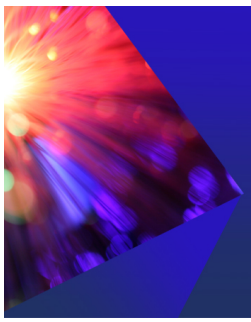
RESEARCH ARTICLE | FEBRUARY 09 2015

Stochastic linear multistep methods for the simulation of chemical kinetics

Manuel Barrio ; Kevin Burrage; Pamela Burrage

J. Chem. Phys. 142, 064101 (2015)
https://doi.org/10.1063/1.4907008



Stochastic linear multistep methods for the simulation of chemical kinetics

Manuel Barrio,[1,a)] Kevin Burrage,[2,3,b)] and Pamela Burrage[3,c)]
[1]*Departamento de Informática, University of Valladolid, Valladolid, Spain*
[2]*Department of Computer Science, University of Oxford, Oxford, United Kingdom*
[3]*School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia*

(Received 7 November 2014; accepted 16 January 2015; published online 9 February 2015)

In this paper, we introduce the Stochastic Adams-Bashforth (SAB) and Stochastic Adams-Moulton (SAM) methods as an extension of the τ-leaping framework to past information. Using the Θ-trapezoidal τ-leap method of weak order two as a starting procedure, we show that the k-step SAB method with $k \geq 3$ is order three in the mean and correlation, while a predictor-corrector implementation of the SAM method is weak order three in the mean but only order one in the correlation. These convergence results have been derived analytically for linear problems and successfully tested numerically for both linear and non-linear systems. A series of additional examples have been implemented in order to demonstrate the efficacy of this approach. © *2015 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4907008]

I. INTRODUCTION

Understanding the effects of intrinsic noise in biochemical systems is a vital component in the burgeoning scientific field of *Computational Cell Biology*. Intrinsic noise is associated with the uncertainty of knowing when a reaction occurs and what that reaction is, and this effect is accentuated when there are small numbers of molecules in some component of the system. In many cases, the biochemical dynamics associated with intrinsic noise can be very different to the dynamics associated with the Law of Mass Action, and therefore, appropriate methods to deal with the stochastic modelling and simulation of these systems are needed.

Fundamental work on modelling intrinsic noise was done by Kurtz[17,18] and Gillespie.[10] Gillespie introduced the Stochastic Simulation Algorithm (SSA) that describes the time evolution of the dynamics of a well-stirred chemical reaction system as a discrete nonlinear Markov process. The key elements are a timestep that is drawn from an exponential distribution and the update of the state vector by the stoichiometric vector associated with the most likely reaction in that timestep. The generated samples are known to be *exact* with respect to the probability density function determined by the chemical master equation.

A drawback of the SSA is that the waiting time can be very small if there are some large rate constants and/or there are some species with large numbers of molecules. As a consequence, Gillespie[11] introduced the Poisson τ-leap method, in which all the reactions are allowed to fire in a given step (no longer a waiting time) with a frequency determined from a Poisson distribution. It is important to note that this is an approximate algorithm and therefore amenable to accuracy analysis as to the rate of convergence of the numerical approximation to the exact solution.

Since the publication of the τ-leap algorithm, many extensions have appeared. These include the implicit[25] and the slow-scale[5] τ-leap methods which are designed to deal with stiff systems, the binomial τ-leap[31] that avoids negative populations, a delayed τ-leap method[19] that incorporates delayed reactions, and higher-order τ-leap methods that achieve higher order accuracy either introducing random corrections,[13] Poisson increments,[15] or extrapolation.[29,30]

This paper presents a novel method that explores further the possibilities of highly accurate τ-leap procedures for discrete simulation of chemical kinetics. The method described here is based on a well-known principle applied in the deterministic setting, namely, the use of information from several previous simulation steps as an attempt to produce more accurate and efficient results. Our approach follows the rationale that is used in linear multistep methods (LMM) for Ordinary Differential Equations (ODEs) where a linear combination of previous points and derivative values is used to calculate a next state solution. In our proposal, the previously calculated information that must be stored comprises the state points and the corresponding reaction propensities. Note that it is now the propensity term that plays the role of integrand since it represents the varying intensity of the Poisson processes that account for the number of reactions fired in each step.

When mimicking strategies that are taken from the deterministic resolution of ODEs and those strategies are *translated* into a corresponding discrete simulation method, there is always a risk of focusing on numerical results without any further analysis of the simulated behaviour. In this paper, we present an accuracy analysis, at least for linear examples, that justifies the behaviour of the algorithms in terms of their order of convergence.

Following the different families of multistep methods that can be found in the literature, we have constructed stochastic versions of the class of explicit k-step Adams-Bashforth (AB) methods and the implicit Adams-Moulton (AM) methods solved with a predictor-corrector scheme. These methods allow a different number of previous steps to be considered. An

a)mbarrio@infor.uva.es.
b)kevin.burrage@cs.ox.ac.uk.
c)pamela.burrage@qut.edu.au.

order method theory is established and numerical results are shown in terms of both accuracy and efficiency. Furthermore, all of them are dependent on the starting procedure and care must be taken in selecting this starting procedure.

This paper is structured as follows: Sec. II gives a survey of linear multistep methods in the deterministic setting and points out some important results we may want to know in deriving new stochastic methods (some topics on the stability of LMMs are given in Appendix A). The rationale of how linear multistep methods can be applied to stochastic chemical kinetics is described in Sec. III along with main results about the order and stability of the methods for linear systems. Section IV presents numerical results of test problems that show the order of convergence of the suggested methods, while Sec. V presents numerical results that show the efficiency of the stochastic linear multistep methods when applied to different chemical examples of varying complexity. The numerical efficiency is discussed in terms of the histogram distances to the *exact* solution as a function of the computational cost measured in terms of the associated runtime. The paper concludes with Sec. VI with conclusions and discussions about the methods and their applicability. The mathematical derivation of the order of convergence, for linear problems and for the proposed methods can be found in Appendix B.

II. LINEAR MULTISTEP METHODS

Consider the initial value problem

$$y' = f(y), \quad y(0) = y_0 \in \mathbb{R}^n. \tag{1}$$

Numerical methods that advance the approximation y_n at $t = t_n$ to a new approximation y_{n+1} at $t = t_{n+1}$, where $t_{n+1} = t_n + h$ can be broadly characterised into two classes: one step (Runge-Kutta) and linear multistep methods. Runge-Kutta methods are constructed by forming s internal approximations, usually within the current step, and then taking a linear combination of the derivatives of these approximations to form the approximation at the end of the step. Runge-Kutta methods can have high order ($2s$) and in the case of implicit methods, have excellent stability properties such as A-stability. In addition, there also exist very efficient high order explicit methods but in these cases, the number of stages can exceed 8.[24] Due to their one-step nature, Runge-Kutta methods can be very simply implemented in a variable stepsize and variable order setting.

LMMs are at the opposite end of the spectrum to Runge-Kutta methods, and a k-step method takes the form

$$y_{n+1} = \sum_{j=1}^{k} \alpha_j y_{n+1-j} + h \sum_{j=0}^{k} \beta_j f_{n+1-j}. \tag{2}$$

Two very important classes of methods are the Adams-type methods in which $\alpha_1 = 1$, $\alpha_j = 0$, $j = 2, \ldots, k$ and Backward Differentiation Formulae (BDF methods) in which $\beta_j = 0$, $j = 1, \ldots, k$. In this paper, we will focus on explicit AB methods of order k ($\beta_0 = 0$) and implicit AM methods of order $k+1$ ($\beta_0 \neq 0$). Some examples of Adams method coefficients are given in Table I.

The order of a LMM can easily be established by assuming perfectly accurate past information and finding the local error in the update step by using a Taylor series information. This leads to the following.

TABLE I. Coefficients of AB and AM methods.

	AB				AM			
k	β_1	β_2	β_3	β_4	β_0	β_1	β_2	β_3
1	1				1/2	1/2		
2	3/2	−1/2			5/12	8/12	−1/12	
3	23/12	−16/12	5/12		9/24	19/24	−5/24	1/24
4	55/24	−59/24	37/24	−9/24				

Theorem 2.1. *A k-step LMM is of order w if* $\sum_{j=1}^{k} \alpha_j = 1$, *consistency condition* $\sum_{j=2}^{k} \alpha_j (1-j)^p + p \sum_{j=0}^{k} \beta_j (1-j)^{p-1} = 1$, $p = 1, \ldots, w$.

Corollary 2.1. An Adams method will be of order w if $p \sum_{j=0}^{k} \beta_j (1-j)^{p-1} = 1, p = 1, \ldots, w$.

In the case of an implicit LMM, y_{n+1} is described by a nonlinear system of equations. This can be solved using linear algebra by a modified Newton approach, but it can also be solved by fixed point iteration that leads to the concept of predictor-corrector methods. In the case of Adams methods, the predictor will be based on the AB method, while the corrector is based on the AM method. It takes the general form

$$\begin{aligned} &\text{P: } y_{n+1}^{P} = y_n + h \textstyle\sum_{j=1}^{k} \hat{\beta}_j f_{n+1-j}, \\ &\text{E: } f_{n+1} = f(y_{n+1}^{P}), \\ &\text{C: } y_{n+1}^{C} = y_n + h \textstyle\sum_{j=1}^{k} \beta_j f_{n+1-j} + h\beta_0 f_{n+1}. \end{aligned}$$

Here, the $\hat{\beta}_j$ are the coefficients of the AB method and the β_j are the coefficients of the AM method. Note that one correction improves the order of the method from k to $k+1$. Further corrections do not improve the order of the procedure but can improve the stability region. A standard approach is to use in advance a fixed number of corrections in the iteration process.

