MULTISCALE MODEL. SIMUL.
Vol. 4, No. 3, pp. 867–895

CONSISTENCY AND STABILITY OF TAU-LEAPING SCHEMES FOR CHEMICAL REACTION SYSTEMS*

MURUHAN RATHINAM†, LINDA R. PETZOLD‡, YANG CAO‡, AND DANIEL T. GILLESPIE§

Abstract. We develop a theory of local errors for the explicit and implicit tau-leaping methods for simulating stochastic chemical systems, and we prove that these methods are first-order consistent. Our theory provides local error formulae that could serve as the basis for future stepsize control techniques. We prove that, for the special case of systems with linear propensity functions, both tau-leaping methods are first-order convergent in all moments. We provide a stiff stability analysis of the mean of both leaping methods, and we confirm that the implicit method is unconditionally stable in the mean for stable systems. Finally, we give some theoretical and numerical examples to illustrate these results.

Key words. stochastic chemical kinetics, tau-leaping

AMS subject classifications. 60H35, 65C30

DOI. 10.1137/040603206

1. Introduction.

1.1. Background on tau-leaping methods.

Chemical reaction systems which involve small numbers of molecules of certain species may not be adequately modeled by a deterministic ordinary differential equation (ODE) model. Systems that involve small numbers of molecules of some species and very large numbers of molecules of other species may exhibit dynamics that range from discrete and stochastic to continuous and deterministic. Important examples of such systems are living cells [13, 12, 1, 4].

Gillespie [5, 6] describes a discrete stochastic microscopic model for well-stirred systems, and an algorithm called the stochastic simulation algorithm (SSA) for its exact simulation in which each chemical reaction event is simulated, one at a time. More recent papers [9, 10] show that starting from this "finer scale" model and making certain approximations, one may carry out a more efficient but approximate simulation called *tau-leaping*, which is valid under certain assumptions. Furthermore, it is shown that under a sequence of assumptions, which generally hold when the number of molecules of every species is very large, successive steps of approximations can be

*Received by the editors January 13, 2004; accepted for publication (in revised form) March 9, 2005; published electronically September 8, 2005. This work was supported by the California Institute of Technology under DARPA award F30602-01-2-0558, by the U.S. Department of Energy under DOE award DE-FG03-00ER25430, by the National Science Foundation under NSF award CTS-0205584, and by the Institute for Collaborative Biotechnologies through grant DAAD19-03-D-0004 from the U.S. Army Research Office.

http://www.siam.org/journals/mms/4-3/60320.html

†Department of Mathematics and Statistics, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250 (muruhan@math.umbc.edu). The research of this author was supported in part by grant NSF DMS-0309647.

‡Department of Computer Science, University of California, Santa Barbara, CA 93106 (petzold@engineering.ucsb.edu, ycao@engineering.ucsb.edu).

§Dan T. Gillespie Consulting, 30504 Cordoba Place, Castaic, CA 91384 (GillespieDT@mailaps.org). The research of this author was supported in part by the Air Force Office of Scientific Research and the California Institute of Technology under DARPA award F30602-01-2-0558 and in part by the Molecular Sciences Institute under contract 244725 with Sandia National Laboratories and the Department of Energy's "Genomes to Life" Program.

made that lead to a stochastic differential equation (SDE) model (i.e., one driven by Gaussian white noise) and ultimately to the ODEs that are well known to chemists as the reaction rate equations. Tau-leaping, together with known methods for SDEs and ODEs, provides a framework for the seamless approximation and solution of *multiscale* chemically reacting systems, where the scales are over a range of populations of chemical species from moderate to very large numbers.

Another sense in which chemically reacting systems are multiscale involves the wide variation of reaction rates. For deterministic systems, this type of multiscale behavior is commonly known as stiffness. In the chemical kinetics context, stiffness occurs due to the presence of slow and fast reactions where two or more fast reactions nearly cancel each other. It is well known that in such situations, explicit numerical methods perform very poorly compared to implicit methods [2]. The stiffness phenomenon can also be observed in the stochastic chemical model. The implicit tau method [15] was proposed as an alternative to the explicit tau method for stiff stochastic systems and was numerically demonstrated to outperform the explicit tau method for those systems.

1.2. Motivation and contributions of this work. The explicit tau-leaping method introduced in [9] uses a piecewise Poisson approximation of the continuous time discrete state Markov process that is described by the chemical master equation [8]. The stochastic chemical model described by the master equation is a Markov process with state transition rates a_j that are functions of the state x. In the context of chemical kinetics these rates are also known as propensity functions $a_j(x)$. These $a_j(x)$ are such that the probability of a type j reaction occurring in the time interval $[t, t+\tau)$, given that the system is in state x at time t, is $a_j(x)\tau + o(\tau)$ as $\tau \to 0$. Intuitively, one might expect that the behavior of explicit tau, as well as that of implicit tau, would approach for small τ that of the model described by the SSA (or equivalently by the chemical master equation [8], which is an infinite system of ODEs for the time evolution of the probabilities). However, a quantitative analysis of the errors incurred by the use of these approximate methods has not been provided in the literature. A quantitative estimate of the local errors as a function of stepsize τ should be useful for developing strategies for automatic stepsize selection, as well as a guide in the search for higher-order accurate tau-leaping methods.

In this paper we develop an analysis of local errors of both the explicit and the implicit tau methods. We do so by observing that the exact probability for any given sequence of reactions to occur in a finite time duration τ can be written as a sequence of iterated integrals involving the propensity functions and exponentials of the propensity functions. Although these integrals cannot be simplified in general, they can be expanded as a Taylor series in powers of τ. Using this approach we show that both the explicit and the implicit tau methods are first-order consistent in τ. In addition we provide explicit formulae for the leading-order local error term of $O(\tau^2)$ for the mean and covariance of explicit tau. We also provide a similar but less explicit formula for implicit tau. More importantly, our approach provides a systematic method for finding the local error terms of all moments up to any power of τ.

It must be noted that unlike an ODE, the discrete stochastic model can be simulated *exactly* since it involves a sequence of discrete events, albeit happening in continuous time. Thus any local error formulae that hold only for *small enough* τ may not be of relevance unless the range of validity includes τ values in which multiple reactions are likely to occur. This way the local error formulae could be applied

in situations where one actually would use the leaping methods. We look at this issue in the context of some examples and show that keeping only the τ^2 term in the local error for the mean and variance often provides reasonable approximations to the local error for τ values that are large enough that the mean number of reactions leapt over is on the order of 10 or more.

Consistency of a leaping method alone does not guarantee convergence over a finite interval. We restrict ourselves to systems with linear propensity functions and show that for both the explicit and the implicit tau methods, all moments converge to those of the correct solution with a global error of $O(\tau)$. The significance of this result is that it rigorously justifies the claim that the tau methods "seamlessly" transition into the SSA. If indeed tau methods diverged or converged to some processes other than the SSA, then one would have to be wary of using very small τ steps. This would also raise concerns about how small is too small. Knowing that they converge assures us that we need only concern ourselves with computational efficiency in deciding when to switch from the tau methods to the SSA.

We also provide finite stepsize asymptotic stability results for the mean of both tau methods for systems with linear propensity functions. These stability results confirm that the mean of the implicit tau is unconditionally stable if the original system is asymptotically stable, while explicit tau requires limitations on its stepsize to ensure stability. These results are very similar to the stiff stability results for explicit and implicit Euler methods applied to ODEs. This also explains why the implicit tau method performs better for stiff systems [15]. We note that a more thorough analysis of stiff stability involves looking at not just the mean but also the behavior of higher-order moments, and this is dealt with in a different paper [3].

1.3. Relationship to other multiscale approaches. One approach to understanding and analyzing multiscale behavior is via singular perturbation methods in which a small parameter ϵ is introduced in the system equations and the asymptotic behavior as $\epsilon \to 0$ of the system is derived. Notions of averaging microscopic variables to obtain macroscopic variables could also be dealt with in this framework. See [11] for a good overview. Such analysis may be used to motivate and analyze numerical methods for "macroscopic time stepping" [18]. These works primarily deal with ODEs and SDEs (driven by Brownian motion), but such analysis may be extended to discrete state continuous time Markov processes as well. We do not perform such an analysis in this paper. However, the convergence results and absolute stability results for the mean in this paper as well as the absolute stability results for all moments in [3] do provide an important first step towards a better understanding of tau-leaping methods, which have shown promise in simulation results [15].

1.4. Outline of the paper. In section 2 we briefly describe the stochastic chemical model and the tau-leaping schemes. We also discuss two issues that impact our analysis, namely extension of the propensity function to real states and bounding the number of reaction events to avoid negative states as well as an arbitrarily large number of firings. In section 3 we present the local error analysis of the tau-leaping methods and prove their first-order consistency. In section 4 we focus on systems with linear propensity functions and prove 0-stability and convergence of all moments for both tau-leaping methods. Furthermore, we provide finite stepsize asymptotic stability analysis of the mean for both tau methods. In section 5 we present some examples and provide both theoretical and numerical results showing first-order convergence of the tau methods, as well as the region of validity of the second-order local error formulae. Finally, we make some concluding remarks and suggestions for future work.

2. The stochastic chemical model and the tau-leaping schemes. In this section we describe the discrete state, continuous time Markov process model for well-stirred chemical reaction systems, as well as an exact simulation algorithm for this model known as the SSA [5, 6]. We also describe the tau-leaping schemes, both explicit and implicit, which were designed for efficient and approximate simulation of this model [9, 10, 15]. Throughout this paper, $\mathbb{Z}_+$ denotes the set of nonnegative integers and $\mathbb{R}_+$ the set of nonnegative real numbers.

Well-stirred chemical reaction systems can be accurately described by a continuous time discrete state Markov process [5]. A system consisting of N molecular species has state $X(t) \in \mathbb{Z}_+^N$, where $X_i(t)$ is the number of molecules of the ith species at time t. Suppose that M chemical reactions R_j $(j = 1, \ldots, M)$ may occur among these species. Associated with each reaction R_j is a nonnegative *propensity function* $a_j : \mathbb{Z}_+^N \to \mathbb{R}_+$ such that the probability that reaction R_j will happen in the next small time interval $[t, t+\tau)$ is $a_j(X(t))\tau + o(\tau)$ as $\tau \to 0$. Occurrence of a reaction R_j leads to a change in state of $\nu_j \in \mathbb{Z}^N$. The ν_j are determined by the type of reaction R_j and are independent of the state as well as time. This model leads to the probability density function $a_0(x)e^{-a_0(x)\tau}$ for the time $t + \tau$ for the next reaction, where x is the state at time t, and $a_0(x) = \sum_{j=1}^{M} a_j(x)$. Also, the probability that the next reaction is R_j is given by $\frac{a_j(x)}{a_0(x)}$. Knowing these two probability densities for the next reaction time and type, we can simulate the system one reaction event at a time. Since the density functions have simple forms, the simulation process is easy. This method is known as the SSA [6]. Given a system of chemical reactions, the form of the propensities $a_j(x)$ may be derived from fundamental principles under certain assumptions, and the a_j turn out to be polynomials [5]. Note that the dynamics of a system is completely characterized by N, M, and a_j, ν_j for $j = 1, \ldots, M$.

