Proceedings of the International Multiconference on
Computer Science and Information Technology pp. 289–296
ISBN 978-83-60810-27-9
ISSN 1896-7094

On the Numerical Analysis of Stochastic Lotka-Volterra Models

Tuğrul Dayar[1], Linar Mikeev, and Verena Wolf
Department of Computer Science, Saarland University, Saarbrücken, Germany

***Abstract*—The stochastic Lotka-Volterra model is an infinite Markov population model that has applications in various life science domains. Its analysis is challenging since, besides an infinite state space with unbounded rates, it shows strongly fluctuating dynamics and becomes unstable in the long-run. Traditional numerical methods are therefore not appropriate to solve the system. Here, we suggest adaptations and combinations of traditional methods that yield fast and accurate solutions for certain parameter ranges of the stochastic Lotka-Volterra model. We substantiate our theoretical investigations with a comparison based on experimental results.**

I. Introduction

MORE than 80 years ago, Lotka and Volterra independently proposed the following three rules to describe the numerical evolution of two populations in interspecific competition [24, 30]:

1) The first population (prey) grows at rate αx, where x represents the population of prey.
2) The second population (predator) "eats" prey and grows at rate βxy, where y represents the population of predator.
3) The predator population decreases through natural death at rate γy.

The solution of the corresponding system of (non-linear) ordinary differential equations (ODEs)

$$\begin{array}{rcl} \frac{dx}{dt} & = & \alpha x - \beta xy \\ \frac{dy}{dt} & = & \beta xy - \gamma y \end{array} \qquad (1)$$

shows sustained oscillations for positive constants α, β, γ along the closed curves

$$\beta x - \gamma \log x + \beta y - \alpha \log y = const$$

except if initially the system does start in the equilibrium point $x = \gamma/\beta, y = \alpha/\beta$. In Fig. 1(a), we plot the solution for $\alpha = \gamma = 1$ and $\beta = 0.01$ across time where we initially started with $x_0 = y_0 = 20$. Since its introduction, the Lotka-Volterra model became one of the most popular models in population dynamics, but it has also been successfully applied to neural networks [25] and game theoretic problems [16].

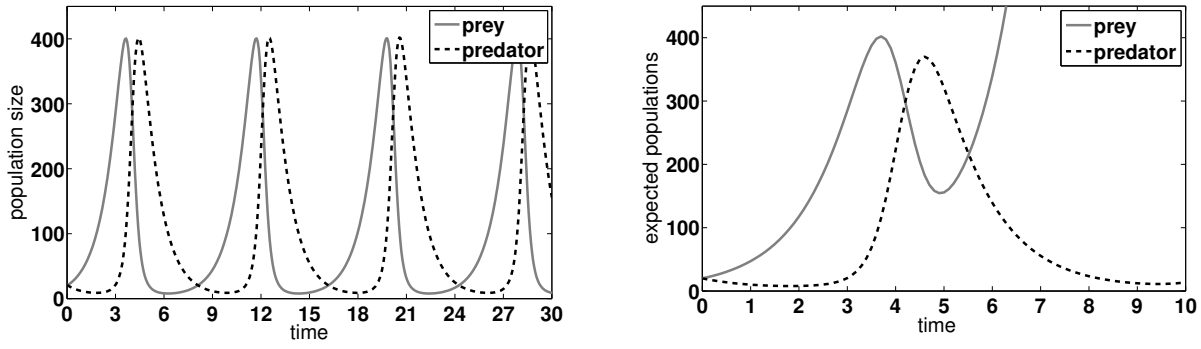
The stochastic Lotka-Volterra model assumes that x and y are represented by discrete random variables X and Y and that their evolution in time is given by a two-dimensional Markov process $\{(X(t), Y(t)), t \geq 0\}$ [11, 13, 20]. As opposed to the original Lotka-Volterra model, the stochastic variant takes into account the discreteness of the populations and their random fluctuations. In the stochastic model, extinction of species is possible and, depending on the initial conditions, the dynamics of the system can deviate drastically from the deterministic model. Recently, generalizations of the stochastic Lotka-Volterra model have been used to investigate the mechanisms that maintain biodiversity in Escherichia coli (E. coli) populations [18, 19, 26]. In such generalizations, the number of species is larger but the rules of interspecific competition remain the same.

Here, we are interested in the transient probability distribution of the stochastic Lotka-Volterra model (cf. Section II). It can be used to derive measures such as the distribution of the time until extinction, the probability of extinction until a certain time instant, the expected populations, etc. For a transient solution of the stochastic Lotka-Volterra model, the underlying Markov process has to be solved; this involves the solution of a system of ordinary differential equations known as the master equation [17]. Standard methods for the transient solution of Markov processes are based on the numerical integration of the differential equations, on uniformization of the process, or on approximations in the Krylov subspace (for an overview see [28]). An alternative approach, called method of moments, has recently been proposed by Engblom and is based on a deterministic approximation of the moments of the Markov process described by the master equation [10].

Already in the case of two species, the analysis of the stochastic Lotka-Volterra model using standard methods, such as numerical integration methods or approximations of the moments become inefficient or even infeasible for several reasons. The state space is infinite in two dimensions and it is non-trivial to truncate the state space in such a way that the number of states in the finite truncation remains tractable. Before one of the species becomes extinct, the expected populations oscillate in a similar way as in the deterministic model. Therefore the dynamics of the system change drastically. If the population of at least one of the two species is high, then a large number of events occur even in small time intervals. Moreover, the rates are unbounded; that is, the system is not uniformizable. In order to describe the strong non-linear dependence between the two populations, higher-order moments are necessary and a deterministic approximation of two or more moments yields poor results.

In this paper, we suggest several adaptations and combinations of standard methods that render an efficient solution of the stochastic Lotka-Volterra model possible. We concentrate on the numerical integration of the master equation (cf. Section III) as well as the method of moments (cf. Section IV).

