Efficient step size selection for the tau-leaping simulation method

Yang Cao[a)]
Department of Computer Science, Virginia Tech, Blacksburg, Virginia 24061

Daniel T. Gillespie
Dan T. Gillespie Consulting, 30504 Cordoba Place, Castaic, California 91384

Linda R. Petzold
Department of Computer Science, University of California, Santa Barbara, Santa Barbara, California 93106

(Received 3 August 2005; accepted 29 November 2005; published online 30 January 2006)

The tau-leaping method of simulating the stochastic time evolution of a well-stirred chemically reacting system uses a Poisson approximation to take time steps that leap over many reaction events. Theory implies that tau leaping should be accurate so long as no propensity function changes its value "significantly" during any time step τ. Presented here is an improved procedure for estimating the largest value for τ that is consistent with this condition. This new τ-selection procedure is more accurate, easier to code, and faster to execute than the currently used procedure. The speedup in execution will be especially pronounced in systems that have many reaction channels. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2159468]

I. INTRODUCTION

Stochastic simulation of chemically reacting systems is a topic of current interest, since discreteness and stochasticity can be important in systems formed by living cells where some key reactant molecules may be present in small numbers.[1–3] Gillespie's stochastic simulation algorithm[4,5] (SSA) is an essentially exact numerical simulation method for well-stirred systems and is widely used in the simulation of biochemical systems. But because the SSA keeps track of every reaction event, it is impractical for many realistic problems, in spite of recent significant improvements.[6,7]

To speed up discrete stochastic simulation, Gillespie[8] proposed the tau-leaping method as an approximate simulation strategy. By using Poisson random numbers, the tau-leaping method can often leap over many reactions without a significant loss of accuracy. Tau leaping provides a natural bridge from the SSA in the discrete stochastic regime, to the explicit Euler method for the chemical Langevin equation (CLE) in the continuous stochastic regime, to the explicit Euler method for the reaction rate equation (RRE) in the continuous deterministic regime.[9] It seems likely that some form of tau leaping will be required to successfully simulate most biological systems.

Several improvements in tau leaping have recently been proposed. Gillespie and Petzold[10] improved Gillespie's original strategy for choosing the size of the tau leap. Rathinam *et al.*[11] developed an *implicit* tau-leaping method to more efficiently simulate stiff systems. Tian and Burrage[12] and Chatterjee *et al.*,[13] noting that the tau-leaping tactic of using Poisson random numbers can sometimes produce negative populations, introduced a *binomial* tau-leaping method to avoid that. Cao *et al.*[14] then modified the original Poisson tau-leaping method in a way that seems to resolve the negativity problem somewhat more adroitly.

For all forms of tau leaping, the procedure for selecting τ has been the one proposed by Gillespie and Petzold.[10] But this Gillespie-Petzold (GP) procedure has two notable shortcomings: First, since it bounds the estimated change in each propensity function during a leap by a specified fraction ϵ of the *sum* of all the propensity functions, any propensity function that has a relatively small value will be allowed to change by a *relatively* large amount. That would seem to violate the leap condition (required by theory for accuracy in tau leaping[8]), which says that all propensity functions should remain "approximately constant" during a leap. Secondly, the GP τ-selection procedure requires an evaluation at each leap on the order of M^2 auxiliary quantities, where M is the number of reaction channels. This can amount to a significant computational burden in realistic systems, where M is typically large.

In this paper, we propose a new τ-selection procedure that addresses both of these shortcomings. Our new τ-selection procedure is more accurate than the GP procedure because it adheres more closely the leap condition; more specifically, it uniformly bounds the *relative* changes in the propensity functions. In addition, our new τ-selection procedure is faster than the GP procedure, because the number of auxiliary computations required to implement it increases linearly with the number of reactant species, rather than quadratically with the number of reaction channels.

The outline of this paper is as follows: In Sec. II we briefly review the SSA, the original and modified Poisson tau-leaping methods, and the GP τ-selection procedure. In Sec. III we discuss the strategy of uniformly bounding the relative changes in the propensity functions. We show that this strategy can be incorporated into the GP τ-selection procedure quite easily, and that although this does not make τ

[a)]Author to whom correspondence should be addressed. Electronic mail: ycao@vt.edu

selection any faster it does make the simulations more accurate. In Sec. IV we develop a different way of uniformly bounding the relative changes in the propensity functions that is easier to code and significantly faster to execute—this is our newly proposed τ-selection procedure. Numerical results illustrating its improved performance are exhibited in Sec. V, and our conclusions are summarized in Sec. VI.

II. BACKGROUND

A. The SSA

We consider a system of N molecular species $\{S_1,\ldots,S_N\}$ interacting through M chemical reaction channels $\{R_1,\ldots,R_M\}$. The state of the system is described by the vector $\mathbf{X}(t)\equiv(X_1(t),\ldots,X_N(t))$, where $X_i(t)$ is the number of molecules of species S_i in the system at time t. We assume that the system is well stirred and in thermal (but not chemical) equilibrium. The dynamics of reaction channel R_j is characterized by a *propensity function* a_j and a *state change vector* $\boldsymbol{\nu}_j\equiv(\nu_{1j},\ldots,\nu_{Nj})$: $a_j(\mathbf{x})dt$ gives the probability, given $\mathbf{X}(t)=\mathbf{x}$, that one R_j reaction will occur in the next infinitesimal time interval $[t,t+dt)$, and ν_{ij} is the change in the S_i molecular population induced by one R_j reaction.

The dynamics of the system obeys the *chemical master equation* (CME),

$$\frac{\partial P(\mathbf{x},t|\mathbf{x}_0,t_0)}{\partial t}=\sum_{j=1}^{M}\left[a_j(\mathbf{x}-\boldsymbol{\nu}_j)P(\mathbf{x}-\boldsymbol{\nu}_j,t|\mathbf{x}_0,t_0)-a_j(\mathbf{x})P(\mathbf{x},t|\mathbf{x}_0,t_0)\right], \tag{1}$$

where $P(\mathbf{x},t|\mathbf{x}_0,t_0)$ denotes the probability that $\mathbf{X}(t)$ will be $\mathbf{x}$ given that $\mathbf{X}(t_0)=\mathbf{x}_0$. The CME is computationally intractable for all but the simplest models, so a recourse is taken to the logically equivalent SSA.[4,5] It is based on the fact that, with

$$a_0(\mathbf{x})\equiv\sum_{j=1}^{M}a_j(\mathbf{x}), \tag{2}$$

then given $\mathbf{X}(t)=\mathbf{x}$, the time τ to the next occurring reaction is the exponentially distributed random variable with mean $1/a_0(\mathbf{x})$, and the index j of that reaction is the integer random variable with point probability $a_j(\mathbf{x})/a_0(\mathbf{x})$. To advance the system from state $\mathbf{x}$ at time t, the SSA generates two random numbers r_1 and r_2 uniformly in the unit interval, and then takes the time of the next reaction to be $t+\tau$ where

$$\tau=\frac{1}{a_0(\mathbf{x})}\ln\left(\frac{1}{r_1}\right), \tag{3}$$

and the index for the next reaction to be the smallest integer j satisfying

$$\sum_{j'=1}^{j}a_{j'}(\mathbf{x})>r_2a_0(\mathbf{x}). \tag{4}$$

The system state is then updated according to $\mathbf{X}(t+\tau)=\mathbf{x}+\boldsymbol{\nu}_j$, and this process gets repeated until some final time or condition is reached. The SSA is exact in the sense that the sample paths it generates are precisely distributed according to the solution of the CME. But its strategy of simulating every reaction event one at a time often makes it too time consuming to implement for real systems.

