicating the order of magnitude of the errors introduced in our approximations.

As a general rule, whenever we introduce an approximated evolution law the predictions deteriorate with time. The approximations presented in this work will not be the exception, they will "improve" with our tolerance and deteriorate with time. The different limits will then represent different balances between time of evolution and precision of the description.

Between the event-by-event realizations of the Monte Carlo method and the "as many as needed" events per unit time (large population limit) of the Langevin method there is room and need for a description that applies to large and small populations indistinctly. Such possibility has been demonstrated recently [14] introducing a naive Poisson approximation. As pointed out in Ref. [14], this naive approximation requires several improvements if it is going to achieve a fully respectable status. Mainly, it should match the Langevin approximation in the large population limits and the *ad hoc* procedure at the boundary of the population space (when one or more populations are zero) should receive careful consideration.

The aim of this paper is to present a Poisson approximation in which the parameters of the (truncated) Poisson distributions obey a set of ordinary differential equations (ODE). We intend to show that such approximation will provide an adequate tool to handle systems with large and small populations indistinctly, being especially suitable when the population is not as large as required for a deterministic description.

The basic description will be performed in the space of events which can be projected down onto the phase space (the population numbers) of the problem. We will show that in the large population limit the Poisson approximation converges to the solutions of the Langevin problem but furthermore, it converges to the Monte Carlo process when the mean number of events per unit time is kept fixed in the limit.

The rest of the work is organized as follows: In Sec. II, we formulate the problem in terms of its probabilities and the class of events involved. In Sec. III, we present the Poisson approximation. Section IV contains the core results of this work since the quality of the approximations is asserted under different potential uses. Section IV B is rather technical as well as parts of Sec. III and it might be skipped in a first reading. Section V works out a simple example where just one class of events is involved. Concluding remarks are left for Sec. VI.

II. STATEMENT OF THE PROBLEM AND BASIC PROPERTIES

A. Phase space and space of events

We shall consider a stochastic process in which the state of the system is described by a vector of integer variables (populations), X_i, $i=1,\ldots,N$.

The evolution of the system is described by the occurrence of E (classes of) events. The event j produces a change $X \to X+\delta_j$ in the state of the system, while $W_j(X)dt$ is the probability of occurrence of the event j in an infinitesimal time interval $[t,t+dt]$ given that the state of the system at the time t is X. We shall often refer to $W_j(X)$ as the "transition probability." Without loss of generality we can consider that $\delta_i \neq \delta_j$ for $i \neq j$.

We shall further assume that the formulation is consistent, meaning that the corresponding transition probability for an event that would eventually produce a meaningless (nonphysical, nonbiological, ...) state is zero. Also we define $W_j(X)=0$, $j=1,\ldots,E$ if X is a meaningless state (negative populations are the most noticeable meaningless states).

We shall consider that at time $t=0$ the system is in the state $X(0)=X_0$. Let $P(n_1,\ldots,n_E;t/X_0)$ be the probability that at time t, exactly n_j, $j=1,\ldots,E$ events of each different kind have occurred, given that the initial state (at time $t=0$) was X_0. Let ζ_i, $i=1,\ldots,E$ be non-negative integer random variables distributed with probabilities $P(n_1,\ldots,n_E;t/X_0)$. Then,

$$X=X_0+\sum_{j=1}^{E}\delta_j\zeta_j, \tag{1}$$

meaning that the random variable X represents the state of the system at time t and its probability distribution is the one resulting from the effects of the different possible events.

Note that the decomposition in terms of events carries more information than the transition probabilities from the state X_0 to the state X since, in principle, several combinations of events can produce the same final state. Such combinations are distinguished in the present formulation (think, for example, of a birth-death process, the k-death k-born events all lead to a transition from X_0 to X_0 and our probability distribution keeps track of the contributions of each situation).

Our interest is then to produce suitable approximations to $P(n_1,\ldots,n_E;t/X_0)$ that can be used in a direct analysis of the problem in question or efficient numerical realizations of the process under consideration.

B. Time evolution of the probabilities and the generating function

The probabilities $P(n_1,\ldots,n_E;t/X_0)$ satisfy the forward Kolmogorov equation (sometimes referred as Master equation in the natural sciences literature)

$$\frac{d}{dt}P(n_1,\ldots,n_E;t/X_0)$$
$$=\sum_{j=1}^{E}[W_j(X-\delta_j)P(n_1,\ldots,n_j-1,\ldots,n_E;t/X_0)]$$
$$-\left(\sum_{j=1}^{E}W_j(X)\right)P(n_1,\ldots,n_E;t/X_0), \tag{2}$$

where $X=X_0+\Sigma_{j=1}^{E}\delta_j n_j$ and $P(n_1,\ldots,n_E;0/X_0)$

$=\Pi_{j=1}^{E}\delta_{0n_j}$ (δ_{ij} is the Kronecker delta). We notice that $P(n_1,\ldots,n_E;t/X_0)=0$, whenever one or more n_j is negative.

The generating function associated to the probabilities $P(n_1,\ldots,n_E;t/X_0)$ is

$$\Phi(z_1,\ldots,z_E;X_0)=\sum_{\{n_j\}}\left(\prod_{j=1}^{E}z_j^{n_j}\right)P(n_1,\ldots,n_E;t/X_0). \tag{3}$$

$\Phi(z_1,\ldots,z_E;X_0)$ can, as a function of z, be regarded as an analytic function defined via a non-negative series on the unit cube R, namely, $R=|z_j|\in[0,1]$, $j=1,\ldots,E$ (actually, we may regard z_j as real and non-negative variables, except for Sec. IV E). Moreover, for all t values, $\Phi(z_1=1,\ldots,z_E=1;X_0)=1$ and its z derivatives can be computed by term-by-term derivation with the possible exception of the border of R.

Proposition. Let $\Phi:R\mapsto\mathbf{R}$ defined as above. Then, Φ is non-negative, has non-negative derivatives with respect to every z_i and $d\Phi/dt$ is nonpositive.

Proof. The only statement which is not obvious is the last one. First, we note that the fact that the time coefficients in $\Phi(z,t;X_0)$ are probabilities (therefore, non-negative and summing one), guarantees that $\Phi(z,t;X_0)$ is uniformly convergent for all $t\geqslant 0$ and all $z\in[0,1]$. Hence, the t derivatives of Φ can be computed term by term.

Hence, we have

$$\frac{d}{dt}\Phi(z,t;X_0)=\sum_{\{n_j\};X_0}\left(\prod_{j=1}^{E}z_j^{n_j}\right)\frac{d}{dt}P(n_1,\ldots,n_E;t/X_0) \tag{4}$$

$$=\sum_{\{n_j\}}\left(\prod_{j=1}^{E}z_j^{n_j}\right)\left(\sum_{j=1}^{E}W_j(X-\delta_j)\times P(n_1,\ldots,n_j-1,\ldots,n_E;t/X_0)-\sum_{j=1}^{E}W_j(X)P(n_1,\ldots,n_E;t/X_0)\right) \tag{5}$$

(where $X=X_0+\Sigma_{j=1}^{E}\delta_j n_j$), changing the index of sum in the first term from $n_j-1\rightarrow n_j$, we obtain

$$\frac{d}{dt}\Phi(z,t;X_0)=\sum_{j=1}^{E}(z_j-1)G_j(z_1,,\ldots,z_E)$$
$$\equiv\mathcal{L}(X_0)\Phi(z,t;X_0), \tag{6}$$

where

$$G_j(z_1,,\ldots,z_E)=\sum_{\{n_j\}}\left(\prod_{j=1}^{E}z_j^{n_j}\right)W_j(X)P(n_1,\ldots,n_E;t/X_0)$$
$$=\mathcal{L}_j(X_0)\Phi(z,t;X_0), \tag{7}$$

$\mathcal{L}_j(X_0)$ is an operator that acts by multiplying each term in the analytical expansion of $\Phi(z,t;X_0)$ by the transition rate associated to the j event in the state $X=X_0+\Sigma_k\delta_k n_k$.