REMARK 2.1. *For any real chemical system starting in a nonnegative state, a reaction event should not lead to a negative state, since nonnegative states are not physically meaningful. The following condition on a_j and ν_j is necessary and sufficient to ensure this. The condition is that for any $x \in \mathbb{Z}_+^N$, if $x + \nu_j \notin \mathbb{Z}_+^N$, then $a_j(x) = 0$. Note that these conditions are automatically captured by the polynomial propensities that come from Gillespie's model.*

ASSUMPTION 2.2. *Throughout this paper we shall assume that given any initial state $X(0) = x \in \mathbb{Z}_+^N$, there is a bounded subset $\Omega_x \subset \mathbb{Z}_+^N$ such that the system remains in Ω_x for all time with probability* 1. *In other words,*

$$P\{X(t) \in \Omega_x \ \forall t \geq 0\} = 1. \tag{2.1}$$

2.1. The explicit and implicit tau methods. The *explicit tau-leaping* method has been described in [9, 10] together with certain stepsize selection criteria. Throughout this paper we shall refer to the state update method to be described below as the explicit tau method, and this terminology shall apply regardless of what stepsize selection criterion is used.

Suppose $X^{(et)}(t) = x$ is the current state (the superfix "*et*" stands for explicit tau). Then for a time step of $\tau > 0$, the state at $t + \tau$ is given by

$$X^{(et)}(t+\tau) = x + \sum_{j=1}^{M} \nu_j K_j^{(et)}(x, \tau). \tag{2.2}$$

Here $K_j^{(et)}(x, \tau) = \mathcal{P}_j(a_j(x), \tau)$ for $j = 1, \ldots, M$ are independent Poisson random variables with mean and variance $a_j(x)\tau$.

The implicit tau (unrounded version) method proposed in [15] proceeds as follows. Given that $X^{(it)}(t) = x$ is the current state, the state at time $t + \tau$ ($\tau > 0$) is taken to be

$$X^{(it)}(t+\tau) = x + \sum_{j=1}^{M} \nu_j a_j(X^{(it)}(t+\tau))\,\tau + \sum_{j=1}^{M} \nu_j \left(\mathcal{P}_j(a_j(x),\tau) - a_j(x)\,\tau\right). \tag{2.3}$$

Here the superfix "*it*" stands for implicit tau.

The unrounded implicit tau has the disadvantage that it leads to state values that are not integers. In order to circumvent this problem, the rounded implicit tau was proposed in [15]. It may be described as follows.

Suppose that at time t we have the state $X^{(itr)}(t) = x$. (The superfix "*itr*" stands for rounded implicit tau.) First, compute the intermediate state value X' according to (2.3). Thus

$$X' = x + \sum_{j=1}^{M} \nu_j a_j(X')\,\tau + \sum_{j=1}^{M} \nu_j \left(\mathcal{P}_j(a_j(x),\tau) - a_j(x)\,\tau\right). \tag{2.4}$$

Then approximate the actual number of firings $K_j(x,\tau)$ of reaction channel R_j in the time interval $(t, t+\tau]$ by the *integer-valued* random variable $K_j^{(itr)}(x,\tau)$ defined by

$$K_j^{(itr)}(x,\tau) = [a_j(X')\,\tau + \mathcal{P}_j(a_j(x),\tau) - a_j(x)\,\tau]. \tag{2.5}$$

Here the $\mathcal{P}_j(a_j(x),\tau)$ for $j = 1,\ldots,M$ are the *same numbers* used in (2.4), and $[z]$ denotes the nearest nonnegative integer to z.

Finally, take the state at time $t+\tau$ to be

$$X^{(itr)}(t+\tau) = x + \sum_{j=1}^{M} \nu_j K_j^{(itr)}(x,\tau). \tag{2.6}$$

If $X^{(itr)}(t) = x$ is an integer vector, then so is $X^{(itr)}(t+\tau)$.

Throughout the rest of this paper, implicit tau shall mean the unrounded implicit tau, unless specified otherwise.

2.2. Extension of propensity function to real states. For the purpose of analysis as well as practical computation of the implicit tau, it is important to extend the definition of propensity function to positive real states. The implicit tau method can produce states that are in $\mathbb{R}_+^N$ but are not necessarily integers. Even though physically this is not meaningful, it nevertheless results in a mathematically well-defined process because the propensity functions $a_j(x)$ are polynomials, and hence have a natural extension to real numbers. However, if one or more of the propensity functions become negative, then the process becomes ill defined. This is easily fixed by setting to zero any propensity function $a_j(x)$ that evaluates to being negative.

Consider the example of a single reaction, single species case

$$S_1 + S_1 \to 0.$$

The propensity function $a(x) = \frac{1}{2}cx(x-1)$ will be negative if $0 < x < 1$. Even when the state is positive, the propensity can be negative. This does not happen with integer states. When x is an integer, $a(x) = 0$ for both $x = 0$ and $x = 1$, and $a(x) > 0$

for $x \geq 2$. Thus it makes sense to set $a(x) = 0$ for $0 < x < 1$. When $1 < x < 2$, $a(x) > 0$, however, the occurrence of a reaction will lead to an unrealistic negative state $x - 2$. This may be avoided by defining $a(x) = 0$ for $1 < x < 2$ as well. But then this leads to a discontinuous propensity function $a(x)$. From the point of view of our analysis, it is far more convenient to have continuity of $a(x)$ to ensure consistency results. We shall adopt the following extension of the propensity function.

DEFINITION 2.3 (extension of propensity function to positive real states). *Given a polynomial propensity function $a_j : \mathbb{Z}_+^N \to \mathbb{R}$, its extension $a_j : \mathbb{R}_+^N \to \mathbb{R}$ to any nonnegative real vector x is defined as follows. If the value of $a_j(x)$ according to the "natural extension" is nonnegative, then $a_j(x)$ is given by the natural extension. Otherwise $a_j(x) = 0$.*

The following lemma is easy to prove.

LEMMA 2.4. *If the original propensity functions $a_j : \mathbb{Z}_+^N \to \mathbb{R}$ are polynomials, then their extensions $a_j : \mathbb{R}_+^N \to \mathbb{R}$ as defined by Definition 2.3 are Lipschitz continuous on any bounded subdomain of $\mathbb{R}_+^N$.*

REMARK 2.5. *None of the consistency and convergence results proven in this paper require a_j to be polynomials. But the results for implicit tau assume them to be Lipschitz continuous on any bounded subdomain of $\mathbb{R}_+^N$.*

2.3. Bounding of the Poisson random variables. All the tau-leaping methods are based on generating Poisson random numbers with given mean and variance. The Poisson distribution with nonzero mean assigns nonzero probability to arbitrarily large numbers. This can cause some practical as well as theoretical problems.

One problem is that it may produce negative state values, which are nonphysical. Then the propensities may become negative for negative states. This can lead to a process which is both mathematically and computationally ill defined, since probabilities, and hence propensities, cannot be negative. For instance, consider the isomerization reaction

$$S_1 \to 0.$$

The propensity function is $a(x) = cx$, where $c > 0$ is some constant. Starting at an initial state x_0, there is a nonzero probability that the first step of explicit tau may produce a negative state. If this happens, then the explicit tau method becomes ill defined. In order to avoid this situation, whenever a negative state is reached we shall replace it with a nonnegative state as follows.

Bounding the K_j to avoid negative states. Suppose the state before the tau-leap is x (nonnegative), the state reached by the leap x_n is a negative state, and the number of reactions that occurred according to the leap is K_j for $j = 1, \ldots, M$. Then

while x_m has negative components
  for $l = 1$ to M
    $K_l \leftarrow K_l - 1$;
    $x_m = \sum_{j=1}^{M} \nu_j K_j$;
    if x_m is nonnegative, then break;
  end for
end while

The resulting state x_m is taken to be the new updated state. Also note that this procedure will terminate (will not result in an infinite loop).

The second problem is that even if the true system satisfies Assumption 2.2, the tau methods using Poisson random numbers will generate arbitrarily large numbers with nonzero probability. This unboundedness leads to difficulties with the implicit tau method because the range of τ values for which the implicit method is well defined could possibly get arbitrarily small. In order to avoid this problem, throughout the rest of the paper we shall assume that the following bounding/truncating procedure is used.

Bounding the K_j to avoid an arbitrarily large number of reactions. We choose a predefined value for $K_{\max}$ such that whenever $K_j = \mathcal{P}(a_j, \tau)$ exceeds $K_{\max}$ we set $K_j = K_{\max}$. In fact, any practical numerical scheme for producing Poisson random numbers always produces a bounded distribution to avoid overflow. So something akin to this bounding is automatically done by any Poisson number generator that one would use in practice.

REMARK 2.6. *We observe that starting at any nonnegative integer state x the bounding procedures described in this section do not alter the probability of events that involve zero or one reaction firing (assuming $K_{\max} > 1$). This is because by Remark* 2.1*, if the current state $x \in \mathbb{Z}_+^N$ is such that the occurrence of one reaction R_j leads to a negative state, the propensity function would be zero, and hence this will not happen.*

3. Local error analysis. In general, the evolution equations for probability densities of continuous time Markov chains on an infinite lattice form an infinite system of ODEs which is known as the chemical master equation in the case of the stochastic chemical kinetic model [6]. In this section we derive recursive integral forms for the solution of these equations and local Taylor expansions of these integrals. This gives us a basis for analyzing the local errors incurred by the tau-leaping methods.

Let the multi-index $k = (k_1, \ldots, k_l)$, where $k_j \in \{1, \ldots, M\}$, denote a sequence of reaction events R_{k_j} happening in that order, and let $|k| = l$ be the number of reaction events. Let $p(k; x, \tau)$ denote the probability that the sequence of reactions that occurred in the interval $(t, t + \tau]$ is precisely k, conditioned on being at state x at time t. Then it follows that

$$p((); x, \tau) = e^{-a_0(x)\tau},$$

where () stands for no reactions happening (the empty sequence) and $a_0(x) = \sum_{j=1}^{M} a_j(x)$. In addition $p((k_1, \ldots, k_l); x, \tau)$ may be written recursively in terms of $p((k_1, \ldots, k_{l-1}); x, \tau)$ by

$$p((k_1, \ldots, k_l); x, \tau) = \int_0^\tau p((k_1, \ldots, k_{l-1}), x; s) \times a_{k_l}(x + \nu_{j_1} + \cdots + \nu_{j_{l-1}}) e^{-a_0(x+\nu_{j_1}+\cdots+\nu_{j_l})(\tau - s)} ds. \tag{3.1}$$

It follows by induction that, for each x and k, $p(k; x, \tau)$ is an analytic function of τ for all $\tau \in \mathbb{R}$ and that $p(k; x, \tau) = O(\tau^{|k|})$ as $\tau \to 0$.

We will compute $p(k; x, \tau)$ for terms up to $|k| = 2$ (i.e., terms of up to $O(\tau^2)$) for general M and N. A numerical scheme must have a Taylor expansion in τ for the transition probabilities that matches that of the true process described above for terms up to $O(\tau)$ for first-order accuracy.