A fundamental difference between a one-step method and a LMM is that the latter needs a starting procedure to generate the approximations $y_1, \ldots, y_{k-1}$ from y_0. This starting procedure is often a Runge-Kutta method. This has additional ramifications in that the order of a LMM must also be assessed in terms of the order of the starting procedure. It is easily shown that if the starting procedure is of order p and a LMM is order p, then the convergence order of the combined implementation is also order p.[4] However, what is less well known is that if the LMM has order p and the starting procedure has order $p-1$ and that all the order p terms in the Taylor expansion for that starting procedure collapse down into a single term, then the overall order can still be p—see pages 72–76.[4] We will exploit this fact when constructing stochastic LMMs of order 3, as then the starting procedure can still be of order 2.

This concludes our summary of important aspects associated with the use of LMM for deterministic problems. Now, we turn our attention to discrete stochastic problems arising from chemical kinetics.

III. LINEAR MULTISTEP METHODS FOR CHEMICAL KINETICS

When modelling the dynamics of chemical kinetics when there are relatively few molecules in the system, a continuous model based on ordinary differential equations and derived from the law of mass action may not be appropriate. The stochastic simulation algorithm[16] is a nonlinear discrete Markov process that updates the reactions, a reaction at a time based on the relative sizes of the propensity functions and an exponential waiting time distribution to the next reaction. Thus, given a state vector $X(t)$ consisting of integer values and a set of m stoichiometric vectors $\nu_1, \ldots, \nu_M$ (the update rules for each of the M reactions) and the relative probabilities of the reactions occurring (propensity functions) $a_1(X(t)), \ldots, a_M(X(t))$, the SSA evolves as

$$X(t+\tau) = X(t) + \nu_j, \tag{3}$$

where τ is exponentially distributed as $e^{-\tau \sum_{j=1}^{M} a_j(X(t))}$.

As the SSA can be very computationally intensive, a number of approaches have been suggested to improve performance at a cost of loss of accuracy. The initial approach was due to Gillespie[10] and takes the form

$$X(t+\tau) = X(t) + \sum_{j=1}^{M} \nu_j \mathcal{P}(\tau a_j(X(t))), \tag{4}$$

where $\mathcal{P}$ is a Poisson random variable. Such a method is called a Poisson τ-leap method. The original stepsize selection procedure has since been modified to improve accuracy[12,6] and can be determined deterministically. It has been shown that this method corresponds to the Euler method for solving ODEs and to the Euler-Maruyama method for solving stochastic differential equations driven by jump processes and accordingly has weak order one.[26,20]

In order to deal with issues of negative populations, Tian and Burrage[31] and Chatterjee *et al.*[8] developed the binomial τ-leap method, which samples from the Binomial distribution. Further extensions are given in Refs. 22 and 23.

One way of improving the efficacy of the τ-leap approach is to try and construct methods with higher weak order. Accordingly, Gillespie[11] introduced the midpoint τ-leap method that has an additional half step. Under the scaling $\tau \to 0$, this method has weak order one in the mean but under large volume scaling $V \to \infty$, it has weak order two,[1,14] although its covariance is still order one.[14] More recently, there has been a search for methods that are truly weak order two and this has led to the unbiased τ-leap,[33] the random-corrected τ-leap,[13] and the Θ-trapezoidal τ-leap.[15] This latter method is truly weak order two and is based on a method for solving Stochastic Differential Equations (SDEs) by Anderson and Mattingly.[2] In the case $\Theta = 1/2$, the trapezoidal τ-leap method is given by

$$\begin{aligned} X^P &= X_n + \sum_{j=1}^{M} \nu_j \mathcal{P}(\frac{1}{2}\tau a_j(X_n)), \\ l_j &= max\{2a_j(X^P) - a_j(X_n), 0\}, \quad j = 1, \ldots, M, \\ X_{n+1} &= X^P + \sum_{j=1}^{M} \nu_j \mathcal{P}(\frac{1}{2}\tau l_j). \end{aligned}$$

This suggests that in order to construct methods of weak order two, there must be samples of the form $\mathcal{P}(\tau a_j(\mathcal{P}(c\tau a_i(X_n))))$. These samples are akin to the double integrals of Wiener processes that are needed to construct higher order methods for SDEs.[3]

Finally, other ideas used in constructing effective numerical methods for ODEs have been applied to developing efficient discrete stochastic numerical methods. These include constructing methods with large stability regions based on a class of Poisson Runge-Kutta methods[28] and developing an approach based on Richardson extrapolation of moments,[27,29] and the construction of a stochastic Bulirsch-Stoer extrapolation technique that allows an efficient adaptive τ-leap and order approach.[30]

In the same light as using ideas from the ODE setting to develop effective methods for discrete nonlinear stochastic problems, we now propose using past information as a means of increasing accuracy and efficiency, and this leads us to develop a class of methods based on the Adams methods. We note in passing that although the stochastic simulation algorithm is exact and a memoryless process, in that its evolution depends exclusively on its present state, this does not preclude developing simulation methods that use past information. Rather, this past information is used to obtain better estimates of certain integrals—see above comments on the Θ-trapezoidal τ-leap method. The underlying LMMs that we will use are the Adams methods and so we will call our methods Stochastic Adams Methods (SAMs).

We will consider two variants of SAMs: the first is a class of methods based on the AB method that we will call Stochastic Adams-Bashforth Methods (SABMs) and the second is a class of methods based on a predictor-corrector implementation of the AM methods that we will call Stochastic Adams-Moulton methods (SAMMs).

An important aspect, as is the case for linear multistep methods applied to ODEs, is the nature of the starting procedures for our methods. As noted in Sec. II, given a LMM method of order p, then a starting procedure must be of order $p-1$ or p if the combined method is to be of order p. In the case of stochastic linear multistep methods, if these methods are to be of weak order three, then the starting procedure must be of weak order two in both mean and variance. This suggests using the one-step Θ-trapezoidal method as the starting procedure as it is weak order two.

Full details of the order analysis for both SABMs and SAMMs in conjunction with the starting procedure are given in the Appendix B.

We now give a brief summary of the results from the Appendix B. The order analysis is based on the linear decay problem $X \xrightarrow{c} \varnothing$, with rate constant c. We let $z = \tau c$. In this case, the SABM takes the form

$$X_{n+1} = X_n - \mathcal{P}(z \sum_{j=1}^{k} \beta_j X_{n+1-j}), \tag{5}$$

where the β_j are the coefficients of the AB method. The SAMM, with a deterministic predictor, has the predictor-corrector form

$$\hat{X}_{n+1} = X_n - z\sum_{j=1}^{k}\beta_j X_{n+1-j}, \tag{6}$$

$$X_{n+1} = X_n - \mathcal{P}(z\sum_{j=1}^{k}\hat{\beta}_j X_{n+1-j} + z\hat{\beta}_0\hat{X}_{n+1}), \tag{7}$$

where the $\hat{\beta}_j$ are the coefficients of the AM method. In both cases, the starting procedure will be the Θ-trapezoidal τ-leap method that has weak order two (in the mean and variance). The following results are proved for linear systems.

If the underlying AM has order p and the order of the starting procedure in the mean is p or $p-1$, then the mean order of the SAM will also be p. Hence, it follows that the mean order of the SABMs ($k \geq 3$) and SAMMs ($k \geq 2$) will be three if the Θ-trapezoidal τ-leap method is used as a starting procedure.

The correlation order of the SABMs ($k \geq 3$) is three if the Θ-trapezoidal τ-leap method is used as a starting procedure. However, the correlation order of the predictor-corrector SAMM is only one.

We note that while the order analysis is for the linear test problem, we conjecture that the above results also hold for the nonlinear problems. The fact that for deterministic LMMs, order can be analysed by just considering the linear equation $y' = q\,y$ lends support to this conjecture. Furthermore, in Sec. IV, we present simulation results for a variety of SAMs based on the linear and nonlinear test equations, $X \xrightarrow{c} \varnothing$ and $X + Y \xrightarrow{c} \varnothing$, respectively. More general results on performance and efficiency of the SAMs on more complex systems are given in Sec. V.

Finally, we note that further study on the form of the predictor-corrector is needed, as the correlation order collapses to one for our particular formulation.

IV. SIMULATION RESULTS CONFIRMING ORDER

In this section, we present simulation results that show numerically the order of convergence of the stochastic simulation algorithm described in Sec. III for the linear and nonlinear test examples, $X \xrightarrow{c} \varnothing$ and $X + Y \xrightarrow{c} \varnothing$, respectively. The method we use to show convergence is a standard procedure[26,13] based on the absolute error of the moments, mean and variance, and their dependence with the stepsize.