B. Tau leaping

The tau-leaping method[8] tries to speed up stochastic simulation by answering the following question: How often does each reaction channel fire in the next *specified* time interval τ? More precisely, let

$$K_j(\tau;\mathbf{x},t)\triangleq\text{the number of times, given }\mathbf{X}(t)=\mathbf{x},\text{ that reaction channel }R_j\text{ will fire in the time interval }[t,t+\tau)\quad(j=1,\ldots,M). \tag{5}$$

For arbitrary values of τ it will be about as difficult to compute $K_j(\tau;\mathbf{x},t)$ as to solve the CME. But if τ is small enough that, during $[t,t+\tau)$, no propensity function suffers an "appreciable change" in its value, a requirement that is called the *leap condition*, then a good *approximation* to $K_j(\tau;\mathbf{x},t)$ will be provided by $P(a_j(\mathbf{x}),\tau)$, where $P(a,\tau)$ is the Poisson random variable with mean (and variance) $a\tau$. So if $\mathbf{X}(t)=\mathbf{x}$ and we choose τ to satisfy the leap condition, we can update the state to time $t+\tau$ according to the *approximate* formula

$$\mathbf{X}(t+\tau)\doteq\mathbf{x}+\sum_{j=1}^{M}\boldsymbol{\nu}_jP_j(a_j(\mathbf{x}),\tau), \tag{6}$$

where $P_j(a_j(\mathbf{x}),\tau)$ for each $j=1,\ldots,M$ denotes an independent sample of the Poisson random variable with mean $a_j(\mathbf{x})\tau$. This computational procedure is known as the *tau-leaping approximation*. If it also happens that $a_j(\mathbf{x})\tau\gg1$ for all $j=1,\ldots,M$, it is easy to show that formula (6) reduces to the simple Euler method for the CLE.[8]

In order for tau leaping to be practical, we need to have a procedure for quickly determining the largest value of τ that is compatible with the leap condition. Gillespie[8] originally proposed that the leap condition could be considered satisfied if the expected change in each propensity function $a_j(\mathbf{x})$ during the leap were bounded by $\epsilon a_0(\mathbf{x})$, where ϵ is an error control parameter $(0<\epsilon\ll1)$. Later, Gillespie and Petzold[10] showed that the largest value of τ that satisfies this requirement can be estimated as follows: First compute the M^2+2M auxiliary quantities

$$f_{jj'}(\mathbf{x})\equiv\sum_{i=1}^{N}\frac{\partial a_j(\mathbf{x})}{\partial x_i}\nu_{ij'},\quad j,j'=1,\ldots,M, \tag{7}$$

$$\mu_j(\mathbf{x})\equiv\sum_{j'=1}^{M}f_{jj'}(\mathbf{x})a_{j'}(\mathbf{x}),\quad j=1,\ldots,M, \tag{8a}$$

$$\sigma_j^2(\mathbf{x})\equiv\sum_{j'=1}^{M}f_{jj'}^2(\mathbf{x})a_{j'}(\mathbf{x}),\quad j=1,\ldots,M; \tag{8b}$$

then take

$$\tau = \min_{j\in[1,M]} \left\{ \frac{\epsilon a_0(\mathbf{x})}{|\mu_j(\mathbf{x})|}, \frac{(\epsilon a_0(\mathbf{x}))^2}{\sigma_j^2(\mathbf{x})} \right\}. \tag{9}$$

The derivation of these formulas[10] shows that $\mu_j(\mathbf{x})\tau$ estimates the *mean* of the expected change in $a_j(\mathbf{x})$ in time τ, $\sqrt{\sigma_j^2(\mathbf{x})\tau}$ estimates the *standard deviation* of the expected change in $a_j(\mathbf{x})$ in time τ, and formula (9) essentially requires that both of those quantities be bounded by $\epsilon a_0(\mathbf{x})$ for all j. We should note that Gillespie's original τ-selection formula[8] was deficient in that it lacked the σ_j^2 argument in Eq. (9).

C. Modified (non-negative) Poisson tau leaping

Because the Poisson random variable is unbounded, it is possible that the Poisson approximation to $K_j(\tau;\mathbf{x},t)$ in Eq. (6) might result in reaction channel R_j firing so many times that the population of one of its reactant species will be driven negative. This has actually been found to happen in the simulation of certain systems in which some consumed reactant species is present in small numbers. To resolve this problem, Tian and Burrage,[12] and independently Chatterjee *et al.*,[13] proposed a binomial tau-leaping method, in which bounded binomial random variables replace the unbounded Poisson random variables. More recently, Cao *et al.*[14] devised a *modified* Poisson tau-leaping procedure that seems to resolve the negativity problem more satisfactorily.

The modified Poisson tau-leaping algorithm[14] is based on the fact that negative populations typically arise from multiple firings of reactions that are only a few firings away from consuming all the molecules of one of their reactants. To focus on those reaction channels, the modified tau-leaping algorithm introduces a second control parameter n_c, a positive integer that is usually set somewhere between 2 and 20. Any reaction channel with a positive propensity function that is currently within n_c firings of exhausting one of its reactants is then classified as a *critical* reaction. The modified algorithm chooses τ in such a way that no more than one firing of *all* the critical reactions can occur during the leap. Essentially, the algorithm simulates the *critical* reactions using an adapted (and thus not quite exact) version of the SSA, and the remaining *noncritical* reactions using the previously described Poisson tau-leaping method. Since no more than one firing of a critical reaction can occur during a leap, the probability of producing a negative population is reduced to nearly zero. On those rare occasions when a negative population does arise (from firings of some noncritical reaction), the leap can simply be rejected and repeated with τ reduced by half, or else the simulation can be started over using a larger value for n_c.

It can be shown[14] that the modified Poisson tau-leaping procedure becomes identical to the SSA if n_c is chosen so large that *every* reaction channel is critical, and becomes identical to the tau-leaping procedure of Sec. II B if $n_c=0$ (and *no* reaction channels are critical). Thus, the modified Poisson tau-leaping algorithm is not only more robust but also potentially more accurate than the earlier tau-leaping algorithm. The explicit steps in the algorithm are as follows.

1. Modified Poisson tau-leaping algorithm

(1) In state $\mathbf{x}$ at time t, identify the currently critical reactions. This is done by first estimating for each reaction R_j with $a_j(\mathbf{x})>0$ the maximum number of times L_j that R_j can fire before exhausting one of its reactants,[12,13]

$$L_j = \min_{i\in[1,N];\nu_{ij}<0} \left[\frac{x_i}{|\nu_{ij}|} \right]. \tag{10}$$

Here the minimum is taken over only those index values i for which $\nu_{ij}<0$, and the brackets denote "greatest integer in." Any reaction R_j with $a_j(\mathbf{x})>0$ is deemed critical if $L_j<n_c$. (We will normally take $n_c=10$.)

(2) With a value chosen for ϵ (we normally take $\epsilon=0.03$), compute a *candidate* time leap τ' by using the following *slightly altered* version of the GP τ-selection formulas: Let J_{ncr} denote the set of indices of the *noncritical reactions*. If J_{ncr} is empty (i.e., there are no noncritical reactions), take $\tau'=\infty$ and go to step 3. Otherwise, compute the auxiliary quantities

$$f_{jj'}(\mathbf{x}) \equiv \sum_{i=1}^{N} \frac{\partial a_j(\mathbf{x})}{\partial x_i} \nu_{ij'}, \quad j\in[1,M]; j'\in J_{\mathrm{ncr}}, \tag{11}$$

$$\mu_j(\mathbf{x}) \equiv \sum_{j'\in J_{\mathrm{ncr}}} f_{jj'}(\mathbf{x}) a_{j'}(\mathbf{x}), \quad j\in[1,M], \tag{12a}$$

$$\sigma_j^2(\mathbf{x}) \equiv \sum_{j'\in J_{\mathrm{ncr}}} f_{jj'}^2(\mathbf{x}) a_{j'}(\mathbf{x}), \quad j\in[1,M]; \tag{12b}$$

then take

$$\tau' = \min_{j\in[1,M]} \left\{ \frac{\epsilon a_0(\mathbf{x})}{|\mu_j(\mathbf{x})|}, \frac{(\epsilon a_0(\mathbf{x}))^2}{\sigma_j^2(\mathbf{x})} \right\}. \tag{13}$$

Note that formulas (11) and (12) differ from formulas (7) and (8) in that the index j' now runs over only the noncritical reactions (but j still runs over all reactions). As thus computed, τ' is the largest permissible tau-leaping timestep for the *noncritical* reactions.

(3) If τ' is less than some small multiple (which we usually take to be 10) of $1/a_0(\mathbf{x})$, abandon tau leaping temporarily, execute some modest number (which we usually take to be 100) of single-reaction SSA steps, and return to step 1. Otherwise, proceed to step 4.

