Report Number 11/40

Higher-order numerical methods for stochastic simulation of chemical reaction systems

by

Tamás Székely Jr, Kevin Burrage, Radek Erban, and Konstantinos C. Zygalakis

UNIVERSITY OF
OXFORD

Oxford Centre for Collaborative Applied Mathematics
Mathematical Institute
24 - 29 St Giles'
Oxford
OX1 3LB
England

Higher-order numerical methods for stochastic simulation of chemical reaction systems

Tamás Székely Jr.,[1, a)] Kevin Burrage,[1, 2] Radek Erban,[3] and Konstantinos C. Zygalakis[3]

[1)] *Department of Computer Science, University of Oxford, Oxford, OX1 3QD, United Kingdom*

[2)] *Department of Mathematics, Queensland University of Technology, Brisbane, Qld 4001, Australia*

[3)] *Mathematical Institute, University of Oxford, Oxford, OX1 3LB, United Kingdom*

In this paper, using the framework of extrapolation, we present an approach for obtaining higher-order τ-leap methods for the Monte Carlo simulation of stochastic chemical kinetics. Specifically, Richardson extrapolation is applied to the expectations of functionals obtained by a fixed-step τ-leap algorithm. We prove that this procedure gives rise to second-order approximations for the first two moments obtained by the chemical master equation for zeroth- and first-order chemical systems. Numerical simulations verify that this is also the case for higher-order chemical systems of biological importance. This approach, as in the case of ordinary and stochastic differential equations, can be repeated to obtain even higher-order approximations. We illustrate the results of a second extrapolation on two systems. The biggest barrier for observing higher-order convergence is the Monte Carlo error; we discuss different strategies for reducing it.

Keywords: stochastic simulation algorithms, τ-leap, high-order methods, Monte Carlo error

---

[a)] Electronic mail: tamas.szekely@cs.ox.ac.uk

I. INTRODUCTION

Biochemical systems with small numbers of interacting components have increasingly been studied in recent years, as they are some of the most basic systems in cell biology[1–3]. They are inherently stochastic; applying deterministic ordinary differential equation (ODE) models, which approximate particle numbers as continuous concentrations, to such systems can lead to confusing results[4,5]. In some cases, even systems with larger populations cannot be accurately modelled by ODEs, for instance when close to a bifurcation regime as in Erban *et al.*[6]. In this model, for some parameter values the ODE approximation cannot reproduce the behaviour of the system. Gillespie's stochastic simulation algorithm (SSA)[7] is an exact method for modelling reactions in a well-stirred stochastic system. Each simulation gives one possible trajectory of the system as it evolves in time.

The SSA can be computationally intensive to run for realistic problems, and alternative methods such as the τ-leap have been developed to speed it up[8]. Instead of simulating each reaction, the τ-leap performs several reactions at once, thus 'leaping' along the history axis of the system. This means that, unlike the SSA, the τ-leap is not exact; accuracy is maintained by only allowing a few reactions per step. The size of each timestep, τ, determines the number of reactions occurring during that step, given by a Poisson random number. This allows one to use fewer steps per simulation, reducing computational time.

This gain in speed must be balanced with loss of accuracy: larger steps mean fewer calculations but reduced accuracy. The original τ-leap proposed by Gillespie[8] was improved by Gillespie and Petzold[9] by the imposition of new conditions to select τ. In addition, Rathinam *et al.*[10] developed an implicit τ-leap method, which is useful for stiff systems. In the original τ-leap the molecular populations can go negative: the binomial τ-leap proposed by Tian and Burrage[11] and Chatterjee *et al.*[12] is one way of solving this problem, as binomial samples are bounded. In addition, Cao *et al.*[13] have avoided the problem by switching to a modified SSA for critical reactions which involve low chemical populations. More recently, Cao *et al.*[14] proposed an updated version of the τ-leap algorithm, where τ is given by molecular populations rather than the propensity functions.

It has been shown that the τ-leap algorithm has weak order of convergence one[15,16]. Recently, Hu and Li[17] introduced the random-corrected τ-leap, where at each timestep a random correction is added to the Poisson random number which determines the number of reactions in that step. Given a suitable random correction, the lowest order errors on the moments can be cancelled. In this way they constructed methods with up to weak order two for both mean and covariance.

The main idea in this paper is to combine Richardson extrapolation, an extrapolation technique, with a fixed-step τ-leap method in order to obtain higher-order approximations to the moments of the chemical master equation. Richardson extrapolation is a well-known technique for improving the order of accuracy of a numerical solver by combining sets of simulations with different stepsizes. It was originally developed for ODEs but was also successfully applied to stochastic differential equations[18] (SDEs). We extend the extrapolation framework to the τ-leap algorithm by evaluating the moments of sets of simulations of a fixed-step τ-leap and extrapolating them to get a weak order two estimator. We show that the first two moments of the extrapolated τ-leap method are always order two in the case of linear propensity functions. We also demonstrate that this is true for some non-linear test systems by using numerical simulations. In addition, we show that higher-order τ-leap methods can be constructed using this framework.

The rest of this paper is organized as follows. In Section II we give an overview of the SSA and τ-leap methods. In Section III we introduce the extrapolated τ-leap and show that it approximates the first two moments with weak order two in the case of zeroth- and first-order reactions, and in Section IV give some numerical results to support this. Section V introduces a useful way of splitting up the total errors to reveal the underlying bias of a numerical method. We discuss our findings and give our conclusions in Section VI.

II. BACKGROUND

Gillespie's SSA[7] is a statistically exact method for simulating paths from a Markov jump process. The two basic assumptions of the SSA are (i) that individual molecules are not explicitly tracked, and (ii) there are frequent non-reactive collisions. Thus we assume that the system is well-mixed and homogeneous.

The SSA models a system of biochemical reactions with N species and M sets of reactions, chemically interacting inside a fixed volume Ω at constant temperature. The populations of chemical species (as molecular numbers, not concentrations) at time t are represented as a state vector $\mathbf{x} \equiv \mathbf{X}(t) \equiv (X_1, \ldots, X_N)^T$. Reactions are represented by a stoichiometric matrix $\boldsymbol{\nu}_j \equiv (\nu_{1j}, \ldots, \nu_{Nj})^T$, where $j = 1, \ldots, M$, composed of M individual stoichiometric vectors. Each stoichiometric vector represents a reaction j occurring and the system changing from state $\mathbf{x}$ to $\mathbf{x} + \boldsymbol{\nu}_j$. Each reaction occurs in an interval $[t, t + dt)$ with relative probability $a_j(\mathbf{x})dt$, where a_j is the propensity function of the j-th reaction. Propensity functions are given by the mass-action kinetics of the reactant chemical species. For more detail, the reader is referred to Ref. 7. Here $\mathbf{X}(t) = \mathbf{x}, \boldsymbol{\nu}_j$ and $a_j(\mathbf{x})$ fully characterise the system at each point in time.

The Direct Method of Gillespie[7] calculates two quantities at each timestep: the time until the next reaction, and its index. The time until the next reaction is given by

$$dt = \frac{1}{a_0} \ln\left(\frac{1}{r_1}\right), \tag{1}$$

and the index j of the next reaction is calculated from

$$\sum_{m=1}^{j-1} a_m(\mathbf{x}) < a_0 r_2 \leq \sum_{m=1}^{j} a_m(\mathbf{x}), \tag{2}$$

where r_1 and r_2 are random numbers drawn from the unit-interval uniform distribution $\mathcal{U}(0, 1)$ and $a_0(\mathbf{x}) = \sum_{j=1}^{M} a_j(\mathbf{x})$. One SSA simulation consists of repeatedly calculating (1) and (2) and modifying $\mathbf{x}$ and $a_j(\mathbf{x})$ accordingly, thus evolving $\mathbf{x}$ through time. The Direct Method requires two newly-generated random numbers at each timestep. Although there are other SSA implementations, such as the Next Reaction Method[19] and the Optimised Direct Method[20], which can be more economical, in general the SSA is computationally costly.