We have implicitly used that the transition probabilities are zero when they imply a transition to a meaningless state and that the probabilities of the meaningless states are identically zero, hence allowing to extend the sums formally up to infinity in all cases. Since $G_i(z_1,\ldots,z_n)\geqslant 0$ and $(z_i-1)\leqslant 0$, we have $(z_i-1)G_i(z_1,\ldots,z_n)\leqslant 0$ and the result follows immediately. ∎

The previous results shows that the time derivative of the generating function can be written as the action of a linear operator $\mathcal{L}$ on Φ. Moreover, $\mathcal{L}(X_0)$ in Eq. (6) can be thought of as applying not only to generating functions but to any analytic function of $z\in R$. The evolution operator can now formally be written as $\exp(\mathcal{L}t)$. In particular, it transforms analytic functions into analytic functions.

Furthermore, when applied to any monomial, the equality

$$\left|\exp(\mathcal{L}t)\prod_{j=1}^{E}z_j^{n_j}\right|=\left|\prod_{j=1}^{E}z_j^{n_j}\right| \tag{8}$$

can be easily verified (it follows immediately from the fact that $\mathcal{L}$ maps generating functions into generating functions).

For every analytic function $\phi(z_1^{n_1},\ldots,z_E^{n_E})=\Sigma_{n_1,\ldots,n_E}a_{n_1,\ldots,n_E}z_1^{n_1},\ldots,z_E^{n_E}$ consider the norm $||\phi||=\Sigma_{n_1,\ldots,n_E}|a_{n_1,\ldots,n_E}|$.

Proposition. The evolution operator $\exp[\mathcal{L}(X)t]$ is bounded by the identity, i.e.,

$$||\exp[\mathcal{L}(X)t]\phi||\leqslant||\phi||. \tag{9}$$

Proof. By definition $\Phi(z,t;X_0)=\exp(\mathcal{L}t)\Phi(z,0;X_0)$ for every generating function $\Phi(z,0;X_0)$. Since the sum of the probabilities of all the states is 1, it is clear that the statement of the theorem is true (the equal sign holds) when $\phi(z,0)$ is a generating function. Moreover, $\exp(\mathcal{L}t)$ is a linear operator and the generating functions $z_1^{n_1},\ldots,z_E^{n_E}$ form a basis of the space of entire functions, hence

$$||\exp[\mathcal{L}(X)t]\sum_{n_1,\ldots,n_E}z_1^{n_1},\ldots,z_E^{n_E}a_{n_1,\ldots,n_E}||$$
$$=||\sum_{n_1,\ldots,n_E}a_{n_1,\ldots,n_E}\exp[\mathcal{L}(X)t](z_1^{n_1},\ldots,z_E^{n_E})||$$
$$\leqslant\sum_{n_1,\ldots,n_E}|a_{n_1,\ldots,n_E}|. \quad ∎ \tag{10}$$

Notice that $\Phi(z,t;X_0)$ satisfies (6) with initial condition $\Phi(z,0;X_0)=1$ according to Eq. (3). We will allow in this paragraph more general initial conditions associated to Eq. (6). Let $\phi(z,t;X_0)$ be a solution of Eq. (6) with initial condition $\phi(z,0;X_0)=z_j^m$ representing a system with probability 1 at $t=0$ of being in the state $X=X_0+m\delta_j$, we have that $\phi(z,t;X_0)=z_j^m\Phi(z,t;X)$, since $\Phi(z,t;X)$ represents the generating function with initial condition at the reference state X. Hence, the following result holds:

$$\frac{d}{dt}[z_j\Phi(z,t;X_0+\delta_j)]=z_j\frac{d}{dt}\Phi(z,t;X_0+\delta_j)$$
$$=z_j\mathcal{L}(X_0+\delta_j)\Phi(z,t;X_0+\delta_j)$$
$$=\mathcal{L}(X_0)[z_j\Phi(z,t;X_0+\delta_j)], \quad (11)$$

which can be independently verified starting from the definition of $\mathcal{L}(X)$ in Eqs. (6) and (7).

C. The border between admissible and meaningless states

We will further characterize the states in the border of the region of admissible states.

Let $\mathcal{B}_j(X_0)$ be the set of events such that

$$X=X_0+\sum_{i=1,\ldots,E}\delta_i n_i \quad (12)$$

is an admissible state, but

$$Y=X_0+\sum_{i=1,\ldots,E}\delta_i n_i+\delta_j \quad (13)$$

is not. In such cases, we shall say that $\{n_i\}i=1,\ldots,E$ belongs to the jth component of the border $B_j(X_0)$ of the admissible region (note that the different components of the border are not necessarily disjoint).

After this preparatory section, we are ready to introduce the Poisson approximation in the following section.

III. APPROXIMATION BY (ALMOST) INDEPENDENT POISSON PROCESSES

A. The Poisson approximation

We will now attempt to approximate the probabilities $P(n_1,\ldots,n_E;t/X_0)$ by a product of probabilities representing independent Poisson processes with parameters $\lambda_i(t)$, $i=1,\ldots,E$ which satisfy a differential equation that we shall prescribe.

The main aim of this paper is to understand the quality (size of errors and convergence properties) of this approximate model.

Consider the event j, we let $P^j_{n_1,\ldots,n_E}(\lambda_j)$ be

$$P^j_{n_1,\ldots,n_E}(\lambda_j)=\exp(-\lambda_j)\frac{\lambda_j^{n_j}}{n_j!}, \quad (14)$$

whenever $\{n_i\}\notin B_j(X_0)$ and

$$P^j_{n_1,\ldots,n_E}(\lambda_j)=\exp(-\lambda_j)\sum_{i=n_j}^{\infty}\frac{\lambda_j^i}{i!}=1-\exp(-\lambda_j)\sum_{i=0}^{n_j-1}\frac{\lambda_j^i}{i!}, \quad (15)$$

if $\{n_i\}\in B_j$.

Notice that for $n_j=0$,

$$\frac{d}{dt}P^j_{n_1,\ldots,n_E}(\lambda_j)=-\frac{d\lambda_j}{dt}P^j_{n_1,\ldots,n_E}(\lambda_j), \quad (16)$$

while for if $\{n_i\}\notin B_j(X_0)$ and $n_j\neq 0$ (the general case)

$$\frac{d}{dt}P^j_{n_1,\ldots,n_E}(\lambda_j)=\frac{d\lambda_j}{dt}[P^j_{n_1,\ldots,n_j-1,\ldots,n_E}(\lambda_j)-P^j_{n_1,\ldots,n_E}(\lambda_j)], \quad (17)$$

and finally,

$$\frac{d}{dt}P^j_{n_1,\ldots,n_E}(\lambda_j)=\frac{d\lambda_j}{dt}P^j_{n_1,\ldots,n_j-1,\ldots,n_E}(\lambda_j), \quad (18)$$

if the state belongs to the boundary set B_j. Further notice that the expressions (16), (17), and (18) resemble the contributions of the jth event to the change of the probabilities (2) for boundary and non boundary states.