(4) Compute the sum $a_0^c(\mathbf{x})$ of the propensity functions of all the critical reactions. Generate a *second candidate* time leap τ'' as a sample of the exponential random variable with mean $1/a_0^c(\mathbf{x})$. As thus computed, τ'' tentatively estimates the time to the next critical reaction.

(5) Take the actual time leap τ to be the smaller of τ' and τ'', and set the number of firings k_j of each reaction R_j accordingly.

(a) If $\tau'<\tau''$, take $\tau=\tau'$. For all critical reactions R_j set $k_j=0$ (no critical reactions will fire during this leap). For all noncritical reactions R_j, generate k_j as a sample of the Poisson random variable with mean $a_j(\mathbf{x})\tau$.

(b) If $\tau'' \leqslant \tau'$, take $\tau = \tau''$. Generate j_c as a sample of the integer random variable with point probabilities $a_j(\mathbf{x})/a_0^c(\mathbf{x})$, where j runs over the index values of the critical reactions only. (The value of j_c identifies the next critical reaction, the only critical reaction that will fire in this leap.) Set $k_{j_c}=1$, and for all other critical reactions R_j set $k_j=0$. For all the noncritical reactions R_j, generate k_j as a sample of the Poisson random variable with mean $a_j(\mathbf{x})\tau$.

(6) If there is a negative component in $\mathbf{x}+\Sigma_j k_j \boldsymbol{\nu}_j$, reduce τ' by half, and return to step 3. Otherwise, leap by replacing $t \leftarrow t+\tau$ and $\mathbf{x} \leftarrow \mathbf{x}+\Sigma_j k_j \boldsymbol{\nu}_j$; then return to step 1, or else stop.

In the following sections, we will focus on improving the procedure for choosing τ' in step 2.

III. BOUNDING THE *RELATIVE* CHANGES IN THE PROPENSITIES

As was noted earlier, the GP tau-selection procedure seeks to bound the change in each propensity function $a_j(\mathbf{x})$ during a time step τ by a small fraction ϵ of the sum $a_0(\mathbf{x})$ of all the propensity functions. Denoting the change in propensity function a_j from time t to time $t+\tau$, given $\mathbf{X}(t)=\mathbf{x}$, by $\Delta_\tau a_j(\mathbf{x})$, this requirement can be stated as

$$|\Delta_\tau a_j(\mathbf{x})| \leqslant \epsilon a_0(\mathbf{x}), \quad j=1,\ldots,M. \tag{14}$$

This bound is explicitly reflected in the numerators of the two fractions in the τ-selection formulas (9) and (13). Although this strategy does indeed limit the changes in the propensities during a leap as required by the leap condition, it usually does not accomplish that task in a *uniform* way. We recall that the aim of the leap condition is to ensure that every propensity function remains "practically constant" during a τ leap, since that is what allows the number of firings of each reaction R_j during τ to be accurately approximated by a statistically independent Poisson random variable with mean $a_j(\mathbf{x})\tau$. But if $a_j(\mathbf{x})$ happens to be very small compared to $a_0(\mathbf{x})$, condition (14) will allow a large *relative* change in $a_j(\mathbf{x})$, and that could result in simulation inaccuracies.

To illustrate this point, consider the simple reaction set

$$S_1 \xrightarrow{c_1} S_2 \xrightarrow{c_2} S_3, \tag{15}$$

with $c_1=1$ and $c_2=1$, and initial populations $x_1=10^4$, $x_2=1$, and $x_3=0$. The Gillespie-Petzold tau-selection procedure with $\epsilon=0.03$ gives $\tau=0.03$. But for this value of τ, the expected relative change in $a_2(\mathbf{x})$ is about 300. That such a large relative change in a propensity function during a tau leap can lead to simulation errors is demonstrated in Fig. 1, which shows 10^6-sample histograms of X_3 at $t=0.1$ as computed using the exact SSA (solid curve with triangles) and the tau-leaping method with $\epsilon=0.03$ (solid curve with circles).

We might expect that the leap condition would be better satisfied if we instead bounded the relative changes in all the propensity functions by the same amount ϵ,

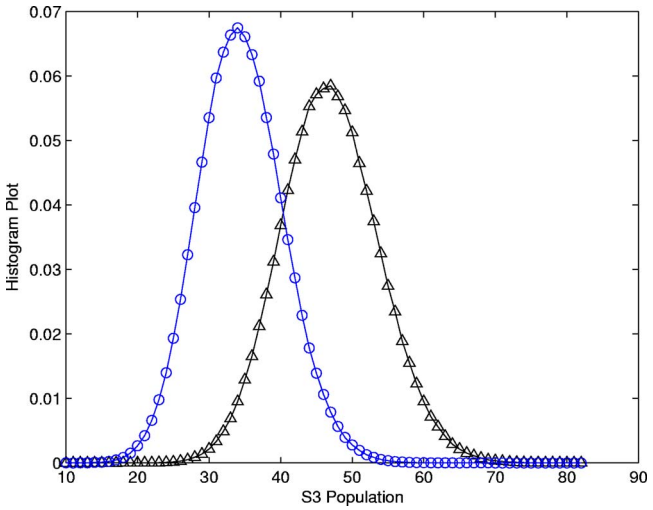

FIG. 1. Histogram plots of $X_3(0.1)$ for reactions (15) computed from 10^6 runs each of the SSA (solid line with triangle) and the tau-leaping method using the original τ'-selection formula (13) with $\epsilon=0.03$ (solid line with circle).

$$|\Delta_\tau a_j(\mathbf{x})| \leqslant \epsilon a_j(\mathbf{x}), \quad j=1,\ldots,M. \tag{16}$$

But doing this can lead to difficulties if $a_j(\mathbf{x})$ happens to approach zero; because then condition (16) will force $|\Delta_\tau a_j(\mathbf{x})|$, and hence also τ, to approach zero, effectively bringing the tau-leaping process to a halt. This difficulty with condition (16) was, in fact, the original motivation for using condition (14) instead.[8] But we can make a simple modification to condition (16) that will avoid this problem. Propensity functions change as reactions occur by discrete amounts, and for every propensity function $a_j(\mathbf{x})$ there will always be a *minimum* amount by which it can change. For example, if R_j is the unimolecular reaction with propensity function $a_j(\mathbf{x})=c_j x_i$, then the minimum (positive) amount by which $a_j(\mathbf{x})$ can change will obviously be c_j. It is not hard to show that if the propensity function of any bimolecular or trimolecular reaction R_j changes at all, it must do so by an amount greater than or equal to c_j. Since it is therefore unreasonable to require any propensity function $a_j(\mathbf{x})$ to change by less than c_j, we should replace the bound on the right-hand side of condition (16) with the *larger* of $\epsilon a_j(\mathbf{x})$ and c_j,

$$\Delta_\tau a_j(\mathbf{x}) \leqslant \max\{\epsilon a_j(\mathbf{x}), c_j\}, \quad j=1,\ldots,M. \tag{17}$$

If the arguments used to derive the GP τ-selection procedure[10] are now applied to the bounding criterion (17), the result is a τ-selection procedure in which formulas (7) and (8) remain unchanged, while in formula (9) the quantity $\epsilon a_0(\mathbf{x})$ gets replaced by $\max\{\epsilon a_j(\mathbf{x}), c_j\}$. Therefore, we can make the modified (non-negative) Poisson tau-leaping algorithm outlined in Sec. II C enforce condition (17) instead of condition (14) simply by *replacing* formula (13) in step 2 with

$$\tau' = \min_{j\in[1,M]} \left\{ \frac{\max\{\epsilon a_j(\mathbf{x}), c_j\}}{|\mu_j(\mathbf{x})|}, \frac{(\max\{\epsilon a_j(\mathbf{x}), c_j\})^2}{\sigma_j^2(\mathbf{x})} \right\}. \tag{18}$$

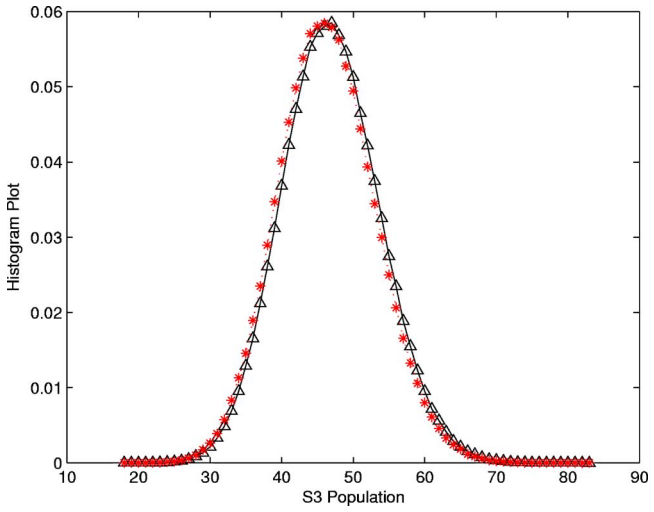

FIG. 2. Histogram plots of $X_3(0.1)$ for reactions (15) computed from 10^6 runs each of the SSA (solid line with triangle) and the tau-leaping method using the improved τ'-selection formula (18) with $\epsilon=0.03$ (dotted line with star).