Using (3.1) once, we obtain

$$p((j);x,\tau) = \int_0^\tau e^{-a_0(x)s} a_j(x) e^{-a_0(x+\nu_j)(\tau-s)} ds, \quad j = 1,\ldots,M.$$

Applying (3.1) once again, we obtain

$$p((j_1,j_2);x,\tau) = \int_0^\tau p((j_1);x,s) a_{j_2}(x+\nu_{j_1}) e^{-a_0(x+\nu_{j_1}+\nu_{j_2})\tau} e^{a_0(x+\nu_{j_1}+\nu_{j_2})s} ds,$$

where $j, j_1, j_2 = 1,\ldots,M$.

We may Taylor expand the equations for $p((j);x,\tau)$ and $p((j_1,j_2);x,\tau)$ to obtain the following result:

$$\begin{aligned} p((j);x,\tau) &= a_j(x)\tau - \frac{1}{2}\tau^2 \sum_{j_1=1}^{M} a_j(x)\{a_{j_1}(x+\nu_j) + a_{j_1}(x)\} + O(\tau^3), \\ p((j_1,j_2);x,\tau) &= \frac{1}{2}\tau^2 a_{j_1}(x) a_{j_2}(x+\nu_{j_1}) + O(\tau^3), \end{aligned} \tag{3.2}$$

where $j, j_1, j_2 = 1,\ldots,M$.

3.1. Local Taylor series for the evolution of the moments. We consider the moments of the increment $X(t+\tau) - X(t)$. The conditional first moment $E(X(t+\tau) - X(t) \mid X(t) = x)$ is given by

$$E(X(t+\tau) - X(t) \mid X(t) = x) = \sum_{l=1}^{\infty} \sum_{k,|k|=l} p(k;x,\tau) \left(\sum_{\alpha=1}^{l} \nu_{k_\alpha} \right). \tag{3.3}$$

More generally, the conditional rth moment $E((X(t+\tau)-X(t))^r \mid X(t) = x)$ is given by

$$E((X(t+\tau) - X(t))^r \mid X(t) = x) = \sum_{l=1}^{\infty} \sum_{k,|k|=l} p(k;x,\tau) \left(\sum_{\alpha=1}^{l} \nu_{k_\alpha} \right)^r. \tag{3.4}$$

We have adopted the simple notation that for a vector $y \in \mathbb{R}^N$, y^r denotes the r-fold tensor product, which is a tensor of rank r in N dimensions. In particular the second moment may be regarded as an $N \times N$ matrix. Any of the above moments exists (is finite) only if the corresponding infinite series (3.4) converges. Since, for any fixed x and k, $p(k;x,\tau)$ is analytic for all τ, and if the infinite series converges uniformly for $\tau \in [0,\delta(x)]$ for any fixed x, then the components of $E((X(t+\tau)-X(t))^r \mid X(t) = x)$ will be analytic functions of τ for $\tau \in [0,\delta(x)]$ for some $\delta(x) > 0$. Assuming that the moments exist, we will compute them up to terms including $O(\tau^2)$. For this we need only to sum terms with $l = 1$ and $l = 2$ in (3.3) and (3.4). Thus we get

$$\begin{aligned} E(X(t+\tau) - X(t) \mid X(t) = x) = &\sum_{j=1}^{M} \nu_j p((j);x,\tau) \\ &+ \sum_{j_1=1}^{M} \sum_{j_2=1}^{M} (\nu_{j_1} + \nu_{j_2}) p((j_1,j_2);x,\tau) + O(\tau^3) \end{aligned}$$

and

$$E((X(t+\tau)-X(t))^r \mid X(t)=x) = \sum_{j=1}^{M} \nu_j^r p((j);x,\tau) + \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} (\nu_{j_1}+\nu_{j_2})^r p((j_1,j_2);x,\tau) + O(\tau^3).$$

Substituting (3.2) and simplifying we obtain

(3.5)
$$E(X(t+\tau)-X(t) \mid X(t)=x) = \tau \sum_{j=1}^{M} \nu_j a_j(x) + \frac{1}{2}\tau^2 \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} \nu_{j_1} a_{j_2}(x)\{a_{j_1}(x+\nu_{j_2}) - a_{j_1}(x)\} + O(\tau^3)$$

for the mean and

(3.6)
$$\begin{aligned} E((X(t+\tau)-X(t))^r \mid X(t)=x) = {} & \tau \sum_{j=1}^{M} \nu_j^r a_j(x) \\ & - \frac{1}{2}\tau^2 \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} \nu_{j_1}^r a_{j_1}(x) a_{j_2}(x) \\ & - \frac{1}{2}\tau^2 \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} \nu_{j_1}^r a_{j_1}(x) a_{j_2}(x+\nu_{j_1}) \\ & + \frac{1}{2}\tau^2 \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} (\nu_{j_1}+\nu_{j_2})^r a_{j_1}(x) a_{j_2}(x+\nu_{j_1}) \\ & + O(\tau^3) \end{aligned}$$

for the general rth moment.

The conditional covariance up to $O(\tau^2)$ may be computed also using (3.2). Using the fact that

$$\begin{aligned} &\mathrm{Cov}(X(t+\tau) \mid X(t)=x) = \mathrm{Cov}(X(t+\tau)-X(t) \mid X(t)=x) \\ &= E((X(t+\tau)-X(t))^2 \mid X(t)=x) - (E(X(t+\tau)-X(t) \mid X(t)=x))^2, \end{aligned}$$

we may obtain

(3.7)
$$\mathrm{Cov}(X(t+\tau) \mid X(t)=x) = \tau \sum_{j=1}^{M} \nu_j^2 a_j(x) + \frac{1}{2}\tau^2 \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} \nu_{j_1}^2 a_{j_2}(x)\left\{a_{j_1}(x+\nu_{j_2}) - a_{j_1}(x)\right\}$$

$$+\frac{1}{2}\tau^2\sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\nu_{j_1}\nu_{j_2}a_{j_1}(x)\left\{a_{j_2}(x+\nu_{j_1})-a_{j_2}(x)\right\}$$
$$+\frac{1}{2}\tau^2\sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\nu_{j_1}\nu_{j_2}a_{j_2}(x)\left\{a_{j_1}(x+\nu_{j_2})-a_{j_1}(x)\right\}+O(\tau^3),$$

where we have not shown the details of algebra.

3.2. The explicit tau method. For the explicit tau method, the rth moment of the increment conditioned on $X^{(et)}(t)=x$ is given by

$$E\left((X^{(et)}(t+\tau)-X^{(et)}(t))^r \;\middle|\; X^{(et)}(t)=x\right)=E\left(\sum_{j=1}^{M}\nu_j\mathcal{P}_j(a_j(x),\tau)\right)^r. \tag{3.8}$$

For convenience we summarize a well-known fact about the moments of a Poisson random variable with mean λ.

LEMMA 3.1. *Suppose P is a Poisson random variable with mean and variance λ. Then for any integer $r\geq 2$*

$$E(P^r)=\lambda+O(\lambda^2),\quad \lambda\to 0. \tag{3.9}$$

The term inside the expectation operator on the right-hand side of (3.8) may be expanded as sums of M^r terms, each of which is an r-fold tensor product. It must be noted that the tensor products do not commute in general. For instance $\nu_1\nu_2$, which may be interpreted as the matrix $\nu_1\nu_2^T$, is not equal to $\nu_2\nu_1$, which may be interpreted as the matrix $\nu_2\nu_1^T$. Since the $\mathcal{P}_j$ are independent Poissons of these M^r terms, those which involve two or more different values of j (two or more different reaction channels) will have expectations that are $O(\tau^2)$ or higher. First, let us consider the terms that involve only one reaction channel. There are M of these, and they are of the form

$$\nu_j^r\mathcal{P}^r(a_j(x),\tau),\quad j=1,\ldots,M.$$

Taking expectations and retaining only terms up to $O(\tau)$, we obtain

$$E(\nu_j^r\mathcal{P}^r(a_j(x),\tau))=\tau\nu_j^r a_j(x)+O(\tau^2),$$

where we have used Lemma 3.1. Thus we conclude that

$$E\left((X^{(et)}(t+\tau)-X^{(et)}(t))^r \;\middle|\; X^{(et)}(t)=x\right)=\tau\sum_{j=1}^{M}\nu_j^r a_j(x)+O(\tau^2). \tag{3.10}$$

REMARK 3.2. *Equation* (3.10) *still holds if the Poisson numbers are bounded according to the procedures in section* 2.3*, because by Remark* 2.6 *(provided x is an integer state) this bounding affects only events with probability $O(\tau^2)$.*

Comparing (3.10) with (3.6) and also considering Remark 3.2 provides the proof of the following theorem.

THEOREM 3.3 (weak consistency of the explicit tau method). *The explicit tau method is weakly consistent to first order in the following sense. Consider a given*

initial state $x \in \mathbb{Z}_+^N$ for the explicit tau. Then for each $r \geq 1$ there exist $C_r > 0$ and $\delta_r > 0$ (depending on x) such that

(3.11)
$$\|E((X^{(et)}(t+\tau)-X^{(et)}(t))^r \mid X^{(et)}(t)=x)-E((X(t+\tau)-X(t))^r \mid X(t)=x)\| < C_r\tau^2 \quad \forall \tau \in [0,\delta_r].$$

Here the norm is any suitable norm such as the induced 2-norm for the tensor. This result holds regardless of whether or not the Poisson numbers are bounded.

Since the components of the moment tensors are expectations of scalar-valued polynomials in $X(t+\tau)$ or $X^{(et)}(t+\tau)$ conditioned on $X(t) = X^{(et)}(t) = x$, Theorem 3.3 leads to the following corollary.

COROLLARY 3.4. *For any multivariate polynomial function $g : \mathbb{R}^N \to \mathbb{R}$ and initial state $x \in \mathbb{Z}_+^N$ there exist $C > 0$ and $\delta > 0$ such that*

(3.12)
$$\left|E\left(g(X^{(et)}(t+\tau)) - g(X(t+\tau)) \mid X^{(et)}(t) = X(t) = x\right)\right| < C\tau^2 \quad \forall \tau \in [0,\delta].$$

Now we will find expressions for the coefficients of the $O(\tau^2)$ error terms for the mean and the covariance matrix. For the mean we get

$$E\left(X^{(et)}(t+\tau) - X^{(et)}(t) \,\middle|\, X^{(et)}(t) = x\right) = E\left(\sum_{j=1}^{M} \nu_j \mathcal{P}_j(a_j(x),\tau)\right),$$

which gives

$$E\left(X^{(et)}(t+\tau) - X^{(et)}(t) \,\middle|\, X^{(et)}(t) = x\right) = \tau \sum_{j=1}^{M} \nu_j a_j(x). \tag{3.13}$$

Equations (3.13) and (3.5) together provide the local error in the mean for the explicit tau method.

Local error formula for the mean of explicit tau.

$$E\left(X^{(et)}(t+\tau) - X(t+\tau) \,\middle|\, X^{(et)} = X(t) = x\right) = -\frac{1}{2}\tau^2 \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} \nu_{j_1} a_{j_2}(x) \left\{a_{j_1}(x+\nu_{j_2}) - a_{j_1}(x)\right\} + O(\tau^3). \tag{3.14}$$

Since the Poisson numbers $\mathcal{P}_j(a_j(x),\tau)$ are independent, we get

$$\mathrm{Cov}(X^{(et)}(t+\tau) \mid X^{(et)}(t) = x) = \tau \sum_{j=1}^{M} \nu_j^2 a_j(x), \tag{3.15}$$

where the tensors ν_j^2 can be represented by the matrices $\nu_j \nu_j^T$. Equations (3.15) and (3.7) together provide us with the local error in the covariance for explicit tau.