The τ-leap algorithm leaps along the history axis of the SSA by evaluating groups of reactions at once[8]. This means significantly fewer calculations, i.e. shorter computational time, per simulation, but simulation accuracy is compromised: we can no longer tell exactly when each reaction occurs, only in which timestep. The *leap condition* defines an upper bound for the size of each timestep τ: it must be so small that the propensities do not change significantly for its duration, i.e. the change in state from time t to $t + \tau$ is very small[8]. Since τ is small, the probability $a(\mathbf{x})\tau$ that a reaction occurs during $[t, t + \tau)$ is also small, so the number of times K_j each reaction fires over one timestep can be approximated[8] by $\mathcal{P}(a_j(\mathbf{x})\tau)$, a Poisson random variable with mean and variance $a_j(\mathbf{x})\tau$. Thus the τ-leap algorithm evolves the system through time as

$$\mathbf{X}(t+\tau) = \mathbf{x} + \sum_{j=1}^{M} \boldsymbol{\nu}_j \mathcal{P}(a_j(\mathbf{x})\tau). \tag{3}$$

An issue with this method is that in some cases the molecular populations can go negative, as Poisson random numbers are not bounded. Tian and Burrage[11] and Chatterjee *et al.*[12] suggested that K_j be sampled from a binomial distribution. The binomial random numbers used have the same mean as the Poisson random numbers they approximate, but are bounded. As long as the upper bound is set correctly the problem of negativity can be avoided in most cases. However, there are complications when there are several reactions which decrease the same molecular population. For a satisfactory solution, the total decrease in number of a molecular species should not exceed the population of that species. Cao *et al.*[13] addressed the problem by introducing a critical reaction parameter n_c which is set to a low integer value. If there are any critical reactions which would fire over $[t, t+\tau)$, they switch to a modified SSA for that step. More recently, Yates and Burrage[21] described an alternative way of eliminating negativity: they calculate the probability of each species going negative and thus classify a species as critical or not.

One way of enforcing the leap condition is to bound the expected change in each propensity function by a small percentage ε of the sum of the propensities,

$$|a_j(\mathbf{x}+\boldsymbol{\Lambda}) - a_j(\mathbf{x})| \leq \varepsilon a_0(\mathbf{x}), \tag{4}$$

where $\boldsymbol{\Lambda}$ is the change in state over $[t, t+\tau)$. $0 < \varepsilon \ll 1$ is generally set very small, with typical values between 0.01 and 0.05. Larger values cause an increase in speed and a concurrent loss in accuracy. Eq. (4) may not be strict enough in enforcing the leap condition if the system contains both fast and slow reaction channels. Since the change in each propensity function is bounded by a fraction of the sum of all propensities, we cannot be sure that all of them stay approximately constant, as required by the leap condition. Gillespie and Petzold[9] proposed an improved procedure for selecting τ based on this condition. $\boldsymbol{\Lambda}$ is a random variable; the original τ-selection procedure of Gillespie[8] approximates it by its mean. This is reasonable in many cases, but when its variance is large (i.e. when the molecular populations fluctuate rapidly), it can allow the selection of a τ that is too large and breaks the leap condition. Gillespie and Petzold[9] also bound τ by the standard deviation of $\boldsymbol{\Lambda}$, ensuring that a suitable τ is chosen even under such conditions. More recently, Cao *et al.*[14]

suggested an alternative method for finding τ based on the relative change in molecular populations rather than propensity functions.

The τ-leap methods described in this section all have weak order one[15,16]. In the next section we introduce the extrapolated τ-leap, which gives second-order weak approximations for the first and second moments.

III. EXTRAPOLATED TAU-LEAP METHOD

The fixed-step τ-leap (which we call fsTL from here on) evolves the chemical system in time as in Eq. (3), but instead of attempting to choose the largest timestep that sticks to the leap condition at each time t, it uses a fixed timestep. Clearly it is not as efficient as some of the methods described in section II, but it has the advantage of being easy to both program and analyse. In fact, it is the obvious choice of τ-leap method to extrapolate: as more complex τ-leap methods vary τ at each step, it is not clear how to perform extrapolation on them. We can extrapolate the fsTL in a similar way to ODEs. The extrapolated τ-leap (exTL) involves running two sets of D fsTL simulations for time $T = n\tau$. These sets of approximations are then combined to give higher-order ones. The exact procedure is as follows:

Algorithm 1. Extrapolated τ-leap method

- Step 1: Run D simulations of the fsTL with stepsize τ.
- Step 2: Calculate the first two moments $\mathbb{E}_D(\mathbf{x}_n^\tau)$ and $\mathbb{E}_D(\mathbf{x}_n^\tau(\mathbf{x}_n^\tau)^T)$, where $\mathbb{E}_D(\boldsymbol{f}(\mathbf{x}))$ are vectors given by taking the moments of the final populations of each of D simulations, i.e. $\mathbb{E}_D(\boldsymbol{f}(\mathbf{x})) = \frac{1}{D}\sum_{s=1}^{D} \boldsymbol{f}_s(\mathbf{x})$.
- Step 3: Repeat steps 1 and 2 using stepsize $\tau/2$ to get $\mathbb{E}_D(\mathbf{x}_{2n}^{\tau/2})$ and $\mathbb{E}_D(\mathbf{x}_{2n}^{\tau/2}(\mathbf{x}_{2n}^{\tau/2})^T)$.
- Step 4: Take $\left(2\mathbb{E}_D(\mathbf{x}_{2n}^{\tau/2}) - \mathbb{E}_D(\mathbf{x}_n^\tau)\right)$ and $\left(2\mathbb{E}_D(\mathbf{x}_{2n}^{\tau/2}(\mathbf{x}_{2n}^{\tau/2})^T) - \mathbb{E}_D(\mathbf{x}_n^\tau(\mathbf{x}_n^\tau)^T)\right)$ as the extrapolated approximations to the first two moments.

If higher-order solutions are desired, Algorithm 1 may be repeated using yet smaller stepsizes $\tau/2^k$, with k successive integers greater than one, and the approximations combined using a Romberg table. In Section IV we describe a double-extrapolated τ-leap method which we constructed by using stepsizes $\tau, \tau/2$ and $\tau/4$.

The exTL has weak second-order accuracy for the first two moments for the case of first-order reactions (corresponding to linear propensity functions). In this case the equations for the moments are closed, so the propensity functions can be written as

$$[a_1(\mathbf{x}), a_2(\mathbf{x}), \ldots, a_M(\mathbf{x})]^T = C\mathbf{x} + \mathbf{d}, \tag{5}$$

where $C \in \mathbb{R}^{M\times N}$ and $\mathbf{d} \in \mathbb{R}^M$, and we define $W = \nu C$, i.e. $W \in \mathbb{R}^{N\times N}$. This means, using matrix notation, we can write Eq. (3) at some timestep m as

$$\mathbf{x}^\tau_{m+1} = \mathbf{x}^\tau_m + \nu\mathcal{P}(\tau(C\mathbf{x}^\tau_m + \mathbf{d})),$$

where $\mathbf{x}^\tau_m$ is an approximation to the true state vector $\mathbf{x}(t)$ at the m-th timestep (i.e. time t) with stepsize τ. Taking the expectation of both sides, we see that the fsTL evolves the mean as

$$\begin{aligned}\mathbb{E}(\mathbf{x}^\tau_{m+1}) &= \mathbb{E}(\mathbf{x}^\tau_m) + \nu\mathbb{E}(\mathcal{P}(\tau(C\mathbf{x}^\tau_m + \mathbf{d}))) \\ &= \mathbb{E}(\mathbf{x}^\tau_m) + \nu\mathbb{E}(\mathbb{E}(\mathcal{P}(\tau(C\mathbf{x}^\tau_m + \mathbf{d}))|\mathbf{x}^\tau_m)) \\ &= (\mathbb{I} + \tau W)\mathbb{E}(\mathbf{x}^\tau_m) + \tau\nu\mathbf{d}.\end{aligned} \tag{6}$$

Note that we cannot evaluate the expectation of $\mathcal{P}(\tau(C\mathbf{x}^\tau_m + \mathbf{d}))$ directly: because $\mathbf{x}^\tau_m$ is a random variable, we do not know the distribution of $\mathcal{P}(\mathbf{x}^\tau_m)$. Using the law of total expectation we can condition on $\mathbf{x}^\tau_m$ taking a specific value; $\mathcal{P}(\mathbf{x}^\tau_m)|\mathbf{x}^\tau_m$ does have a Poisson distribution.