*T. Dayar is currently on sabbatical leave from Bilkent University, Turkey.



(a) Sustained oscillations of the deterministic Lotka-Volterra model.

(b) Expected populations in the stochastic Lotka-Volterra model.

Fig. 1. ODE solution and stochastic solution of the Lotka-Volterra model where $\alpha = \gamma = 1$, $\beta = 0.01$, and $x_0 = y_0 = 20$.

We extend these methods in such a way that the cheap but inaccurate method of moments becomes more accurate and the expensive but accurate numerical integration becomes faster. In Section V, we propose a stochastic hybrid approach that dynamically switches between a moment-based representation and a distribution-based representation as well as having the ability to simultaneously use a combination of both. We implemented all methods in C++ and performed extensive simulations of the stochastic Lotka-Volterra model in order to provide comparisons in terms of accuracy and run time. Numerical experiments on a number of problems indicate that the stochastic hybrid approach is overall the best method. It is at least an order faster than the numerical integration of the master equation and yields no more than a 5 % relative error in the first moment of the transient distribution.

II. Stochastic Lotka-Volterra model

We describe the transitions of the stochastic Lotka-Volterra model in a guarded command style [1, 23]. A guarded command has the form `guard |- rate -> update;` and it describes a set of transitions in the underlying Markov process. The `guard` is a Boolean predicate that determines in which states the transition is possible, the `rate` is the real-valued function that evaluated in the current state gives the positive transition rate in the Markov process, and `update` is the function that calculates the successor state of the transition. Let x and y be the positive integers that represent the populations of prey and predator, respectively. The three different possible transitions are given as follows:

$$\begin{array}{lllll} \mathtt{x} > 0 & \mid\!- & \alpha \cdot \mathtt{x} & \text{->} & \mathtt{x} := \mathtt{x} + 1; \\ \mathtt{x} > 0, \mathtt{y} > 0 & \mid\!- & \beta \cdot \mathtt{x} \cdot \mathtt{y} & \text{->} & \mathtt{x} := \mathtt{x} - 1, \mathtt{y} := \mathtt{y} + 1; \\ \mathtt{y} > 0 & \mid\!- & \gamma \cdot \mathtt{y} & \text{->} & \mathtt{y} := \mathtt{y} - 1; \end{array}$$

The above guarded command model specifies the elements of the infinitesimal generator matrix Q of a Markov process $\{(X(t), Y(t)), t \geq 0\}$ which takes states (x, y) in $\mathbb{N}^2$. More precisely, if for a pair (x, y) the guard is true, then the element that belongs to (x, y) and the update of (x, y) is equal to the rate. E.g. the row of state $(2, 2)$ contains three positive entries. One for the transition to state $(3, 2)$ at rate 2α, one for the transition to state $(1, 3)$ at rate 4β, and one for the transition to state $(2, 1)$ at rate 2γ. The diagonal element is then given by the negative sum of the off-diagonal elements. It is easy to see that $\{(X(t), Y(t)), t \geq 0\}$ is a regular Markov process [8]. We remark that the outflow rate from state (x, y) with $x, y > 0$ is $(\alpha x + \gamma y + \beta xy)$, where $\alpha, \gamma, \beta > 0$, and this rate increases unboundedly with increasing values of x and y.

Let us fix the initial state of the system as (x_0, y_0). Then the transient probability distribution is given by the master equation

$$\frac{d}{dt} p^{(t)}(x, y) = \mathcal{M}(p^{(t)}(x, y)), \tag{2}$$

where $p^{(t)}(x, y) = P\{X(t) = x, Y(t) = y | X(0) = x_0, Y(0) = y_0\}$. The master operator $\mathcal{M}$ is defined for any real-valued function $g : \mathbb{N}^2 \to \mathbb{R}$ such that $\mathcal{M}(g)$ is the function that maps a state (x, y) to the value[1]

$$\begin{aligned} \mathcal{M}(g(x, y)) = {} & \alpha(x - 1) g(x - 1, y) \\ & + \beta(x + 1)(y - 1) g(x + 1, y - 1) \\ & + \gamma(y + 1) g(x, y + 1) \\ & - (\alpha x + \beta xy + \gamma y) g(x, y), \end{aligned} \tag{3}$$

where α, β, γ are the rate constants as defined before and $x, y > 0$. If $x = 0$ and/or $y = 0$, then the terms involving $g(x - 1, y)$ and/or $g(x + 1, y - 1)$ are removed. The ordinary first-order differential equation in (2) is a direct consequence of the Kolmogorov forward equation and describes the change of the probability distribution as the difference between inflow of probability from direct predecessors (first three terms) and outflow of probability in state (x, y) (last term).

III. Direct numerical approximation

Since no analytical solution is known for Eq. (2) and the number of equations is infinite in two dimensions, a numerical solution is only possible if appropriate bounds for the variables x and y are found.

[1]We assume that all terms with a negative argument are zero (e.g. $\alpha(x - 1)g(x - 1, y) = 0$ if $x < 1$).

A. Dynamic numerical integration

Here, we suggest to construct the state space in a dynamic manner up to a certain bound. We discretize time and integrate over small time steps. In the first step, state (x_0, y_0) has probability 1 and, for a time step $h > 0$, we integrate Eq. (2) by considering only those states that have a non-zero probability during the next h time units. For the numerical integration, we use an explicit fourth-order Runge-Kutta method and thus in each step we add the states within a distance of four transitions from the current set of states.

