Michaelis–Menten speeds up tau-leaping under a wide range of conditions

Sheng Wu,[1,a)] Jin Fu,[1,b)] Yang Cao,[2,c)] and Linda Petzold[1,d)]

[1]*Department of Computer Science, University of California, Santa Barbara, Santa Barbara, California 93106, USA*
[2]*Department of Computer Science, Virginia Tech, Blacksburg, Virginia 24061, USA*

This paper examines the benefits of Michaelis–Menten model reduction techniques in stochastic tau-leaping simulations. Results show that although the conditions for the validity of the reductions for tau-leaping remain the same as those for the stochastic simulation algorithm (SSA), the reductions result in a substantial speed-up for tau-leaping under a different range of conditions than they do for SSA. The reason of this discrepancy is that the time steps for SSA and for tau-leaping are determined by different properties of system dynamics. © *2011 American Institute of Physics.* [doi:10.1063/1.3576123]

I. INTRODUCTION

Biochemical systems typically involve complex networks with many reactions and many molecular species. While studying such systems, reduced models can not only benefit simulation speed but also aid in the understanding of complex models. The Michaelis–Menten (M–M) approximation[1,2] of enzyme–substrate reactions, which replaces the set of three reactions

$$E + S \underset{c_2}{\overset{c_1}{\rightleftharpoons}} ES \overset{c_3}{\rightarrow} E + P, \qquad (1)$$
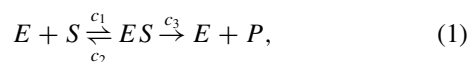
with the single reaction

$$S \overset{c}{\rightarrow} P, \qquad (2)$$

is a widely used model reduction technique for ordinary differential equation (ODE) chemical kinetics models, and has been the subject of refinements over the years.[3–5]

Recent studies have shown that the M–M approximation can also benefit the simulation of stochastic chemical kinetics models. With regard to the use of the M–M approximation in the stochastic simulation algorithm (SSA),[6] Rao and Arkin[7] derived a stochastic M–M approximation from the quasi-steady-state assumption (QSSA); and Mastny *et al.*[8] verified the QSSA using perturbation analysis. Gillespie *et al.*[9] examined the validity of the same abridgment for a similar but simpler reaction set

$$S_1 \underset{c_2}{\overset{c_1}{\rightleftharpoons}} S_2 \overset{c_3}{\rightarrow} S_3, \qquad (3)$$

abridged to

$$S_{12} \overset{c}{\rightarrow} S_3. \qquad (4)$$

They showed that the abridgment is valid for SSA under four sets of conditions, and that for only one of these cases does

it result in a substantial speed-up for SSA. Sanft *et al.*[10] extended the analysis to the M–M approximation and obtained similar results.

However, little attention has been paid to the abridgment when it is used in the context of tau-leaping. The tau-leaping algorithm[11] is an approximate strategy to accelerate SSA simulations. This paper examines the conditions for validity and for speed-up of the abridgment applied to both reaction set (3) and reaction set (1) in the context of tau-leaping, and shows that the abridgment results in a substantial speed-up for tau-leaping under a different range of conditions than is the case for SSA.

II. BACKGROUND

Consider a well-stirred chemical reaction system with n molecular species $S_1, \ldots, S_n$ and m reaction channels $R_1, \ldots, R_m$. Let $x_i(t)$ denote the population of species S_i at time t, and $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^T$ the state vector of the system at time t. Let each reaction R_j be characterized by its propensity function $a_j(\mathbf{x})$ and stoichiometry vector v_j: the probability that one R_j reaction will occur in the next infinitesimal time interval $[t, t + dt)$, given $\mathbf{x}(t) = \mathbf{x}$, is $a_j(\mathbf{x}) dt$, and the change to the system's state vector induced by one R_j reaction is v_j. Then the dynamics of the whole system can be described by the chemical master equation.[12] The SSA is an exact method to numerically solve the chemical master equation by simulating a large number of trajectories of the system.

The tau-leaping algorithm is an approximate method to accelerate SSA. Instead of simulating one reaction at a time, tau-leaping steps the system by a selected time interval, τ, during which many reactions may fire. The idea is that if the propensity functions are nearly constant during the interval, the number of times that each reaction fires can be approximated by a Poisson random number $\mathcal{P}(a_j(\mathbf{x})\tau)$. Then the state of the system can be advanced by the formula

$$\mathbf{x}(t + \tau) \approx \mathbf{x}(t) + \sum_{j=1}^{m} \mathcal{P}(a_j(\mathbf{x})\tau) \, v_j. \qquad (5)$$

a)Electronic mail: sheng@cs.ucsb.edu.
b)Electronic mail: fujin@cs.ucsb.edu.
c)Electronic mail: ycao@cs.vt.edu.
d)Electronic mail: petzold@cs.ucsb.edu.