The probability density function of the SSA at time t is given by the Chemical Master Equation (CME)[22,23]. The mean $\boldsymbol{\mu}(t) = \mathbb{E}(\mathbf{x}(t))$ can be found from the CME[24]; it evolves as

$$\frac{d\boldsymbol{\mu}(t)}{dt} = W\boldsymbol{\mu}(t) + \nu\mathbf{d}. \tag{7}$$

The solution of this is

$$\boldsymbol{\mu}(t) = e^{Wt}\boldsymbol{\mu}(0) + e^{Wt}\int_0^t e^{-Ws}\nu\mathbf{d}\, ds. \tag{8}$$

At $t = \tau$, this evaluates to

$$\left(\mathbb{I} + \tau W + \tfrac{1}{2}\tau^2W^2 + \ldots\right)\boldsymbol{\mu}(0) + W^{-1}\left(\tau W + \tfrac{1}{2}\tau^2W^2 + \ldots\right)\nu\mathbf{d}.$$

Thus we find the local error on the first timestep to be

$$\boldsymbol{\mu}(\tau) - \mathbb{E}(\mathbf{x}^\tau_1) = \left(\tfrac{1}{2}\tau^2W^2 + \mathcal{O}(\tau^3)\right)\boldsymbol{\mu}(0) + \left(\tfrac{1}{2}\tau^2W + \mathcal{O}(\tau^3)\right)\nu\mathbf{d}. \tag{9}$$

This implies that the global weak error on the mean of the fsTL is $\mathcal{O}(\tau)$. Extrapolating (6) at time $t=\tau$, we find

$$2\mathbb{E}(\mathbf{x}_2^{\tau/2})-\mathbb{E}(\mathbf{x}_1^{\tau})=\left(2(\mathbb{I}+\tfrac{1}{2}\tau W)^2-(\mathbb{I}+\tau W)\right)\boldsymbol{\mu}(0)+\tfrac{1}{2}\tau^2 W\nu\mathbf{d}$$

which gives a local error of

$$\boldsymbol{\mu}(\tau)-\left(2\mathbb{E}(\mathbf{x}_2^{\tau/2})-\mathbb{E}(\mathbf{x}_1^{\tau})\right)=\left(\tfrac{1}{6}\tau^3W^3+\mathcal{O}(\tau^4)\right)\boldsymbol{\mu}(0)+\left(\tfrac{1}{6}\tau^3W^2+\mathcal{O}(\tau^4)\right)\nu\mathbf{d}. \tag{10}$$

The mean of the exTL has local error $\mathcal{O}(\tau^3)$. As we have assumed that all the reactions are zeroth or first order, the moment equations are closed and reduce to ODEs. In this case, looking at local extrapolation is enough to verify higher-order convergence. Thus (10) implies that the global error of the exTL is $\mathcal{O}(\tau^2)$. However to show this for higher-order reactions, we would need an expression for the global error of the form

$$\mathbf{x}(T)-\mathbf{x}_n^{\tau}=\mathbf{e}_k\tau^k+\mathbf{e}_{k+1}\tau^{k+1}+\mathbf{e}_{k+2}\tau^{k+2}+\ldots, \tag{11}$$

where k is a known constant and the $\mathbf{e}_k$ are unknown ones. We derive this expansion for the case of the mean. The solution of the fsTL at time $t=T$ is

$$\mathbb{E}(\mathbf{x}_n^{\tau})=\left(\mathbb{I}+\tau nW+\binom{n}{2}\tau^2W^2+\ldots\right)\boldsymbol{\mu}(0)+W^{-1}(\tau nW+\tfrac{1}{2}\tau^2n^2W^2+\ldots)\nu\mathbf{d}.$$

As $T=n\tau$, the global error is

$$\boldsymbol{\mu}(T)-\mathbb{E}(\mathbf{x}_n^{\tau})=\left(\tfrac{1}{2}\tau TW^2+\mathcal{O}(\tau^2)\right)\boldsymbol{\mu}(0)+\left(\tfrac{1}{2}\tau TW+\mathcal{O}(\tau^2)\right)\nu\mathbf{d}. \tag{12}$$

The extrapolated error is

$$\boldsymbol{\mu}(T)-\left(2\mathbb{E}(\mathbf{x}_{2n}^{\tau/2})-\mathbb{E}(\mathbf{x}_n^{\tau})\right)=\left(\tfrac{1}{6}\tau^2TW^3+\mathcal{O}(\tau^3)\right)\boldsymbol{\mu}(0)+\left(\tfrac{1}{6}\tau^2TW^2+\mathcal{O}(\tau^3)\right)\nu\mathbf{d}, \tag{13}$$

as expected from the local error analysis.

The fsTL evolves the second moment in time as

$$\mathbb{E}(\mathbf{x}_{m+1}^{\tau}(\mathbf{x}_{m+1}^{\tau})^T)=\mathbb{E}\left[\left(\mathbf{x}_m^{\tau}+\nu\mathcal{P}(\tau(C\mathbf{x}_m^{\tau}+\mathbf{d}))\right)\left(\mathbf{x}_m^{\tau}+\nu\mathcal{P}(\tau(C\mathbf{x}_m^{\tau}+\mathbf{d}))\right)^T\right]$$

Using the law of total expectation and writing $S_m^{\tau}=\mathbb{E}(\mathbf{x}_m^{\tau}(\mathbf{x}_m^{\tau})^T)$, $B_m^{\tau}=\mathrm{diag}(C\mathbb{E}(\mathbf{x}_m^{\tau}))$ and $D=\mathrm{diag}(\mathbf{d})$, this gives

$$\begin{aligned}S_{m+1}^{\tau}=&S_m^{\tau}+\tau WS_m^{\tau}+\tau S_m^{\tau}W^T+\tau\boldsymbol{\mu}_m^{\tau}\mathbf{d}^T\nu^T+\tau\nu\mathbf{d}(\boldsymbol{\mu}_m^{\tau})^T+\tau^2WS_m^{\tau}W^T\\&+\tau^2W\boldsymbol{\mu}_m^{\tau}\mathbf{d}^T\nu^T+\tau^2\nu\mathbf{d}(\boldsymbol{\mu}_m^{\tau})^TW^T+\tau^2\nu\mathbf{d}\mathbf{d}^T\nu^T+\tau\nu B_m^{\tau}\nu^T+\tau\nu D\nu^T.\end{aligned} \tag{14}$$

The behaviour of the second moment in time as given by the CME, $S(t) = \mathbb{E}(\mathbf{x}(t)\mathbf{x}^T(t))$, is[25,26]

$$\frac{dS(t)}{dt} = WS(t) + S(t)W^T + \boldsymbol{\mu}(t)\mathbf{d}^T\nu^T + \nu\mathbf{d}\boldsymbol{\mu}^T(t) + \nu B(t)\nu^T + \nu D\nu^T, \tag{15}$$

where $B(t) = \mathrm{diag}(C\boldsymbol{\mu}(t))$. Making the approximation $\frac{dS}{dt}\big|_{t+\tau} = \frac{1}{\tau}\left(S(t+\tau) - S(t)\right)$, (15) becomes

$$S(t+\tau) = S(t) + \tau WS(t) + \tau S(t)W^T + \tau\boldsymbol{\mu}(t)\mathbf{d}^T\nu^T + \tau\nu\mathbf{d}\boldsymbol{\mu}^T(t) + \tau\nu B(t)\nu^T + \nu D\nu^T + \mathcal{O}(\tau^2). \tag{16}$$

It is easy to see that the local error $S(t+\tau) - S^\tau_{m+1}$ is $\mathcal{O}(\tau^2)$. However, a higher-order method is necessary to show the higher-order errors of extrapolating. We show the calculations in the Appendix. The extrapolated error of the second moment is indeed $\mathcal{O}(\tau^3)$. As in the case of the mean, this implies a global error on the second moment of $\mathcal{O}(\tau^2)$.