We propose the following expression as an approximation:

$$P(n_1,\ldots,n_E;t/X_0)\sim\left(\prod_{j=1}^{E}P^j_{n_1,\ldots,n_E}\right)=\overline{P}(n_1,\ldots,n_E;t/X_0), \quad (19)$$

where the expression is valid for all $\{n_j\}$ such that $X=X_0+\Sigma_j\delta_j n_j$ is an admissible state and the coefficient functions $\lambda_j(t)$ are still to be determined.

Once again, the probabilities of inadmissible states are zero and the evolution of Eq. (19) is completely decoupled from the probabilities of the inadmissible states as can be seen from the expression (18). We could formally extend the expression to all possible values of $\{n\}$ without introducing errors.

The generating function associated with $\overline{P}$ reads

$$\Psi(z_1,\ldots,z_E,t;X_0)=\sum\left(\prod_{j=1}^{E}z_j^{n_j}\right)\overline{P}(n_1,\ldots,n_E;t/X_0), \quad (20)$$

where the summation is extended to all the admissible states and can be formally extended to all combinations of non-negative integer numbers $\{n_j\}$.

Proposition. $\Psi(0,t;X_0)=1$.

Proof. Notice first that this is equivalent to say that the probability of being in any admissible state is 1. The proposition is proved realizing that it is true for $t=0$ and that Eqs. (16), (17), and (18) assure that $d\Psi(0,t;X_0)/dt=0$.

To completely specify the proposed approximation, we have to establish the dependence of the Poisson parameters $\{\lambda_j\}$ with respect to time. We propose that the parameters satisfy the following ODE with initial condition $\lambda_j(0)=0$:

$$\mathcal{N}_j(t,X_0)\frac{d\lambda_j}{dt}=\sum_{\{n\}\notin B_j}W_j(X)\overline{P}(n_1,\ldots,n_E,t;X_0)\equiv f_j(\lambda)\mathcal{N}_j(t,X_0), \quad (21)$$

where the functions $f_j(\lambda)$ are defined by the right-hand side of the equations, $\mathcal{N}_j(t,X_0)=\Sigma_{\{n\}\notin B_j}\overline{P}(n_1,\ldots,n_E,t;X_0)$ denotes the sum of approximated probabilities *excluding* the border of the admissible region and $X = X_0 + \Sigma_j(\delta_j n_j)$, this choice will render the errors of the approximation more manageable, as we will see in Eq. (51).

Having introduced the Poisson approximation the remaining task ahead is to evaluated the quality of the approximation. We shall address this central issue in the following section.

B. Evaluation of the Poisson approximation

We shall proceed in two steps. In the first step, we will find a suitable expression for the difference between the exact generating function and the approximated one; while in the second step, we will find some bounds and limiting behavior (“order of” relations) for the difference between generating functions.

We begin by writing a formal solution to Eq. (6) in terms of our guess $\overline{P}(t)$ and the correction $\Delta(z,t;X_0)$ in the form

$$\Phi(z,t;X_0) = \Psi(z,t;X_0) + \Delta(z,t;X_0), \tag{22}$$

where

$$\begin{aligned}\Delta(z,t;X_0) &= -\int_0^t \frac{d\{\exp[\mathcal{L}(X_0)(t-s)]\Psi(z,s;X_0)\}}{ds} ds \\ &= \int_0^t \exp[\mathcal{L}(X_0)(t-s)]\left(\mathcal{L}(X_0) - \frac{d}{ds}\right) \\ &\quad \times \Psi(z,s;X_0)ds,\end{aligned} \tag{23}$$

which holds for any $\Psi(z,t;X_0)$ provided that $\Phi(z,0;X_0) = \Psi(z,0;X_0)$.

The formal expression involves the evaluation of the exponential of an operator (which might be a formidable task). However, as soon as we realize that the exponential propagates in time an initial condition in the state $X_0 + \Sigma_{j=1,\ldots,E}\delta_j n_j$, we understand that it can be written in terms of solutions of Eq. (2) with the appropriate value for the initial state X.

We first notice that according to Eqs. (16), (17), and (18), we have

$$\begin{aligned}\frac{d}{dt}\Psi(z,t;X_0) &= \sum_{j=1,\ldots,E} f_j\Bigg(\sum_{\{n\}\notin B_j, n_j\neq 0} z^n[\overline{P}(n_1,n_j \\ &\quad -1,\ldots,n_E,t;X_0) - \overline{P}(n_1,\ldots,n_E,t;X_0)] \\ &\quad - \sum_{\{n\}\notin B_j, n_j=0} z^n \overline{P}(n_1,\ldots,n_E,t;X_0) \\ &\quad + \sum_{\{n\}\in B_j} z^n[\overline{P}(n_1,n_j-1,\ldots,n_E,t;X_0)]\Bigg) \\ &= \sum_{j=1,\ldots,E}(z_j-1)f_j \\ &\quad \times\left(\sum_{\{n\}\notin B_j} z^n\overline{P}(n_1,\ldots,n_E,t;X_0)\right),\end{aligned} \tag{24}$$

where $z^n = \Pi_j z_j^{n_j}$. The expression for the error $\Delta(z,t;X_0)$ can be rewritten using Eqs. (6), (11), and (24) as follows.

First, we compute the s time derivative with Eq. (24) and the action of $\mathcal{L}(X_0)$ on Ψ with Eq. (6). We have

$$\begin{aligned}\left(\mathcal{L}(X_0) - \frac{d}{ds}\right)\Psi(z,s;X_0) &= \sum_{j=1,\ldots,E}(z_j-1)\sum_{\{n\}\notin B_j}[W_j(X) \\ &\quad - f_j]z^n\overline{P}(n_1,\ldots,n_E,s;X_0).\end{aligned} \tag{25}$$

Second, the exponential operator $\exp[\mathcal{L}(X_0)(t-s)]$ acting on powers of z can be recasted as the time evolution of shifted reference states $X = X_0 + \Sigma_{j=1}^E \delta_j n_j$ as in Eq. (11):

$$\begin{aligned}\exp[\mathcal{L}(X_0)(t-s)]z_j z^n &= z_j z^n \exp[\mathcal{L}(X+\delta_j)(t-s)1] \\ &= z^n z_j \exp[\mathcal{L}(X+\delta_j)(t-s)] \\ &\quad \times \Phi(z,0,X+\delta_j) \\ &= z^n z_j \Phi(z,t-s,X+\delta_j).\end{aligned}$$

Finally, we rearrange Eq. (23) as follows:

$$\begin{aligned}\Delta(z,t;X_0) &= \int_0^t\Bigg(\sum_{j=1,\ldots,E}\sum_{\{n\}\notin B_j} z^n\overline{P}(n,s;X_0)[W_j(X)-f_j] \\ &\quad \times[z_j\Phi(z,t-s;X+\delta_j) - \Phi(z,t-s;X)]\Bigg)ds,\end{aligned} \tag{26}$$

where $X = X_0 + \Sigma_{j=1}^E \delta_j n_j$ as usual.

We can now proceed to find upper bounds for the expression (26).

IV. RESULTS

In this section, we will present our main results. Some of them will make contact with previous results: the deterministic approximation to density dependent jump Markov processes described in Ref. [11], see also Refs. [13] Chap. 5 and [10] Chap. 11 (also known as the *law of large numbers*) and the diffusion or Langevin approximation directly connected to the *central limit theorem* expressions for the deviations from the deterministic limit obtained by Kurtz in Ref. [12], see also Refs. [10,13].