Figure 2 shows what Fig. 1 would have looked like if the tau-leaping simulations had been carried out using formula (18) instead of formula (13). Obviously, this gives a much better agreement with the exact SSA results.

IV. A NEW TAU-SELECTION PROCEDURE

Although τ selection using formula (18) results in a more accurate simulation than τ selection using formula (13), the evaluation of the functions $\mu_j(\mathbf{x})$ and $\sigma_j^2(\mathbf{x})$ in Eqs. (11) and (12) prior to each leap tends to be very time consuming, especially if both M and N are large. In this section we shall develop a new τ-selection procedure that *approximately* enforces condition (17), but does so in a way that is easier to implement and faster to execute than the procedure specified by formulas (11), (12), and (18).

The underlying strategy of this new τ-selection procedure is to bound the relative changes in the *molecular populations* in such a way that the relative changes in the propensity functions will all be approximately bounded by a specified value ϵ $(0<\epsilon<1)$. Let

$$\Delta_\tau X_i \equiv \Delta_\tau X_i(\mathbf{x}) \triangleq X_i(t+\tau) - x_i \quad \text{given } \mathbf{X}(t)=\mathbf{x}. \tag{19}$$

Instead of basing τ selection on condition (17), we shall base it on the condition

$$\Delta_\tau X_i \leqslant \max\{\epsilon_i x_i, 1\}, \quad \forall\, i \in I_{\mathrm{rs}}. \tag{20}$$

The values of $\epsilon_i=\epsilon_i(\epsilon,x_i)$ are assigned in a way that will be specified shortly, and I_{rs} denotes the set of indices of all *reactant species* (so $i\in I_{\mathrm{rs}}$ if and only if x_i is an argument of at least one propensity function). Condition (20) evidently requires the relative change in X_i to be bounded by ϵ_i, except that X_i will never be required to change by an amount less than 1.

To determine how ϵ_i in condition (20) should be chosen so that the relative changes in all the propensity functions will be bounded by ϵ, we have to examine individually all the possible types of reactions.

Consider first the case in which reaction R_j is the *first-order* reaction $S_i\rightarrow$ products. Its propensity function then has the form $a_j(\mathbf{x})=c_j x_i$. Since the change in a_j is related to the change in X_i by $\Delta a_j=c_j\Delta x_i$, it follows that the relative change in a_j is related to the relative change in X_i by

$$\frac{\Delta a_j}{a_j}=\frac{\Delta x_i}{x_i}. \tag{21}$$

Therefore, if we bound the relative change in X_i by $\epsilon_i=\epsilon$, we will also bound the relative change in a_j by ϵ.

Consider next the case in which reaction R_j is the *second-order* reaction $S_1+S_2\rightarrow$ products, so that its propensity function has the form $a_j(\mathbf{x})=c_j x_1 x_2$. In that case we have, to a reasonably good approximation,

$$\Delta a_j \approx c_j x_2 \Delta x_1 + c_j x_1 \Delta x_2,$$

where we have neglected on the right the usually small term $c_j\Delta x_1\Delta x_2$. To that approximation, we have

$$\frac{\Delta a_j}{a_j}\approx\frac{\Delta x_1}{x_1}+\frac{\Delta x_2}{x_2}. \tag{22}$$

If we bound the relative change in X_1 by $\epsilon_1=\epsilon/2$, and the relative change in X_2 by $\epsilon_2=\epsilon/2$, then *to a first approximation* the relative change in a_j will be bounded by ϵ. The approximate nature of this result arises not only from our neglect of terms nonlinear in the small changes Δx_1 and Δx_2 (an approximation that the GP procedure makes as well) but also from our neglect of any correlation between those changes. Such a correlation would not affect the mean of Eq. (22); however, it could make the variance of the left side of Eq. (22) a little larger or a little smaller than the sum of the variances of the terms on the right side. So the relative change in a_j will actually be bounded by $f\epsilon$, where f is something "close" to 1—close enough for our limited purpose of satisfying the leap condition.

If the second-order reaction R_j has the form $S_i+S_i\rightarrow$ products, then its propensity function will be $a_j(\mathbf{x})=c_j\frac{1}{2}x_i(x_i-1)$, and ignoring terms proportional to $(\Delta x_i)^2$ we will have

$$\Delta a_j \approx c_j\tfrac{1}{2}(x_i-1)\Delta x_i + c_j\tfrac{1}{2}x_i\Delta x_i.$$

Then

$$\frac{\Delta a_j}{a_j}\approx\frac{\Delta x_i}{x_i}+\frac{\Delta x_i}{x_i-1}=\frac{\Delta x_i}{x_i}\left(2+\frac{1}{x_i-1}\right). \tag{23}$$

If we choose the bound ϵ_i on the relative change in X_i to be ϵ divided by the factor in parentheses on the right, then this equation shows that the relative change in a_j will be approximately bounded by ϵ. Note that since this reaction cannot occur unless $x_i\geqslant 2$, the factor in parentheses will always be between 2 (when $x_i=\infty$) and 3 (when $x_i=2$).

Finally, although *third-order* reactions are rare (they are really approximations to sets of coupled first- and second-order reactions), we should for the sake of completeness allow them. They come in three different forms, namely, with all three reacting molecules being different species, or with two of the reacting molecules the same species, or with all

three of the reacting molecules the same species. In the first case we have $a_j(\mathbf{x})=c_j x_1 x_2 x_3$, and ignoring terms quadratic and cubic in the Δx_i's we find

$$\frac{\Delta a_j}{a_j} \approx \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2} + \frac{\Delta x_3}{x_3}. \tag{24}$$

Therefore, by bounding the relative change in each of the reactant species X_i by $\epsilon_i=\epsilon/3$, we can be assured that the relative change in a_j will be approximately bounded by ϵ. For the case in which two of the reactant species are the same, the propensity function will have the form $a_j(\mathbf{x}) = c_j x_1 \frac{1}{2} x_2(x_2-1)$, and ignoring terms quadratic and cubic in the Δx_i's we find

$$\frac{\Delta a_j}{a_j} \approx \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2}\left(2 + \frac{1}{x_2-1}\right). \tag{25}$$

Assuming the relative change in X_1 is bounded by $\epsilon_1=\epsilon/3$, then if we choose the bound ϵ_2 on the relative change in X_2 to be ϵ divided by $3/2$ times the factor in parentheses on the right, the relative change in a_j will be approximately bounded by ϵ. As before, the factor in parentheses will necessarily be between 2 and 3. Finally, for the case in which all three reacting molecules are the same species, the propensity function will have the form $a_j(\mathbf{x})=c_j \frac{1}{6} x_i(x_i-1)(x_i-2)$, and ignoring terms quadratic and cubic in the Δx_i's we find

$$\frac{\Delta a_j}{a_j} \approx \frac{\Delta x_i}{x_i}\left(3 + \frac{1}{x_i-1} + \frac{2}{x_i-2}\right). \tag{26}$$

If we choose the bound ϵ_i on the relative change in X_i to be ϵ divided by the factor in parentheses on the right, then the relative change in a_j will be approximately bounded by ϵ. Note that since this reaction cannot occur unless $x_i \geqslant 3$, the factor in parentheses will always be between 3 (when $x_i=\infty$) and $11/2$ (when $x_i=3$).