The main requirement for applying Richardson extrapolation to improve the order of accuracy of an estimator is the existence of a global error expansion of the form (11). Obtaining such an expansion is not a trivial task[27], even for ODE numerical solvers. Talay and Tubaro[18] calculated it for the case of the Euler and the Milstein methods for solving SDEs. More recently, Liu and Li[28] have done this for the case of SDEs with Poisson jumps, allowing them to construct higher approximations for $\mathbb{E}(\boldsymbol{f}(\mathbf{x}(T)))$. Their analysis can also be used for the τ-leap method, since in the absence of drift and diffusion the numerical scheme they study is the same as the τ-leap. This is important as it implies that Richardson extrapolation can successfully be applied to numerical solutions of the τ-leap algorithm, as we have shown for zeroth- and first-order reactions.

IV. NUMERICAL RESULTS

We now illustrate the accuracy of the exTL method with some example systems. We simulate the systems with both the fsTL and the exTL for various stepsizes τ over time $t = [0, T]$. We plot the absolute weak errors in the mean and second moment, i.e.

$$|\mathbb{E}(\mathbf{x}(T)) - \mathbb{E}_D(\mathbf{x}^\tau_n)|, \quad |\mathbb{E}(\mathbf{x}(T)\mathbf{x}^T(T)) - \mathbb{E}_D(\mathbf{x}^\tau_n(\mathbf{x}^\tau_n)^T)| \tag{17a}$$

for the fsTL and

$$\begin{gathered} |\mathbb{E}(\mathbf{x}(T)) - 2\mathbb{E}_D(\mathbf{x}^{\tau/2}_{2n}) + \mathbb{E}_D(\mathbf{x}^\tau_n)|, \\ |\mathbb{E}(\mathbf{x}(T)\mathbf{x}^T(T)) - 2\mathbb{E}_D(\mathbf{x}^{\tau/2}_{2n}(\mathbf{x}^{\tau/2}_{2n})^T) + \mathbb{E}_D(\mathbf{x}^\tau_n(\mathbf{x}^\tau_n)^T)| \end{gathered} \tag{17b}$$

for the exTL. Here $\mathbf{x}(T)$ is the analytical solution at time T and $\mathbb{E}_D(\boldsymbol{f}(\mathbf{x}_n^\tau))$ are the moments of its approximations given by D simulations of the fsTL method with stepsize τ run for n steps. For the linear systems, the true solution is calculated analytically; for the non-linear systems we use the value given by 10^6 or 10^7 repeats of the SSA (depending on the system). The error of a weak order α method with stepsize τ is approximately $C\tau^\alpha$, where C is an unknown constant. To easily see the order of accuracy of our results, we plot all the errors on log-log plots. Gradients are calculated using a least squares fit through the points.

A. Particle decay system

A simple test system is a particle decay,

$$X \xrightarrow{k} \emptyset, \quad k = 0.1. \tag{18}$$

The initial particle number $X(0) = 10^4$ and the simulation time $T = 10.4$. The units here, and in the systems below, are non-dimensional. This system is useful only as a test problem, but is first-order and easily tractable. The analytical mean and second moment are

$$\begin{aligned}\mathbb{E}(X(T)) &= X(0)e^{-kT},\\ \mathbb{E}(X(T)^2) &= X(0)e^{-kT} - X(0)e^{-2kT} + (X(0))^2e^{-2kT}.\end{aligned}$$

Fig. 1 shows the errors in the mean and second moment of this system for 10^8 simulations using $\tau = 0.05, \dots, 0.8$. In both cases, the fsTL gives first-order errors and the exTL gives approximately second-order errors.

B. Chain decay system

This system consists of five chemical species, each undergoing a first-order reaction. It forms a closed chain of linear reactions:

$$\begin{aligned}X_1 &\xrightarrow{k_1} X_2, \quad k_1 = 0.3,\\ X_2 &\xrightarrow{k_2} X_3, \quad k_2 = 0.3,\\ X_3 &\xrightarrow{k_3} X_4, \quad k_3 = 0.3,\\ X_4 &\xrightarrow{k_4} X_5, \quad k_4 = 0.3,\\ X_5 &\xrightarrow{k_5} X_1, \quad k_5 = 0.3.\end{aligned} \tag{19}$$

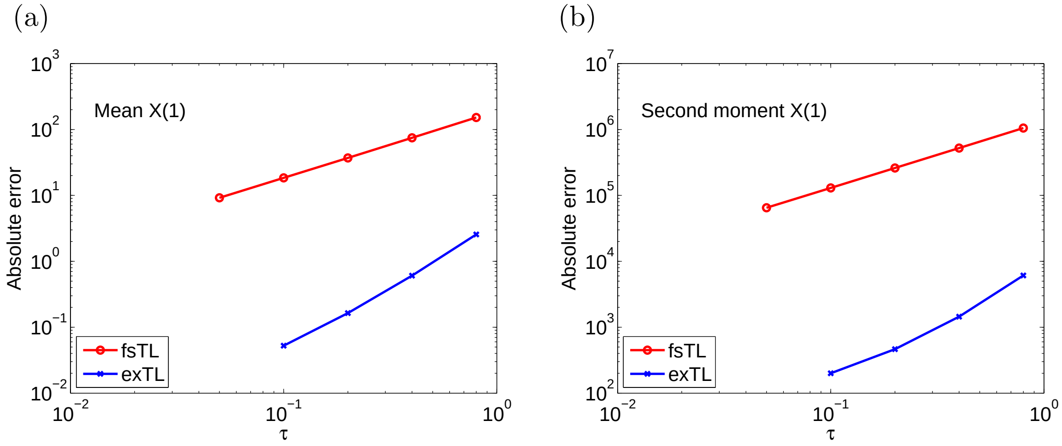

FIG. 1: **Particle decay system.** Absolute error (17) in **(a)** the mean for a fixed-step τ-leap (fsTL, gradient 1.0) and an extrapolated τ-leap (exTL, gradient 1.9); **(b)** the second moment for fsTL (gradient 1.0) and exTL (gradient 1.6). Results from 10^8 simulations, each run for $T = 10.4$.

The simulation time was $T = 16$. The initial populations were

$$\mathbf{X}(0) = (2500, 1875, 1875, 1875, 1875)^T.$$

Since this system is linear, the expectation and covariance of this system can be calculated analytically, as shown in Jahnke and Huisinga[29]; we used these to calculate the true second moment. Fig. 2 shows the errors for the mean and second moment of X_1 using $\tau = 0.1, \ldots, 1.6$. The fsTL is again approximately weak order one and the exTL weak order two for 10^7 simulations.

C. Michaelis-Menten system

The Michaelis-Menten is a common non-linear test system, and represents an enzyme converting a substrate to a product. It is made up of four chemical species in three reactions:

$$\begin{aligned} X_1 + X_2 &\xrightarrow{k_1} X_3, & k_1 &= 10^{-5}, \\ X_3 &\xrightarrow{k_2} X_1 + X_2, & k_2 &= 0.2, \\ X_3 &\xrightarrow{k_3} X_2 + X_4, & k_3 &= 0.2. \end{aligned} \tag{20}$$

Simulation time was $T = 16$ and the initial populations were $\mathbf{X}(0) = (10^4, 2\times10^3, 2\times10^4, 0)^T$. There is no analytical solution, so in this case we approximated it with 10^7 SSA simulations.

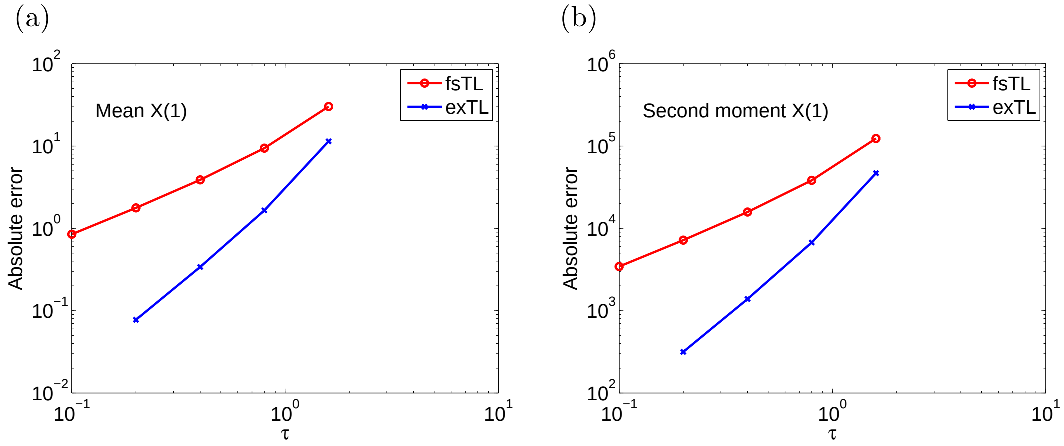

FIG. 2: **Chain decay system.** Absolute error (17) of X_1 in **(a)** the mean for fsTL (gradient 1.3) and exTL (gradient 2.4); **(b)** the second moment for fsTL (gradient 1.3) and exTL (gradient 2.4). Results from 10^7 simulations, each run for $T = 16$.