Other results are completely new (to our knowledge) such as the Poisson limit and most importantly the explicit error bounds and, hence, range of applicability associated to all the approximations which are needed in natural sciences. In addition, we must emphasize that all the limit cases and error bounds are obtained within the framework of the Poisson approximation hence unifying several limit cases as different realizations of the same approximation method.

We shall show in this section the main results of this work, namely: convergence of several approximated moment generating functions to the exact moment generating functions as well as convergence for short times of the probability generating function.

A. The generalized mass-action law

In what follows, we will consider that the system is restricted to (or at least unlikely to escape) a region of the phase space

$$||X||_1 \leqslant \Omega. \tag{27}$$

We shall call Ω the size of the system. In most physical, chemical, and biological systems such restriction appears naturally as a consequence of the finiteness of the available energy, number of molecules, total population, or carrying capacity of the environment.

We shall further consider when taking limits that

$$W_j(X) = \Omega w_j(X/\Omega), \tag{28}$$

a generalized mass-action law which renders explicit the idea that if the system is, for example, made twice as large duplicating each population as well as the environment, then the number of interactions per unit time will be roughly twice the original figure.

B. Inequalities

Before we proceed further, we shall notice that the expression

$$K_j(z,t-s) = [z_j \Phi(z,t-s;X+\delta_j) - \Phi(z,t-s;X)],$$

that is, part of Eq. (26), satisfies the equation

$$\frac{d}{dt} K_j(z,t) = \mathcal{L}(X) K_j(z,t), \tag{29}$$

with initial condition $z_j - 1$. The function K_j can be written as

$$K_j(z,t) = \sum_l (z_l - 1) g_j^l(z,t). \tag{30}$$

After making use of Eqs. (6), (7), and (11), we have

$$\frac{d}{dt} g_j^k = \mathcal{L}(X) g_j^k + \sum_l z_l [\mathcal{L}_k(X+\delta_l) - \mathcal{L}_k(X)] g_j^l, \tag{31}$$

with initial condition $g_j^k(z,0) = \delta_{jk}$.

Writing Eq. (31) in integral form as

$$g_j^k(z,t) = \delta_{jk} \Phi(z,t;X) + \int_0^t \exp[\mathcal{L}(X)(t-s)] \times \sum_l z_l [\mathcal{L}_k(X+\delta_l) - \mathcal{L}_k(X)] g_j^l(z,s) ds \tag{32}$$

and considering the difference $[g_k^j - \Phi(z,t;X)\delta_{kj}]$ as well as the inequality (9), we can use Gronwall's inequality to estimate the errors:

$$\sup_k ||g_j^k - \delta_{jk}\Phi(z,t;X)|| \leqslant t \sup_k |W_k(X+\delta_j) - W_k(X)| \times \exp\left(t \sum_l \sup_{kY} |W_k(Y+\delta_l) - W_k(Y)| \right), \tag{33}$$

which as we shall see is approximately independent of the size of the system when the transition probabilities satisfy a mass-action law.

There is a more obvious bound for K_j. Since $0 \leqslant |z_j|$, $|\Phi(z,t;X)| \leqslant 1$, we have

$$|K_j(z,t)| \leqslant 1. \tag{34}$$

The expression (33) has the advantage that produces the exact result for $z_i = 1$ and, as such it is better suited for the study of expectation values which correspond to expansions around $z_i = 1$ of the generating function.

The correction to the generating function (26) is then bounded by

$$\left| \Delta(z,t;X_0) - \int_0^t \left(\sum_{j=1,\ldots,E} \sum_{\{n\} \notin B_j} z^n \overline{P}(n,s) [W_j(X) - f_j] \times (z_j - 1) \Phi(z,t-s;X) \right) ds \right| \leqslant \left| \int_0^t (t-s) ds \left(\sum_{j=1,\ldots,E} \sum_{\{n\} \notin B_j} z^n \overline{P}(n,s) |W_j(X) - f_j| \left\{ \sum_k (1-z_k) \right\} \sup_{kY} |W_k(Y+\delta_j) - W_k(Y)| \times \exp\left[(t-s) \sum_l \sup_{kY} |W_k(Y+\delta_l) - W_k(Y)| \right] \right) \right|. \tag{35}$$

C. Deterministic limit

The standard process to obtain a deterministic limit for the random variable X proceeds by noting that under adequate conditions, in particular, for $\Omega \to \infty$, the variable $x = X/\Omega$ obeys a deterministic differential equation, up to deviations going to zero with $h(\Omega) = 1/\sqrt{\Omega}$ [10,13] The crucial matter with this limit is that in order to disregard the deviations, Ω must be *sufficiently* large, i.e., *as large as needed*. In applications to natural sciences, Ω is often fixed and whether it is large enough or not depends on the specific problem. Our approximations aim to cast some light on the frequent situation where Ω is *not* large enough to accept the deterministic limit as a good approximation.

We shall consider the behavior of the different momenta of the stochastic variables representing the changes in the populations as a function of time in the scale $1/\Omega$. For such purpose it is convenient to introduce *moment generating functions*. The generating function $\Phi(z)$ is one of such functions since its (left) derivatives at $z=1$ are linear combinations of the moments of the distribution P. We will for the moment write $\Phi_X(z)$ to render explicit which stochastic variable is in action. The moment generating function $H_X(w)$ is defined via $H_X(w)=\Phi_X(\exp(w))$, i.e., by just replacing z with $\exp(w)$ in Φ. The name arises by the fact that the moments of X are given by the derivatives of H_X at $w=0$ (provided they exist), in other words, if M_s are the moments of X, we have $H_X(w)=\Sigma_0^\infty M_s w^s/s!$.

At this point, the reader can verify that the moment generating function for the random variable $x=X/\Omega$ is obtained by replacing w by ν/Ω in H_X, i.e., $H_x(\nu)=H_X(\nu/\Omega)$ or in other words, $H_x(\nu)=\Phi_X(\exp(\nu/\Omega))$. The moments of x are now given by the derivatives of H_x at $\nu=0$ (assuming existence). We have then the following theorem.

Theorem 1 (truncated Poisson approximation). Let V_j be the minimum distance to the j-border states, i.e., the minimum of all n such that $X_0+\Sigma_{k\neq j}n_k\delta_k+n\delta_j\in B_j$. Assume also that the generalized mass-action law (28) holds and that $|W_j(X)-W_j(Y)|\leqslant C_j|X-Y|$ with C_j $(j=1,\ldots,E)$ finite. Then, for $\epsilon>0$ sufficiently small and $-\epsilon<\nu_i\leqslant 0$, $\Psi_X(\exp[\Sigma_j\nu_j/\Omega])$ converges uniformly to $H_x(\nu)$ in the limit $\Omega\to\infty$, provided that $\forall j,\ \lambda_j/V_j<1$.

Proof. We will show that under the conditions of the theorem, $\Psi_X(\exp(\nu/\Omega))$ converges to $H_x(\nu)=\Phi_X(\exp(\nu/\Omega))$ by checking that the difference $\Delta(\exp(\nu/\Omega))$ goes to zero in the limit $\Omega\to\infty$ (ν stands for $\Sigma_j\nu_j$ throughout the proof when needed).