In order to calculate the error of the moments, the system state is simulated from an initial state up to a predefined time, T, using a specific stepsize, h. The simulation of a large number of trajectories allows the calculation of a sample mean (respectively, variance) that can be compared with the *true* mean value (respectively, variance). There are examples, for instance, the linear test problem $X \xrightarrow{c} \varnothing$, where this latter value can be analytically derived and therefore exactly calculated. For those where this is not possible, we will consider the sample mean (respectively, variance) of SSA simulations to be an unbiased estimation of the moment provided the sample size is large enough.

The absolute error of moments is calculated using different stepsizes and then represented in a log-log plot of error vs stepsize, in which the slope of the plot gives the order of the method for that particular moment. Note that this technique is subject to a certain degree of variability due to a number of different issues, including the number of samples generated, the intrinsic Monte Carlo error that might be noticeable for small stepsizes, and an order of convergence that might not be as expected if there are not enough simulation steps when using large stepsizes. Therefore, the stepsizes used in the test problems have been selected so that this variability is reduced as much as possible. Additionally, a fitting function has been applied to each plot in order to give a better estimate of the order of convergence.

A. Linear test problem

This is the simplest linear chemical system consisting of a single species, X, and a single degradation reaction, $X \xrightarrow{c} \varnothing$. For our particular example, the reaction rate constant is $c = 0.1$ and the initial condition $X_0 = 10\,000$.

Figure 1 shows the absolute error of the moments comparing the results obtained using the Θ-trapezoidal method (TT), the SAB with $k = 3$ (SAB3), and the SAM with $k = 3$ (SAM3) methods applied to the linear test problem. In the case of the Θ-trapezoidal method, both moments, and mean and variance show weak order two as described in Ref. 15. The simulation results for the SAB3 method using the Θ-trapezoidal method as the starting procedure show the combined weak order of convergence to be approximately three for both mean and variance. However, it can be seen that the variance order of the SAM3 method is approximately one. This confirms the theoretical analysis of Sec. III and Appendix B.

B. Nonlinear test problem

The nonlinear test problem is similar to the linear one but with a second order reaction degrading two reactants, $X + Y \xrightarrow{c} \varnothing$. The nonlinearity of this system is reflected in the propensity function of the reaction that multiplies the reaction rate constant and the number of molecules of both reactants. In our example, the rate constant is 10^{-5} and the initial state $(1000,1000)^{\top}$.

Figure 2 shows the order results of the stochastic Adams methods for this example and compares them with those obtained for the Θ-trapezoidal τ-leap method. Here, as in all examples in this paper, the latter method has been used as the starting procedure of the proposed multistep methods. It is shown that, up to a certain degree of variability, the order results are consistent with those obtained for the linear test problem and this supports our claim that the order of convergence of the SAB and SAM methods, as stated in Sec. III, also holds for nonlinear problems. However, we note it is for the nonlinear test problem that the order results show a noticeable variability that mainly depends on the stepsize and number of steps. The use of small stepsizes in the simulation procedure leads to very small errors and therefore to frequent Monte Carlo error bias. Conversely, larger stepsizes imply a smaller number of steps, and this may alter the cumulative effect of the method on the order terms and therefore affect the resulting order.

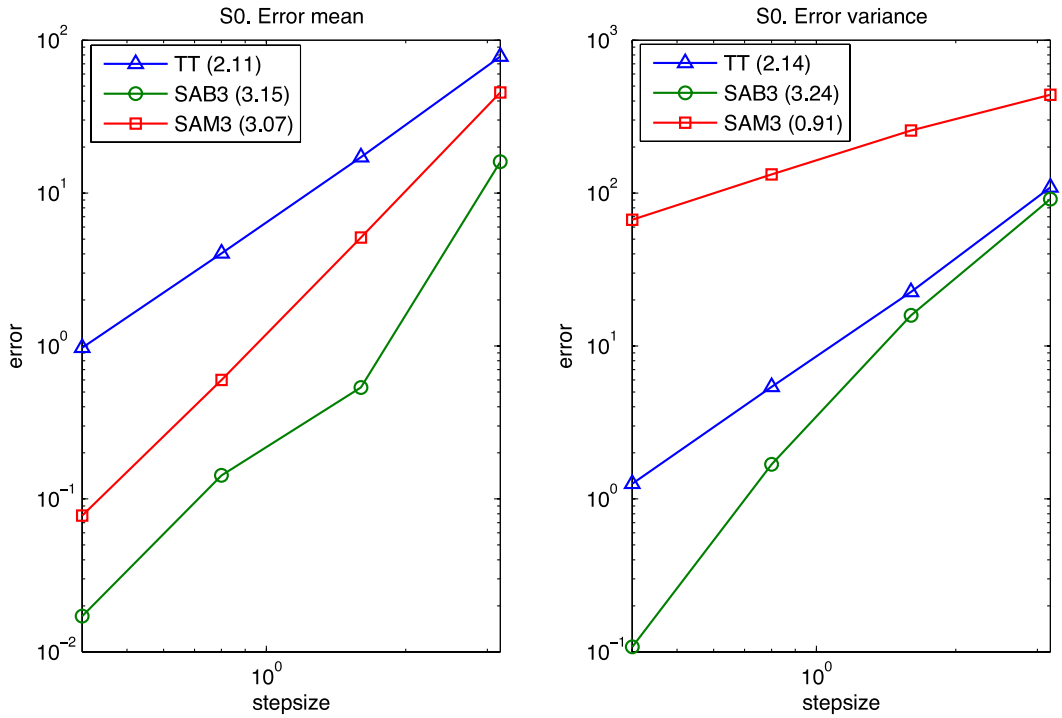


FIG. 1. Linear test problem: absolute error of mean and variance. Compared results of the TT, SAB with $k = 3$ (SAB3), and SAM with $k = 3$ (SAM3) methods using stepsizes $[0.4, 0.8, 1.6, 3.2]$, $T = 12.8$, and sample size 10^8. Both SAB3 and SAM3 use TT as the starting procedure as described in Sec. III. The numbers accompanying each method in the legend correspond to the fitted slope and therefore to the estimated order.

V. COMPARISON OF METHODS ON SOME TEST PROBLEMS

The order of convergence is a very important framework to compare stochastic simulation procedures: the higher the order, the faster the error is reduced with the stepsize. However, a stepsize reduction will imply an increment of the total simulation time and this is very often a key issue when deciding which simulation method and parameter setting (mainly the stepsize) to use. Consequently, it is very important to give a measure of efficiency for the methods we have described.

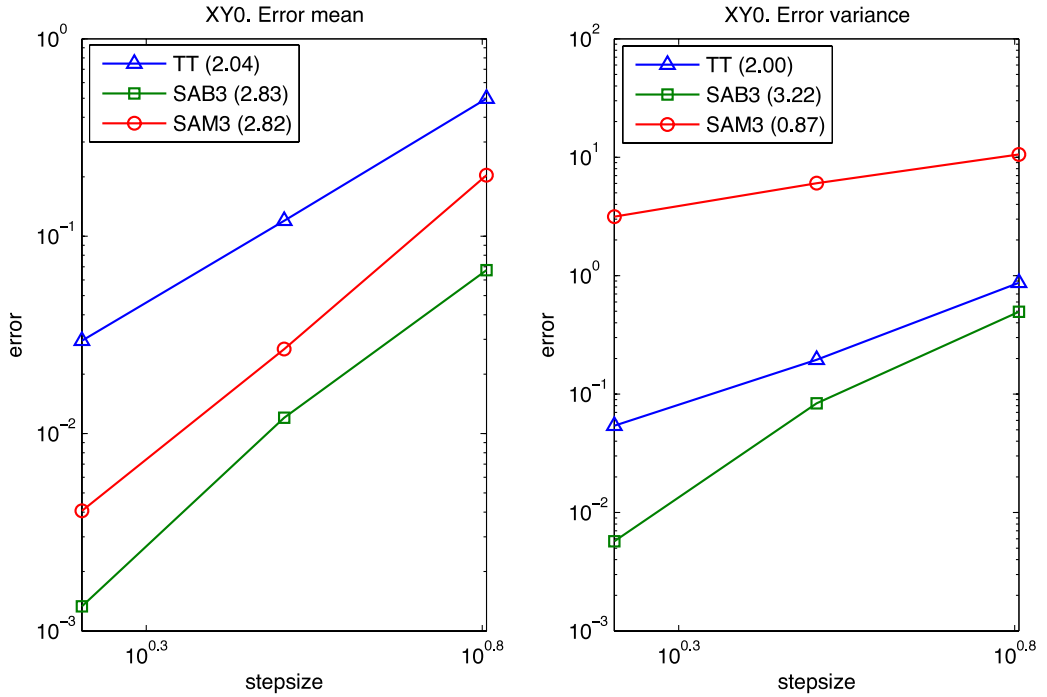


FIG. 2. Nonlinear test problem: absolute error of mean and variance. Compared results of the TT, SAB with $k = 3$ (SAB3), and SAM with $k = 3$ (SAM3) methods using stepsizes $[1.6, 3.2, 6.4]$, $T = 51.2$, and sample size 10^9. The numbers accompanying each method in the legend correspond to the estimated order.

A very common measure of efficiency for stochastic simulation schemes is to represent the relationship between histogram distance and execution time.[6,7] The histogram distance is calculated as the L^1-distance of two distribution functions obtained from an ensemble of simulated and *exact* values, respectively. The simulated values are the result of executing the simulation procedure up to $t = T$. Since the true distribution of $X(t)$ is not known for many examples, we will use SSA samples to generate the *exact* distribution. This is an appropriate approximation provided the sample size is big enough.