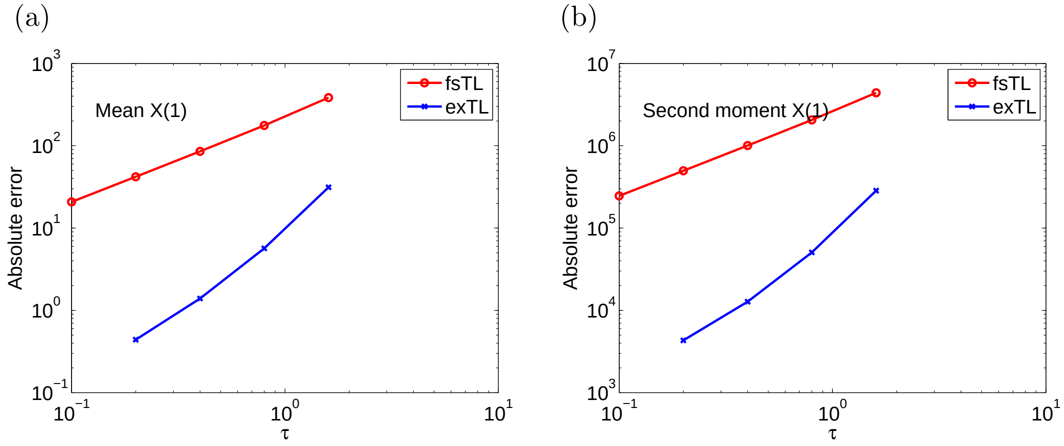

FIG. 3: **Michaelis-Menten system.** Absolute error (17) of X_1 in **(a)** the mean for fsTL (gradient 1.0) and exTL (gradient 2.0); **(b)** the second moment for fsTL (gradient 1.0) and exTL (gradient 2.0). Results from 10^8 simulations, each run for $T = 16$.

Again, we used $\tau = 0.1, \ldots, 1.6$. The errors in the mean and second moment for X_1 shown in Fig. 3. We see that the fsTL converges with order one for 10^8 simulations, and the exTL with order two.

D. Transcriptional regulatory system

This is a more realistic system involving 8 chemical species and 12 reactions[30]. It is a modified version of the transcriptional regulatory system used by Goutsias[31], and is bistable in species A and B. The reactions are

$$\begin{aligned}
E_A &\xrightarrow{k_1} E_A + A, && k_1 = 15,\\
E_B &\xrightarrow{k_2} E_B + B, && k_2 = 15,\\
E_A + B &\underset{k_4}{\overset{k_3}{\rightleftharpoons}} E_AB, && k_3 = 5\times 10^{-4}, k_4 = 2,\\
E_AB + B &\underset{k_6}{\overset{k_5}{\rightleftharpoons}} E_AB_2, && k_5 = 10^{-3}, k_6 = 6,\\
A &\xrightarrow{k_7} \emptyset, && k_7 = 5,\\
E_B + A &\underset{k_9}{\overset{k_8}{\rightleftharpoons}} E_BA, && k_8 = 5\times 10^{-4}, k_9 = 2,\\
E_BA + A &\underset{k_{11}}{\overset{k_{10}}{\rightleftharpoons}} E_BA_2, && k_{10} = 10^{-3}, k_{11} = 6,\\
B &\xrightarrow{k_{12}} \emptyset, && k_{12} = 5.
\end{aligned} \tag{21}$$

We simulated this system for $T = 3.2$ using initial populations of

$$\mathbf{X}(0) = (2\times 10^4, 1.5\times 10^4, 9500, 9500, 2000, 500, 2000, 500)^T,$$

where $\mathbf{X} = (A, B, E_A, E_B, E_AB, E_AB_2, E_BA, E_BA_2)^T$. 10^6 SSA simulations were used as an approximation to the analytical values. We ran 10^7 simulations of the τ-leap for $\tau = 0.005, \ldots, 0.08$ (corresponding to a range of approximately 1000 to 30000 SSA steps). The errors for X_1 in this system are shown in Fig. 4. The fsTL and exTL gave errors of approximately order one and two respectively.

E. Higher-order extrapolation

So far we have only used the most basic extrapolation, going from first to second order. As mentioned in Section III, this can be generalised, leading to higher-order approximations. The usefulness of repeated extrapolations is debatable, as each adds extra overhead and the extra accuracy can be obscured by statistical fluctuations. However it might be reasonable to create up to third or fourth-order methods in this way. We tried double-extrapolating our test systems, with reasonable success. Fig. 5 shows the fsTL, exTL and double-extrapolated

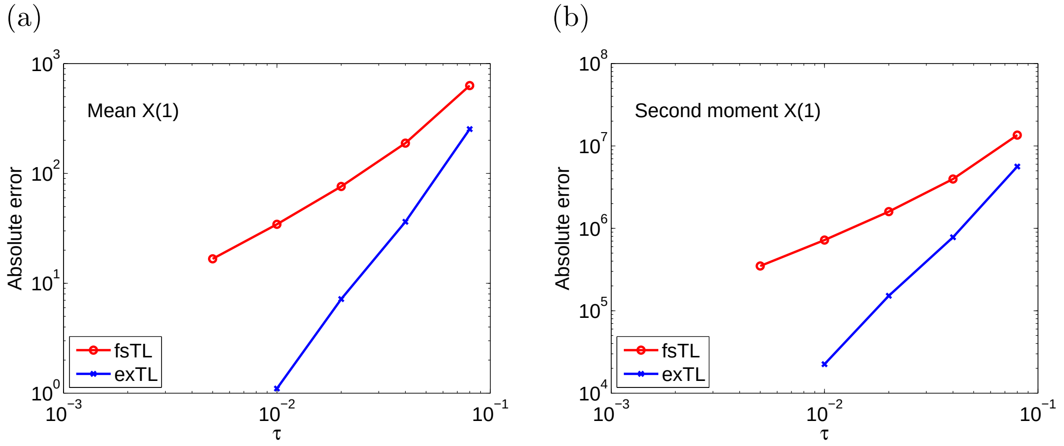

FIG. 4: **Transcriptional regulatory system.** Absolute error (17) of X_1 in **(a)** the mean for fsTL (gradient 1.3) and exTL (gradient 2.6); **(b)** the second moment for fsTL (gradient 1.3) and exTL (gradient 2.6). Results from 10^7 simulations, each run for $T = 3.2$.

τ-leap (2exTL) errors on species X_1 of the chain decay system; in this case we can see that the 2exTL results are approximately order three (or higher). Double-extrapolating gave substantially lower errors in almost every system, but it was not always easy to determine the order of accuracy of the method. A good example of this is the Michaelis-Menten system in Fig. 6. In such systems with relatively high Monte Carlo error, the bias of the double-extrapolated τ-leap was obscured by these fluctuations. The order of accuracy of the 2exTL could be successfully seen for most molecular species of the chain decay system, but it was not possible for the other systems as we would need even more simulations.