We begin by noting from Eq. (35) that $|\Delta|$ is bounded by the sum of two terms. We study each term separately. For the first term, we note that $|z_j-1|=\epsilon/\Omega[1+O(\epsilon/\Omega)]$. Therefore,

$$\left|\int_0^t\sum_{j=1,\ldots,E}\sum_{\{n\}\notin B_j}z^n\overline{P}(n,s)[W_j(X)-f_j](z_j-1)\Phi(z,t-s;X)ds\right|\leqslant\frac{\epsilon}{\Omega}[1+O(\epsilon/\Omega)]\sum_{j=1,\ldots,E}R_j, \tag{36}$$

where $z=\exp(\nu/\Omega)$ and the contribution R_j reads

$$R_j=\int_0^t ds\sum_{\{n\}\notin B_j}|z^n\Phi(z,t-s;X)|\quad|W_j(X)-f_j|\overline{P}(n,s). \tag{37}$$

Letting $\mathcal{N}_j(t,X_0)$ as in Eq. (21) and using Schwartz' inequality with the weights $Q(n,s)=\overline{P}(n,s)/\mathcal{N}_j(s,X_0)$, we have

$$R_j\leqslant\int_0^t ds\mathcal{N}_j(s,X_0)\sqrt{\sum_{\{n\}\notin B_j}|W_j(X)-f_j|^2Q(n,s)}, \tag{38}$$

after noting that $|z^n\Phi(z,t-s;X)|\leqslant 1$ for $|z|\leqslant 1$ and any t.

Recalling Eq. (21) $f_j=\Sigma_{\{n\}\notin B_j}Q(n,s)W_j(X)$ so the argument of the root can be written as $\langle|W_j-\langle W_j\rangle|^2\rangle$ averaged with the weights $Q(n,s)$. Moreover, $\langle|W_j-\langle W_j\rangle|^2\rangle=(1/2)\langle|W_j(n)-W_j(m)|^2\rangle$. In the same way $(1/2)\langle|n-m|^2\rangle=\langle|n-\langle n\rangle|^2\rangle$.

Let us now fulfill the transition to the Ω-scaled variables using the generalized mass-action law: $|W_j(n)-W_j(m)|=\Omega|w_j(n/\Omega)-w_j(m/\Omega)|\leqslant C_j|n-m|=\Omega C_j|\delta_j||(n/\Omega)-(m/\Omega)|$. Thus,

$$\langle|W_j-\langle W_j\rangle|^2\rangle\leqslant(\Omega C_j\delta_j)^2\langle|(n/\Omega)-(m/\Omega)|^2\rangle=(C_j\delta_j)^2\langle|n-\langle n\rangle|^2\rangle,$$

the average always taken with the weights $Q(n,s)$.

Hence, since $|\mathcal{N}_j|\leqslant 1$ and noticing that $\langle n_j\rangle\leqslant\lambda_j$ while $\langle(n_j-\langle n_j\rangle)^2\rangle\leqslant\lambda_j$ using that the weights $Q(n,s)$ correspond to a truncated Poisson and that $V_j-\lambda_j>0$,

$$R_j\leqslant t\sqrt{\lambda_j}C_j|\delta_j| \tag{39}$$

and the first contribution goes to zero as

$$\frac{\epsilon}{\Omega}[1+O(\epsilon/\Omega)]\sum_{j=1,\ldots,E}R_j=O(\epsilon t\sqrt{\hat{\lambda}}/\sqrt{\Omega}). \tag{40}$$

Turning to the second term, we proceed as above, namely,

$$\begin{aligned}&\left|\int_0^t(t-s)ds\left(\sum_{j=1,\ldots,E}\sum_{\{n\}\notin B_j}z^n\overline{P}(n,s)|W_j(X)-f_j|\right.\right.\\&\times\left\{\sum_k(1-z_k)\right\}\sup_{kY}|W_k(Y+\delta_j)-W_k(Y)|\\&\left.\left.\times\exp\left[(t-s)\sum_l\sup_{kY}|W_k(Y+\delta_l)-W_k(Y)|\right]\right)\right|\\&\leqslant\sum_{j=1,\ldots,E}\frac{E\epsilon}{\Omega}[1+O(\epsilon/\Omega)]\int_0^t ds\left((t-s)\right.\\&\left.\times\exp\left[(t-s)C\sum_l|\delta_l|\right]C|\delta_j|\right)(\sqrt{\lambda_j}C_j|\delta_j|),\end{aligned} \tag{41}$$

where $C=\Sigma_kC_k$.

Putting everything together, we obtain

$$\begin{aligned}&|\Delta(\exp(\nu/\Omega),t,X_0)-O(\epsilon t\sqrt{\hat{\lambda}}/\sqrt{\Omega})|\\&\leqslant\sum_{j=1,\ldots,E}(\sqrt{\hat{\lambda}_j}C_j|\delta_j|)O(t^2\epsilon/\sqrt{\Omega}).\end{aligned} \tag{42}$$

It is hence clear that $\Psi_X(\exp(\nu/\Omega))$ converges to $H_x(\nu)=\Phi_X(\exp(\nu/\Omega))$ as fast as $\epsilon t\sqrt{\hat{\lambda}/\Omega}$. ■

Theorem 2 (large-size limit). Under the conditions of the previous theorem and if, additionally, $\lim_{\Omega\to\infty}\lambda_j/V_j=b_j<1$ then $\Psi_X(\exp(\nu/\Omega))$ converges uniformly to $\exp(\Sigma_j\nu_j\hat{\lambda}_j)$, where $\hat{\lambda}_j=\lim_{\Omega\to\infty}\lambda_j/\Omega$ satisfies the differential equation

$$\frac{d\hat{\lambda}_j}{dt}=\lim_{\Omega\to\infty}\frac{f_j}{\Omega}.$$

Proof.

$$\Psi_X(\exp(\nu/\Omega))=\sum_n \exp(\nu n/\Omega)\overline{P}(n,t)$$
$$=\sum_{n,n_i<V_i}\exp(\nu n/\Omega)\overline{P}(n,t)$$
$$+\sum_{n,n_i\geq V_i}\exp(\nu n/\Omega)\overline{P}(n,t), \qquad (43)$$

where both sums include only admissible states. Note that each term takes its largest value for $\nu=0$ and that (43) can be recasted more explicitly as

$$\Psi_X(\exp(\nu/\Omega))=\prod_j^E \sum_{n_j\leq V_j}[\lambda_j\exp(\nu_j/\Omega)]^{n_j}/n_j!\exp(-\lambda_j)$$
$$+\sum_{n,n_i\geq V_i}\exp(\nu n/\Omega)\overline{P}(n,t), \qquad (44)$$

while

$$\sum_{n,n_i\geq V_i}\exp(\nu n/\Omega)\overline{P}(n,t)\leq\prod_j^E\sum_{n_j\geq V_j}[\lambda_j^{n_j}/n_j!\exp(-\lambda_j)]. \qquad (45)$$

After rearranging the right-hand side, using known properties of the Γ function and the Stirling approximation for factorials, it is easy to verify that the right-hand side of Eq. (45) is bounded from above by $C\exp(-\Sigma_j c_j V_j)$ for some positive constants C,c_j (see the Appendix for the explicit computation). The hypothesis of the theorem imply that $V_j\to\infty$ with Ω at least linearly and hence the right-hand side goes to zero with Ω exponentially fast, regardless of ν. Reasoning along the same lines, we note that the error introduced by continuing the sums of the first term of Eq. (43) beyond V_i also decreases roughly exponentially fast (see also the Appendix for an illustration) in the limit $\Omega\to\infty$. Hence, we have exponentially fast convergence of

$$\Psi_X(\exp(\nu/\Omega))-\exp\left(\sum_j \lambda_j(e^{\nu_j/\Omega}-1)\right)\to 0,\quad \Omega\to\infty. \qquad (46)$$

To conclude the proof, we note that under the assumption of the mass-action law

$$\exp\left(\sum_j \lambda_j(e^{\nu_j/\Omega}-1)\right)=\exp\left(\sum_j \hat{\lambda}_j\nu_j\right)+O(\nu^2/\Omega). \quad \blacksquare \qquad (47)$$

Corollary (deterministic limit). In the conditions of the above theorems, the fluctuation of the variables x_i are zero; i.e., the variables have a deterministic behavior in the limit $\Omega\to\infty$.