We note that the efficiency comparison should not be based on the implementation but on the algorithm definition itself. To do so, all the methods have been implemented in Matlab using a common scheme with the only differences being the state update block that is specific for each of them.

Before introducing new test problems, we present in Fig. 3 the efficiency results for the linear and nonlinear test problems described in Sec. IV. The resulting conclusion is that the SAB3 method is a very efficient procedure that outperforms the other two in almost any simulation setting. According to the simulation results, if we set a certain time value as available runtime, the error reduction of the SAB3 over the TT method can be as high as 180-fold (minimum 6, average 66) for the linear test problem and 30-fold (minimum 5, average 17) for the nonlinear one. Should we set the error level instead, the maximum speed-up of the SAB3 over the TT method is about 4-fold (min. 2, avg. 3) for both the linear and nonlinear test problems.

A. Michaelis-Menten system

The Michaelis-Menten formulation accounts for the kinetics of many enzymes. It includes four molecular species and three reactions transforming a reactant into a product with the intervention of an enzyme. The reaction rate constants and initial state are as in Ref. 15, where the Θ-trapezoidal τ-leap method is defined to facilitate method comparison,

$$\begin{aligned} R_1&: S + E \xrightarrow{k_1} SE, \quad k_1 = 10^{-4}, \\ R_2&: SE \xrightarrow{k_{-1}} S + E, \quad k_{-1} = 0.5, \\ R_3&: SE \xrightarrow{k_2} E + P, \quad k_2 = 0.5, \\ X_0 &= (1000,200,2000,0)^\top. \end{aligned}$$

Figure 4 shows the histogram distance vs runtime plots obtained from the simulations of the Michaelis-Menten system using the TT, SAB with $k = 3$ (SAB3), and SAM with $k = 3$ (SAM3) methods. Clearly, the SAM3 method has the worst performance for almost any simulation setting, and this is probably related to the lack of higher order of convergence for the covariance and the extra calculation of the predictor. The SAB3 method achieves better efficiency results than the TT method for most levels of error or runtime, with the exception of small or large runtime values (corresponding to large and small stepsizes, respectively) where their performance tends to be similar. The higher error of the SAB3 method when using large steps is probably related to the issue described Sec. IV, namely, an insufficient number of steps required to achieve the expected order. On the other hand, the similar performance of the SAB3 and TT methods when using a high number of small steps must be due to the specific contributions of two opposing effects, the higher order and the more time-consuming step update of the SAB3 method.

Besides these considerations, the overall efficiency of the SAB3 over the TT method for this example is 3 (max), 0.5 (min), and 1.6-fold (avg) for the error reduction, and 2 (max), 0.8 (min), and 1.2-fold (avg) for the speed-up. We conclude that even though the SAB3 performs worse than the TT method for very specific settings, it is on average 20% more efficient if we set the runtime and measure the error and 60% if we set the error and measure the runtime.

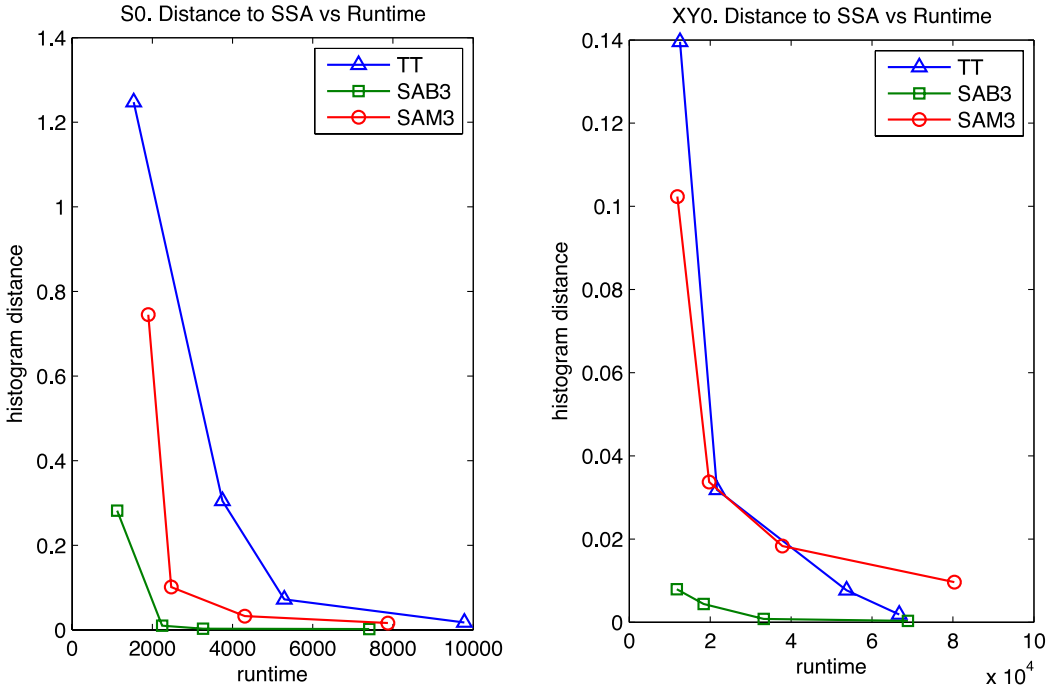


FIG. 3. Linear and nonlinear test problems: histogram distance vs runtime. Simulation setting as described in Figs. 1 and 2.

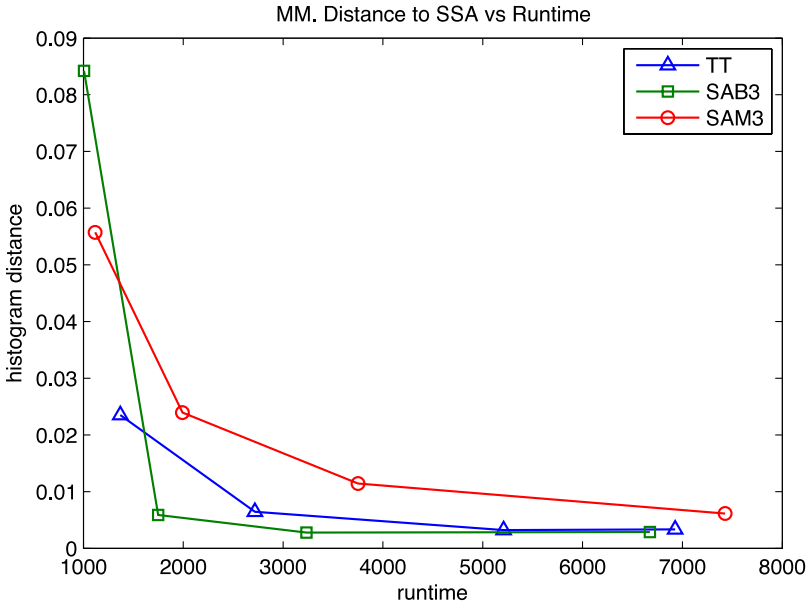


FIG. 4. Michaelis-Menten system: histogram distance vs runtime. Compared results of the TT, SAB with $k = 3$ (SAB3), and SAM with $k = 3$ (SAM3) methods using stepsizes $[0.05, 0.1, 0.2, 0.4]$, simulation time $T = 6.4$ and sample size 10^9. The runtime has been recorded with standard Matlab stopwatch timer functions.

We have also obtained results on the order of convergence of the stochastic Adams methods when applied to more complicated systems such as the Michaelis-Menten example. Specifically, the estimated order is 2.96 (mean), 2.65 (variance) for the SAB3 method and 2.48 (mean), 1.11 (variance) for the SAM3 method (plots not included).

B. EGFR system

This is a simplified model of transcription factor formation in the epidermal growth factor receptor (EGFR). This model has been described in several papers[21,15] and used to compare the efficiency of different stochastic simulation methods. We include it in this paper as a more complicated test problem using the same settings as in Ref. 15. The model is described in terms of 8 species,

$$(A, B, E_A, E_B, E_AB, E_AB_2, E_BA, E_BA_2)^\top,$$

and 12 reactions

$$\begin{aligned}
E_A &\to E_A + A, & c_1 &= 15,\\
E_B &\to E_B + B, & c_2 &= 15,\\
E_A + B &\to E_AB, & c_3 &= 0.0001,\\
E_AB &\to E_A + B, & c_4 &= 0.6,\\
E_AB + B &\to E_AB_2, & c_5 &= 0.0001,\\
E_AB_2 &\to E_AB + B, & c_6 &= 0.6,\\
A &\to \varnothing, & c_7 &= 0.5,\\
E_B + A &\to E_BA, & c_8 &= 0.0001,\\
E_BA &\to E_B + A, & c_9 &= 0.6,\\
E_BA + A &\to E_BA_2, & c_{10} &= 0.0001,\\
E_BA_2 &\to E_BA + A, & c_{11} &= 0.6,\\
B &\to \varnothing, & c_{12} &= 0.5.
\end{aligned}$$

Figure 5 shows the results of the histogram distance for this example using the set of methods TT, SAB3, and SAM3 methods. The results are consistent with those obtained for the Michaelis-Menten system, with the exception of the SAM3 method that performs very similar, if not better, than the SAB3 method. Clearly, this category of stochastic Adams methods shows an important efficiency gain over the TT method. Specifically, the error reduction of the SAB3 over the TT method is 1.8 (max), 1.1 (min), and 1.4-fold (avg) when we set the runtime and the speed-up is 2.5 (max), 1.5 (min), and 1.7-fold (avg) when we set the error level.