Local error formula for the covariance of explicit tau.

(3.16)
$$\begin{aligned}\mathrm{Cov}(X^{(et)}(t+\tau) \mid X^{(et)}(t)=x) - \mathrm{Cov}(X(t+\tau) \mid X(t)=x)\\ = -\frac{1}{2}\tau^2 \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} \nu_{j_1}^2 a_{j_2}(x)\left\{a_{j_1}(x+\nu_{j_2}) - a_{j_1}(x)\right\}\\ - \frac{1}{2}\tau^2 \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} \nu_{j_1}\nu_{j_2} a_{j_1}(x)\left\{a_{j_2}(x+\nu_{j_1}) - a_{j_2}(x)\right\}\\ - \frac{1}{2}\tau^2 \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} \nu_{j_1}\nu_{j_2} a_{j_2}(x)\left\{a_{j_1}(x+\nu_{j_2}) - a_{j_1}(x)\right\} + O(\tau^3).\end{aligned}$$

REMARK 3.5. *Note that the $O(\tau^2)$ local error terms derived for the mean and covariance hold regardless of whether or not the Poisson numbers are truncated at $K_{\max}$, since this affects only events with probability $O(\tau^{K_{\max}})$. However, these formulae do not hold if x is one reaction away from a state with at least one zero component. This is because the probabilities of two reaction events (which is $O(\tau^2)$) will be altered by the bounding procedure of section* 2.3*, which avoids negative states.*

3.3. The implicit tau method. Given that $X^{(it)}(t) = x$, for a step size $\tau > 0$ the implicit tau method involves finding $X^{(it)}(t+\tau)$, which is the unique solution of (2.3). Writing $X^{(it)}(t+\tau) = X'$ and comparing (2.3) with (2.2) shows that X' may be written as the unique solution of

$$X' = X'_e + \tau \sum_{j=1}^{M} \nu_j \left\{a_j(X') - a_j(x)\right\}, \tag{3.17}$$

where

$$X'_e = x + \sum_{j=1}^{M} \nu_j K_j^{(et)}(x,\tau).$$

Thus given an initial state $x \in \mathbb{R}_+^N$ we can think of the implicit tau method as involving first the computation of an intermediate state X'_e according to explicit tau and then solving for X' in the implicit equation (3.17). Thus X' is a deterministic function of X'_e.

We rewrite (3.17) for convenience as

$$F(X', X'_e, x, \tau) = 0, \tag{3.18}$$

where the C^∞-smooth function $F(X', X'_e, x, \tau) : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^N$ is given by

$$F(X', X'_e, x, \tau) = X' - X'_e - \tau \sum_{j=1}^{M} \nu_j \left\{a_j(X') - a_j(x)\right\}. \tag{3.19}$$

In order to ensure that X' is well defined for τ sufficiently small, we use the implicit function theorem. We note that

$$F(X'_e, X'_e, x, 0) = 0.$$

The Jacobian $\frac{\partial F}{\partial X'}$ is given by

$$\frac{\partial F}{\partial X'} = I - \tau \sum_{j=1}^{M} \nu_j \frac{\partial a_j}{\partial x}(X'), \tag{3.20}$$

where I is the $N \times N$ identity matrix. Thus at $\tau = 0$ the Jacobian is the identity matrix, and hence is full rank. Therefore, by the implicit function theorem there exists $\delta > 0$, a region $D_e \subset \mathbb{R}^N$, and a C^∞-smooth function $G : D_e \times [0, \delta] \to \mathbb{R}^N$ such that X' is given by

$$X' = G(X'_e, \tau) \quad \forall (X'_e, \tau) \in D_e \times [0, \delta]. \tag{3.21}$$

Note that $G(X'_e, 0) = X'_e$. (The function G also depends on x, but since we have fixed the x we do not show the dependence on it explicitly.) A careful look at the proof of the implicit function theorem (see [14] for instance) tells us that since the Jacobian is independent of X'_e, we may choose D_e to be arbitrarily large but bounded. However, the size of δ may be smaller the larger D_e gets. Since by the bounding procedure of Poisson random numbers X'_e takes bounded values, we can choose D_e to be a bounded region that contains all possible values of X'_e. For this choice, we are still assured of the existence of $\delta > 0$ such that (3.21) holds.

REMARK 3.6. *While the theoretical consideration mentioned above indicates that the larger $K_{\max}$ is, the smaller δ may be, in practice we have never encountered problems in finding a solution to the implicit equation using the Newton iteration when using practical stepsizes. It must be noted that a similar theoretical concern exists in the application of the implicit Euler method to SDEs driven by Gaussian white noise.*

We will compute the values of $\frac{\partial G}{\partial \tau}$ and $\frac{\partial^2 G}{\partial \tau^2}$ by differentiating (3.18). Differentiating (3.18) with respect to τ, we obtain

$$\frac{\partial F}{\partial X'} \frac{\partial G}{\partial \tau} + \frac{\partial F}{\partial \tau} = 0,$$

which upon using the expression (3.19) for F yields

$$\left(I - \tau \sum_{j=1}^{M} \nu_j \frac{\partial a_j}{\partial x}(X') \right) \frac{\partial G}{\partial \tau} - \sum_{j=1}^{M} \nu_j \{a_j(X') - a_j(x)\} = 0. \tag{3.22}$$

Substituting $\tau = 0$ in (3.22) and also using (3.21) and the fact that $G(X'_e, 0) = X'_e$ yields

$$\frac{\partial G}{\partial \tau}(\tau = 0) = \tau \sum_{j=1}^{M} \nu_j \{a_j(X'_e) - a_j(x)\}. \tag{3.23}$$

Differentiating (3.22) with respect to τ, we obtain

$$\frac{\partial^2 F}{\partial X'^2} \left(\frac{\partial G}{\partial \tau} \right)^2 + \frac{\partial F}{\partial X'} \frac{\partial^2 G}{\partial \tau^2} + \frac{\partial^2 F}{\partial \tau \partial X'} \frac{\partial G}{\partial \tau} + \frac{\partial^2 F}{\partial \tau^2} = 0.$$

Noting that $\frac{\partial^2 F}{\partial \tau^2} = 0$ and that when $\tau = 0$, $\frac{\partial^2 F}{\partial X'^2} = 0$ and $\frac{\partial F}{\partial X'} = I$, we obtain

$$\frac{\partial^2 G}{\partial \tau^2}(\tau = 0) = \sum_{j_1=1}^{M} \sum_{j_2=1}^{M} \nu_{j_1} \frac{\partial a_{j_1}}{\partial x}(X'_e) \nu_{j_2} \{a_{j_2}(X'_e) - a_{j_2}(x)\}, \tag{3.24}$$

where we have used the fact that $X' = X'_e$ at $\tau = 0$ and (3.23). Since G is jointly C^∞-smooth in X'_e and τ, we obtain using Taylor's formula and (3.23) and (3.24)

$$\begin{aligned} X' &= X'_e + \tau \sum_{j=1}^{M} \nu_j \{a_j(X'_e) - a_j(x)\} \\ &\quad + \frac{\tau^2}{2} \sum_{j_1=1}^{M} \sum_{j_2=1}^{M} \nu_{j_1} \frac{\partial a_{j_1}}{\partial x}(X'_e) \nu_{j_2} \{a_{j_2}(X'_e) - a_{j_2}(x)\} + O(\tau^3). \end{aligned} \tag{3.25}$$

To show the consistency of the implicit tau method we need certain lemmas which are stated and proven in Appendix A.

The following lemma asserts that for sufficiently small stepsizes the rounded version of the implicit tau coincides with explicit tau.

LEMMA 3.7. *Assuming the bounding procedures of section* 2.3*, for any given initial state* $x \in \mathbb{Z}_+^N$ *there exists* $\delta > 0$ *such that* $X^{(itr)}(t+\tau) = X^{(et)}(t+\tau)$ *with probability* 1*, conditioned on* $X^{(itr)}(t) = X^{(et)}(t) = x$*, for all* $\tau \in [0, \delta]$.

Proof. It follows from (2.5) and Lemma A.4 that for $\tau > 0$ small enough, $K_j^{(itr)} = K_j^{(et)}$ (with probability 1). ∎

THEOREM 3.8 (consistency of rounded implicit tau). *Assuming the bounding procedures of section* 2.3*, for any multivariate polynomial function* $g : \mathbb{R}^N \to \mathbb{R}$ *and initial state* $x \in \mathbb{Z}_+^N$ *there exist* $C > 0$ *and* $\delta > 0$ *such that*

$$\left| E\left(g(X^{(itr)}(t+\tau)) - g(X(t+\tau)) \mid X^{(itr)}(t) = X(t) = x \right) \right| < C\tau^2 \quad \forall \tau \in [0, \delta]. \tag{3.26}$$

Proof. The proof follows from Lemma 3.7 and Corollary 3.4. ∎

THEOREM 3.9 (consistency of unrounded implicit tau). *Assuming the bounding procedures of section* 2.3*, for any multivariate polynomial function* $g : \mathbb{R}^N \to \mathbb{R}$ *and initial state* $x \in \mathbb{Z}_+^N$ *there exist* $C > 0$ *and* $\delta > 0$ *such that*

$$\left| E\left(g(X^{(it)}(t+\tau)) - g(X(t+\tau)) \mid X^{(it)}(t) = X(t) = x \right) \right| < C\tau^2 \quad \forall \tau \in [0, \delta]. \tag{3.27}$$

Proof. The proof follows from Lemma A.3 and Corollary 3.4. ∎

Moreover, we can derive the formula for the local error of the implicit tau method.

Local error formulae for implicit tau. We note that the term with double summation in (3.25) is $O(\tau^3)$. This follows from Lemma A.1. Hence the following equation may be used to relate local error in implicit tau to that of explicit tau:

$$X^{(it)}(t+\tau) = X^{(et)}(t+\tau) + \tau \sum_{j=1}^{M} \nu_j \{a_j(X^{(et)}(t+\tau)) - a_j(x)\} + O(\tau^3). \tag{3.28}$$

Note that this equation assumes $X^{(it)}(t) = X^{(et)}(t) = x$. In order to compute the rth moment of $X^{(it)}(t+\tau)$, one has to raise this equation to the power r and take expectations of both sides. If the a_j are polynomials, the expectation on the right-hand side will involve taking expectations of various powers of Poisson random variables, the formulae for which are well known.

4. Stability and convergence for systems with linear propensity functions. In this section we will investigate the stability properties of the explicit and implicit tau methods for systems with linear propensity functions and prove that they converge with first-order accuracy. We focus on chemical reaction systems where N and M are arbitrary, but the propensity functions $a_j(x)$ take the form

$$a_j(x) = c_j^T x, \quad x \in \mathbb{R}^N; \quad j = 1, \ldots, M, \tag{4.1}$$

where $c_j \in \mathbb{R}^N$ are constant vectors. An important form of stability relevant for convergence is that of 0-stability [2]. The following definition is a generalization of this notion to stochastic systems that is adequate for our purpose of establishing the convergence of all moments.