V. SEPARATION OF ERRORS

We have looked at the absolute weak error, given by

$$\text{error} = \mathbb{E}(\boldsymbol{f}(\mathbf{x}(t))) - \mathbb{E}_D(\boldsymbol{f}(\mathbf{x}_n^\tau)),$$

for any polynomial $\boldsymbol{f}$. We can separate this into two parts:

$$\text{error} = \left(\mathbb{E}(\boldsymbol{f}(\mathbf{x}(t))) - \mathbb{E}_\infty(\boldsymbol{f}(\mathbf{x}_n^\tau))\right) + \left(\mathbb{E}_\infty(\boldsymbol{f}(\mathbf{x}_n^\tau)) - \mathbb{E}_D(\boldsymbol{f}(\mathbf{x}_n^\tau))\right), \tag{22}$$

where $\mathbb{E}_\infty(\boldsymbol{f}(\mathbf{x}_n^\tau))$ are the theoretical values of the moments calculated by an infinite number of fsTL realizations with stepsize τ. The first term is the truncation error of the moments of the τ-leap from their analytical solutions, i.e. the bias of the method. The second term

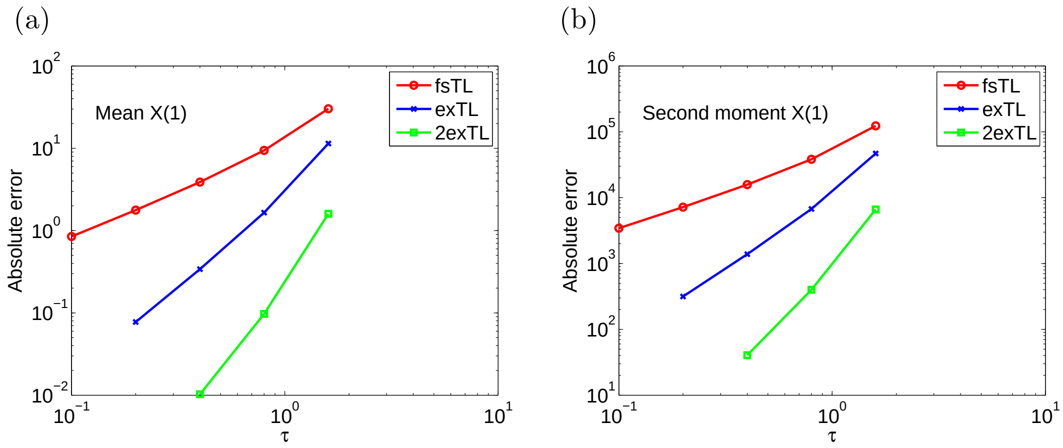

FIG. 5: **Chain decay system.** Absolute errors of X_1 for a fixed-step τ-leap (fsTL), an extrapolated τ-leap (exTL) and a double-extrapolated τ-leap (2exTL) algorithm for **(a)** the mean, gradients 1.3 (fsTL), 2.4 (exTL), 3.6 (2exTL), and **(b)** the second moment, gradients 1.3 (fsTL), 2.4 (exTL), 3.7 (2exTL). Results from 10^7 simulations. Double-extrapolation produces consistently higher-order results.

is the Monte Carlo error. This scales as the inverse of the square root of the number of simulations; sometimes reducing it sufficiently would mean using hundreds of millions of simulations, which can be unfeasible.

Therefore, instead of plotting only the absolute error, it is useful to be able to see the contributions of each of these terms. We can calculate the first term easily for linear systems: the analytical values of the moments are given by Eqs. (7) and (15) and their τ-leap approximations by Eqs. (6) and (14). Unfortunately, this is not possible for non-linear systems, as we must be able to write the propensities in the form in Eq. (5). However this is a useful technique for linear systems, as it allows us to quantify the relative level of Monte Carlo error on the results. It also reveals the true order of accuracy of the algorithm, without it being obscured by Monte Carlo errors.

Such a breakdown of the errors of the particle decay system is shown in Fig. 7, using results from 10^8 simulations. We see that for the case of the fsTL and exTL, the bias can easily be seen as the Monte Carlo errors are very low. However, when we extrapolate a second time, the bias of the resulting estimator is so low that Monte Carlo fluctuations completely obscure it, even using 10^8 simulations. This results in an ‘erroneous’ total error. The only way to reduce Monte Carlo error is to run more simulations. Thus, in general it

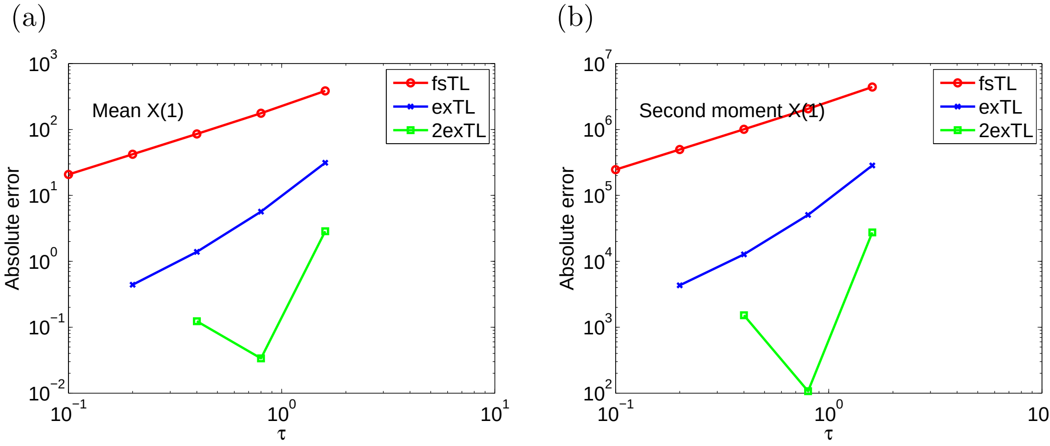

FIG. 6: **Michaelis-Menten system.** Absolute errors of X_1 for fsTL, exTL and 2exTL for **(a)** the mean, gradients 1.0 (fsTL), 2.0 (exTL) and 2.3 (2exTL), and **(b)** the second moment, gradients 1.0 (fsTL), 2.0 (exTL), 2.1 (2exTL). Results from 10^8 simulations. In this case, the 2exTL errors do not show a clear order of accuracy, but are nonetheless smaller compared to the other methods.

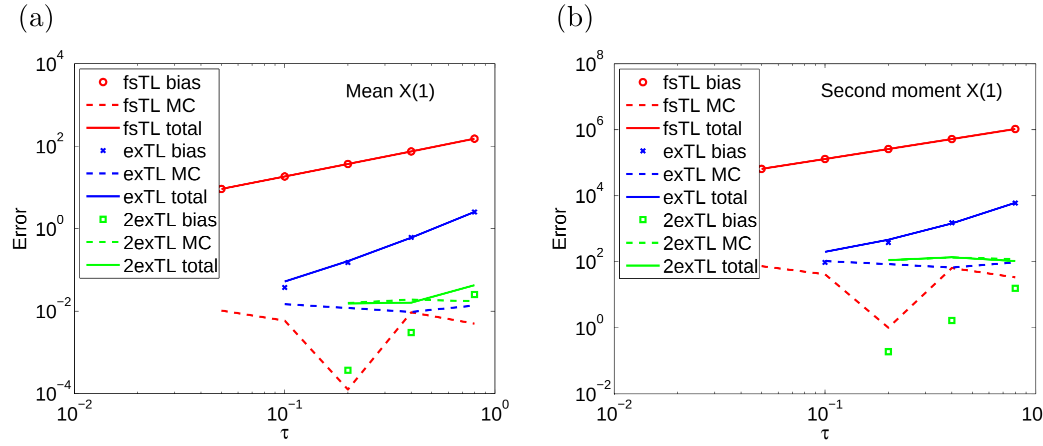

FIG. 7: Split errors for the particle decay system for (a)the mean and (b)the second moment. The bias can be easily seen with a fixed-step τ-leap (fsTL) or only a single extrapolation (exTL), but is obscured when we extrapolate a second time (2exTL). Results are from 10^8 simulations.

is important to perform a suitable number of simulations to get accurate estimates of the moments.

TABLE I: Processing times (in seconds) of 1000 simulations of the particle decay system using a fixed-step τ-leap (fsTL), an extrapolated τ-leap (exTL) and an SSA. Absolute errors in the mean from the analytical values are included in brackets. For comparison, our implementation of 1000 simulations of an SSA using the Direct Method took 1.73 seconds.