As with the previous example, we have also used this more complicated one to estimate the order of convergence of the stochastic Adams methods. Thus, fitted orders show to be 3.65 (mean), 2.96 (variance) and 2.90 (mean), 0.60 (variance) for the SAB3 and SAM3 methods, respectively (plots not included).

C. Other results

In addition to the examples shown in this paper, we have implemented a bistable chemical system, namely, the Schlögl system,[32] an autocatalytic scheme consisting of a trimolecular state that evolves to a steady state that is distributed according to a bimodal distribution. This is also a very common test problem since it allows a histogram distance test, even though the mean and variance are not very meaningful measures of the steady state probability distribution.

We have not been able to perform a clear characterization of the order of convergence of the SAB and SAM methods using this example. This has also been the case with the TT method, and this is a key issue since it is the starting procedure of the SAB and SAM methods and hence would directly affect their order of convergence.

The efficiency results (not included in this paper) of the SAB3 and SAM3 methods are very similar to those of the TT method. For this particular example, there is not a noticeable efficiency gain as seen in the examples described in Sec. V. Further work would be needed to understand the specific

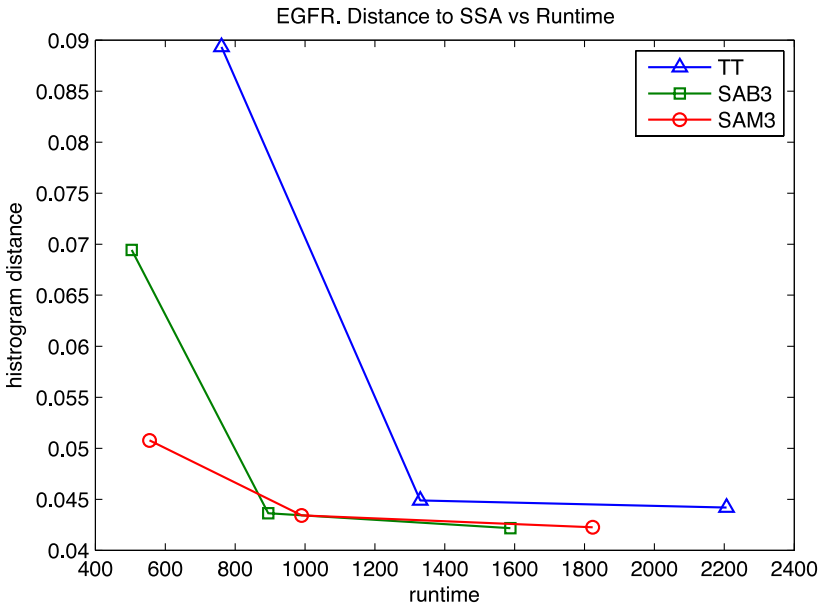


FIG. 5. EGFR system: histogram distance vs runtime. Compared results of the TT, SAB with $k=3$ (SAB3), and SAM with $k=3$ (SAM3) methods using stepsizes $[0.05, 0.1, 0.2]$, initial state $(2000, 1500, 950, 950, 200, 50, 200, 50)^{\top}$, simulation time $T=3.2$, and sample size 10^6.

details of why this happens and whether this is directly related to the bistability property.

Another result worth noting is that while the stochastic Adams-Moulton methods developed in this paper use a deterministic predictor as described in Eq. (6), there is a straightforward implementation of a stochastic predictor by just drawing from a Poisson distribution with parameter the expected value that appears in the deterministic version. We had conjectured that the stochastic predictor might improve the variance order of convergence of the SAM3 method since it would lead to a double Poisson distribution (see discussion on this issue in Sec. III). However, this stochastic predictor has not shown any improvement on the order of convergence of the stochastic Adams-Moulton methods. In terms of efficiency, the results are not any better since the stochastic predictor requires an extra random number generation per step. Again, further study would be needed to understand the specific effects of the stochastic predictor on the proposed methods.

There is another source of variability on the examples we have simulated, namely, the reaction rate constants. In order to keep our analysis simple, we have focused on the linear and nonlinear test problems since they include just one reaction with an associated rate constant. This is a parameter that can be modified in order to test the robustness of our findings. The examples are given with a specific rate constant, but we have also simulated them for different reaction rate constants multiplying and dividing by 2 successively. In general, we have found that there is a wide range of values where order is not affected. However, for very small reactions rates, the order seems to be noticeably reduced suggesting stability issues here. For the test problems, this effect starts to become pronounced at approximately 1/4 of the values given in the description of the examples.

VI. DISCUSSION AND CONCLUSIONS

This paper brings ideas from the numerical solution of ordinary differential equations and develops a class of stochastic linear multistep methods for solving problems in discrete chemical kinetics. The class is based on Adams Bashforth methods (explicit Adams methods) and Adams Moulton methods (predictor corrector implementation).

An order theory is developed, and, as in the deterministic case, the order of the starting procedure is a vital aspect in establishing the order of a stochastic linear multistep method. Thus, we use the Θ-trapezoidal τ-leap method that is known to be of weak order two and based on that we show that the k-step SAB methods with $k \geq 3$ have order three in both the mean and the correlation. In the case of the k-step predictor corrector SAM methods with $k \geq 3$, the mean order is shown to be three but the correlation order is only one.

Numerical simulations support this theoretical analysis, while simulations on some important application problems show the efficacy of these new methods compared with methods in the literature.

Further work is needed in modifying the predictor-corrector implementation to try to improve the correlation order beyond one.

ACKNOWLEDGMENTS

This research was partially supported by the Australian Research Council Grant No. DP120103770.

APPENDIX A: TOPICS ON THE STABILITY OF LMMS

LMMs are ordinary differential equations solvers that take the form

$$y_{n+1} = \sum_{j=1}^{k} \alpha_j y_{n+1-j} + h \sum_{j=0}^{k} \beta_j f_{n+1-j}. \tag{A1}$$

An important property of a LMM is zero-stability.

Definition A.1. A LMM is said to be zero-stable if, when (A1) is applied to $y' = 0$, it gives bounded solutions for all n.

Remark. This is equivalent to the polynomial $p(z) = z^k - \sum_{j=1}^{k} \alpha_k z^{k-j}$ *having zeros in the unit disk with simple zeros on the boundary.*

Remark. There exist LMMs of order $2k$ *but these are not zero-stable. In fact, the maximum order of any practical zero-stable method is* $k+1$. *(In fact, there exist zero-stable methods of order* $k+2$, *but these have stability regions that only contain a single point, namely, the origin—see later).*

The stability of a numerical method can be assessed by applying the method to the linear test problem

$$y' = qy, \quad Re(q) \le 0. \tag{A2}$$

The stability region, S, is defined as $S = \{z \in \mathbb{C}^{-1}, z = hq, Re(z) \le 0\}$. A LMM gives stable results when applied to (A2).

In the case of LMM, (A1) let

$$\rho(v) = 1 - \sum_{j=1}^{k} \alpha_j v^j, \qquad \sigma(v) = \sum_{j=0}^{k} \beta_j v^j$$

and define the stability polynomial

$$\Pi(v,z) = \rho(v) - z\sigma(v). \tag{A3}$$

Then, the stability region of any LMM is the set of z for which $\Pi(v,z)$ satisfies the root condition. Note that if a LMM satisfies the consistency condition then $z = 0 \in S$. The following stability results are important.

Definition A.2. A LMM is said to be A-stable if $\mathbb{C}^{-1} \subseteq S$.

Theorem A.1. *The maximum order of an A-stable LMM is two.*[9]

Table II[4] gives the real stability interval, I, for predictor-corrector Adams pairs.

APPENDIX B: WEAK ANALYSIS OF THE ADAMS METHODS

Given a predictor (explicit) formulation, our SABM applied to the problem $X \xrightarrow{c} \varnothing$ with rate constant c takes the form

$$X_{n+1} = X_n - \mathcal{P}(z \sum_{j=1}^{k} B_j X_{n+1-j}), \quad z = \tau c. \tag{B1}$$

In the case of a deterministic predictor, then our stochastic predictor-corrector formulation based on the SAMM can be written as

$$\hat{X}_{n+1} = X_n - z \sum_{j=1}^{k} B_j X_{n+1-j}$$
$$X_{n+1} = X_n - \mathcal{P}(z \sum_{j=1}^{k} \hat{B}_j X_{n+1-j} + z B_0 \hat{X}_{n+1}).$$

This simplifies to

$$X_{n+1} = X_n - \mathcal{P}(z \sum_{j=1}^{k} \alpha_j X_{n+1-j}), \tag{B2}$$

where

$$\alpha_1 = \hat{B}_0 + \hat{B}_1 - z\hat{B}_0 B_1,$$
$$\alpha_j = \hat{B}_j - z\hat{B}_0 B_j, \quad j = 2, \ldots, k.$$

Note that the $k = 2$ Adams-Bashforth method has $B_1 = \frac{3}{2}$, $B_2 = -\frac{1}{2}$, while the $k = 2$ Adams-Moulton has $\hat{B}_0 = \frac{5}{12}$, $\hat{B}_1 = \frac{8}{12}$, $\hat{B}_2 = -\frac{1}{12}$, in which case

$$\alpha_1 = \frac{13}{12} - z\frac{5}{8}, \quad \alpha_2 = -\frac{1}{12} + z\frac{5}{24}.$$

In order to study the order of these stochastic Adams methods we will find it useful to define the vector $Z_n = (X_{n+k-1}, \ldots, X_n)^\top$ and let A be a $k \times k$ matrix given by

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

then, formulation (B1) and (B2) can be written as

$$Z_{n+1} = AZ_n - F(Z_n), \tag{B3}$$

where $F(Z_n) = (\mathcal{P}(b^\top Z_n), 0, \ldots, 0)^\top$, with $b^\top = (B_1, \ldots, B_k)$, or $(\alpha_1, \ldots, \alpha_k)$ if using formulation (B2).