DEFINITION 4.1 (0-stability). *Denote by $\hat{X}$ the discrete time numerical approximation of a stochastic process. We shall say that the numerical method is* 0*-stable up to r moments if, for a fixed time interval $[0, T]$, there exist $\delta > 0$ and $K_{lj} > 0$ for $l = 1, \ldots, r$ and $j = 1, \ldots, l$, such that*

$$\|E((\hat{X}_1(t))^l) - E((\hat{X}_2(t))^l)\| \leq \sum_{j=1}^{l} K_{lj} \|E((\hat{X}_1(0))^j) - E((\hat{X}_2(0))^j)\|$$

for all $l = 1, \ldots, r$, and $t = \sum_{i=1}^{n} \tau_i \leq T$, where n is arbitrary, $0 < \tau_i \leq \delta$ for $i = 1, \ldots, n$, and $\hat{X}_1, \hat{X}_2$ correspond, respectively, to the numerical solutions obtained from any pair of arbitrary initial conditions $\hat{X}_1(0), \hat{X}_2(0)$, which are random variables assumed to have finite moments.

For systems with linear propensities of the form (4.1), the explicit tau method is given by

$$X^{(et)}(t + \tau) = X^{(et)}(t) + \sum_{j=1}^{M} \nu_j \mathcal{P}_j(c_j^T X^{(et)}(t), \tau). \tag{4.2}$$

Taking expectations conditioned on $X^{(et)}(t)$, we obtain

$$E(X^{(et)}(t + \tau) \mid X^{(et)}(t)) = X^{(et)}(t) + \tau \sum_{j=1}^{M} \nu_j c_j^T X^{(et)}(t),$$

which may be written as

$$E(X^{(et)}(t + \tau) \mid X^{(et)}(t)) = X^{(et)}(t) + \tau A X^{(et)}(t), \tag{4.3}$$

where the $N \times N$ matrix A is given by

$$A = \sum_{j=1}^{M} \nu_j c_j^T. \tag{4.4}$$

REMARK 4.2. *We remark for the unfamiliar reader that if X and Y are random variables, then $E(Y \mid X)$ is the conditional expectation of Y conditioned on X and is a random variable which is a deterministic function of X that takes the value $E(Y \mid X = x)$ when X takes the value $X = x$. Then it follows that $E(E(Y \mid X)) = E(Y)$. We have used these concepts in the above derivations. For a basic treatment of this notion see* [16] *for instance.*

Taking the expectation of (4.3), we obtain

$$E(X^{(et)}(t+\tau)) = (I+\tau A)E(X^{(et)}(t)). \tag{4.5}$$

Similarly we obtain the following for implicit tau:

$$\begin{aligned} X^{(it)}(t+\tau) &= X^{(it)}(t) + \tau \sum_{j=1}^{M} \nu_j c_j^T X^{(it)}(t+\tau) \\ &\quad + \sum_{j=1}^{M} \nu_j \mathcal{P}_j(c_j^T X^{(it)}(t), \tau) - \tau \sum_{j=1}^{M} \nu_j c_j^T X^{(it)}(t), \end{aligned} \tag{4.6}$$

which leads to

$$E(X^{(it)}(t+\tau)) = (I-\tau A)^{-1} E(X^{(it)}(t)), \tag{4.7}$$

where A is the same matrix given by (4.4).

4.1. Asymptotic stability of the mean. For a constant stepsize τ,

$$E(X^{(et)}(t+n\tau)) = (I+\tau A)^n E(X^{(et)}(t)) \tag{4.8}$$

and

$$E(X^{(it)}(t+n\tau)) = (I-\tau A)^{-n} E(X^{(it)}(t)). \tag{4.9}$$

Thus the mean of the explicit tau simulation evolves in a geometric progression just the same as the explicit Euler simulation of the system of equations $\dot{x} = Ax$, while the mean of implicit tau evolves just the same as the implicit Euler simulation. Thus the mean value of the explicit tau simulation is asymptotically stable if

$$|1+\lambda_i(A)\tau| < 1, \quad i = 1, \ldots, N, \tag{4.10}$$

while the mean value of the implicit tau simulation is asymptotically stable if

$$1/|1-\lambda_i(A)\tau| < 1, \quad i = 1, \ldots, N, \tag{4.11}$$

where $\lambda_i(A)$ is the ith largest eigenvalue of A.

REMARK 4.3. *Note that implicit tau is unconditionally stable in the mean for stable systems (i.e., systems whose A matrix has all eigenvalues in the left complex half plane). This clearly confirms its advantage over explicit tau for stiff systems, as observed in* [15]. *However, a detailed study of the asymptotic behavior of the covariance and higher-order moments is necessary in order to assess the suitability of implicit tau for stiff systems* [3].

4.2. 0-stability of the mean.

LEMMA 4.4. *Both the explicit and the implicit tau methods are* 0*-stable in the mean.*

Proof. These follow from standard procedures known in numerical analysis of ODEs. They essentially follow from bounding $\|I+\tau A\| < e^{\tau\|A\|}$ for explicit tau and the bound $\|I-\tau A\| < e^{K\tau}$ which holds for sufficiently small τ. Here K is any number greater than $\|A\|$. ∎

4.3. Weak convergence. We will derive equations for the evolution of the rth moment $E((X^{(et)}(t+n\tau))^r)$. From (4.2) it follows that

$$E((X^{(et)}(t+\tau))^r \mid X^{(et)}(t)) = (X^{(et)}(t))^r + \sum_{k=1}^{r} \frac{r!}{k!(r-k)!}(X^{(et)}(t))^{(r-k)} \sum_{j=1}^{M} \nu_j^k \{c_j^T X^{(et)}(t)\tau + O(\tau^2)\},$$

where we have used Lemma 3.1. We rewrite this equation as

(4.12)
$$E((X^{(et)}(t+\tau))^r \mid X^{(et)}(t)) = (X^{(et)}(t))^r + \tau \sum_{k=1}^{r}\sum_{j=1}^{M} \frac{r!}{k!(r-k)!}(X^{(et)}(t))^{(r-k)}\, \nu_j^k\, (c^T X^{(et)}(t)) + O(\tau^2).$$

The following lemma is useful.

LEMMA 4.5. *If $z, \nu, c \in \mathbb{R}^N$, then the components of the rank r tensor $z^{(r-k)}\nu^k(c^T z)$ are linear combinations of the components of the rank $(r-k+1)$ tensor $z^{(r-k+1)}$.*

Proof. Each component of the tensor $z^{(r-k)}\nu^k(c^T z)$ (which has N^r components) is a linear combination of multivariate monomials in z with degree $r-k+1$. The tensor $z^{(r-k+1)}$ consists of all the multivariate monomials in z of degree $r-k+1$. ∎

Combining Lemma 4.5 with (4.12) enables us to write

$$E((X^{(et)}(t+\tau))^r \mid X^{(et)}(t)) = X^{(et)}(t)^r + \tau B_{rr}(X^{(et)}(t))^r + \tau \sum_{l=1}^{r-1} B_{lr}(X^{(et)}(t))^l + O(\tau^2),$$

where the linear operators B_{lr} map symmetric tensors of rank l into symmetric tensors of rank r. Taking expectations of both sides, we obtain

(4.13)
$$E((X^{(et)}(t+\tau))^r) = (I+\tau B_{rr})E((X^{(et)}(t))^r) + \tau \sum_{l=1}^{r-1} B_{lr}E((X^{(et)}(t))^l) + O(\tau^2).$$

Similarly we shall derive equations for the evolution of the rth moment $E((X^{(it)}(t+n\tau))^r)$. From (4.6) it follows that

$$X^{(it)}(t+\tau) = X^{(it)}(t) + \sum_{j=1}^{M} (I-\tau A)^{-1}\nu_j \mathcal{P}(c_j^T X^{(it)}(t), \tau).$$

Hence it follows that

$$E((X^{(it)}(t+\tau))^r \mid X^{(it)}(t)) = (X^{(it)}(t))^r + \tau \sum_{k=1}^{r} \frac{r!}{k!(r-k)!}(X^{(it)}(t))^{(r-k)} \sum_{j=1}^{M} \left((I-\tau A)^{-1}\nu_j\right)^k \{c_j^T X^{(it)}(t)\tau + O(\tau^2)\},$$

where we have used Lemma 3.1. Taylor expanding $(I-\tau A)^{-1}$ and rearranging terms, we rewrite this equation as

$$E((X^{(it)}(t+\tau))^r \mid X^{(it)}(t)) = (X^{(it)}(t))^r + \tau \sum_{k=1}^{r} \sum_{j=1}^{M} \frac{r!}{k!(r-k)!} (X^{(it)}(t))^{(r-k)} \nu_j^k \, (c^T X^{(it)}(t)) + O(\tau^2). \tag{4.14}$$

This equation is similar to (4.12) up to $O(\tau)$ terms. Hence, as before, combining Lemma 4.5 with (4.14) enables us to write

$$E((X^{(it)}(t+\tau))^r) = E((X^{(it)}(t))^r) + \tau B_{rr}^{(it)} E((X^{(it)}(t))^r) + \tau \sum_{l=1}^{r-1} B_{lr} E((X^{(it)}(t))^l) + O(\tau^2), \tag{4.15}$$

where the linear operators B_{lr} are defined as in the case of explicit tau.

LEMMA 4.6. *Both the explicit tau and the implicit tau methods are* 0*-stable up to the* r*th moment for all integers* $r > 0$.

Proof. This lemma is true for $r = 1$ by Lemma 4.4. The proof follows by induction. Use (4.13) or (4.15) and apply standard bounding techniques (see Lemma 7.2.2.2 of [17] for instance). ∎

REMARK 4.7. *The bounding procedures of section* 2.3 *do not affect the* 0*-stability of moments, since bounding affects only events of* $O(\tau)$. *For explicit tau and rounded implicit tau, because of integer states it will affect only* $O(\tau^2)$ *terms according to Remark* 2.6. *For the (unrounded) implicit tau, since noninteger states occur the* $O(\tau)$ *term may be modified. Thus, in any event, bounding amounts to modifying the matrix* A *as well as the tensors* B_{lr} *in* (4.13) *and* (4.15)*, but their norms will still be finite since they are finite dimensional.*

THEOREM 4.8 (weak convergence of explicit tau). *Consider a system with linear propensity functions. For any given initial state* $X(t_0) = x \in \mathbb{Z}_+^N$, *let the explicit tau method be applied with the initial state* $X^{(et)}(t_0) = x$ *and with time steps* $\tau_1, \tau_2, \ldots, \tau_n$, *where* $\sum_{j=1}^{n} \tau_j = T$ *and* $\tau = \max\{\tau_j\}$. *Then for each positive integer* r *there exist constants* $C > 0$ *and* $\delta > 0$ *(which depend on* x *and* r*) independent of* n *and* τ *such that*

$$\|E((X^{(et)}(t_0+T))^r) - E((X(t_0+T))^r)\| < C\tau \quad \forall \tau \in (0, \delta).$$

Proof. The proof follows from a modification of standard techniques in numerical analysis that show that consistency and 0-stability imply convergence. See Appendix B. ∎

THEOREM 4.9 (weak convergence of implicit tau). *Consider a system with linear propensity functions. For any given initial state* $X(t_0) = x \in \mathbb{Z}_+^N$, *let the implicit tau method be applied with the initial state* $X^{(it)}(t_0) = x$ *and with time steps* $\tau_1, \tau_2, \ldots, \tau_n$, *where* $\sum_{j=1}^{n} \tau_j = T$ *and* $\tau = \max\{\tau_j\}$. *Assume the bounding procedures of section* 2.3 *are used. Then for each positive integer* r *there exist constants* $C > 0$ *and* $\delta > 0$ *(which depend on* x *and* r*) independent of* n *and* τ *such that*

$$\|E(X^{(it)}(t_0+T)^r) - E(X(t_0+T)^r)\| < C\tau \quad \forall \tau \in (0, \delta).$$

Proof. See Appendix B. ∎

5. Examples. In this section we consider several examples and illustrate via theory and numerical experiments that the local error formulae and the linear convergence behavior may be observed for a range of stepsize-τ values in which the expected number of reactions occurring is considerably larger than 1.