The standard explicit fourth-order Runge-Kutta method applied to Eq. (2) yields the integration step [28]

$$\begin{aligned} p^{(t+h)}(x,y) = p^{(t)}(x,y) + \tfrac{h}{6}\Big(k^{(1)}(x,y) \\ +2k^{(2)}(x,y)+2k^{(3)}(x,y)+k^{(4)}(x,y)\Big), \end{aligned} \tag{4}$$

where $h > 0$ is the time step of the method. For $i \in \{1,2,3,4\}$ the values $k^{(i)}(x,y)$ are defined recursively as

$$\begin{aligned} k^{(1)} &= \mathcal{M}(p^{(t)}), \\ k^{(2)} &= k^{(1)} + \tfrac{h}{2}\mathcal{M}(k^{(1)}), \\ k^{(3)} &= k^{(1)} + \tfrac{h}{2}\mathcal{M}(k^{(2)}), \\ k^{(4)} &= k^{(1)} + h\mathcal{M}(k^{(3)}). \end{aligned} \tag{5}$$

Let $\tilde{p}^{(t)}$ be the approximation of $p^{(t)}$ at time t. Given $\tilde{p}^{(t)}$, we integrate Eq. (2) for h time units as follows. Let $S \subseteq \mathbb{N}^2$ be the set of states (x,y) with $\tilde{p}^{(t)}(x,y) > 0$. Each state $(x,y) \in S$ is represented as an array with entries $k^{(1)}(x,y), \ldots, k^{(4)}(x,y)$, $p(x,y)$. The former four entries are initialized with 0 while $p(x,y) = \tilde{p}^{(t)}(x,y)$ due to the previous integration step and, in the first integration step, $p(x,y) = 1$ if $x = x_0$, $y = y_0$ and $p(x,y) = 0$ otherwise. We have five substeps in which we go over all elements of S to compute for each state the four k-values as well as $\tilde{p}^{(t+h)}(x,y)$. In the first substep, each state (x,y) adds the three outflow probabilities $\alpha x p(x,y)$, $\beta xy p(x,y)$, and $\gamma y p(x,y)$ to the array element $k^{(1)}$ of the corresponding successor states $(x+1,y)$, $(x-1,y+1)$, and $(x,y-1)$ (see Eqs. (3) and (5)). Whenever a successor state is not in S, we add it to S. For $i \in \{2,3\}$, in the i-th substep we first add $k^{(1)}(x,y)$ to the element for $k^{(i)}(x,y)$ and then add $\frac{h}{2}\alpha x k^{(i-1)}(x,y)$, $\frac{h}{2}\beta xy k^{(i-1)}(x,y)$, and $\frac{h}{2}\gamma y k^{(i-1)}(x,y)$ to the $k^{(i)}$-field of the corresponding successors $(x+1,y)$, $(x-1,y+1)$, and $(x,y-1)$ (again we add successors to S whenever they do not yet belong to S). The fourth substep is identical to the second and third except that $\frac{h}{2}$ is replaced by h (cf. Eq. (5)). In the fifth substep, we compute for each state (x,y) the probability $\tilde{p}^{(t+h)}(x,y)$ according to Eq. (4).

The dynamic numerical integration procedure described above yields accurate approximations of the transient probability distribution $p^{(t)}$ of the infinite Markov process $\{(X(t), Y(t)), t \geq 0\}$. Its drawback, however, is that the size of set S becomes very large since in each integration step states within a distance of four transitions are added. E.g. after a computer time of five hours and a time horizon of $t = 0.317$ we run out of memory on a 64-bit machine with 8 GB of main memory where we used the same parameters as in Fig. 1. At that time instant, the state space contained about 35 million states.

TABLE I
RESULTS OF AN APPROXIMATE DIRECT SOLUTION OF EQ. (2).

δ	run time	$\lvert S\rvert$	error
1e-15	51h 39min	9e5	2e-7
1e-14	40h 51min	7e5	2e-6
1e-12	22h 46min	4e5	9e-5
1e-10	10h 12min	2e5	6e-3

B. Inexact numerical integration

Similar to the approach in [9], we modify the numerical integration procedure described above as follows. During the first four substeps of the integration step, we only add new states to set S if their probability is greater than a small positive threshold δ. This has shown to lead to a significant reduction of the size of S while, for most systems, the approximation error is small. For instance, for $\delta = 10^{-15}$ the size of set S at time $t = 0.317$ is just 0.004% of the size of S for $\delta = 0$, while the total probability that got "lost", when states with probability smaller than δ were ignored, is only 7×10^{-12}.

In Table I we list our results of approximate direct solution of Eq. (2) for different values of δ. We analyzed the system for a time horizon of $t = 10$ and used the same parameters as in Fig. 1. For other parameters, such as those that we used for our experiments in Section VI, the results are similar (not shown). The column labeled $|S|$ lists the average size of set S, i.e. average number of states considered during one integration step. The last column gives the total probability that got "lost" due to the state space truncation. If the numerical integration were exact, the values in the last column would contain the total error of the approximation, i.e. the sum of all absolute errors of the state probabilities. Note that the resulting probability distributions provide full information about the system, i.e. arbitrary moments of the distribution as well as probability of certain events (such as extinction until time t) can be derived.

With the proposed numerical approach, we are able to solve the stochastic Lotka-Volterra model and accurate results are obtained. The method is, however, rather slow whenever the expected populations become large. This is because large populations are represented by a large number of discrete states, even though the relative variance is small. This problem gets worse when the populations oscillate in an even higher range as, for instance, in Fig. 2. On the other hand, a discrete stochastic representation is necessary whenever the populations become small.

Also, if the underlying system is stiff the explicit Runge-Kutta method will perform poorly and should be replaced by implicit finite difference methods. For instance, if the transition rates in the Markov process have very different orders of magnitude, the time step of explicit methods will

be proportional to the fastest time scale. Besides implicit methods, it is possible to apply aggregation techniques that are designed for stiff systems [6, 7].

IV. Deterministic approximation of moments

Deterministic approximations are, besides Monte-Carlo simulation, the analysis techniques with most widespread use. The mathematical justification of the simplest deterministic approximation, mean-field analysis, has first been provided by Kurtz in the context of chemical kinetics [21] where it is applied extensively.