Furthermore, let S be a starting procedure that generates $X_1, \ldots, X_{k-1}$ from X_0 and let $Z_0 = (X_{k-1}, \ldots, X_0)^\top$.

In order to analyse the order of our linear multistep method with respect to the starting procedure, we need the following results.

Lemma 1. The conditioned mean and variance satisfy

$$\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X|Y]], \tag{B4}$$
$$\mathrm{Var}(X) = \mathrm{Var}(\mathrm{E}[X|Y]) + \mathrm{E}[\mathrm{Var}(X|Y)]. \tag{B5}$$

Furthermore, given a Poisson process $\mathcal{P}(\lambda)$ *then*

$$\mathrm{E}[\mathcal{P}(\lambda)] = \lambda, \quad \mathrm{E}[\mathcal{P}^2(\lambda)] = \lambda^2 + \lambda.$$

Lemma 2. For the linear test problem $X \xrightarrow{c} \varnothing$, *the following results hold for the exact solution* $X(t)$:

$$\mathrm{E}[X(t)] = \mathrm{e}^{-ct} X_0,$$
$$\mathrm{Var}[X(t)] = (\mathrm{e}^{-ct} - \mathrm{e}^{-2ct}) X_0,$$
$$\mathrm{E}[X(t+\tau)X(t)] = \mathrm{e}^{-c\tau}\mathrm{E}[X^2(t)]. \tag{B6}$$

TABLE II. Effect of correction on the stability region.

Order	Mode	I
3	P	[−0.55,0]
3	PECE	[−1.8,0]
4	P	[−0.3,0]
4	PECE	[−1.3,0]

Proof. The solution of the linear test problem satisfies

$$X(t+\tau) = X(t) - \mathcal{P}(\int_t^{t+\tau} cX(s)ds). \tag{B7}$$

Taking expectations conditioned on $X(t)$ and using Eq. (B4) in Lemma 1 gives

$$\mathrm{E}[X(t+\tau)] = \mathrm{E}[X(t)] - c\int_t^{t+\tau} \mathrm{E}[X(s)]ds.$$

Let $m(t) = \mathrm{E}[X(t)]$, then $m(t)$ satisfies the ODE $m'(t) = -cm(t)$ and hence, $m(t) = \mathrm{e}^{-ct}X_0$.

The second result comes from using Eqs. (B5) and (B7), while the third result comes form Eq. (B7) and noting that

$$X(t+\tau)X(t) = X^2(t) - X(t)\mathcal{P}(\int_t^{t+\tau} cX(s)ds)$$

so that

$$\mathrm{E}[X(t+\tau)X(t)] = \mathrm{E}[X^2(t)] - c\int_t^{t+\tau} \mathrm{E}[X(s)X(t)]ds,$$

that is,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{E}[X(t+\tau)X(t)] = -c\,\mathrm{E}[X(t+\tau)X(t)]$$

or

$$\mathrm{E}[X(t+\tau)X(t)] = \mathrm{e}^{-c\tau}\mathrm{E}[X^2(t)].$$

□

Now define

$$m_n = \mathrm{E}[Z_n], \quad n = 0,1,\ldots,$$

and taking expectation of Eq. (B3) conditioned on Z_n gives since

$$\mathrm{E}[\mathrm{E}[\mathcal{P}(b^\top Z_n)|Zn]] = b^\top \mathrm{E}Z_n, \text{ (from Lemma 1)},$$
$$m_{n+1} = GM_n, \tag{B8}$$

where

$$G = A - zB, \quad B = e_1 b^\top, \quad e_1 = (1,0,\ldots,0)^\top.$$

Hence,

$$m_N = G^N m_0. \tag{B9}$$

Our intention is to use the trapezoidal τ-leap method as a starting procedure and so the following Lemma will prove useful.

Lemma 3. For the test problem $X \xrightarrow{c} \varnothing$, the trapezoidal τ-leap method satisfies the following recurrence relation, with $z = c\tau$:

$$\mathrm{E}[X_{n+1}] = (1 - z + \frac{1}{2}z^2)\mathrm{E}[X_n]. \tag{B10}$$

Proof. For the linear test problem, the trapezoidal τ-leap method has the form

$$X^* = X_n - \mathcal{P}(X_n z/2), \quad X_{n+1} = X^* - \mathcal{P}((2X^* - X_n)z/2)$$

or equivalently,

$$X_{n+1} = X_n - \mathcal{P}(\theta_n) - \mathcal{P}(\theta_n - z\mathcal{P}(\theta_n)), \quad \theta_n = X_n z/2. \tag{B11}$$

Taking expectations conditioned on X_n and using Lemma 1 immediately gives

$$\mathrm{E}[X_{n+1}] = (1 - z + \frac{1}{2}z^2)\mathrm{E}[X_n],$$

where we have used the fact that $\mathrm{E}[\mathcal{P}(\lambda)] = \lambda$. □

Remark. It is immediately clear from Lemma 1, 2, and 3 that the trapezoidal τ-leap method is order 2 in the mean for the linear test problem and indeed for all linear problems as well.

We will first investigate the mean order of the stochastic Adams methods with respect to an arbitrary starting procedure.

We first note the following Lemma, which can be obtained from Corollary 2.1 by expanding the right-hand side around j^{l-1} rather than $(1-j)^{l-1}$.

Lemma 4. A k-step linear multistep method of Adams-type of order s satisfies the order conditions

$$\sum_{j=0}^{k} j^{l-1} b_j = \frac{1}{l}, \quad l = 1,\ldots,s. \tag{B12}$$

This can also be written as

$$(\mathrm{e}_1^\top - zb^\top)d = \mathrm{e}^{-kz} + \mathcal{O}(z^{s+1}),$$
$$d = (\mathrm{e}^{-(k-1)z},\ldots,1)^\top.$$

Hence,

$$b^\top d = \frac{1}{z}(\mathrm{e}^{-(k+1)z} - \mathrm{e}^{-kz}) + \mathcal{O}(z^s). \tag{B13}$$

Proof. This is a consequence of the order conditions (see Ref. 4, for example). □

Consider now the k-step AB method of order s and suppose that we have a starting procedure that is of order $p-1$ in the mean, then the starting procedure satisfies

$$\mathrm{E}[X_0] = X_0,$$
$$\mathrm{E}[X_l] = (\sum_{j=0}^{k} \frac{(-lz)^j}{j!} + l(C - \frac{1}{p!})(-z^p) + \mathcal{O}(z^{p+1}))X_0,$$
$$l = 1,\ldots,k-1,$$

where C is the leading error coefficient term.

Thus, we can write the vector $m_0 = \mathrm{E}[Z_0]$ as

$$m_0 = \begin{pmatrix} \mathrm{e}^{-(k-1)z} \\ \vdots \\ \mathrm{e}^{-z} \\ 1 \end{pmatrix} X_0 + (C - \frac{1}{p!}) \begin{pmatrix} k-1 \\ \vdots \\ 1 \\ 0 \end{pmatrix} (-z)^p X_0 + \mathcal{O}(z^{p+1}). \tag{B14}$$

We now have the elements to prove the following result.

Theorem B.1. *Given a k-step Adams method of order p or more and a starting procedure that has order $p-1$ in the mean then the global mean order of the combined method is p.*

Proof. First, let $e = (1,\ldots,1)^\top$ and $u = (k-1,\ldots,1,0)^\top$. Now from Eq. (B9),

$$m_N = G^N m_0$$

with

$$G = A - ze_1 b^\top$$

and

$$m_0 = dX_0 + (C - \frac{1}{p!})u(-z)^p X_0 + O(z^{p+1}).$$

Now, it is easily seen that with $Ae = e$, then for $N \geq k-1$,

$$G^N u = (k-1)e + O(z).$$

Furthermore, Eq. (B13) and Lemma 4 with $s = p$ imply

$$G^N d = (\mathrm{e}^{-(N+k-1)z},\ldots,\mathrm{e}^{-Nz})^\top + O(z^p).$$

Hence,

$$m_N = (\mathrm{e}^{-(N+k-1)z},\ldots,\mathrm{e}^{-Nz})^\top X_0 + (C - \frac{1}{p!})(k-1)e(-z^p)X_0 + O(Nz^{p+1}) + O(z^p)$$

and the result is proved since $O(Nz^{p+1}) = O(z^p)$. □

Corollary B.1. The SABM and SAMM based on the underlying Adams method of order three or more will be order three in the mean if the trapezoidal τ-leap starting procedure is used.