			τ		
	0.05	0.1	0.2	0.4	0.8
fsTL	0.4(9.2)	0.3(18.5)	0.2(37.1)	0.1(74.7)	0.1(152.0)
exTL	-	0.7(0.05)	0.5(0.16)	0.3(0.61)	0.2(2.6)

VI. CONCLUSIONS

Although twice as many τ-leap simulations must be run to extrapolate, the loss in computational time from this is more than compensated by the signficant reduction in error. This technique is very easy to use with the fsTL and considerably improves its accuracy. One advantage of this is that a larger τ can be used for a given level of error, shortening the computational time. This is often a limitation of stochastic methods, as many simulations must be run to build up an accurate picture of the system statistics. Computational times and the corresponding errors in the mean (in brackets) for the particle decay and transcriptional regulatory systems are shown in Tables I and II. The exTL takes $2-3$ times longer to run, but the errors it gives are tens to hundreds of times lower for the particle decay system, and several to tens of times lower for the transcriptional regulatory system. Thus the extrapolation is worth carrying out for both these systems. The pattern was similar for the other systems.

When we look for the convergence of a stochastic method numerically, Monte Carlo error is an unavoidable problem. The statistical fluctuations inherent in stochastic systems can obscure the bias error (i.e. order of convergence) of the numerical method. A large number D of simulations must be run, as the Monte Carlo error scales as $1/\sqrt{D}$. This error varies for each system. Figs. 5 and 6 show this clearly: both of the 2exTLs have total errors of similar size for the same τ, but the chain decay system has relatively low Monte Carlo error, allowing us to see the bias of that system. However, we believe the Michaelis-Menten system has relatively high Monte Carlo error compared to its bias, implying that the 2exTL errors

TABLE II: Processing times (in seconds) of 1000 simulations of the transcriptional regulatory system using a fixed-step τ-leap (fsTL), an extrapolated τ-leap (exTL) and an SSA. Absolute errors in the mean of X_1 from the SSA values are included in brackets. Our implementation of 1000 simulations of an SSA using the Direct Method took 1030 seconds.

	τ				
	0.005	0.010	0.020	0.040	0.08
fsTL	17.1(16.7)	9.2(34.5)	4.9(76.1)	2.5(188.5)	1.4(630.1)
exTL	-	24.0(1.1)	14.3(7.2)	7.5(36.3)	4.5(253.1)

we see in the figure are all due to statistical fluctuations.

A similar idea to the work in this paper is the multilevel Monte Carlo method, developed for SDEs[32,33]. The multilevel Monte-Carlo method, using a formulation of the error similar to Eq. (22), aims to cut the total error by reducing the Monte Carlo error. This is done in a way similar to Richardson extrapolation: approximate solutions are generated using a variety of different timesteps and combined, but by combining many coarse-scale solutions with few fine-scale ones, the method gets results with an accuracy comparable to that of the fine-scale solutions. In contrast, extrapolation uses the same number of coarse and fine-scale solutions and gives results which are more accurate than the fine-scale solution, by reducing the bias instead of the Monte Carlo error. In cases where the bias is obscured by statistical errors, using a combination of both extrapolation and the multilevel Monte Carlo method would be ideal, as it would reduce both sources of error.

The fsTL is the most basic and least accurate τ-leap algorithm. It would be ideal to extrapolate a more powerful τ-leap method such as that of Cao *et al.*[14], but these are more efficient because the timestep is varied, allowing the use of the largest possible step at each time point. This also means there is no obvious way of extrapolating such algorithms. However the power of our approach is that it can be applied to any numerical method as long as there exists an expression for the global error of the form (11). We can make higher-order versions of methods useful for specific classes of problems. For instance, Rué *et al.*[34] constructed a class of Runge-Kutta τ-leap methods, useful for simulating stiff systems. These methods have an extended stability region to allow larger stepsizes but are still weak order one. Extrapolating them would give methods which have higher-order accuracy as well as being very stable.

ACKNOWLEDGMENTS

T.Sz. is supported by the Engineering and Physical Sciences Research Council through the Systems Biology Doctoral Training Centre, University of Oxford. This publication was based on work supported in part by Award No. KUK-C1-013-04, made by King Abdullah University of Science and Technology (KAUST). The research leading to these results has received funding from the European Research Council under the *European Community's* Seventh Framework Programme (FP7/2007-2013) / ERC *grant agreement* No. 239870. R.E. would also like to thank Somerville College, University of Oxford for his Fulford Junior Research Fellowship.

Appendix: Second moment extrapolated local error

In order to see the effect of extrapolation on the local error, a sufficiently accurate approximation to (15) is necessary. Using the Heun method, a second-order method, we can write (15) as

$$\begin{aligned} S(t+\tau) &= S(t) + \tau W S(t) + \tau S(t) W^T + \tfrac{1}{2}\tau\nu\mathbf{d}(\boldsymbol{\mu}^T(t) + \boldsymbol{\mu}^T(t+\tau)) \\ &\quad + \tfrac{1}{2}(\boldsymbol{\mu}(t) + \boldsymbol{\mu}(t+\tau))\mathbf{d}^T\nu^T + \tfrac{1}{2}\tau\nu(B(t) + B(t+\tau))\nu^T + \tau\nu D\nu^T + \tfrac{1}{2}\tau^2 W^2 S(t) \\ &\quad + \tfrac{1}{2}\tau^2 S(t)(W^T)^2 + \tau^2 W S(t) W^T + \tfrac{1}{2}\tau^2 W\nu\mathbf{d}\boldsymbol{\mu}^T(t) + \tfrac{1}{2}\tau^2 W\boldsymbol{\mu}(t)\mathbf{d}^T\nu^T \\ &\quad + \tfrac{1}{2}\tau^2\nu\mathbf{d}\boldsymbol{\mu}^T(t)W^T + \tfrac{1}{2}\tau^2\boldsymbol{\mu}(t)\mathbf{d}^T\nu^T W^T + \tfrac{1}{2}\tau^2 W\nu B(t)\nu^T + \tfrac{1}{2}\tau^2\nu B(t)\nu^T W^T \\ &\quad + \tfrac{1}{2}\tau^2 W\nu D\nu^T + \tfrac{1}{2}\tau^2\nu D\nu^T W^T + \mathcal{O}(\tau^3). \end{aligned}$$

Extrapolating (14), we arrive at

$$\begin{aligned} 2S_{2m+2}^{\tau/2} - S_{m+1}^{\tau} &= S_{2m}^{\tau/2} + \tau W S_{2m}^{\tau/2} + \tau S_{2m}^{\tau/2} W^T + \tau^2 W S_{2m}^{\tau/2} W^T + \tau\nu D\nu^T \\ &\quad + \tfrac{1}{2}\tau^2 W^2 S_{2m}^{\tau/2} + \tfrac{1}{2}\tau^2 W\nu\mathbf{d}(\boldsymbol{\mu}_{2m}^{\tau/2})^T + \tfrac{1}{2}\tau^2 W\nu B_{2m}^{\tau/2}\nu^T + \tfrac{1}{2}\tau^2 W\nu D\nu^T \\ &\quad + \tfrac{1}{2}\tau^2 S_{2m}^{\tau/2}(W^T)^2 + \tfrac{1}{2}\tau^2\boldsymbol{\mu}_{2m}^{\tau/2}\mathbf{d}^T\nu^T W^T + \tfrac{1}{2}\tau^2\nu B_{2m}^{\tau/2}\nu^T W^T + \tfrac{1}{2}\tau^2\nu D\nu^T W^T \\ &\quad + \tau\boldsymbol{\mu}_{2m+1}^{\tau/2}\mathbf{d}^T\nu^T + \tau\nu\mathbf{d}(\boldsymbol{\mu}_{2m+1}^{\tau/2})^T + \tfrac{1}{2}\tau^2 W\boldsymbol{\mu}_{2m+1}^{\tau/2}\mathbf{d}^T\nu^T + \tfrac{1}{2}\tau^2\nu\mathbf{d}(\boldsymbol{\mu}_{2m+1}^{\tau/2})^T W^T \\ &\quad + \tau\nu B_{2m+1}^{\tau/2}\nu^T. \end{aligned}$$