Proof. The momenta of order k generated by $H_x(\nu)$ or, equivalently, by $\exp(\Sigma_j\hat{\lambda}_j\nu_j)$ in the limit $\Omega\to\infty$ are $\hat{\lambda}_j^k$. Hence, in the limit, $\langle\Pi_j(n_j/\Omega-\hat{\lambda}_j)^{k_j}\rangle=0$ for any collection of integers k. ■

Notice that when projected onto the phase space, the populations obey the following deterministic equation in the limit $\Omega\to\infty$:

$$\frac{dx}{dt}=\sum_j \delta_j w_j(x), \qquad (48)$$

where, letting $x_0=\lim_{\Omega\to\infty}X_0/\Omega$, we have

$$x=x_0+\sum_j \delta_j\hat{\lambda}_j, \qquad (49)$$

a result that corresponds with the main theorem in Ref. [11]: the law of large numbers.

D. Poisson limit

A different limit is obtained when $\Omega\to\infty$ and $t\to 0$, while λ_j is kept bounded.

Theorem 3 (Poisson limit). Under the assumption of the mass-action law and if $|W_j(X)-W_j(Y)|\leq C_j|X-Y|$ with C_j $(j=1,\ldots,E)$ finite, then $\Phi(z,t;X_0;\Omega)-\Psi(z,t;X_0;\Omega)$ converges uniformly to zero as a function of z in $[0,1]$ in the limit $\Omega\to\infty$, $t\to 0$, while λ_j is kept bounded.

Proof. The proof proceeds along the lines followed in Theorem 1 with the only difference that we now bound z and $(1-z)$ by 1. Then, the two contributions to the error in Eq. (35) computed in Eqs. (39) and (42) become of order $O(t\sqrt{\lambda})$ and $O(t^2\sqrt{\lambda})$, respectively. The latter is negligible in front of the first one for $t\to 0$. Both contributions go to zero in the conditions of the theorem when $t\to 0$ as $1/\Omega$. ■

E. Fluctuations around the deterministic limit

We shall now estimate the distribution function for the fluctuations around the deterministic limit. To accomplish this end, we shall proceed in several steps: First, we will improve the error bound for the deterministic limit; second, we will estimate the error for the mean values, and finally we will proceed to estimate bounds for the errors in the generating function of centered momenta.

1. Improved error estimates

The relation between the generating function corresponding to the Poisson approximation and the one corresponding to the exact process has been formulated in Eq. (35) as an estimation for Δ and an error bound for this estimation.

The estimation reads

$$\Delta(z,t;X_0)\sim\int_0^t\left(\sum_{j=1,\ldots,E}\sum_{\{n\}\notin B_j} z^n\overline{P}(n,s)[W_j(X)-f_j](z_j-1)\Phi(z,t-s;X)\right)ds, \qquad (50)$$

and it actually goes to zero quadratically when $\sqrt{\Sigma_j(z_j-1)^2}\to 0$. To realize this fact it is enough to notice that every term in the right-hand side of Eq. (50) includes a factor (z_j-1) and the evaluation of the remaining factors at $z=1$ produces

$$\Delta(z,t;X_0)\sim\int_0^t\left(\sum_{j=1,\ldots,E}\sum_{\{n\}\notin B_j}\bar{P}(n,s)[W_j(X)-f_j]\times(z_j-1)\right)ds+O\left(\sum_j (z_j-1)^2\right)=O\left(\sum_j (z_j-1)^2\right). \quad (51)$$

Recalling the definition of f_j in Eq. (21) it follows that the first term sums to zero.

This observation has two immediate consequences, first, the convergence in Theorem 1 is improved to

$$|\Delta(\exp(\nu/\Omega),t,X_0)-O(\epsilon^2 t\sqrt{\hat{\lambda}}/\Omega^{3/2})|\leqslant\sum_{j=1,\ldots,E}(\sqrt{\hat{\lambda}_j}C_j|\delta_j|)O(t^2\epsilon/\sqrt{\Omega}). \quad (52)$$

Second, the estimation in Eq. (35) has no terms linear in ν_j and hence, its correction to the mean value is zero. In other words, the exact mean values of n_j relate to the approximate ones as follows:

$$[n_j]-[n_j]_P=\sum_{j=1,\ldots,E}(\sqrt{\hat{\lambda}_j}C_j|\delta_j|)O(t^2\epsilon\sqrt{\Omega}), \quad (53)$$

where $[n_j]_P$ is the mean value of n_j using the Poisson weights $\bar{P}$.

2. Fluctuations around the mean value

In this section, we shall consider the stochastic variables $\zeta_i=(n_i-[n_i])/\sqrt{\Omega}$ that represent properly scaled fluctuations around the expectation values of the stochastic variables n_i ($[n_i]$ is the exact mean value of n_i).

The moment generating function reads [15]

$$\Xi(\nu,t)=\exp\left(-\sum_j \nu_j[n_j]/\sqrt{\Omega}\right)\times\Phi\left(\exp\left(\sum_j \nu_j/\sqrt{\Omega}\right),t,X_0\right), \quad (54)$$

with $-\epsilon'\leqslant\nu_j\leqslant 0$. We shall now analyze the different contributions to Eq. (54) taking into account the expression (26).

First, notice that

$$\Xi(\nu,t)=\exp\left(-\sum_j \nu_j([n_j]-[n_j]_P)/\sqrt{\Omega}\right)\times\exp\left(-\sum_j \nu_j[n_j]_P/\sqrt{\Omega}\right)\times\{\Psi(\exp(\nu/\sqrt{\Omega}))+\Delta(\exp(\nu/\sqrt{\Omega}))\}, \quad (55)$$

with $[n_j]_P$ as above.

Second, we realize that in order to produce error bounds for our approximations we can simply use those developed for the deterministic approximation changing the range of the variable from $|\nu|\leqslant\epsilon$ to $|\nu|\leqslant\epsilon'\sqrt{\Omega}$. Hence, according to Eqs. (52) and (53),

$$|\Delta(\exp(\nu/\sqrt{\Omega})-O(\epsilon'^2 t\sqrt{\hat{\lambda}}/\Omega^{1/2})|\leqslant\sum_{j=1,\ldots,E}(\sqrt{\hat{\lambda}_j}C_j|\delta_j|)O(t^2\epsilon'), \quad (56)$$

$$([n_j]-[n_j]_P)/\sqrt{\Omega}=\sum_{j=1,\ldots,E}(\sqrt{\hat{\lambda}_j}C_j|\delta_j|)O(t^2\epsilon'). \quad (57)$$

We can state the result by the following.