We now turn our attention to the covariance order. Given Eq. (B3), we will define C_{n+1} to be the correlation matrix at time point $n+1$. Thus,

$$C_{n+1} = \mathrm{E}[Z_{n+1}Z_{n+1}^\top] - (\mathrm{E}[Z_{n+1}])(\mathrm{E}[Z_{n+1}])^\top = \mathrm{E}[Z_{n+1}Z_{n+1}^\top] - (m_{n+1})(m_{n+1})^\top.$$

From Eq. (B3),

$$Z_{n+1} = AZ_n - F(Z_n) = AZ_n - \mathcal{P}(b^\top Z_n)e_1,$$
$$Z_{n+1}^\top = Z_n^\top A^\top - e_1\mathcal{P}(b^\top Z_n).$$

First define $S_{n+1} = \mathrm{E}[Z_{n+1}Z_{n+1}^\top]$. We will now take means conditioned on Z_n. Thus,

$$S_{n+1} = AS_nA^\top - zAS_nB^\top - zBS_nA^\top + \mathrm{E}[F(Z_n)F(Z_n)^\top],$$

where we have used the fact that $\mathrm{E}[Z_nF(Z_n)^\top] = zS_nB^\top = zS_nbe_1^\top$.

Furthermore, from Lemma 1,

$$\mathrm{E}[F(Z_n)F(Z_n)^\top] = \mathrm{E}[P^2(zb^\top Z_n)]e_1e_1^\top = (z^2b^\top S_nb + zb^\top m_n)e_1e_1^\top,$$

then

$$S_{n+1} = (A - zB)S_n(A - zB)^\top + zb^\top m_nE, \quad E = e_1e_1^\top. \tag{B15}$$

Hence, with $G = A - zB$ and Eq. (B9),

$$C_{n+1} = S_{n+1} - m_{n+1}m_{n+1}^\top = S_{n+1} - Gm_nm_n^\top G^\top,$$

so that Eqs. (B15) and (B8) give

$$C_{n+1} + m_{n+1}m_{n+1}^\top = G(C_n + m_nm_n^\top)G^\top + zb^\top m_nE \quad \text{or}$$
$$C_{n+1} = GC_nG^\top + zb^\top m_nE. \tag{B16}$$

The solution to this simple linear recurrence relation is

$$C_N = G^NC_0(G^N)^\top + z\sum_{j=0}^{N-1} b^\top m_{N-1-j}G^jE(G^j)^\top \tag{B17}$$

which from Eq. (B9) becomes

$$C_N = G^NC_0(G^N)^\top + z\sum_{j=0}^{N-1} b^\top G^{N-1-j}m_0G^je_1(e_1G^j)^\top. \tag{B18}$$

As before, we can assume the starting procedure is of order $p-1$ in the mean and covariance. Thus, as before, we can write m_0 as

$$m_0 = dX_0 + (C - \frac{1}{p!})u(-z)^pX_0 + O(z^{p+1})$$

and from Lemma 2, the starting procedure will have

$$C_0 = \begin{pmatrix} V_{k-1} & \cdots & \cdots & \cdots & R^{k-1}V_0 \\ RV_{k-2} & V_{k-2} & \cdots & \cdots & \vdots \\ R^2V_{k-3} & RV_{k-3} & V_{k-3} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R^{k-1}V_0 & R^{k-2}V_0 & R^{k-3}V_0 & \cdots & V_0 \end{pmatrix}, \tag{B19}$$

where

$$R(z) = (\sum_{j=0}^{p}\frac{(-z)^j}{j!} + (C - \frac{1}{p!})(-z)^p)X_0 = \mathrm{e}^{-z}X_0 + O(z^p)$$

and

$$V_l(z) = (\mathrm{e}^{-lz} - \mathrm{e}^{-2lz})X_0 + O(z^p), \quad l = 1,\ldots,k-1,$$
$$V_0(z) = 0.$$

Noting what happens when multiplying $R(z)$ by $V_l(z)$, we can write C_0 as

$$C_0 = C_{00} - C_{01} + O(z^p),$$

where

$$C_{00} = \begin{pmatrix} \mathrm{e}^{-(k-1)z} & \mathrm{e}^{-(k-1)z} & \mathrm{e}^{-(k-1)z} & \cdots & \mathrm{e}^{-(k-1)z} \\ \mathrm{e}^{-(k-1)z} & \mathrm{e}^{-(k-2)z} & \mathrm{e}^{-(k-2)z} & \cdots & \mathrm{e}^{-(k-2)z} \\ \cdots & \cdots & \cdots & \ddots & \vdots \\ \mathrm{e}^{-(k-1)z} & \mathrm{e}^{-(k-2)z} & \mathrm{e}^{-(k-3)z} & \cdots & 1 \end{pmatrix}$$

and

$$C_{01} = \begin{pmatrix} \mathrm{e}^{-(2k-2)z} & \mathrm{e}^{-(2k-3)z} & \cdots & \mathrm{e}^{-(k-1)z} \\ \mathrm{e}^{-(2k-3)z} & \mathrm{e}^{-(2k-4)z} & \cdots & \mathrm{e}^{-(k-2)z} \\ \cdots & \cdots & \ddots & \vdots \\ \mathrm{e}^{-(k-1)z} & \mathrm{e}^{-(k-2)z} & \cdots & 1 \end{pmatrix}.$$

Defining $D = \mathrm{Diag}(\mathrm{e}^{-(k-1)z},\ldots,1)$, $d = De$, $e = (1,\ldots,1)^\top$ then clearly

$$C_{01} = Dee^\top D = dd^\top. \tag{B20}$$

An elegant representation of C_{00} needs more consideration. Define the vectors $f_j \in \mathbb{R}^k$, $j = 1, \ldots, k$ as

$$f_1 = (1,1,\ldots,1,1)^\top,\ f_2 = (0,1,\ldots,1,1)^\top, \ldots,$$
$$f_k = (0,0,\ldots,0,1)^\top.$$

Furthermore, let $f_{k+1} = (0,0,\ldots,0,0)^\top$ then, we can write C_{00} as

$$C_{00}(z) = \mathrm{e}^{-(k-1)z} f_1 f_1^\top + \sum_{j=2}^{k} (\mathrm{e}^{-(k-j)z} - \mathrm{e}^{-(k+1-j)z}) f_j f_j^\top, \quad \text{(B21)}$$

which is a sum of rank one updates.

Now, from Eq. (B18), C_N consists of two terms. First note that

$$G^N C_0 (G^N)^\top = G^N C_{00} (G^N)^\top - G^N C_{01} (G^N)^\top + O(z^p).$$

But from Lemma 4 and with $d_N = (\mathrm{e}^{-(k+N-1)z}, \ldots, \mathrm{e}^{-Nz})^\top$, $N = 1,2,\ldots,$

$$Gd = d_1 + O(z^{p+1}) \mathrm{e}_1^\top,$$

and by induction and Eq. (B20),

$$G^N C_{01} (G^N)^\top = (G^N d)(G^N d)^\top = d_N d_N^\top + O(z^{2p+2}). \quad \text{(B22)}$$

Now consider $L = G^N C_{00} (G^N)^\top$, which from Eq. (B21) is

$$L = \mathrm{e}^{-(k-1)z} (G^N f_1)(G^N f_1)^\top + \sum_{j=2}^{k} (\mathrm{e}^{-(k-j)z} - \mathrm{e}^{(k+1-j)z}) (G^N f_j)(G^N f_j)^\top.$$

In order to make further progress with the global error, we should first understand the local error. Over one step, the correlation matrix is given by $C_1 = G C_0 G^\top + z b^\top m_0 e_1 e_1^\top$, which from Lemma 4 is

$$C_1 = G C_0 G^\top + (\mathrm{e}^{-(k-1)z} - \mathrm{e}^{-kz}) e_1 e_1^\top + O(z^{s+1}). \quad \text{(B23)}$$

Now write $C_0 = C_{00} - C_{01} + O(z^p)$ and note from Eq. (B22) that

$$G C_{01} G^\top = d_1 d_1^\top + O(z^{2p+2})$$

and from Eq. (B21) that

$$G C_{00} G^\top = \mathrm{e}^{-(k-1)z} (G f_1)(G f_1)^\top + \sum_{j=2}^{k} (\mathrm{e}^{-(k-j)z} - \mathrm{e}^{-(k+1-j)z}) (G f_j)(G f_j)^\top.$$

It is easily seen that

$$G f_1 = f_2 + (1 - z\gamma) e_1,$$
$$G f_j = f_{j+1} - (z \sum_{i=j}^{k} b_i) e_1, \quad j = 2, \ldots, k,$$

where

$$\gamma = \sum_{j=1}^{k} b_j.$$

Hence,

$$G C_{00} G^\top = \mathrm{e}^{-(k-1)z} f_2 f_2^\top + \sum_{j=2}^{k} (\mathrm{e}^{-(k-j)z} - \mathrm{e}^{-(k+1-j)z}) f_{j+1} f_{j+1}^\top + \mathrm{e}^{-(k-1)z} (1 - z\gamma)(f_2 + f_2^\top) - z \sum_{j=2}^{k} (\mathrm{e}^{-(k-j)z} - \mathrm{e}^{-(k+1-j)z}) (\sum_{l=j}^{k} b_l)(f_{j+1} f_{j+1}^\top) + (\mathrm{e}^{-(k-1)z} (1 - z\gamma)^2 + z^2 \sum_{j=2}^{k} (\mathrm{e}^{-(k-j)z} - \mathrm{e}^{-(k+1-j)z}) (\sum_{l=j}^{k} b_l)^2) e_1 e_1^\top.$$