5.1. Example 1: Isomerization reaction. Consider the isomerization reaction

$$X \longrightarrow \emptyset, \tag{5.1}$$

with propensity function $a(x) = cx$, $X(0) = X_0$ in the time interval $[0, T]$.

Linear convergence of mean and variance. For this problem we can derive explicit analytical solutions $X(t)$ for the mean and variance of the exact process (according to the SSA), as well as for the explicit and implicit tau method (unrounded) with constant stepsize τ. It may be shown that $X(T)$ is a binomial random variable with parameters $p = e^{-cT}$ and $N = X_0$ [9]. Hence the expected value and variance are given by

$$E(X(T)) = X_0 e^{-cT} \tag{5.2}$$

and

$$\mathrm{Var}(X(T)) = X_0 e^{-cT} \left(1 - e^{-cT}\right). \tag{5.3}$$

Note that (5.2) is the same as the solution of the deterministic ODE model $\dot{X} = -cX$ with $X(0) = X_0$. This is due to the linearity of the propensity function. In general the true mean of the model does not evolve the same way as the ODE model.

It follows from the results of section 4 that

$$E(X^{(et)}(t + \tau)) = (1 - c\tau)E(X^{(et)}(t))$$

and

$$\mathrm{Var}(X^{(et)}(t + \tau)) = (1 - c\tau)^2 \,\mathrm{Var}(X^{(et)}(t)) + c\tau E(X^{(et)}(t)).$$

From these recursive relationships we obtain $E(X^{(et)}(T))$ and $\mathrm{Var}(X^{(et)}(T))$:

$$E(X^{(et)}(T)) = (1 - c\tau)^n X_0 \tag{5.4}$$

and

$$\mathrm{Var}(X^{(et)}(T)) = (1 - \tau c)^{n-1} \left(1 - (1 - \tau c)^n\right) X_0. \tag{5.5}$$

We omit the details of the derivation. We have assumed $X^{(et)}(0) = X_0$ to be deterministic. Thus $\mathrm{Var}(X_0) = 0$. Note that the evolution of the mean is the same as for the explicit Euler applied to the ODE model $\dot{X} = -cX$ with $X(0) = X_0$ (again due to linearity of the propensity function).

It follows from the results of section 4 that

$$E(X^{(it)}(t + \tau)) = \frac{E(X^{(et)}(t))}{(1 + \tau c)}$$

and

$$\mathrm{Var}(X^{(it)}(t + \tau)) = \frac{\mathrm{Var}(X^{(it)}(t))}{(1 + \tau c)^2} + \frac{c\tau E(X^{(it)}(t))}{(1 + \tau c)^2}.$$

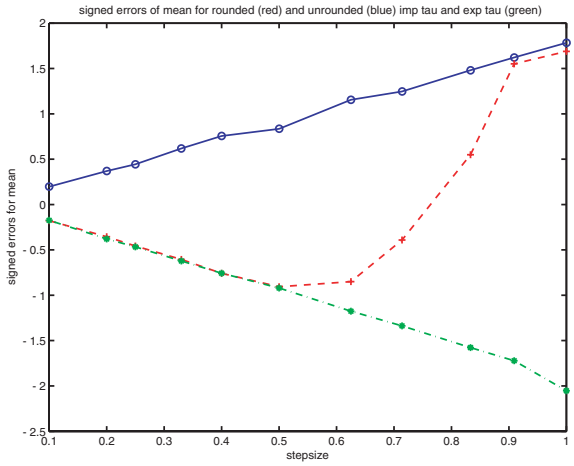

FIG. 5.1. *The errors for the mean vs. the stepsize for Example* 1 *estimated with* 10^5 *samples. The solid line plot with "o" is for implicit tau, while the dashed line with "+" is for rounded implicit tau, and the dotted-dashed line with "*" is for explicit tau.*

From these recursive relationships we obtain $E(X^{(it)}(T))$ and $\mathrm{Var}(X^{(it)}(T))$:

$$E(X^{(it)}(T)) = \frac{X_0}{(1+\tau c)^n} \tag{5.6}$$

and

$$\mathrm{Var}(X^{(it)}(T)) = \frac{(1-\frac{1}{(1+\tau c)^n})}{(1+\tau c)^n} X_0. \tag{5.7}$$

We have assumed $X^{(it)}(0) = X_0$ to be deterministic. Thus $\mathrm{Var}(X_0) = 0$.

For this example we show both numerical estimates as well as theoretical computations. We chose the specific values, $X(0) = 100$ and $c = 0.1$, with a time interval $[0, 10]$.

First, we refer to Figure 5.1, where estimated errors in the mean of explicit tau, implicit tau, and the rounded implicit tau are shown. The linear convergence of explicit and (unrounded) implicit tau is clear from the graphs. In this plot, for the largest value $\tau = 1$, starting at $X(0) = 100$ we expect $100 \times 0.1 \times 1 = 10$ reactions to occur on average per step. The linear convergence behavior is apparent in a region of τ values in which up to 10 expected number of reactions occur. The deviations from the linear curve are due to the effects of the finite sample size of 10^5.

The error of the rounded implicit tau is worth commenting on. For larger tau values it is similar to the unrounded implicit tau. This is intuitively expected, since for large τ values the expected number of reactions according to (unrounded) implicit tau is large, and rounding a large number produces a negligible change. However, for small τ values, as pointed out in Lemma 3.7, the rounded implicit tau should produce the same results as explicit tau. This behavior is observed in Figure 5.1.

Figure 5.2 shows the errors in the variance for the explicit and (unrounded) implicit tau methods as computed from (5.3), (5.5), and (5.7). Note that the curves are almost linear up to $\tau = 1$, which again corresponds to 10 reactions firing on average when starting at $X_0 = 100$. Finally, Figure 5.3 shows the error in variance of implicit tau for a larger range of τ values.

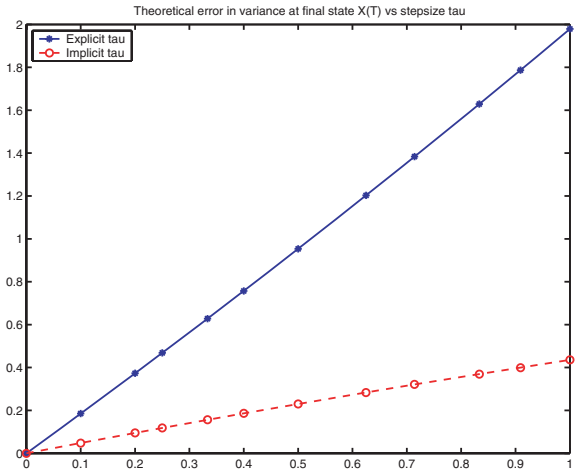

FIG. 5.2. *The errors for the variance vs. the stepsize for Example* 1 *computed using theory. The solid line plot with "*" is for explicit tau, while the dashed line with "o" is for (unrounded) implicit tau.*

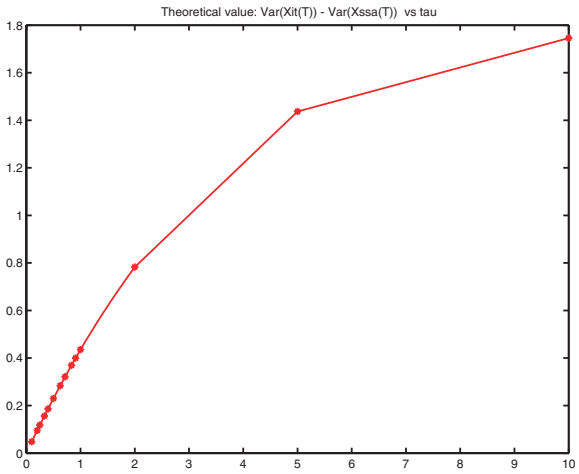

FIG. 5.3. *The errors for the variance of implicit tau vs. the stepsize for Example* 1*, computed using theory.*

Second-order local error formulae. In this example it is also possible to compare the actual local error with its approximation by the $O(\tau^2)$ term, since the former can be derived analytically. This exercise is rather uninteresting in the case of the mean since the behavior is similar to what one expects from explicit and implicit Euler local error formulae for the ODE $\dot{X} = -cX$. Instead we show a plot of the local error in the variance and its approximation by the $O(\tau^2)$ term. Figures 5.4 and 5.5 show these plots for $x = 100$. From these plots it is apparent that the second-order formula is valid up to $\tau = 1$ for explicit tau and up to $\tau = 0.5$ for implicit tau. In fact, the graphs in Figures 5.4 and 5.5 scale proportional to x^2 (in the y-axis) retaining the same shape. Thus the region of validity of the local error formula holds up to $\tau = 0.5$ for implicit tau, regardless of x. The expected number of reactions leapt over will be

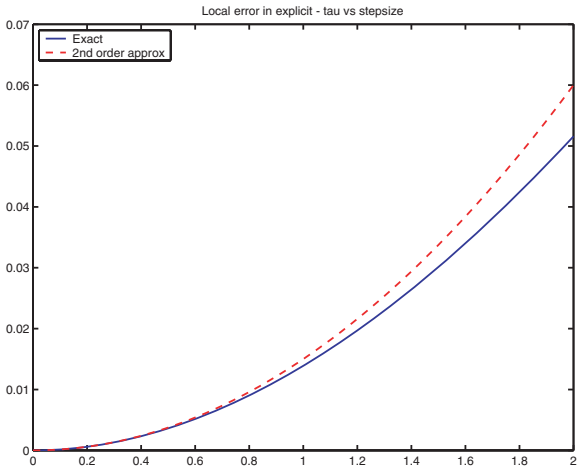

FIG. 5.4. *Local errors for the variance of explicit tau vs. the stepsize for Example* 1*, computed using theory. The solid line is the exact local error. The dashed line is the* $O(\tau^2)$ *approximation.*

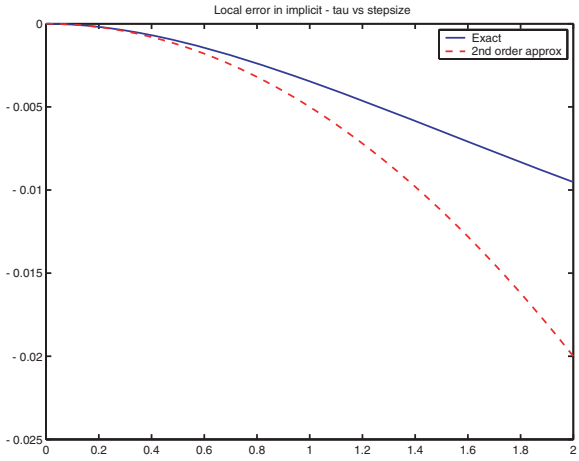

FIG. 5.5. *Local error for the variance of implicit tau vs. the stepsize for Example* 1*, computed using theory. The solid line is the exact local error. The dashed line is the* $O(\tau^2)$ *approximation.*

given by $x \times 0.1 \times 0.5 = x/20$. For example, if the current state is $x = 1000$ molecules, we may leap over 50 reactions and still use the same local error formula. If $x = 10^5$, we can leap over 5000 reactions on average. In this case the system will be behaving nearly deterministically, and implicit tau will practically be implicit Euler.