Mean-field analysis relies on the assumption that the expected populations can be approximated by variables that change continuously and deterministically in time. This idea can be generalized to higher moments where the accuracy of the approximation increases as higher moments are included [10].

A. Mean-field approximation

Let $Z(t) = (X(t), Y(t))$ denote the two-dimensional population vector at time t and, for $d \in \mathbb{N}$ let $f : \mathbb{N}^2 \to \mathbb{R}^d$ be a function that is independent of t. Assume that the expectation $E[f(Z(t))]$ exists. From Eq. (2), it is straightforward to derive the relationship[2]

$$\begin{aligned}\tfrac{d}{dt}E[f(Z)] &= \textstyle\sum_{(x,y)\in\mathbb{N}^2} f(x,y)\tfrac{d}{dt}p^{(t)}(x,y)\\ &= E[\alpha X(f(X+1,Y)-f(Z))\\ &\quad +\beta XY(f(X-1,Y+1)-f(Z))\\ &\quad +\gamma Y(f(X,Y+1)-f(Z))].\end{aligned} \tag{6}$$

If $f(x,y) = (x,y)$ we get as a special case that

$$\begin{aligned}\tfrac{d}{dt}E[X] &= \alpha E[X] - \beta E[XY]\\ \tfrac{d}{dt}E[Y] &= \beta E[XY] - \gamma E[Y].\end{aligned} \tag{7}$$

Thus, if $X(t)$ and $Y(t)$ were uncorrelated (which means that $E[XY] = E[X]E[Y]$), we could compute the expected number of individuals using Eq. (1). But if $X(t)$ and $Y(t)$ are highly correlated, the approximation of the mean in Eq. (1) is rather inaccurate. Fig. 1(a), for instance, shows the mean-field approximation of the Markov process whose true expectations are plotted in Fig. 1(b).

For finite time instants t, the mean-field approximation becomes exact for the scaled process $\{N^{-1}Z(t), t \geq 0\}$ as the scaling constant N approaches infinity under the assumption that the rate functions are density dependent [22]. In the context of the stochastic Lotka-Volterra model, the latter assumption means that the constant β depends on the scaling parameter N (which has a natural interpretation such as the total number of individuals) while α and γ are independent of N. Thus, the accuracy of the approximation increases if the populations become large. In Fig. 2 we plot the mean-field approximation versus the true expected populations for $\alpha = \gamma = 1$, $\beta = 0.003$ and initial state $(180, 200)$. For these parameters, both populations stay above 150 at all times in the deterministic model. Here, the mean-field approximation remains very accurate at least until time $t = 15$ (and the same holds for the expected number of preys).

[2]To improve readability, we omit the dependence on t and write Z instead of $Z(t)$ and so on.

B. Method of moments

Eq. (7) suggests that we would obtain a more accurate approximation of the expected number of individuals if we could accurately approximate the value $E[XY]$. The first idea, of course, is to set $f(x,y) = xy$ in Eq. (6) in order to extend the system of differential equations in Eq. (7) by additional equations for $E[XY]$. It is easy to see that with $f(x,y) = xy$ Eq. (6) yields a differential equation that involves $E[X^2Y]$, $E[X^2Y^2]$, and $E[XY^2]$. To approximate these quantities as well, additional equations are necessary that involve even higher moments. This argument repeats and yields an infinite set of differential equations. The idea of method of moments is to truncate this infinite set after a finite number of equations. Here, we improve the quality of the deterministic approximation by enriching the system in Eq. (7) with deterministic approximations of the second moments as proposed by Engblom for Markov population models [10]. Based on a Taylor series expansion of the rate function, the following deterministic approximation of second order can be shown:

$$\begin{aligned}\tfrac{d}{dt}E[X] &= \alpha E[X] - \beta E[X]E[Y] - \beta c_{XY}\\ \tfrac{d}{dt}E[Y] &= \beta E[X]E[Y] + \beta c_{XY} - \gamma E[Y]\\ \tfrac{d}{dt}c_{XX} &= 2\alpha c_{XX} + \alpha E[X] - 2\beta E[X]c_{XY}\\ &\quad -2\beta E[Y]c_{XY} + \beta E[X]E[Y] + \beta c_{XY}\\ \tfrac{d}{dt}c_{XY} &= \alpha c_{XY} - \beta E[Y]c_{XY} - \beta E[X]c_{YY}\\ &\quad +\beta E[Y]c_{XX} + \beta E[X]c_{XY}\\ &\quad -\beta E[X]E[Y] - \beta c_{XY} - \gamma c_{XY}\\ \tfrac{d}{dt}c_{YY} &= 2E[Y]c_{XY} + 2\beta E[X]c_{YY} + \beta E[X]E[Y]\\ &\quad +\beta c_{XY} - 2\gamma c_{YY} - \gamma E[Y],\end{aligned} \tag{8}$$

where c_{XX}, c_{XY}, c_{YY} approximate the covariances $COV[X,X]$, $COV[X,Y]$, $COV[Y,Y]$, respectively. Note that the first two equations are as in Eq. (7) if we use the relationship $COV[V,W] = E[VW] - E[V]E[W]$ for the two random variables V and W.