Combining this with Eq. (B23), we only have to study the expansion in z of the element (1,1) of $G C_{00} G^\top$. It can be shown that this is

$$1 - z(k - 2(1-\gamma)) + z^2 (\frac{k^2}{2} - 2k(1-\gamma) + (1-\gamma)^2) - z^3 (\frac{k^3}{6} + k^2(\gamma - 1) + k(\gamma - 2)^2 - \frac{1}{2} + \gamma - \gamma^2 - \sum_{j=2}^{k} (\sum_{l=j}^{k} b_l)^2) + O(z^4).$$

Now, if $\gamma = 1$, this becomes

$$\mathrm{e}^{-kz} + z^3 (\frac{1}{2} + \sum_{j=2}^{k} (\sum_{l=j}^{k} b_l)^2) + O(z^4).$$

Since the z^3 term cannot be made zero, the local error in the correlation matrix is at most $O(z^3)$. Furthermore, if $\gamma = 1 + \theta z$, $\theta \neq 0$, which is the case with the predictor-corrector formulation, this gives

$$\mathrm{e}^{-kz} - \theta z^2 + O(z^3)$$

and so, the local error in the correlation matrix is at most $O(z^2)$ for the SAM method.

Thus, we have proved the following result.

Theorem B.2. *Given a starting procedure of order $p \geq 2$ in the mean and correlation, the SAB method of order $s \geq 2$ will have local correlation order $O(z^3)$, while the SAM method is $O(z^2)$.*

All that remains to be seen is whether the global convergence error is one more than the local order or not. Now from Eq. (B18), repeated use of Lemma 4 gives that the second term in Eq. (B18) is

$$z \sum_{j=0}^{N-1} b^\top G^{N-1-j} m_0 (G^j e_1)(e_1 G^j)^\top = \sum_{j=0}^{N-1} (\mathrm{e}^{-(k+N-2-j)z} - \mathrm{e}^{-(k+N-1-j)z}) (G^j e_1)(e_1 G^j)^\top. \quad \text{(B24)}$$

Furthermore, if $N \geq k+2$, it can be shown that

$$G^N f_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} - \gamma z \begin{pmatrix} N \\ N-1 \\ \vdots \\ N-k \end{pmatrix} - \frac{1}{2} \gamma z^2 \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_k \end{pmatrix} + O(z^3),$$

where

$$l_i = \sum_{j=1}^{k} b_j(N+2-j-i)(N+1-j-i), \quad i=1,\ldots,k$$
$$= N^2 \sum_{j=1}^{k} b_j - N \sum_{j=1}^{k} b_j(2j+2i-3)$$
$$+ \sum_{j=1}^{k} b_j(j+i-2)(j+i-1).$$

Furthermore,

$$G^N f_l = -z\mathrm{e} \sum_{j=l}^{k} (j+1-l) b_j + O(z^2), \quad l=2,\ldots,k.$$

Putting these elements together, we see that in the case of the predictor-corrector SAM method, the (1,1) element of C^N is

$$1 - z(k+N-1+2N(\gamma-1)) + O(z^2),$$

but if $\gamma = 1+\theta z$, this gives

$$1 - z(k+N-1) - 2\theta N z^2,$$

and so the predictor-corrector method only has global correlation order 1. On the other hand for the SAB method of order 3 or more, the global correlation order is still $O(z^3)$ and hence, the correlation order is three.

[1]D. F. Anderson, A. Ganguly, and T. G. Kurtz, "Error analysis of tau-leap simulation methods," Ann. Appl. Probab. **21**(6), 2226–2262 (2011).
[2]D. F. Anderson and J. C. Mattingly, "A weak trapezoidal method for a class of stochastic differential equations," Commun. Math. Sci. **9**(1), 301-318 (2011).
[3]K. Burrage and P. M. Burrage, "Order conditions of stochastic Runge–Kutta methods by B-series," SIAM J. Numer. Anal. **38**(5), 1626–1646 (2000).
[4]K. Burrage, *Parallel and Sequential Methods for Ordinary Differential Equations* (Clarendon Press, New York, NY, USA, 1995).
[5]Y. Cao, D. T. Gillespie, and L. R. Petzold, "The slow-scale stochastic simulation algorithm," J. Chem. Phys. **122**(1), 14116 (2005).
[6]Y. Cao, D. T. Gillespie, and L. R. Petzold, "Efficient step size selection for the tau-leaping simulation method," J. Chem. Phys. **124**(4), 044109 (2006).
[7]Y. Cao and L. Petzold, "Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems," J. Comput. Phys. **212**(1), 6–24 (2006).
[8]A. Chatterjee, D. G. Vlachos, and M. A. Katsoulakis, "Binomial distribution based τ-leap accelerated stochastic simulation," J. Chem. Phys. **122**(2), 024112 (2005).
[9]G. Dahlquist, "Convergence and stability in the numerical integration of ordinary differential equations," Math. Scand. **4**, 33–53 (1956).
[10]D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," J. Phys. Chem. **81**(25), 2340–2361 (1977).
[11]D. T. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems," J. Chem. Phys. **115**(4), 1716–1733 (2001).
[12]D. T. Gillespie and L. R. Petzold, "Improved leap-size selection for accelerated stochastic simulation," J. Chem. Phys. **119**(16), 8229–8234 (2003).
[13]Y. Hu and T. Li, "Highly accurate tau-leaping methods with random corrections," J. Chem. Phys. **130**(12), 124109 (2009).
[14]Y. Hu, T. Li, and B. Min, "The weak convergence analysis of tau-leaping methods: Revisited," Commun. Math. Sci. **9**, 965–996 (2011).
[15]Y. Hu, T. Li, and B. Min, "A weak second order tau-leaping method for chemical kinetic systems," J. Chem. Phys. **135**(2), 024113 (2011).
[16]T. G. Kurtz, "Limit theorems for sequences of jump Markov processes approximating ordinary differential processes," J. Appl. Probab. **8**(2), 344–356 (1971).
[17]T. G. Kurtz, "The relationship between stochastic and deterministic models for chemical reactions," J. Chem. Phys. **57**(7), 2976–2978 (1972).
[18]T. G. Kurtz, "Strong approximation theorems for density dependent Markov chains," Stochastic Processes Appl. **6**(3), 223–240 (1978).
[19]A. Leier, T. T. Marquez-Lago, and K. Burrage, "Generalized binomial tau-leap method for biochemical kinetics incorporating both delay and intrinsic noise," J. Chem. Phys. **128**(20), 205107 (2008).
[20]T. Li, "Analysis of explicit tau-leaping schemes for simulating chemically reacting systems," Multiscale Model. Simul. **6**(2), 417–436 (2007).
[21]T. T. Marquez-Lago and K. Burrage, "Binomial tau-leap spatial stochastic simulation algorithm for applications in chemical kinetics," J. Chem. Phys. **127**(10), 104101 (2007).
[22]X. Peng, W. Zhou, and Y. Wang, "Efficient binomial leap method for simulating chemical kinetics," J. Chem. Phys. **126**(22), 224109 (2007).
[23]M. F. Pettigrew and H. Resat, "Multinomial tau-leaping method for stochastic kinetic simulations," J. Chem. Phys. **126**(8), 84101 (2007).
[24]P. J. Prince and J. R. Dormand, "High order embedded Runge–Kutta formulae," J. Comput. Appl. Math. **7**(1), 67–75 (1981).
[25]M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, "Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method," J. Chem. Phys. **119**(24), 12784–12794 (2003).
[26]M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, "Consistency and stability of tau-leaping schemes for chemical reaction systems," Multiscale Model. Simul. **4**(3), 867–895 (2005).
[27]L. F. Richardson, "The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam," Philos. Trans. R. Soc., A **210**, 307–357 (1911).
[28]P. Rué, J. Villà-Freixa, and K. Burrage, "Simulation methods with extended stability for stiff biochemical kinetics," BMC Syst. Biol. **4**(1), 110 (2010).
[29]T. Székely, K. Burrage, R. Erban, and K. C. Zygalakis, "A higher-order numerical framework for stochastic simulation of chemical reaction systems," BMC Syst. Biol. **6**(1), 85 (2012).
[30]T. Székely, Jr., K. Burrage, K. C. Zygalakis, and M. Barrio, "Efficient simulation of stochastic chemical kinetics with the Stochastic Bulirsch-Stoer extrapolation method," BMC Syst. Biol. **8**(1), 71 (2014).
[31]T. Tian and K. Burrage, "Binomial leap methods for simulating stochastic chemical kinetics," J. Chem. Phys. **121**(21), 10356–10364 (2004).
[32]M. Vellela and H. Qian, "Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: The schlögl model revisited," J. R. Soc., Interface **6**(39), 925–940 (2009).
[33]Z. Xu and X. Cai, "Unbiased tau-leap methods for stochastic simulation of chemically reacting systems," J. Chem. Phys. **128**(15), 154112–154112 (2008).