5.2. Example 2: Decaying-dimerizing reaction set. This reaction system was studied in [9] and [15]. It consists of $N = 3$ species undergoing $M = 4$ different types of chemical reactions:

$$\begin{aligned} S_1 &\xrightarrow{c_1} 0, \\ S_1 + S_1 &\xrightarrow{c_2} S_2, \\ S_2 &\xrightarrow{c_3} S_1 + S_1, \end{aligned} \tag{5.8}$$

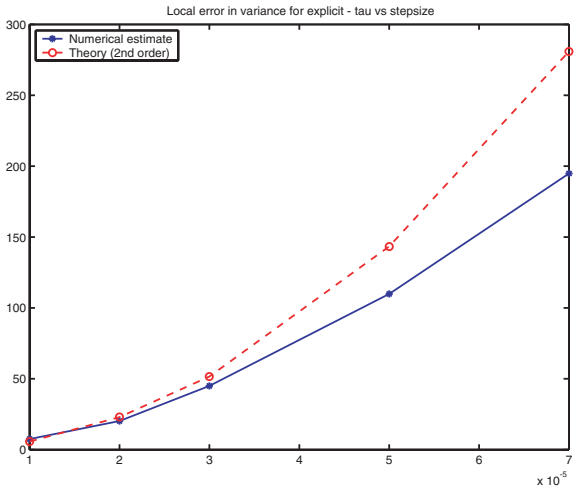

FIG. 5.6. *Local error for the variance of explicit tau vs. the stepsize for Example* 2*, computed using theory and numerical estimation. The solid line is the numerical estimation using* 10^4 *samples. The dashed line is the* $O(\tau^2)$ *theoretical approximation.*

$$S_2 \xrightarrow{c_4} S_3.$$

We chose values for the parameters

$$c_1 = 1, \quad c_2 = 10, \quad c_3 = 1000, \quad c_4 = 0.1.$$

Hence the propensity functions are given by

$$a_1 = x_1, \quad a_2 = 5x_1(x_1 - 1), \quad a_3 = 1000x_2, \quad a_4 = 0.1x_2,$$

x_i being the number of species S_i. We chose the initial conditions $X_1(0) = 400$, $X_2(0) = 798$, and $X_3(0) = 0$ and the final time $T = 0.2$ (the same as in [15]). We computed the second-order local error formula for the variance of $X_1^{(et)}(\tau)$ after a time step of τ, starting at $X_1(0) = 400$. Using (3.16) we find that the local error is

$$\mathrm{Var}(X_1^{(et)}(\tau)) - \mathrm{Var}(X_1(\tau))) = 5.7340 \times 10^{10}\tau^2.$$

We estimated the local error in the variance by running 10^4 samples with both the SSA and explicit tau. Figure 5.6 shows that the second-order local error formula is reasonable up to a stepsize of $\tau = 3 \times 10^{-5}$. At the initial state $X_1(0) = 400$, $X_2(0) = 798$, and $X_3(0) = 0$, the total propensity is 1.60×10^6. Hence the number of reactions leapt over for $\tau = 3 \times 10^{-5}$ is about 48 on average.

To compute the local error in implicit tau we use the formula (3.28). Substituting all the appropriate values into this equation gives

$$X_1^{(it)}(\tau) = X_1^{(et)}(\tau) + 9\tau X_1^{(et)}(\tau) + 400\tau - 10\tau(X_1^{(et)}(\tau))^2 + 2000\tau X_2^{(et)}(\tau) + O(\tau^3).$$

Taking expectations directly on both sides of this formula, as well as squaring the formula and then taking expectations, gives us the values of $E(X_1^{(it)}(\tau))$ and $E((X_1^{(it)}(\tau))^2)$ as functions of τ. To keep the presentation short we do not show the

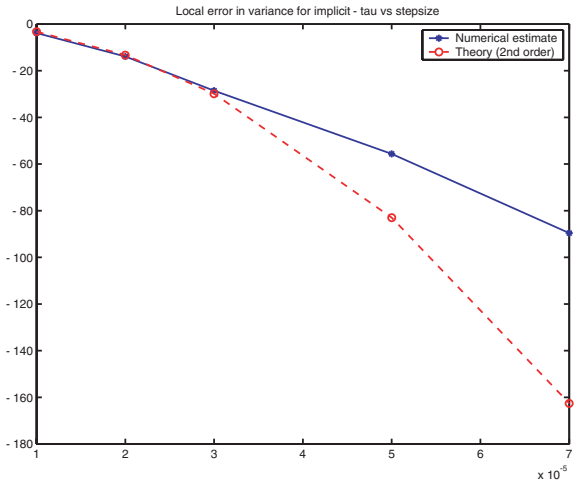

FIG. 5.7. *Local error for the variance of implicit tau vs. the stepsize for Example* 2*, computed using theory and numerical estimation. The solid line is the numerical estimation using* 10^4 *samples. The dashed line is the* $O(\tau^2)$ *theoretical approximation.*

details of the computations here. We obtained with the aid of MAPLE the following formula for the $O(\tau^2)$ error term for the variance of implicit tau (for state X_1):

$$\mathrm{Var}(X_1^{(it)}(\tau)) - \mathrm{Var}(X_1(\tau)) = -3.3174 \times 10^{10}\tau^2.$$

As in the case of explicit tau we estimated the local errors numerically, using a sample size of 10^4. Figure 5.7 shows that the second-order local error formula is a good approximation up to a stepsize of $\tau = 3 \times 10^{-5}$. Hence the number of reactions leapt over for $\tau = 3 \times 10^{-5}$ is about 48 on average.

5.3. Example 3: Schlögl reaction. The Schlögl reaction [7] is a famous example of a reaction with a bimodal stationary distribution. Note that a probability density function is bimodal if it has two local maxima:

$$\begin{aligned} B_1 + 2X &\longleftrightarrow 3X, \\ B_2 &\longleftrightarrow X. \end{aligned} \tag{5.9}$$

B_1 and B_2 denote buffered species whose molecular populations N_1 and N_2 are assumed to remain essentially constant over the time interval of interest. Let

$$X(t) = \text{number of } X \text{ molecules in the system at time } t.$$

The propensity functions are

$$\begin{aligned} a_1(x) &= \frac{c_1}{2} N_1 x(x-1), \\ a_2(x) &= \frac{c_2}{6} x(x-1)(x-2), \\ a_3(x) &= c_3 N_2, \\ a_4(x) &= c_4 x. \end{aligned} \tag{5.10}$$

The state change vectors are $\nu_1 = \nu_3 = 1$, $\nu_2 = \nu_4 = -1$. Note that this reaction does not satisfy Assumption 2.2.

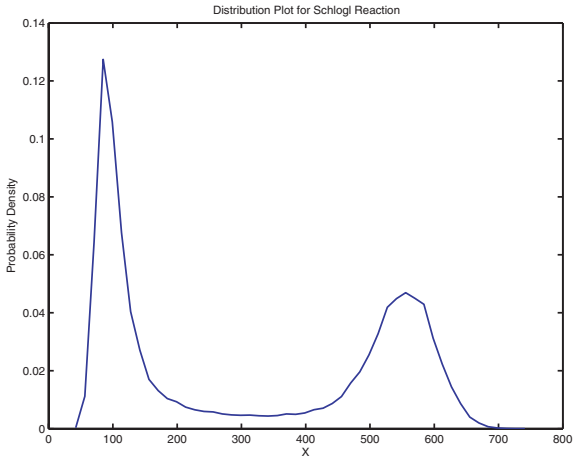

FIG. 5.8. *Histogram distribution of the Schlögl reaction at* $T = 4$ *with* 10^6 *SSA samples.*

For some parameter values, the probability distribution of this reaction is bimodal. The special parameter set that we used in our simulation is

$$c_1 = 3 \times 10^{-7}, \quad c_2 = 10^{-4}, \quad c_3 = 10^{-3}, \quad c_4 = 3.5; \quad N_1 = 1 \times 10^5, \quad N_2 = 2 \times 10^5. \tag{5.11}$$

We ran the simulation from $t = 0$ with initial state $x(0) = 250$ to time $T = 4$.

The distribution of the final state is plotted in Figure 5.8. Since the Schlögl solution has a bimodal distribution, the mean and variance alone do not provide a good measure of weak convergence or convergence in distribution. Hence we use an estimate of the total variation distance between two distribution functions as our measure of error. Suppose X and Y are scalar random variables with probability density functions p_X and p_Y. The total variation between X and Y is defined as

$$d(X, Y) = \int |p_X(s) - p_Y(s)| ds. \tag{5.12}$$

In the simulation, the distance (5.12) can be approximated in an efficient way. Suppose we have realizations $x_1, \ldots, x_N$ for X and $y_1, \ldots, y_M$ for Y, both bounded in the interval $I = [0, L)$. Divide the interval I into K subintervals and denote the subintervals as $z_i = [\frac{(i-1)L}{K}, \frac{iL}{K})$. Now define

$$\begin{aligned} f_X(z_i) &= \tfrac{K}{NL} \cdot \#\{x_j : x_j \in z_i\}, \\ f_Y(z_i) &= \tfrac{K}{ML} \cdot \#\{y_j : y_j \in z_i\}. \end{aligned} \tag{5.13}$$

This yields an estimate $D_K(X, Y)$ of $d(X, Y)$ given by

$$D_K(X, Y) = \sum_{i=1}^{K} \frac{L}{K} |f_x(z_i) - f_y(z_i)|. \tag{5.14}$$

Figure 5.9 shows the distribution error vs. the stepsize. Again we see linear convergence behavior.