In Fig. 2 we plot the solution of Eq. (8) together with the mean-field solution as well as the true expectations. For these parameters both the mean-field approximation and the second-order method of moments become inaccurate after $t = 23$. The second-order method of moments predicts the dynamics of the expected predator population only slightly better. The situation is similar for the expected number of preys. Additional equations for higher moments give a more accurate approximation (not shown) but become very stiff once the prey population grows too fast. The main advantages of the method of moments are the low computational cost (the ODE can be solved in a few seconds) and the fact that for many models a second-order approximation is sufficient. If,

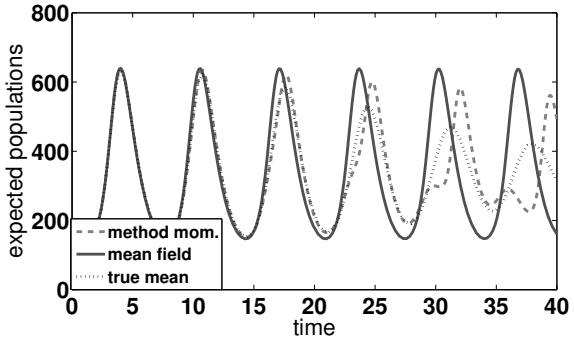


Fig. 2. The deterministic approximations of the expected number of predators remain accurate until $t = 23$. Here, we used the parameters $\alpha = \gamma = 1$, $\beta = 0.003$, and initial state $(180, 200)$.

however, besides the moments, certain probabilities are of interest (such as the probability of extinction) other solution techniques need to be chosen. Also, for the model under study higher order equations are necessary but lead to numerical instabilities. E.g. for the parameters considered in Fig. 1, the fourth-order approximation gives a matrix that is singular to working precision (while the second-order approximation yields poor accuracy). More advanced numerical techniques have to be considered in this case such as implicit finite difference methods and integration methods that are particularly designed for oscillatory systems [2–5, 29].

V. Stochastic hybrid method

In this section, we propose a hybrid solution technique for the stochastic Lotka-Volterra model that is motivated by two observations. First, the approximate direct numerical method presented in Section III-B is able to provide accurate results. However, if only one of the expected populations (say $E[Y(t)]$) is small and the other population is high, then it becomes inefficient because the number of "significant" states is large. It treats $X(t)$ and $Y(t)$ as discrete stochastic (DS) random variables and considers all possible values up to a predefined accuracy $\delta > 0$, i.e. the infinite ranges of $X(t)$ and $Y(t)$ are truncated with respect to δ. Second, the method of moments with a truncation of moments higher than order two provides a fast continuous deterministic (CD) approximation that is accurate whenever the expectations of $X(t)$ and $Y(t)$ are high.

A successful hybrid approach could therefore be based on the following two properties:
(a) It should be able to switch dynamically between a DS representation and a CD representation depending on population thresholds for $E[X(t)]$ and $E[Y(t)]$.
(b) Whenever $X(t)$ is represented as a DS variable and $Y(t)$ as a CD variable (or vice versa), a hybrid approach must be able to take into account the dependencies between $X(t)$ and $Y(t)$ in an appropriate way.

While requirement (a) is easy to realize, (b) turns out to be much more challenging. Assume that at a certain point in time, the predator population is represented by the DS variable $Y(t)$ and the prey population is represented by the CD variable $x(t) \approx E[X(t)]$; that is, we have a (truncated) probability distribution for $Y(t)$ and a single real value $x(t)$ as well as approximations of the covariances c_{XX} and c_{XY}. A straightforward idea is to consider the "global" ODE in Eq. (8) for the computation of $x(t+h) \approx E[X(t+h)]$ as well as the covariances at time $t+h$, where h is a small time step and $E[Y(t)]$ as well as c_{YY} are computed from the current approximation of the distribution of $Y(t)$. The equations for $E[Y(t)]$ and c_{YY} are removed in Eq (8) and the distribution of $Y(t+h)$ is computed by solving a "reduced" version of Eq. (2) given by

$$\begin{aligned} \tfrac{d}{dt}p^{(t)}(y) = {} & \beta(y-1)x(t)p^{(t)}(y-1) \\ & + \gamma(y+1)p^{(t)}(y+1) \\ & - (\beta y x(t) + \gamma y)p^{(t)}(y), \end{aligned} \tag{9}$$

where $y > 0$. For state $y = 0$, we have $\frac{d}{dt}p^{(t)}(0) = \gamma p^{(t)}(1)$. Note that the transition that corresponds to the growth of prey is not included in Eq. (9) but is taken into account in Eq. (8). We can integrate Eq. (9) using the method described in Section III-B. In this way, the possible values for y and thus the number of equations would remain small (since we assumed that $E[Y(t)]$ and h are small) which makes the approach computationally cheap. For instance, the run time needed to approximate the probability distribution for the parameters used in Section I is only about 1 minute. However, it turns out that the approach yields bad accuracy. For instance, at the final time instant $t = 10$, the relative errors of the first moments of prey and predator are above 20%. In particular, the approach fails to detect the steep increase of the expected number of preys after $t = 6$ even though the results are accurate within $[0, 6]$. The reason is that when $E[Y(t)]$ is small, the (relative) variance of $Y(t)$ is important for the evolution of the CD variable $x(t)$ (and vice versa). For instance, if $Y(t) = 0$, then $X(t)$ grows exponentially at rate α. In this case, the total rate of change will therefore deviate largely from the rate $\frac{d}{dt}E[X]$ used in Eq. (8).

A more accurate approach is to consider "local" ODEs, i.e. if $X(t)$ has a CD representation and $Y(t)$ has a DS representation, then we consider the conditional expectation $E[X(t)|Y(t) = y]$ for each state y. Thus, in each integration step, we represent the current state of the system by the probabilities $p^{(t)}(y) \approx P\{Y(t) = y\}$ (but we still neglect probabilities smaller than or equal to δ) and real values $x_y(t)$ that approximate $E[X(t)|Y(t) = y]$. For a small time step h, we integrate the distribution $p^{(t)}$ and the values $x_y(t)$ in three substeps.
(1) We first integrate $p^{(t)}$ according to Eq. (9) to approximate the probabilities $P\{Y(t+h) = y\}$ by $p^{(t+h)}(y)$.
(2) For each state y with $p^{(t)}(y) > \delta$, we compute $\tilde{x}_y(t+h)$ by numerical integration of the ODE

$$\tfrac{d}{dt}\tilde{x}_y(t) = \alpha\tilde{x}_y(t) - \beta y\tilde{x}_y(t)$$

with initial condition $\tilde{x}_y(t) = x_y(t)$. Note that $\tilde{x}_y(t+h)$ is *not* an approximation of $E[X(t+h)|Y(t+h) = y]$ since the above

differential equation does not take into account that Y may leave state y within $[t, t+h]$. E.g. for newly discovered states with $p^{(t)}(y) < \delta$ and $p^{(t+h)}(y) > \delta$ the value $x_y(t)$ (and thus $\tilde{x}_y(t+h)$) does not exist. Thus, a third substep is necessary to approximate $E[X(t+h)|Y(t+h) = y]$.