On the basis of the foregoing results, we can now infer the following procedure for choosing values for the parameters $\{\epsilon_i\}$ so that condition (20) will ensure that the relative changes in the propensity functions will all be bounded, at least approximately, by ϵ. For each $i \in I_{\mathrm{rs}}$, first determine by inspection the value of $\mathrm{HOR}(i)$, the *highest order of reaction* in which species S_i appears as a reactant. Then take

$$\epsilon_i = \frac{\epsilon}{g_i}, \tag{27}$$

where $g_i=g_i(x_i)$ is defined as follows.

(i) If $\mathrm{HOR}(i)=1$, take $g_i=1$.

(ii) If $\mathrm{HOR}(i)=2$, take $g_i=2$, *except* if any second-order reaction requires *two* S_i molecules take instead

$$g_i = \left(2 + \frac{1}{x_i-1}\right).$$

(iii) If $\mathrm{HOR}(i)=3$, take $g_i=3$, *except* if some third-order reaction requires *two* S_i molecules take instead

$$g_i = \frac{3}{2}\left(2 + \frac{1}{x_i-1}\right),$$

except if some third-order reaction requires *three* S_i molecules take instead

$$g_i = \left(3 + \frac{1}{x_i-1} + \frac{2}{x_i-2}\right).$$

Notice that g_i will remain constant throughout the simulation run if there is only one reactant S_i molecule in the highest-order reaction in which S_i is a reactant. If there are two or more reactant S_i molecules in the highest-order reaction in which S_i is a reactant, g_i will depend on the current value of x_i; fortunately, the form of that dependence, as prescribed in the three formulas above, is computationally simple. The extra effort required to handle the parameters $\epsilon_i=\epsilon/g_i$ will usually be more than compensated by the fact that finding the largest value of τ that satisfies condition (20) can be done much more easily and quickly than finding the largest value of τ that satisfies condition (17). To see that this is so, we shall now derive the procedure for computing the largest value of τ that satisfies condition (20).

Recalling the basic tau-leaping formula (6), we see that the quantity defined in (19) will essentially be given by

$$\Delta_\tau X_i = \sum_{j \in J_{\mathrm{ncr}}} \nu_{ij} P_j(a_j(\mathbf{x}), \tau), \quad \forall\, i \in I_{\mathrm{rs}}. \tag{28}$$

The restriction of the summation index j here to the *noncritical reactions* is motivated by the same logic used in step 2 of the modified (non-negative) tau-leaping algorithm, where the index j' in formulas (11) and (12) is similarly restricted. This is done because in any tau leap there will be at most one firing among all the critical reactions, and to a first approximation any changes induced in the propensity functions by that one firing can be ignored. It is only the changes caused by *multiple* firings of the noncritical reactions that give us concern for the integrity of the leap condition.

Since the Poisson random variables $P_j(a_j(\mathbf{x}), \tau)$ on the right-hand side of Eq. (28) are statistically independent and have means and variances $a_j(\mathbf{x})\tau$, the mean and variance of that linear combination can be computed straightforwardly,

$$\langle \Delta_\tau X_i \rangle = \sum_{j \in J_{\mathrm{ncr}}} \nu_{ij}[a_j(x)\tau], \quad \forall\, i \in I_{\mathrm{rs}}, \tag{29a}$$

$$\mathrm{var}\{\Delta_\tau X_i\} = \sum_{j \in J_{\mathrm{ncr}}} \nu_{ij}^2[a_j(x)\tau], \quad \forall\, i \in I_{\mathrm{rs}}. \tag{29b}$$

Using the same reasoning that was used in deriving the GP τ-selection procedure,[10] we may consider the bound (20) on $\Delta_\tau X_i$ to be "substantially satisfied" if it is simultaneously satisfied by the absolute mean and the standard deviation of $\Delta_\tau X_i$:

$$|\langle \Delta_\tau X_i \rangle| \leqslant \max\{\epsilon_i x_i, 1\}, \quad \sqrt{\mathrm{var}\{\Delta_\tau X_i\}} \leqslant \max\{\epsilon_i x_i, 1\},$$

$$\forall\, i \in I_{\mathrm{rs}}. \tag{30}$$

Substituting formulas (29) into conditions (30), we obtain the following bounds on τ:

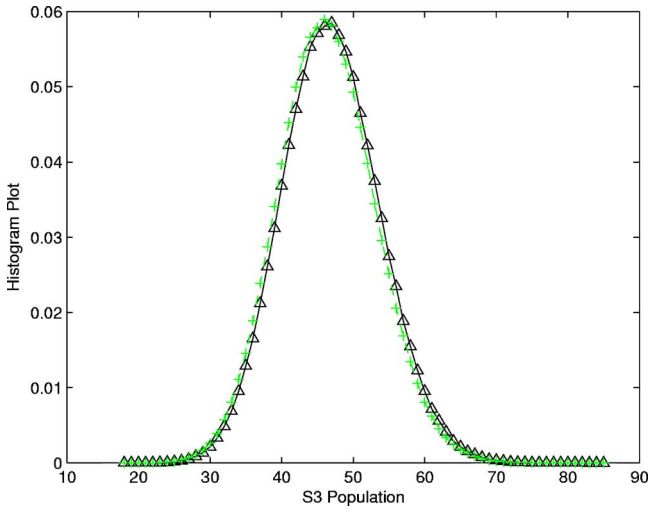

FIG. 3. Histogram plots of $X_3(0.1)$ for reactions (15) computed from 10^6 runs each of the SSA (solid line with triangle) and the tau-leaping method using the new τ'-selection formula (33) with $\epsilon=0.03$ (dashed line with plus).

$$\tau \leqslant \frac{\max\{\epsilon_i x_i, 1\}}{\left|\sum_{j\in J_{\rm ncr}} \nu_{ij} a_j(\mathbf{x})\right|}, \quad \tau \leqslant \frac{\max\{\epsilon_i x_i, 1\}^2}{\sum_{j\in J_{\rm ncr}} \nu_{ij}^2 a_j(\mathbf{x})},$$

$$\forall\, i \in I_{\rm rs}. \tag{31}$$

Recalling formula (27) for ϵ_i, we can now make make the following *change* to the modified (non-negative) Poisson tau-leaping algorithm outlined in Sec. II C: In step 2, compute τ' by first computing the auxiliary quantities

$$\hat{\mu}_i(\mathbf{x}) \triangleq \sum_{j\in J_{\rm ncr}} \nu_{ij} a_j(\mathbf{x}), \quad \forall\, i \in I_{\rm rs}, \tag{32a}$$

$$\hat{\sigma}_i^2(\mathbf{x}) \triangleq \sum_{j\in J_{\rm ncr}} \nu_{ij}^2 a_j(\mathbf{x}), \quad \forall\, i \in I_{\rm rs}, \tag{32b}$$

where $J_{\rm ncr}$ is the set of indices of all noncritical *reactions* and $I_{\rm rs}$ is the set of indices of all reactant species, and then taking

$$\tau' = \min_{i\in I_{\rm ncr}} \left\{ \frac{\max\{\epsilon x_i/g_i, 1\}}{|\hat{\mu}_i(\mathbf{x})|}, \frac{\max\{\epsilon x_i/g_i, 1\}^2}{\hat{\sigma}_i^2(\mathbf{x})} \right\}, \tag{33}$$

where g_i is given by the rules following Eq. (27). As before, if the set $I_{\rm rs}$ is empty (i.e., if there are no noncritical reactions), we instead set $\tau'=\infty$.

The τ'-selection procedure of formulas (32) and (33) will obviously be simpler to program and faster to execute than the τ'-selection procedure of formulas (11), (12), and (18). Note, in particular, that the required number of computational operations increases quadratically with the number of reaction channels in the old formulas, but only linearly with the number of species in the new formulas. Since τ selection has to be performed prior to every tau leap, using these new formulas should lead to substantially faster simulations when the system has many reactions and species.

Figure 3 shows how the results in Fig. 2 would have looked if the τ'-selection procedure in step 2 of the modified (non-negative) Poisson tau-leaping algorithm had been carried out according to the procedure just described. Not surprisingly for this completely first-order system, the new τ'-selection procedure based on condition (20) gives the same excellent agreement with the SSA results as does the τ'-selection procedure based on condition (17), and both are more accurate than the old τ'-selection procedure based on condition (14) (see Fig. 1). Still, in view of the considerable differences between formulas (32) and (33) and formulas (11), (12), and (18), this agreement is very reassuring. In the next section we shall examine more closely the performance of our new τ-selection procedure on some more complicated reaction sets.