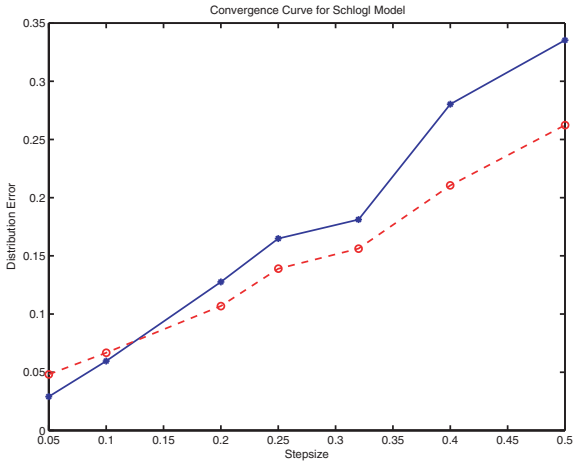

FIG. 5.9. *The distribution error vs. the stepsize for Example* 3 *with* 10^6 *samples. The solid line plot with "*" is for explicit tau, while the dashed line with "o" is for implicit tau.*

6. Conclusions and future work. We have presented consistency results for the explicit and implicit tau leaping methods for simulation of stochastic chemical reaction systems. Our results show that both the explicit and the implicit tau methods are first-order consistent. Our local error analysis provides a systematic way to compute the local error in all the moments up to any power of τ. We have explicitly provided the $O(\tau^2)$ error formulae for the mean and covariance of explicit tau as well as implicit tau. These results hold for general systems under physically reasonable assumptions. We have also verified for several examples that the range of validity of this formula is large enough to be practically applicable in tau-leaping situations.

We showed first-order convergence of all the moments of both tau methods for systems with linear propensity functions. The extension of this result to general nonlinear systems is a subject of ongoing research.

In addition, we have given an asymptotic stability analysis for the mean of both tau methods. This shows that implicit tau is unconditionally asymptotically stable, provided the system is asymptotically stable, while explicit tau has restrictions on its stepsize. These results are identical to those for explicit and implicit Euler methods applied to ODEs.

Appendix A. Some lemmas relevant for the analysis of implicit tau.

LEMMA A.1. *Assuming the bounded Poisson numbers are used, for any function* $g : \mathbb{R}^N \to \mathbb{R}$ *that is Lipschitz continuous on any bounded domain of* $\mathbb{R}^N$ *and initial state* x *there exist* $K > 0$ *and* $\delta > 0$ *such that*

$$\left| E\left(g(X^{(et)}(t+\tau)) - g(x) \mid X^{(et)}(t) = x \right) \right| < K\tau \quad \forall \tau \in [0, \delta].$$

Proof. Given any initial state $X^{(et)}(t) = x$, the possible values for $X^{(et)}(t+\tau)$ lie in a bounded subdomain of $\mathbb{R}^N$. By the stated Lipschitz property of g, there exists $\gamma > 0$ such that

$$|g(X^{(et)}(t+\tau)) - g(x)| \leq \gamma\, |X^{(et)}(t+\tau) - x|.$$

Note that

$$|X^{(et)}(t+\tau) - x| \le \sum_{j=1}^{M} |\nu_j| \, \mathcal{P}_j(a_j(x), \tau),$$

where $|\nu_j|$ denotes the absolute value of each component of vector ν_j. Thus

$$E\left(|X^{(et)}(t+\tau) - x| \;\middle|\; X^{(et)}(t) = x\right) \le \tau \sum_{j=1}^{M} |\nu_j| \, a_j(x).$$

By the reasoning in Remark 3.2, we conclude that

$$E\left(|X^{(et)}(t+\tau) - x| \;\middle|\; X^{(et)}(t) = x\right) \le \tau C_1(x)$$

for all $\tau > 0$ sufficiently small, where $C_1(x) > 0$ is a constant independent of τ. Thus

$$\begin{aligned}
&\left|E\left(g(X^{(et)}(t+\tau)) - g(x) \mid X^{(et)}(t) = x\right)\right| \\
&\qquad \le E\left(|g(X^{(et)}(t+\tau)) - g(x)| \;\middle|\; X^{(et)}(t) = x\right) \\
&\qquad \le \gamma \, E\left(|X^{(et)}(t+\tau) - x| \;\middle|\; X^{(et)}(t) = x\right) \\
&\qquad \le \gamma \, C_1(x)\tau
\end{aligned}$$

for τ sufficiently small. □

LEMMA A.2. *For any positive integer l and initial state x there exist $K > 0$ and $\delta > 0$ such that*

$$\left|E\left((X' - X'_e)^l\right)\right| < K\tau^{l+1} \quad \forall \tau \in [0, \delta].$$

Proof. From (3.25) it follows that

$$(X' - X'_e)^l = \tau^l \left(\sum_{j=1}^{M} \nu_j \{a_j(x) - a_j(X'_e)\}\right)^l + \tau^{l+1} H(X'_e, x),$$

where H is bounded for bounded values of X'_e. Taking expectations and using Lemma A.1 (and also using the fact that a_j are Lipschitz on any bounded domain by Lemma 2.4), we obtain

$$\left|E\left((X' - X'_e)^l\right)\right| \le \tau^l \left(\sum_{j=1}^{M} |\nu_j| \tau K_j\right) + \tau^{l+1} E(|H(X'_e, x)|),$$

where K_j are the constants in Lemma A.1 corresponding to the functions a_j. Since X'_e takes values in a bounded domain with probability 1, we may bound $E(|H(X'_e, x)|) \le H_0(x)$, where $H_0(x)$ is the maximum value of H on this domain. This proves the result. □

LEMMA A.3. *For any multivariate polynomial function $g : \mathbb{R}^N \to \mathbb{R}$ and initial state x there exist $C > 0$ and $\delta > 0$ such that*

$$|E(g(X') - g(X'_e))| \le C\tau^2 \quad \forall \tau \in [0, \delta].$$

Proof. This is an immediate consequence of Lemma A.2. Taylor expanding and then taking expectations, we obtain

$$E(g(X') - g(X'_e)) = \sum_{l=1}^{d} E\left(\frac{d^l g}{dx^l}(X'_e)\,(X' - X'_e)^l\right),$$

where d is the degree of g. Since X'_e takes values on a bounded domain, we obtain the bound

$$\left|E\left(\frac{d^l g}{dx^l}(X'_e)\,(X' - X'_e)^l\right)\right| \le C_l \left|E((X' - X'_e)^l)\right|,$$

where C_l are upper bounds for the absolute values of the derivatives on the bounded domain. Using Lemma A.2 we obtain the result. ∎

LEMMA A.4. *For any function $g : \mathbb{R}^N \to \mathbb{R}$ that is Lipschitz continuous on any bounded domain of $\mathbb{R}^N$ and initial state x there exist $K > 0$ and $\delta > 0$ such that*

$$|E\,(g(X') - g(x))| < K\tau \quad \forall \tau \in [0, \delta].$$

Proof. This follows directly from Lemmas A.1 and A.3. Note that by definition, conditioned on $X^{(et)}(t) = x$, $X^{(et)}(t + \tau) = X'_e$. ∎

Appendix B. Proof of convergence. Let $\Phi_\tau(X_0)$ denote the result of one step of size τ of either of the tau-leaping schemes applied to the initial condition X_0 (which is a random variable assumed to have finite moments). Let $\hat{X}$ denote the numerical solution obtained by the tau-leap method with stepsizes $\tau_1, \ldots, \tau_n$ starting at $t = t_0$ over an interval of length $T = \tau_1 + \cdots + \tau_n$. Let $t_i = t_0 + \tau_1 + \cdots + \tau_i$, and let $\tau = \max \tau_i$. Let X be the true solution process. Then

$$\begin{aligned}
&\|E(\hat{X}(t_0 + T)^r) - E(X(t_0 + T)^r)\| \\
&= \Bigg\| \sum_{i=1}^{n} \big(E((\Phi_{\tau_n} \circ \Phi_{\tau_{n-1}} \circ \cdots \circ \Phi_{\tau_i}(X(t_{i-1})))^r) \\
&\qquad - E((\Phi_{\tau_n} \circ \Phi_{\tau_{n-1}} \circ \cdots \circ \Phi_{\tau_{i+1}}(X(t_i)))^r) \big) \Bigg\| \\
&\le \sum_{i=1}^{n} \sum_{j=1}^{r} K_{rj} \left\| E((\Phi_{\tau_i}(X(t_{i-1})))^j) - E((X(t_i))^j) \right\| \\
&\le \sum_{i=1}^{n} \sum_{j=1}^{r} K_{rj} C_{ij} \tau_i^2 \le \sum_{j=1}^{r} K_{rj} \left(\sum_{i=1}^{n} C_{ij} \tau_i \right) \delta \\
&\le \sum_{j=1}^{r} K_{rj} C \tau \left(\sum_{i=1}^{n} \tau_i \right) \le K\tau.
\end{aligned}$$

Here we have used 0-stability as well as consistency. Constants K_{rj} relate to 0-stability, and C_{ij} relate to consistency. The quantity $\delta > 0$ is small enough for both to hold, provided $\tau < \delta$. An important point to note is that the consistency is needed only for initial states $X(t)$ for $t \in [t_0, t_0 + T]$. Since the true process is integer valued, it is adequate to prove consistency for integer initial states. For 0-stability, however, noninteger states need to be considered as well since $\Phi_{\tau_i}(X(t_i))$ need not be integer valued for (unrounded) implicit tau.

Acknowledgment. We would like to thank the reviewers for their constructive comments.

REFERENCES

[1] A. Arkin, J. Ross, and H. H. McAdams, *Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected* e. coli *cells*, Genetics, 149 (1998), pp. 1633–1648.
[2] U. M. Ascher and L. R. Petzold, *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, SIAM, Philadelphia, 1998.
[3] Y. Cao, L. R. Petzold, M. Rathinam, and D. T. Gillespie, *The numerical stability of leaping methods for stochastic simulation of chemically reacting systems*, J. Chem. Phys., 121 (2004), pp. 12169–12178.
[4] N. Fedoroff and W. Fontana, *Small numbers of big molecules*, Science, 297 (2002), pp. 1129–1131.
[5] D. T. Gillespie, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, J. Computational Phys., 22 (1976), pp. 403–434.
[6] D. T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem., 81 (1977), pp. 2340–2361.
[7] D. T. Gillespie, *Markov Processes: An Introduction for Physical Scientists*, Academic Press, Boston, 1992.
[8] D. T. Gillespie, *The chemical Langevin equation*, J. Chem. Phys., 113 (2000), pp. 297–306.
[9] D. T. Gillespie, *Approximate accelerated stochastic simulation of chemically reacting systems*, J. Chem. Phys., 115 (2001), pp. 1716–1733.
[10] D. T. Gillespie and L. Petzold, *Improved leap-size selection for accelerated stochastic simulation*, J. Chem. Phys., 119 (2003), pp. 8229–8234.
[11] D. Givon, R. Kupferman, and A. M. Stuart, *Extracting macroscopic dynamics: Model problems and algorithms*, Nonlinearity, 17 (2004), pp. 55–127.
[12] H. H. McAdams and A. Arkin, *Stochastic mechanisms in gene expression*, Proc. Natl. Acad. Sci. USA, 94 (1997), pp. 814–819.
[13] H. H. McAdams and A. Arkin, *It's a noisy business!*, Trends in Genetics, 15 (1999), pp. 65–69.
[14] C. C. Pugh, *Real Mathematical Analysis*, 1st ed., Springer-Verlag, New York, 2002.
[15] 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 (2003), pp. 11784–11794.
[16] S. Ross, *Introduction to Probability Models*, 8th ed., Academic Press, San Diego, 2002.
[17] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, 2nd ed., Springer-Verlag, New York, 1993.
[18] E. Vanden-Eijnden, *Analysis of multiscale methods for stochastic differential equations*, Commun. Math. Sci., 1 (2003), pp. 377–384.