Theorem 4 (Fluctuations around the deterministic limit). The fluctuations of n_j around its mean value, $\langle n_j\rangle$, in the scale $\sqrt{\Omega}$, i.e., $(n_j-\langle n_j\rangle)/\sqrt{\Omega})$ converge towards a Brownian process under the conditions of Theorems (1 and 2) in the limit $\Omega\to\infty$ for any fixed $t<t^*$ and the proper does the motion in phase space for the variable $(X-\langle X\rangle)/\sqrt{\Omega}$.

Proof. According to the preceding discussion, the fluctuations have a moment generating function

$$\Xi(\nu)=\exp\left(-\sum_j \nu_j[n_j]/\sqrt{\Omega}\right)\Psi(\exp(\nu/\sqrt{\Omega}))+\sum_{j=1,\ldots,E}(\sqrt{\hat{\lambda}_j}C_j|\delta_j|)O(t^2\epsilon'). \quad (58)$$

Using the standard form of the generating function for a Poisson distribution verify that for fixed $t\leqslant t^*$, Eq. (58) has limit

$$\Xi(\nu)=\exp\left(\sum_j (\nu_j^2\hat{\lambda}_j)/2+O(|\nu|^3\hat{\lambda}/\sqrt{\Omega})\right)+\sum_{j=1,\ldots,E}(\sqrt{\hat{\lambda}_j}C_j|\delta_j|)O(t^2\epsilon'), \quad (59)$$

which converges to the generating function of the centered momenta in a normal process.

The increments of the variable $x=X/\Omega=[X_0+\Sigma_j(\delta_j n_j)]/\Omega$ in a time interval dt are then (using Theorems 1 and 2)

$$dx=\left(\sum_j \delta_j w_j(x)+\zeta(t)/\sqrt{\Omega}\right)dt, \quad (60)$$

where ζ is a normally distributed variable with zero mean and covariance matrix $[\zeta^l(t)\zeta^m(t')]$, which can be estimated as $[\zeta^l(t)\zeta^m(t')]\sim\delta(t-t')\Sigma_j\delta_j^l\delta_j^m w_j(x)$. ■

The expression (60) corresponds to a Langevin process or Brownian motion in phase space with a noise amplitude which is state dependent, this process is also known as the *diffusion approximation* [10] and, as far as we know, it has always been introduced heuristically rather than as a limit case like in the present work ([10] p. 459). The linearization of Eq. (60) assuming small departures of the solution of the stochastic process with respect to the deterministic solution leads to the central Limit Theorem stated by Kurtz [12]. However, such linearization will drastically shorten the time range where the approximation can be used, a fact that is not apparent when Ω can be taken as large as needed but matters when Ω and the time are large but finite.

V. EXAMPLE

Let us consider as an example a Markov process describing the extinction of an isolated species. Let V be the initial population and further $E=1$ denotes the only event-class present in the system, namely, the death of an individual. As above, let $P_n(t)$ denote the probability of n events taking place up to time t, so the natural limits of this problem are $P_n(t)=0$, for $n<0$ and $n>V$ and also $P_n(0)=\delta_{0,n}$.

The Kolmogorov forward equation for the generating function of the problem reads,

$$\frac{\partial\Phi(z,t)}{\partial t}=(z-1)\left(V-z\frac{\partial}{\partial z}\right)\Phi(z,t) \tag{61}$$

and after a standard computation, we can obtain the exact probabilities as

$$P_n(t)=\binom{V}{n}[\exp(-t)]^{V-n}[1-\exp(-t)]^n. \tag{62}$$

The corresponding Poisson approximation to the probabilities reads

$$\Psi(z,t;X_0)=z^V+\sum_{z=0}^{V-1}\exp(-\lambda)\frac{\lambda^n}{n!}(z^n-z^V), \tag{63}$$

where the time-dependent Poisson parameter λ satisfies $\lambda(0)=0$ and

$$\frac{d\lambda}{dt}=V-\lambda\left(1-\frac{\lambda^{V-1}/(V-1)!}{\sum_{k=0}^{V-1}\lambda^k/k!}\right).$$

The deterministic limit of n/V for $V\to\infty$ is then

$$\frac{n}{V}\to\hat{\lambda}=[1-\exp(-t)],$$

which is the solution of $d\hat{\lambda}/dt=1-\hat{\lambda}=1-\lim_{V\to\infty}\langle n\rangle_\lambda/V$.

To illustrate the meaning of Theorem 4, we consider the variable

$$x=\{n-V[1-\exp(-t)]\}/\sqrt{V\hat{\lambda}(t)\exp(-t)},$$

for $|x|\leq A$ and A arbitrary but fixed and independent of V. This variable has mean value zero and dispersion one. Furthermore, a well-known limit relation between the binomial and the normal distribution (see Ref. [15]) states that x is approximately distributed by a normal distribution $N(0,1)$ in $|x|\leq A$ when $V\to\infty$ and the convergence is of order $O(1/\sqrt{V})$ for fixed A.

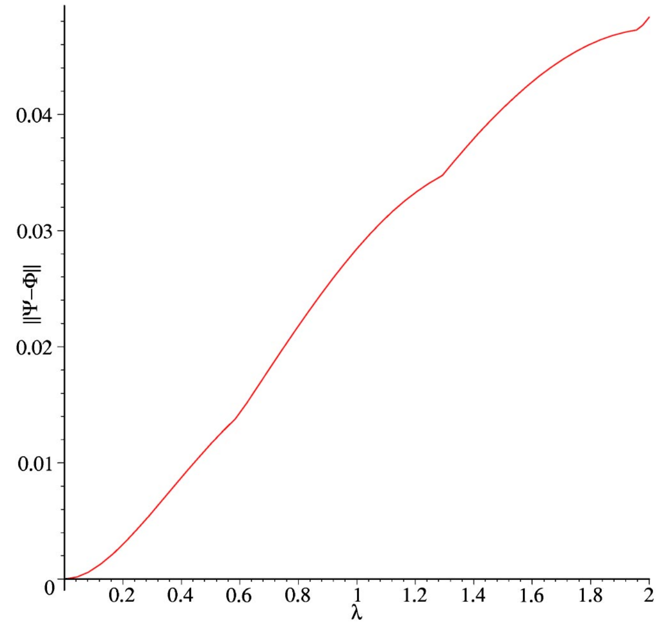

FIG. 1. Norm of the difference between the approximated and exact generating functions as a function of λ for $V=20$.

Hence, $(n-V[1-\exp(-t)])/\sqrt{V}$ is distributed as $N(0,\hat{\lambda}\exp(-t))$. According to Theorem 4 , the variable $(n-V[1-\exp(-t)])/\sqrt{V}$ is distributed approximately as $N(0,\hat{\lambda})$ in the limit $V\to\infty$ for fixed t^* sufficiently small. This result is consistent with the normal approximation of the binomial when both limits coincide.

Both approximations are based in considering a time interval where there is a large number of events (as large as required to keep the approximations within a prescribed tolerance). For the approximation described in Theorem 4 it may be also necessary to make the time interval sufficiently small to achieve the desired tolerance.

Although Theorem 4 is more restrictive that the approximation of a binomial by (integrals of) a normal distribution, it can be applied to problems where exact solutions in terms of binomial distributions are not available, i.e., to general problems.

Recalling the standard approximation of the Binomial (p,q) by a Poisson distribution (see Ref. [15]) when $Vp\to b$ and realizing that if $V[1-\exp(-t)]\to b$ and $V\to\infty$ then $t\to 0$ as $1/V$, we verify in this example the result announced in Theorem 3 as a short-time approximation. Finally, we show the evolution of the norm of the error for $V=20$ (Fig. 1).