V. NUMERICAL EXPERIMENTS

To test the accuracy and efficiency of the new τ-selection formula (33), we have applied the old GP τ-selection formula (13), the improved GP τ-selection formula (18), and the new τ-selection formula (33) to three test problems: the LacZ/LacY model,[12,15] the Schlögl model,[16] and the decaying-dimerizing model.[8,10] For a given value of the error control parameter ϵ, tau-leaping simulations made using different τ-selection formulas exhibit different accuracies and different execution times. To assess the relative accuracies, we first made histograms of final-state populations obtained in a series of repeated SSA runs. Then we made the same number of tau-leaping runs over the same time interval using each of the three τ-selection procedures each with various values for ϵ. The "histogram distances"[17] between the SSA results and the respective tau-leaping results provide a measure of the errors in the tau-leaping methods, assuming enough runs are made that the "self-distances"[17] are small. Since the tau-leaping runs were all made using the same code except for changes in the τ-selection formulas, the relative CPU times for these runs should fairly reflect the relative computational costs of the τ-selection formulas. To verify that the new τ-selection formula (33) generates similar step sizes as those given by the improved GP τ-selection formula (18), we also plotted the step sizes given by two different formulas in one single simulation for each model.

A. LacZ/LacY model

This model was first proposed by Kierzek,[15] and later used by Tian and Burrage[12] to test their binomial tau-leaping procedure. A detailed description of this model, which has 22 reactions and 19 species, can be found in those two references. Since a single SSA simulation from $t=0$ to $t=2100$ took about an hour on our computer, obtaining a large number of SSA samples posed a challenge. We ran the SSA from time $t=0$ to time $t=1000$ to obtain an "initial" state; then we made 10^5 SSA runs from time $t=1000$ to time $t=1001$ (which required about 3.5 h of computer time) and histogrammed the resulting populations. Finally, we made the same number of tau-leaping runs over the same time interval using each of the three τ-selection procedures, each for a range of values for ϵ. Figure 4 shows the plot of histogram distance or "error" for each of the three tau-leaping runs as a function of ϵ. We note that in each case the error increases roughly linearly with ϵ, although more quickly for the original GP τ-selection formula. For a given value of ϵ, the improved GP τ-selection formula (18) and the new τ-selection

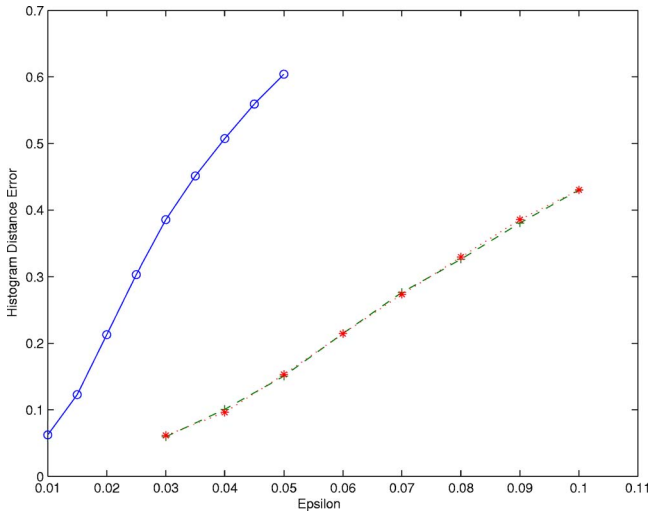

FIG. 4. Plot of the histogram distance errors corresponding to different ϵ values for the three τ'-selection formulas for the LacZ/LacY model. Histogram distance errors are measured between the population distributions of LacZlactose in 10^5 runs of the SSA and the tau-leaping method using different τ-selection formulas. The solid line with "o" is for the original GP τ-selection formula (13), the dotted line with " $*$" is for the improved GP τ-selection formula (18), and the dashed line with "+" is for the new τ-selection formula (33).

formula (33) gave smaller errors, and they were equally accurate. A comparison of the speeds or efficiencies of the three τ-selection procedures is afforded by the plots in Fig. 5 of the error against the CPU run time for an ensemble. Evidently, the new τ-selection formula (33) gave accurate results in less time than either the original or the improved GP τ-selection formula, and the latter two formulas scored about the same on this test. We conclude that for this moderately large reaction set, the new τ-selection formula (33) is the most efficient. Figure 6 shows the step sizes given by the

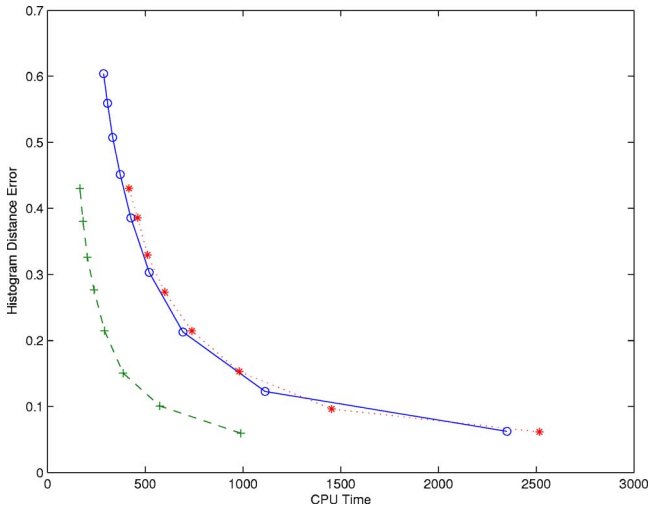

FIG. 5. Plot of the histogram distance errors vs CPU time for the three τ'-selection formulas for the LacZ/LacY model. Histogram distance errors are measured between the population distributions of LacZlactose in 10^5 runs of the SSA and the tau-leaping method using different τ-selection formulas for a range of ϵ values. The solid line with "o" is for the original GP τ-selection formula (13), the dotted line with " $*$" is for the improved GP τ-selection formula (18), and the dashed line with "+" is for the new τ-selection formula (33).

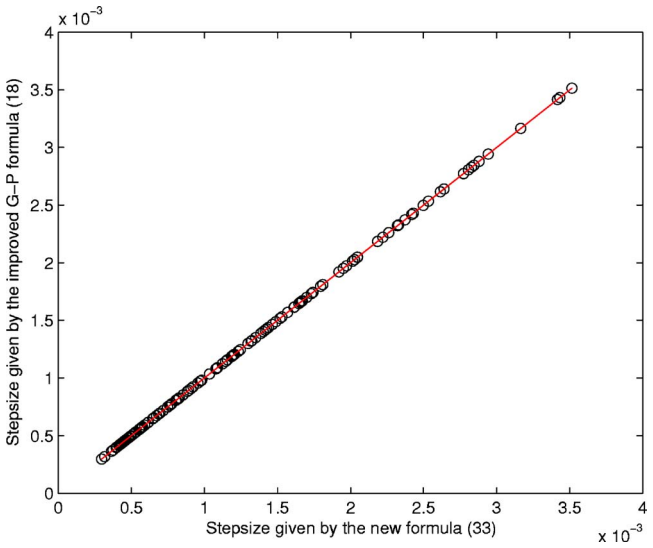

FIG. 6. Plot of the step sizes generated by the improved GP τ-selection formula (18) (y axis) vs the new efficient τ-selection formula (33) (x axis) in a single simulation for the LacZ/LacY model. Since there are many steps in one simulation, the step size is plotted at every 10 000 steps. The data are shown as "o" points. A straight line $y=x$ is plotted for comparison. The maximum of the relative difference is 3.16%.

improved GP τ-selection formula (18) and the new τ-selection formula (33) in a single simulation. Figure 6 shows that, in the course of a typical simulation run, the step sizes given by our new tau-selection formula (33) are practically the same as the step sizes given by the improved GP tau-selection formula (18). But of course, as is shown by Fig. 5, formula (33) gives those step sizes more rapidly.