(3) We compute $x_y(t+h)$ by "distributing" $\tilde{x}_y(t+h)$ according to the change in the distribution of Y as explained below. If we assume that $[t, t+h]$ is an infinitesimal time interval and, for $y \neq y'$, $q(y, y', h)$ is the probability to enter y from y' within $[t, t+h]$, then

$$P\{Y(t+h)=y\} = \sum_{y' \neq y} q(y, y', h) P\{Y(t)=y'\} + (1 - \sum_{y' \neq y} q(y', y, h)) P\{Y(t)=y\}. \quad (10)$$

Thus, we approximate $E[X(t+h)|Y(t+h) = y]$ as

$$\sum_{y' \neq y} \tilde{x}_{y'}(t+h) q(y, y', h) P\{Y(t)=y'|Y(t+h)=y\} + \tilde{x}_y(t+h)(1 - \sum_{y' \neq y} q(y', y, h)) P\{Y(t)=y|Y(t+h)=y\}. \quad (11)$$

The rationale behind this approximation is that, for $t' \in [t, t+h)$, the variance $VAR[X(t')|Y(t') = y]$ is small and the conditional distributions $P\{X(t') = \hat{x} \mid Y(t') = y\}$ can be approximated by a normal distribution with mean $E[X(t')|Y(t') = y]$.

Obviously, we make use of the current approximations $p^{(t)}$ and $p^{(t+h)}$ to compute the conditional probabilities $P\{Y(t)=y'|Y(t+h) = y\}$. For a small time step h, $q(y, y+1, h) \approx h\beta y \tilde{x}_y(t)$ and $q(y, y-1, h) \approx h\gamma y$. Using Eq. (11), we define $x_y(t+h) \approx E[X(t+h)|Y(t+h) = y]$ as

$$x_y(t+h) = \Big(h\beta(y-1) p^{(t)}(y-1) \tilde{x}_{y-1}(t+h) \tilde{x}_{y-1}(t) + h\gamma(y+1) p^{(t)}(y+1) \tilde{x}_{y+1}(t+h) + (1 - hy(\beta \tilde{x}_y(t) + \gamma)) p^{(t)}(y) \tilde{x}_y(t+h) \Big) / p^{(t+h)}(y).$$

We illustrate the three substeps above by means of a simple example.

Assume that the distribution of $Y(t)$ is such that $p^{(t)}(0) = \frac{1}{3}$, $p^{(t)}(1) = \frac{2}{3}$ and all other states y have probability 0 at time t. Assume further that the CD variables $x_y(t)$ are such that $x_0(t) = 600$ and $x_1(t) = 500$. For the parameters $\alpha = \gamma = 1$, $\beta = 0.01$, $h = 0.1$ substep (1) yields $p^{(t+h)}(0) = \frac{2}{5}$, $p^{(t+h)}(1) = \frac{4}{15}$, $p^{(t)}(2) = \frac{1}{3}$ where, for simplicity, we integrate Eq. (9) with the Euler method. Substep (2) yields $\tilde{x}_0(t+h) = 660$ and $\tilde{x}_1(t+h) = 549.5$ and substep (3) gives $x_0(t+h) \approx 641.58$ and $x_1(t+h) = x_2(t+h) = 549.5$. Note that state $y = 2$ "inherits" the x-value of state $y = 1$ since its probability inflow of $\frac{1}{3}$ originated from $y = 2$.

The same strategy as explained above can be used for the case where $X(t)$ is a DS variable and $y(t)$ is a CD variable.

We performed experiments using different parameters to test the accuracy and run time if the representation is hybrid (coexistence of CD and DS variables) and also if switching occurs often. For the parameters used in Section I, for instance, we start with a purely stochastic representation and use a population threshold of 350 to switch the representation. At time $t \approx 3.3$ we switch the representation of $X(t)$ from DS to CD since the expectation reaches 350 (see also Fig. 1(b)). Around time $t = 4$, we switch back to a DS representation. Then, around $t = 4.3$ we switch the representation of $Y(t)$ from DS to CD because $E[Y(t)] > 350$, etc.

TABLE II
RESULTS OF THE STOCHASTIC HYBRID SOLUTION.

pop. thresh.	δ	run time time	$\lvert S\rvert$	relative error of moments 1st	2nd	3rd
100	$1e$-15	12s	$9e2$	0.04	0.11	0.22
	$1e$-10	7s	$6e2$	0.04	0.11	0.22
200	$1e$-15	1h 43min	$5e4$	0.03	0.11	0.23
	$1e$-10	11min	$9e3$	0.03	0.11	0.23
400	$1e$-15	2h 35min	$1e5$	0.03	0.10	0.20
	$1e$-10	25min	$3e4$	0.03	0.11	0.21

We also performed experiments using different population thresholds and different values for δ. We summarize these results in Table II. We always compared our results with the full stochastic solution described in Section III-B to estimate the accuracy of the hybrid method. We list the relative errors of the first three moments when the hybrid method is compared to the full stochastic solution. For these parameters it turned out that the accuracy of the hybrid approach is less sensitive to population threshold than expected. In the range of $100-400$, the relative error of the first three moments is at most 23%. Similarly, the moments are accurately approximated even if δ is about 10^{-10}. It should, however, be noted that if a higher value is chosen for δ, then small event probabilities will be approximated as 0.