VI. CONCLUDING REMARKS

The art of approximating the dynamics of the populations consists in providing algorithms which are simpler than the defining process and which produce approximated results within a prescribed tolerance for a given problem.

In this work, we have introduced and studied in detail a Poisson approximation to population dynamics recovering the standard limits for the description of the fractions of large populations and their fluctuations and establishing the order of magnitude of upper bounds for the errors introduced.

In general, we can state that the approximation improves decreasing the time interval and decreasing the total number of events in the time interval, as manifested by the occurrence of factors proportional to t, t^2 and $\sqrt{\lambda_j}$ in the error estimates.

We have studied three limit cases: the deterministic limit, the fluctuations around the deterministic limit (both of these limits were already known to a large extent) and we have established a third limit case in which the population dynamics corresponds exactly with the approximated dynamics. Most importantly, the three limit cases are nothing but particular situations within one, unifying, approximation where the errors are controlled by the combinations of several variables as stated in the Eqs. (35),(52), and (53).

The occurrence of different limit cases and the explicit dependency with Ω (rather than λ_j) is, up to some point, artificial. It results, obvious, that the validity of *any* approximation "improves" by increasing our tolerance. Whether to consider the variable X or $x = X/\Omega$ or $(X - \langle X \rangle)/\sqrt{\Omega}$ is a decision that can only be taken knowing the accuracy required for useful answers to our questions and as such is foreign to the approximation.

Having these considerations in mind, the main result of this work consists in having established that the errors introduced by the Poisson approximation can always be controlled with sufficiently small time intervals, hence establishing a suitable (eventually numerical) integration scheme for the class of stochastic processes we are analyzing. Note, however, that the time intervals are still quite "large." In other words, the time-intervals required (see Theorem 4) are not so small as to consider that the probability of two events occurring in the same time interval is negligible (a situation in which the Poisson approximation trivially converges to the exact process).

The present integration scheme can be used even when the size of the involved populations fluctuates strongly, as in epidemic outbreaks where the number of infected people rises from zero at the beginning of the outbreak and reaches zero after the outbreak, i.e., when one or more populations can go permanently or temporarily extinct.

Finally, this work introduces an appealing relation between deterministic dynamics and stochastic dynamics through the differential equations satisfied by the parameters of the Poisson approximation.

ACKNOWLEDGMENTS

We would like to acknowledge fruitful discussions with Juan Aparicio and Ingemar Nåsell. This collaboration was partially financed by Fundacin Antorchas and STINT. H.G.S. acknowledges support from the University of Buenos Aires under Grant No. X208/00.

APPENDIX: INCOMPLETE GAMMA FUNCTIONS

We perform the computations in Theorem 2 explicitly for the case $E = 1$. We drop hereafter all labels j. The left-hand side (lhs) of Eq. (46) consists of two terms, one of them is the lhs of Eq. (45), while the other is

$$\sum_{n \geqslant V} \{[\lambda \exp(\nu/\Omega)]^n/n! \exp(-\lambda)\} \leqslant \sum_{n \geqslant V} [\lambda^n/n! \exp(-\lambda)]. \tag{A1}$$

Hence,

$$-[\text{lhs}(46)] = \sum_{k=V+1}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} (e^{\nu V/\Omega} - e^{\nu k/\Omega})$$
$$= \frac{e^{\nu V/\Omega}}{V!} \int_0^{\lambda} e^{-t} t^V dt - \frac{e^{\lambda(e^{\nu\Omega}-1)}}{V!} \int_0^{\lambda e^{\nu/\Omega}} e^{-t} t^V dt \tag{A2}$$

$$= \frac{e^{\nu V/\Omega}}{V!} \int_{\lambda e^{\nu/\Omega}}^{\lambda} e^{-t} t^V dt - \frac{(e^{\nu V/\Omega} - e^{\lambda(e^{\nu\Omega}-1)})}{V!} \int_0^{\lambda e^{\nu/\Omega}} e^{-t} t^V dt. \tag{A3}$$

Since the integrand has its maximum at $t = V$ and the integration interval lies on the region $t < V$ the change of variables $t = V + u$ and Stirling's approximation $V! = (V/e)^V \sqrt{2\pi V} e^{-\theta/12V}$ for $\theta \in [0,1]$, yield

$$|\text{lhs}(46)| = \left| \frac{e^{-\theta/12V}}{\sqrt{2\pi V}} \left(e^{\nu V/\Omega} \int_{\lambda e^{\nu/\Omega} - V}^{\lambda - V} e^{-u} (1 + u/V)^V du - (e^{\nu V/\Omega} - e^{\lambda(e^{\nu\Omega}-1)}) \int_{-V}^{\lambda e^{\nu/\Omega} - V} e^{-u} \times (1 + u/V)^V du \right) \right|$$
$$\leqslant \frac{e^{-\theta/12V}}{2} (e^{\nu V/\Omega} \sqrt{e^{-(V-\lambda)^2/V} - e^{-(V-\lambda e^{\nu/\Omega})^2/V}} + |e^{\nu V/\Omega} - e^{\lambda(e^{\nu\Omega}-1)}| \sqrt{e^{-(V-\lambda e^{\nu/\Omega})^2/V} - e^{-V}}).$$

Under the conditions of the theorem ν is nonpositive and $V - \lambda e^{\nu/\Omega} \geqslant V - \lambda \sim V(1-b)$ and hence, both roots are bounded from above by $\exp[-V(1-b)]$. Therefore,

$$|\text{lhs}(46)| \leqslant e^{-(1-b)V/2}. \tag{A4}$$

Notice further that Eq. (A2) is the difference of two positive terms and each one goes exponentially to zero in the limit considered. Moreover, lhs (46) is exactly zero for $\nu = 0$.

[1] P. Glansdorff and I. Prigogine, *Thermodynamic Theory of Structure, Stability and Fluctuations* (Wiley, London, 1971).
[2] R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, New York, 1975).
[3] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems* (Wiley, New York, 1977).
[4] D.T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[5] J. Güémez and M.A. Matías, Phys. Rev. E **51**, 3059 (1995).
[6] C.H. Henry, IEEE J. Quantum Electron. **18**, 259 (1982).
[7] C.H. Henry, IEEE J. Quantum Electron. **19**, 1391 (1983).
[8] W.W. Chow, S.W. Koch, and M. Sargent, *Semiconductor-Laser Physics* (Springer-Verlag, Berlin, 1994).
[9] E. Renshaw, *Modelling Biological Populations in Space and Time* (Cambridge University Press, Cambridge, 1991).
[10] S.N. Ethier and T.G. Kurtz, *Markov Processes* (Wiley, New York, 1986).
[11] T.G. Kurtz, J. Appl. Probab. **7**, 49 (1970).
[12] T.G. Kurtz, J. Appl. Probab. **8**, 344 (1971).
[13] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*, Lecture Notes in Statistics Vol. 151 (Springer-Verlag, Berlin, 2000).
[14] J.P. Aparicio and H.G. Solari, Phys. Rev. Lett. **86**, 4183 (2001).
[15] A. Renyi, *Cálculo de Probabilidades* (Editorial Reverté, Madrid, 1976).
[16] Situations in which the probability of occurrence of an event depends on the history of the individuals are usually handled by introducing subpopulations. For example, if the time required to reach sexual maturity is relevant, we can introduce an age-structured population: only the sexually mature subpopulation will be involved in reproductive events.