B. Schlögl model

This model is famous for its bistable steady-state distribution. The reactions are

$$B_1 + 2X \underset{c_2}{\overset{c_1}{\rightleftharpoons}} 3X,$$

$$B_2 \underset{c_4}{\overset{c_3}{\rightleftharpoons}} X, \tag{34}$$

where B_1 and B_2 denote buffered species whose respective molecular populations N_1 and N_2 are assumed to remain essentially constant over the time interval of interest. There is only one time-varying species, X; the state change vectors are $\nu_1=\nu_3=1$ and $\nu_2=\nu_4=-1$; and the propensity functions are

$$a_1(x) = (c_1/2)N_1x(x-1),$$

$$a_2(x) = (c_2/6)x(x-1)(x-2),$$

$$a_3(x) = c_3N_2,$$

$$a_4(x) = c_4x. \tag{35}$$

For some values of the parameters this model has two stable states, and that is the case for the parameter values we have chosen here,

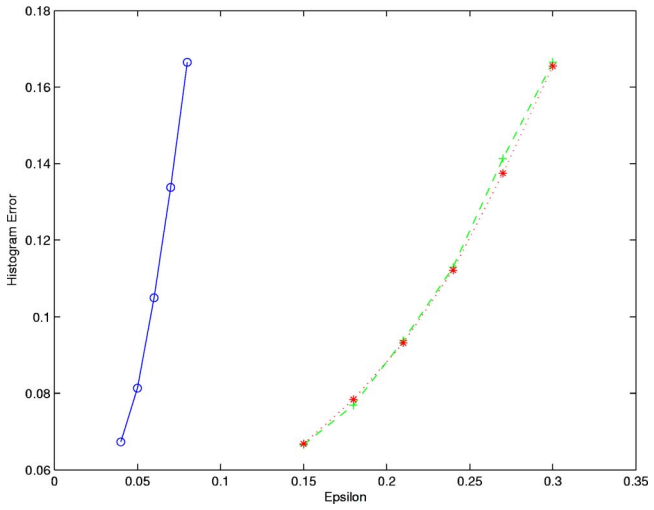

FIG. 7. Plot of the histogram distance errors corresponding to different ϵ values for the three τ-selection formulas for the Schlögl model. Histogram distance errors are measured by 10^6 samples generated from the SSA method and the tau-leaping method using different τ-selection formulas. The solid line with "o" is for the original GP τ-selection formula (13), the dotted line with " $*$" is for the improved GP τ-selection formula (18), and the dashed line with "+" is for the new τ-selection formula (33).

$$c_1 = 3 \times 10^{-7}, \quad c_2 = 10^{-4}, \quad c_3 = 10^{-3}, \quad c_4 = 3.5,$$

$$N_1 = 1 \times 10^5, \quad N_2 = 2 \times 10^5. \tag{36}$$

We made ensembles of 10^6 simulation runs from the initial state $X(0)=250$ to time $t=4$ using the SSA and tau leaping, the latter for each of the three τ-selection procedures and over a range of ϵ-values. Figure 7 shows the histogram distance or error between each tau-leaping ensemble and the SSA ensemble as a function of ϵ. Again, all three τ-selection formulas give errors that increase roughly linearly with ϵ, but this increase is evidently much faster for the original GP τ-selection formula. The improved GP τ-selection formula (18) and the new τ-selection formula (33) give nearly the same accuracy and they evidently give more accurate results than the original GP τ-selection formula for a given value of ϵ. Figure 8 plots the errors in the tau-leaping simulations as a function of CPU time. Although the results show that the improved GP τ-selection formula is less efficient than the original GP τ-selection formula, the new τ-selection formula (33) still shows its highest efficiency. Figure 9 shows that, in the course of a typical simulation run, the step sizes given by our new tau-selection formula (33) are practically the same as the step sizes given by the improved GP tau-selection formula (18). But of course, as is shown by Fig. 8, formula (33) gives those step sizes more rapidly.

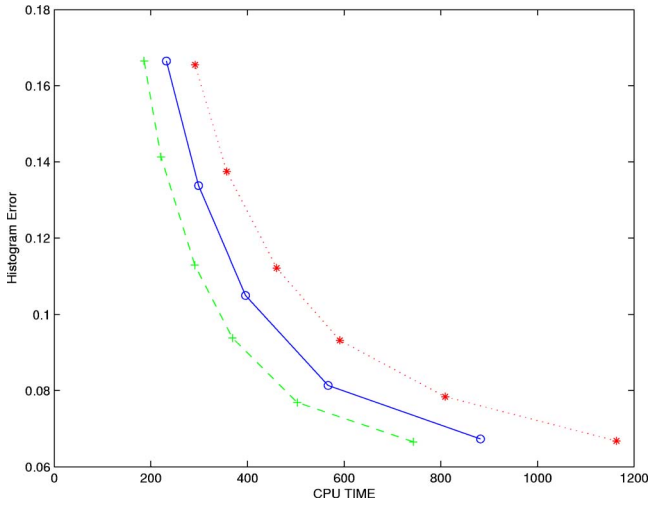

FIG. 8. Plot of the histogram distance errors corresponding to different CPU times for the three τ-selection formulas for the Schlögl model. Histogram distance errors are measured by 10^6 samples generated from the SSA method and the tau-leaping method using different τ-selection formulas and different ϵ values. The solid line with "o" is for the original GP τ-selection formula (13), the dotted line with " $*$" is for the improved GP τ-selection formula (18), and the dashed line with "+" is for the new τ-selection formula (33).

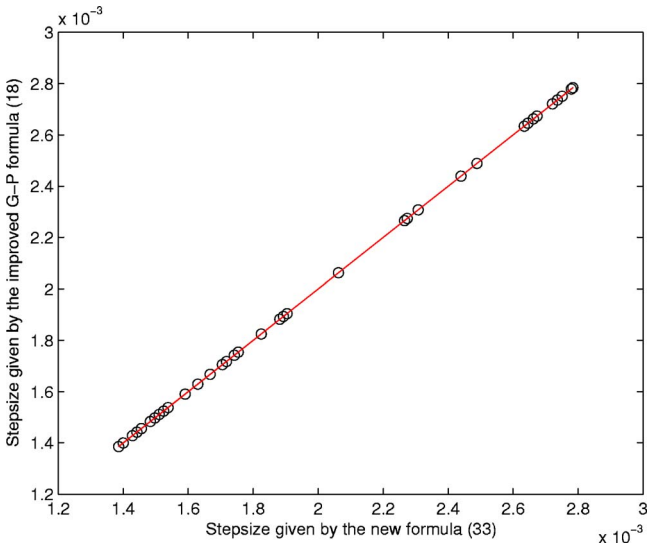

FIG. 9. Plot of the step sizes generated by the improved GP τ-selection formula (18) (y axis) vs the new τ-selection formula (33) (x axis) in a single simulation for the Schlögl model. The step size is plotted at each step. The data are shown as "o" points. A straight line $y=x$ is plotted for comparison. The maximum of the relative difference is 0.01%.

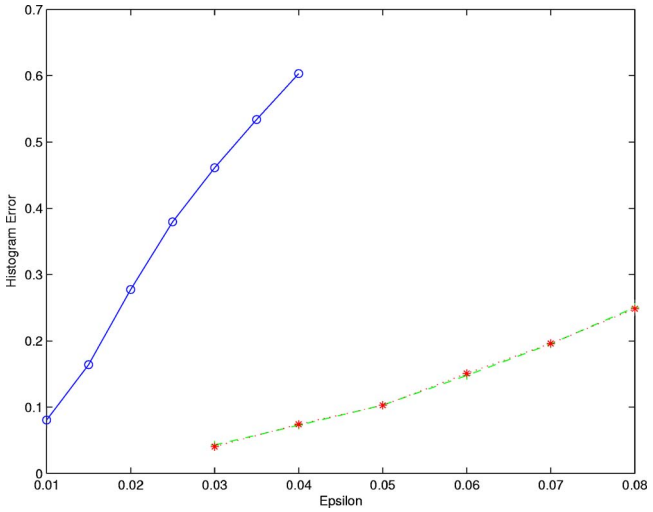

FIG. 10. Plot of the histogram distance errors corresponding to different ϵ values for the three τ-selection formulas for the decaying-dimerizating model. Histogram distance errors are measured by 10^5 samples generated from the SSA method and the tau-leaping method using different τ-selection formulas. The solid line with "o" is for the original GP τ-selection formula (13), the dotted line with " $*$" is for the improved GP τ-selection formula (18), and the dashed line with "+" is for the new τ-selection formula (33).