The accuracy of the hybrid approach becomes worse whenever both variables are represented deterministically. The reason is that the second-order method of moments relies heavily on an accurate prediction of the (co-)variances. But during a preceding hybrid phase where, say, X has CD representation and Y has DS representation, the variance of X is much smaller than the real variance $VAR[X(t)]$, i.e. the approximation

$$VAR[X(t)] \approx \sum_{y:p^{(t)}(y)>\delta} p^{(t)}(y)(x_y(t) - E[X(t)])^2$$

is not accurate enough if it is used as an initial value for c_{xx} in Eq. (8) even though this has only a minor effect on accuracy during the hybrid phase. The hybrid method exploits the fact that, if X is large, then the *relative* variance is small and it is sufficient to record the evolution of $E[X(t) \mid Y(t) = y]$ for each value y that has a significant probability. This seems to contradict with the approximation in Eq. (8) that relies on a good estimate for the (co-)variances. For this reason, we suppressed the configuration CD, CD for our results in Table II and always kept at least one variable DS. We plan to extend the stochastic hybrid approach such that the distribution $P\{X(t) = x \mid Y(t) = y\}$ is approximated by a normal distribution which we represent by two values (mean and variance) instead of the single value $x_y(t)$. This, however,

complicates the interdependencies between X and Y further and its study is beyond the scope of this paper.

VI. Experimental results

We performed experiments using five different parameter combinations for the stochastic Lotka-Volterra model and four different solution methods. In Table III we list our experimental results. Each row corresponds to one combination and is enumerated in the first column. We used the first and the second combination of parameters also in Section III and IV (see also Fig. 1 and Fig. 2). Note that for the latter combination, the expected populations oscillate in a much higher range. The remaining parameter combinations in rows 3 through 5 are taken from the literature [14] (third row), [15] (fourth row), [12] (fifth row). In columns two to four we list the rate constants, and the columns with labels x_0 and y_0 refer to the initial population sizes. In the column with label t_f, we list the final time instant at which we compute the transient distribution. We abbreviate the solutions methods as follows:

- IRK4(δ) refers to the inexact explicit fourth-order Runge-Kutta method as described in Subsection III-B using a significance threshold of δ. We list the run time, the average number of significant states and the error, i.e., the probability mass lost during the computation for the case $\delta = 10^{-15}$.
- MF refers to the mean field analysis discussed in Subsection IV-A. We list the run time and the relative error of the first moment (expected populations) at the final time instant where we compare to IRK4(10^{-15}). We took the average over the two species.
- MM refers to the method of moments that we discussed in Subsection IV-B. We list the run time and the relative error of the first and second moments where we again compare to IRK4(10^{-15}) and average over the two species.
- Finally, SH(δ,K) refers to the stochastic hybrid method where a significance threshold of δ and a population threshold of K is used. Here, we give results for $\delta = 10^{-15}$ and $K = 200$. The column with label $|S|$ lists the average number of significant states in each step and the last three columns list the relative error of the first, second, and third moments compared to those of IRK4(10^{-15}).

For the first combination of parameters the two deterministic approxmations MF and MM yield fast but poor results. The stochastic hybrid method gives a significant speed-up compared to IRK4(10^{-15}) (about 51 times faster) and the approximation is accurate. The parameters used in the second row yield a system where the expected populations oscillate in a high range (between 150 and 650, see also Fig. 2). Therefore, MF and MM give accurate solutions. The stochastic hybrid solution is fast since most of the time both populations are represented as continuous deterministic values. The third combination of parameters gives similar results as the first one. In the last two rows, only IRK4(10^{-15}) gives an accurate solution. The stochastic hybrid method accurately predicts the first moments of the species but for the second and third moments the approximation error is high. The reason is that, for both parameter combinations, at the final time instant $E[Y] \approx 0$ and $|S| = 1$ (fourth row) and $|S| = 3$ (fifth row). During the hybrid solution, Y is represented as a discrete stochastic variable and X becomes continuous deterministic. Only a small number of states y remain and the variance between their conditional expectations $E[X|Y = y]$ is small. Thus, the second moment $E[X^2]$ yields a relative error of around 90% whereas $E[Y^2] \approx 0$ is very accurately approximated. The average relative error is then around 50%.

VII. Conclusion

The popularity of the Lotka-Volterra model for the description of population dynamics relies on the fact that this model is able to reflect the strongly changing dynamics of interacting populations. Its stochastic variant includes the possibility of extinction, which is, for instance, important for modeling biodiversity and coevolution [18, 19, 26].

The analysis of the stochastic Lotka-Volterra model is challenging because events such as extinction of predator require a costly discrete representation of the predator population. This and other characteristics of the stochastic Lotka-Volterra model (change between small and large populations, highly correlated variables, infinite state space, unbounded rates) are prototypical for Markov population models. Therefore, its analysis is not only interesting from an application-oriented viewpoint but also from a methodological viewpoint.

We discussed different analysis approaches for the stochastic Lotka-Volterra model and modified them in order to improve accuracy and run time. It turned out that the stochastic hybrid approach with local ODEs performs best. It is at least an order faster than the numerical integration of the master equation and yields no more than a 5 % relative error in the first moment of the transient distribution. If high populations are approximated by continuous deterministic variables, then error estimates are hard to obtain. The benefits of an approximate numerical solution, however, are highly dependent on accurate error estimates. Here, we compared our hybrid solution with a full stochastic solution. In the future, we will try to estimate the accuracy of the hybrid approach without a comparison to a more accurate but costly solution. Moreover, we plan to additionally consider numerical analysis approaches based on a diffusion approximation of the stochastic Lotka-Volterra model such as the one suggested by Ferm et al. [27] and compare them with the methods discussed here.

Acknowledgment

The authors wish to thank Stefan Engblom, Royal Institute of Technology in Stockholm, Sweden, for providing an implementation for the method of moments.