The local error is $S(t+\tau) - \left(2S_{2m+2}^{\tau/2} - S_{m+1}^{\tau}\right)$. For simplicity, we evaluate it at time $t = \tau$. Most of the terms cancel, leaving

$$\begin{aligned}
&\tau\nu\mathbf{d}\left(\tfrac{1}{2}(\boldsymbol{\mu}^T(0) + \boldsymbol{\mu}^T(\tau)) - (\boldsymbol{\mu}_1^{\tau/2})^T\right) + \tau\left(\tfrac{1}{2}(\boldsymbol{\mu}(0) + \boldsymbol{\mu}(\tau)) - (\boldsymbol{\mu}_1^{\tau/2})\right)\mathbf{d}^T\nu^T \\
&+ \tau\nu\left(\tfrac{1}{2}(B(0) + B(\tau)) - B_1^{\tau/2}\right)\nu^T + \tfrac{1}{2}\tau^2 W(\boldsymbol{\mu}(0) - \boldsymbol{\mu}_1^{\tau/2})\mathbf{d}^T\nu^T \\
&+ \tfrac{1}{2}\tau^2\nu\mathbf{d}(\boldsymbol{\mu}^T(0) - (\boldsymbol{\mu}_1^{\tau/2})^T)W^T + \mathcal{O}(\tau^3).
\end{aligned}$$

The first term is equal to $\tau\nu\mathbf{d}\left(\boldsymbol{\mu}^T(0)(\tfrac{1}{4}\tau^2W^2 + \mathcal{O}(\tau^3)) + \mathbf{d}^T\nu^T(\tfrac{1}{4}\tau^2 W + \mathcal{O}(\tau^3))\right)$ and similarly for the next two terms. The penultimate term is $-\frac{1}{2}\tau^2 W(\frac{1}{2}\tau W\boldsymbol{\mu}(0) + \frac{1}{2}\tau\nu\mathbf{d})\mathbf{d}^T\nu^T$, and similarly for the last term. Thus all terms in the local error are $\mathcal{O}(\tau^3)$, implying a global error of $\mathcal{O}(\tau^2)$.

REFERENCES

[1] H. H. McAdams and A. Arkin, Trends in Genetics **15**, 65 (1999).

[2] M. Elowitz, A. Levine, E. Siggia, and P. Swain, Science **297**, 1183 (2002).

[3] N. Fedoroff and W. Fontana, Science **297**, 1129 (2002).

[4] M. Kaern, T. Elston, W. Blake, and J. Collins, Nature Reviews Genetics **6**, 451 (2005).

[5] D. J. Wilkinson, Nature Reviews Genetics **10**, 122 (2009).

[6] R. Erban, S. J. Chapman, I. Kevrekidis, and T. Vejchodsky, SIAM Journal on Applied Mathematics **70**, 984 (2009).

[7] D. T. Gillespie, Journal of Physical Chemistry **81**, 2340 (1977).

[8] D. T. Gillespie, Journal of Chemical Physics **115**, 1716 (2001).

[9] D. T. Gillespie and L. R. Petzold, Journal of Chemical Physics **119**, 8229 (2003).

[10] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, Journal of Chemical Physics **119**, 12784 (2003).

[11] T. H. Tian and K. Burrage, Journal of Chemical Physics **121**, 10356 (2004).

[12] A. Chatterjee, D. G. Vlachos, and M. A. Katsoulakis, Journal of Chemical Physics **122** (2005).

[13] Y. Cao, D. T. Gillespie, and L. R. Petzold, Journal of Chemical Physics **123**, 054104 (2005).

[14] Y. Cao, D. T. Gillespie, and L. R. Petzold, Journal of Chemical Physics **124**, 044109 (2006).

[15] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, Multiscale Modeling & Simulation **4**, 867 (2005).

[16] T. Li, Multiscale Modeling & Simulation **6**, 417 (2007).

[17] Y. Hu and T. Li, Journal of Chemical Physics **130** (2009).

[18] D. Talay and L. Tubaro, Stochastic Analysis and Applications **8**, 483 (1990).

[19] M. Gibson and J. Bruck, Journal of Physical Chemistry A **104**, 1876 (2000).

[20] Y. Cao, H. Li, and L. R. Petzold, Journal of Chemical Physics **121**, 4059 (2004).

[21] C. A. Yates and K. Burrage, Journal of Chemical Physics **134** (2011).

[22] D. A. McQuarrie, Journal of Applied Probability **4**, 413 (1967).

[23] D. T. Gillespie, Physica A **188**, 404 (1992).

[24] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 2007), 3rd ed.

[25] C. Gadgil, C. Lee, and H. Othmer, Bulletin of Mathematical Biology **67**, 901 (2005).

[26] B. Melykuti, K. Burrage, and K. C. Zygalakis, Journal of Chemical Physics **132** (2010).

[27] R. D. Skeel, Numerische Mathematik **48**, 1 (1986).

[28] X. Q. Liu and C. W. Li, SIAM Journal of Numerical Analysis **37**, 1747 (2000).

[29] T. Jahnke and W. Huisinga, Journal of Mathematical Biology **54**, 1 (2007).

[30] T. T. Marquez-Lago and K. Burrage, Journal of Chemical Physics **127** (2007).

[31] J. Goutsias, Journal of Chemical Physics **122** (2005).

[32] A. Kebaier, Annals of Applied Probability **15**, 2681 (2005).

[33] M. B. Giles, Operations Research **56**, 607 (2008).

[34] P. Rué, J. Villa-Freixà, and K. Burrage, BMC Systems Biology **4**, 110 (2010).

RECENT REPORTS

17/11	Energy versus electron transfer in organic solar cells: a comparison of the photophysics of two indenofluorene: fullerene blend films	Soon Clarke Zhang Agostinelli Kirkpatrick Dyer-Smith McCulloch Nelson Durrant
18/11	Asymptotic analysis of a pile-up of edge dislocation	Hall
19/11	A perturbation analysis of spontaneous action potential initiation by stochastic ion channels	Keener1 Newby
20/11	Hybrid modelling of individual movement and collective behaviour	Franz Erban
21/11	A novel model for one-dimensional morphoelasticity. Part I: Theoretical foundations	Hall Menon McCue McElwain
22/11	A novel model for one-dimensional morphoelasticity. Part II: Application to the contraction of fibroblast-populated collagen lattices	Hall Menon McCue McElwain
23/11	Positive or negative Poynting efect? The role of adscititious inequalities in hyperelastic materials	Mihai Goriely McCue McElwain
24/11	On approaches to modelling lattice dislocations	Hall Markenscoff
25/11	Nonlinear waves in heterogeneous elastic rods via homogenization	de Luna Emptage Goriely Bressloff
26/11	Synaptic bistability due to nucleation and evaporation of receptor clusters	Burlakov Duričković Goriely
27/11	Particle trapping and banding in rapid solidification	Elliot Peppin
28/11	Growth of confined cancer spheroids: a combined experimental and mathematical modelling approach	Loessner Flegg Byrne Hall Moroney Clements McElwain Hutmacher
29/11	Floating carpets and the delamination of elastic sheets	Wagner Vella

32/11	Wrinkling of pressurized elastic shells	Vella Ajdari Vaziri Boudaoud
33/11	Data assimilation using bayesian filters and B-spline geological models	Duan Farmer Hoteit Lu Moroz
34/11	Review of nonlinear Kalman, ensemble and particle filtering with application to the reservoir history matching problem	Luo Hoteit Duan Wang
35/11	Modelling a Tethered Mammalian Sperm Cell undergoing Hyperactivation	Curtis Kirkman-Brown Connolly Gaffney
36/11	A simple mathematical model for investigating the effect of cluster roots on plant nutrient uptake	Zygalakis Roose
37/11	Frequency jumps in the planar vibrations of an elastic beam	Neukirch Frelat Goriely Maurini
38/11	Ice-lens formation and confinement-induced supercooling in soils and other colloidal materials	Style Cocks Peppin Wettlaufer
39/11	An asymptotic theory for the re-equilibration of a micellar surfactant solution	Griffiths Bain Breward Chapman Howell Water

Copies of these, and any other OCCAM reports can be obtained from:

Oxford Centre for Collaborative Applied Mathematics
Mathematical Institute
24 - 29 St Giles'
Oxford
OX1 3LB
England

www.maths.ox.ac.uk/occam