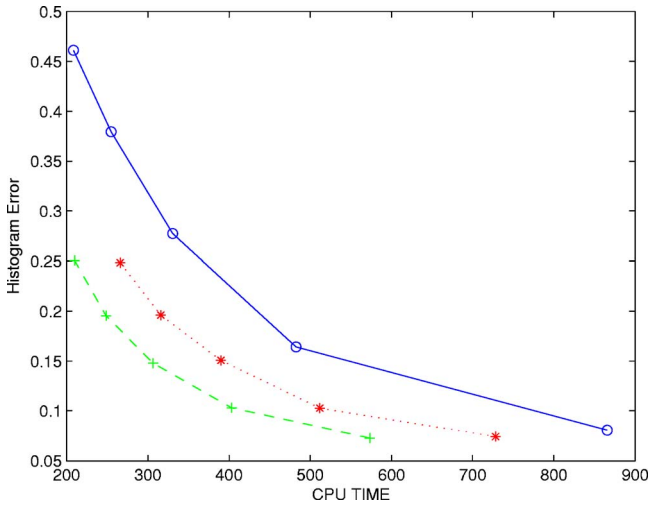

FIG. 11. Plot of the histogram distance errors corresponding to CPU times for the three τ-selection formulas for the decaying-dimerizating model. Histogram distance errors are measured by 10^5 samples generated from the SSA method and the tau-leaping method using different τ-selection formulas and different ϵ values. The solid line with "o" is for the original GP τ-selection formula (13), the dotted line with "$*$" is for the improved GP τ-selection formula (18), and the dashed line with "+" is for the new τ-selection formula (33).

C. Decaying-dimerizing model

This simple model has been used in earlier papers on tau leaping,[8,10] and the reader is referred to those papers for particulars of the model. Suffice it here to say that the model has three time-varying species and four reactions, and one of the reactions is bimolecular with a single reactant species while the other three reactions are unimolecular. The error-versus-ϵ plots for the three τ-selection procedures are shown in Fig. 10, and the error-versus-CPU time plots are shown in Fig. 11. The step size comparison is shown in Fig. 12. The results are very similar to those obtained for the LacZ/LacY model and the Schlögl model: The error increases roughly linearly with ϵ in all cases, but more quickly for the original GP τ-selection formula. The improved GP τ-selection formula (18) yields similar accuracy as the new τ-selection formula (33). The error drops off with CPU time similarly for all three formulas. The new tau-selection formula (33) is just as accurate as, but more efficient than the improved tau-selection formula (18), and both are more accurate and more efficient than the original GP tau-selection formula (13).

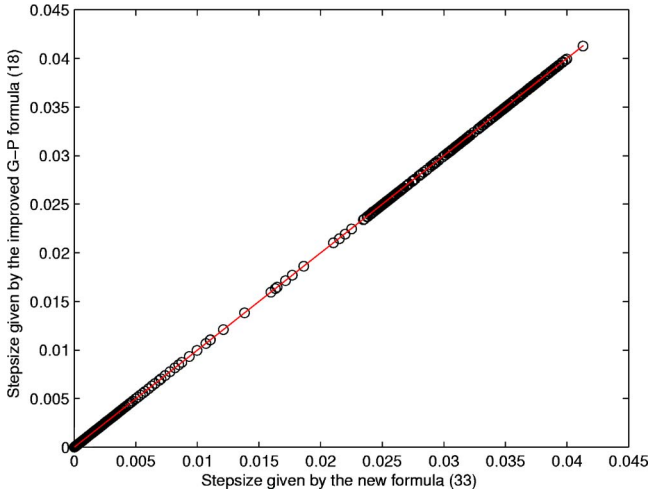

FIG. 12. Plot of the step sizes generated by the improved GP τ-selection formula (18) (y axis) vs the new τ-selection formula (33) (x axis) in a single simulation for the decaying-dimerizating model. The step size is plotted at each step. The data are shown as "o" points. A straight line $y=x$ is plotted for comparison. The maximum of the relative difference is 0.36%.

The new τ-selection formula (33) generates almost identical step sizes as those given by the improved GP τ-selection formula (18) while the new τ-selection formula (33) remains the most efficient.

VI. CONCLUSIONS

The presently used GP τ-selection procedure[10] has two drawbacks: First, it does not enforce a *uniform* bound on the *relative* changes in the propensity functions during a leap. This can lead to simulation inaccuracies and/or computational inefficiencies. Second, the GP τ-selection procedure can be very time consuming to execute. We have proposed here a new τ-selection procedure that mitigates *both* of these problems. It bounds the relative changes in the populations of the reactant species in such a way that the relative changes in all the propensity functions are (approximately) bounded by some prescribed value ϵ. It does this using substantially fewer computational operations than the presently used GP procedure. Test simulations on three different model reaction sets support the conclusion that our new tau-selection procedure is not only faster but also more accurate than the original GP tau-selection procedure.

Our new τ-selection procedure can be incorporated into the modified Poisson tau-leaping algorithm outlined in Sec. II C by making two changes in that algorithm: First, *prior* to step 1, define the functions $g_i(x_i)$ according to the rules set forth in the itemized list following Eq. (27). Second, replace the τ'-selection procedure in step 2 with the τ'-selection procedure that is specified by formulas (32) and (33). Note that in the resulting algorithm, the "accuracy control parameter" ϵ acquires a simple interpretation: It is the approximate upper bound on the *relative* change in *any* propensity function during a leap, allowing for the fact that the minimum nonzero change in any reactant population is 1.

Since the number of computations required to implement the GP τ-selection procedure increases quadratically with the number of reaction channels, whereas the number of computations required to implement the new τ-selection procedure increases linearly with the number of reactant species, the efficiency gain afforded by this new procedure will be most significant for systems with many reaction channels.

ACKNOWLEDGMENTS

This work was supported in part by the California Institute of Technology under DARPA Award No. F30602-01-2-0558, by the U.S. Department of Energy under DOE Award No. DE-FG02-04ER25621, by the National Science Foundation under NSF Award Nos. CCF-0326576 and ACI00-86061, by the Institute for Collaborative Biotechnologies through Grant No. DAAD19-03-D-0004 from the U.S. Army Research Office, and by the Molecular Sciences Institute un-

der Contract No. 244725 with Sandia National Laboratories and the Department of Energy's "Genomes to Life" program.

[1] H. H. McAdams and A. Arkin, Proc. Natl. Acad. Sci. U.S.A. **94**, 814 (1997).
[2] A. Arkin, J. Ross, and H. H. McAdams, Genetics **149**, 1633 (1998).
[3] N. Fedoroff and W. Fontana, Science **297**, 1129 (2002).
[4] D. Gillespie, J. Comput. Phys. **22**, 403 (1976).
[5] D. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[6] M. Gibson and J. Bruck, J. Phys. Chem. A **104**, 1876 (2000).
[7] Y. Cao, H. Li, and L. Petzold, J. Chem. Phys. **121**, 4059 (2004).
[8] D. Gillespie, J. Chem. Phys. **115**, 1716 (2001).
[9] D. Gillespie, in *Handbook of Materials Modeling*, edited by S. Yip (Springer, Dordrecht, 2005), Sec. 5.11, pp. 1735–1752.
[10] D. Gillespie and L. Petzold, J. Chem. Phys. **119**, 8229 (2003).
[11] M. Rathinam, L. Petzold, Y. Cao, and D. Gillespie, J. Chem. Phys. **119**, 12784 (2003).
[12] T. Tian and K. Burrage, J. Chem. Phys. **121**, 10356 (2004).
[13] A. Chatterjee, D. G. Vlachos, and M. A. Katsoulakis, J. Chem. Phys. **122**, 024112 (2005).
[14] Y. Cao, D. Gillespie, and L. Petzold, J. Chem. Phys. **123**, 054104 (2005).
[15] A. M. Kierzek, Bioinformatics **18**, 470 (2002).
[16] D. T. Gillespie, *Markov Processes: An Introduction for Physical Scientists* (Academic, New York, 1992).
[17] Y. Cao and L. Petzold, J. Comput. Phys. **212**, 6 (2006).