References

[1] R. Alur and T. Henzinger. Reactive modules. *Formal Methods in System Design*, 15:7–48, 1999.

[2] Z. Anastassi and T. Simos. A family of two-stage two-step methods for the numerical integration of the

TABLE III
COMPARISON OF THE DIFFERENT SOLUTION METHODS.

	parameters						IRK4(10^{-15})			MF		MM			SH(10^{-15},200)				
	α	β	γ	x_0	y_0	t_f	run time	$\|S\|$	error	run time	1st	run time	1st	2nd	run time	$\|S\|$	1st	2nd	3rd
1	1	0.01	1	20	20	10	51h 39min	$9e5$	$2e$-7	$< 1s$	> 1	$< 1s$	0.54	0.91	1h 43min	$5e4$	0.03	0.11	0.23
2	1	0.003	1	180	200	10	20h 56min	$8e5$	$5e$-8	$< 1s$	0.07	$< 1s$	0.07	0.27	2min 46s	$1e4$	0.05	0.11	0.16
3	1	0.1	1	5	20	10	46h 32min	$7e4$	$3e$-9	$< 1s$	> 1	$< 1s$	> 1	> 1	1h 30min	$2e4$	0.02	0.16	0.31
4	1	1	6	10	20	10	67h 43min	$5e4$	$4e$-11	$< 1s$	> 1	$< 1s$	> 1	> 1	50min 32s	$6e3$	< 0.01	0.45	0.50
5	0.5	1.3	0.67	2	1	20	56h 15min	$6e4$	$3e$-11	$< 1s$	> 1	$< 1s$	> 1	> 1	5min 39s	$1e3$	< 0.01	0.42	0.49

Schrodinger equation and related IVPs with oscillating solution. *Journal of Mathematical Chemistry*, 45(4):1102–1129, 2008.

[3] Z. Anastassi and T. Simos. A six-step p-stable trigonometrically-fitted method for the numerical integration of the radial Schrodinger equation. *MATCH Commun. Math. Comput. Chem.*, 60(3):803–830, 2008.

[4] Z. Anastassi and T. Simos. Numerical multistep methods for the efficient solution of quantum mechanics and related problems. *Phys. Rep.*, 482-483:1–240, 2009.

[5] Z. Anastassi, T. Simos, and G. Panopoulos. Two optimized symmetric eight-step implicit methods for initial-value problems with oscillating solutions. *Journal of Mathematical Chemistry*, 46(2):604–620, 2009.

[6] A. Bobbio and K. S. Trivedi. An aggregation technique for the transient analysis of stiff Markov chains. *IEEE Transactions on Computers*, C-35(9):803–814, 1986.

[7] H. Busch, W. Sandmann, and V. Wolf. A numerical aggregation algorithm for the enzyme-catalyzed substrate conversion. In *Proc, of CMSB*, volume 4210 of *LNCS*, pages 298–311. Springer, 2006.

[8] E. Çinlar. *Introduction to Stochastic Processes*. Prentice-Hall, 1975.

[9] F. Didier, T. A. Henzinger, M. Mateescu, and V. Wolf. Fast adaptive uniformization of the chemical master equation. In *Proc. of HIBI*, pages 118–127. IEEE Computer Society, 2009.

[10] S. Engblom. Computing the moments of high dimensional solutions of the master equation. *Appl. Math. Comput.*, 180:498–515, 2006.

[11] P. Erdi and J. Toth. *Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models*. Manchester University Press, 1989.

[12] C. Evans and G. Findley. A new transformation for the Lotka-Volterra problem. *J. Math. Chem.*, 25:105–110, 1999.

[13] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361, 1977.

[14] N. S. Goel, S. C. Maitra, and E. W. Montroll. On the Volterra and other nonlinear models of interacting populations. *Rev. Mod. Phys.*, 43(2):231–276, 1971.

[15] S. E. Hitchcock. Extinction probabilities in predator-prey models. *J. Appl. Prob.*, 23(1):1–13, 1986.

[16] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.

[17] N. G. v. Kampen. *Stochastic Processes in Physics and Chemistry.* Elsevier, 3rd edition, 2007.

[18] B. Kerr, M. A. Riley, M. W. Feldman, and B. J. M. Bohannan. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature*, 418:171–174, 2002.

[19] B. Kirkup and M. A. Riley. Antibiotic-mediated antagonism leads to a bacterial game of rock-paper-scissors in vitro. *Nature*, 428:412–414, 2004.

[20] F. C. Klebaner. *Introduction to stochastic calculus with applications*. Imperial College Pr., 2005.

[21] T. G. Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.*, 57(7):2976 –2978, 1972.

[22] T. G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes Appl.*, 6(3):223–240, 1977/78.

[23] M. Kwiatkowska, G. Norman, and D. Parker. Prism: Probabilistic model checking for performance and reliability analysis. *ACM SIGMETRICS Performance Evaluation Review*, 36(4):40–45, 2009.

[24] A. J. Lotka. *Elements of Mathematical Biology*. Williams and Wilkins Company, 1924.

[25] M. Rabinovich, R. Huerta, and G. Laurent. Transient dynamics for neural processing. *Science*, 321(5885):48–50, 2008.

[26] T. Reichenbach, M. Mobilia, and E. Frey. Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model. *Phys. Rev. E*, 74:51907–51918, 2006.

[27] P. Sjöberg, P. Lötstedt, and J. Elf. Fokker-Planck approximation of the master equation in molecular biology. Technical Report 2005-044, Department of Information Technology, Uppsala University, 2005.

[28] W. J. Stewart. *Introduction to the Numerical Solution of Markov Chains.* Princeton University Press, 1995.

[29] H. Van de Vyver. A fourth-order symplectic exponentially fitted integrator. *Comput. Phys. Commun.*, 174(4):255–262, 2006.

[30] V. Volterra. Fluctuations in the abundance of a species considered mathematically. *Nature*, 118:558–560, 1926.