The requirement that the propensities are nearly constant during the time interval $[t, t + \tau]$ is called the leap condition: for some $\varepsilon \ll 1$,

$$|\Delta_\tau a_j(\mathbf{x})/a_j(\mathbf{x})| \leq \varepsilon, \quad \text{for all } j = 1, \ldots, m. \quad (6)$$

The step size τ is dictated by the need to satisfy the leap condition.

It is easy to see that the tau-selection strategy is crucial to the accuracy and speed of tau-leaping simulations. Several strategies have been proposed. The most widely used strategy for mass action reactions is due to Cao *et al.*[13] In that strategy, the leap condition is written in terms of the changes in species populations rather than the changes in propensities, and the expression for τ is given by

$$\tau = \min_i \left\{ \frac{\max\{\varepsilon x_i/g_i, 1\}}{\left|\sum_j v_{ij} a_j(x)\right|}, \frac{\max\{\varepsilon x_i/g_i, 1\}^2}{\left|\sum_j v_{ij}^2 a_j(x)\right|} \right\}, \quad (7)$$

where $\varepsilon \ll 1$ is the preset accuracy control parameter, v_{ij} are the stoichiometric terms, and g_i is the highest order of reaction in which species S_i appears as a reactant. (For further details, see Ref. [13].)

III. RESULTS

A. Abridgment of reaction set (3): Conditions for validity

Both SSA and tau-leaping are based on the same assumption that the dynamics of the chemical reaction system is governed by the propensity functions as defined above. In the case of the unimolecular reaction R_3 in reaction set (3), given the population x_2 of species S_2, the probability for reaction R_3 to fire in the next infinitesimal time dt is given by the product of the corresponding propensity function a_3 and dt: $a_3 dt = c_3 x_2 dt$. Gillespie *et al.*[9] showed that this condition is mathematically equivalent to the following condition: in the absence of competing reactions, the time to the next firing of reaction R_3 is an exponential random variable with mean $1/(c_3 x_2)$. Therefore, the validity condition of the abridgment is that the time to the next firing of reaction R_3 (the event of a S_3 molecule being generated) must be approximately an exponential random variable with mean $1/a_a(x_{12})$, where $x_{12} = x_1 + x_2$ and $a_a(x_{12})$ is the propensity function of the abridged reaction. Note that a_a is a function of $x_{12} = x_1 + x_2$, rather than a_3 being a function of x_2. Since both SSA and tau-leaping share the same assumption above of how the system dynamics is governed by propensity functions, this validity condition for the abridgment holds for both SSA and tau-leaping. The additional assumption in tau-leaping that the propensity functions are nearly constant during each time step does not affect the validity condition.

For reaction system (3), Gillespie *et al.*[9] showed that the validity condition can be satisfied if and only if

$$\left(\frac{c_1}{c_1 + c_2 + c_3}\right)\left(\frac{c_3}{c_1 + c_2 + c_3}\right) \ll 1, \quad (8)$$

which is satisfied if and only if at least one of the four conditions holds:

$$c_2 \gg c_1, \quad (9a)$$

$$c_3 \gg c_1, \quad (9b)$$

$$c_1 \gg c_3, \quad (9c)$$

$$c_2 \gg c_3, \quad (9d)$$

and the propensity function of the abridged model is given by

$$a_a(x_{12}) = \frac{c_1 c_3 x_{12}}{c_1 + c_2 + c_3}, \quad (10)$$

where $x_{12} = x_1 + x_2$.

Because SSA and tau-leaping share the same validity condition, (8) is also the condition for the validity of the abridgment for tau-leaping, and the propensity function is given by Eq. (10).

B. Abridgment of reaction set (3): Conditions for speed-up

Gillespie *et al.*[9] also showed that under only one condition will the abridgment speed up the SSA simulation significantly: $c_2 \gg c_3$. This is because the SSA simulates every reaction event of the system; thus the speed of an SSA simulation is determined by the number of reaction firings. When $c_2 \gg c_3$, many fewer reaction firings take place in the abridged system than in the original system.

For tau-leaping, since the speed of the simulation is determined by the size of the leap step τ, we must re-examine the conditions for speed-up. The leap condition (7) requires that all propensity functions remain almost constant during a time step. Thus, the time step of tau-leaping is restricted by how fast the propensity functions change and is implicitly restricted by how fast the populations of the reactants change. The key to a significant acceleration of tau-leaping simulation by abridgment is to remove the intermediate highly reactive species with the following properties:

1. Their populations change rapidly, thus the step size of tau-leaping is restricted by those species.
2. The distributions of their populations remain almost constant for a much longer time than their populations do, thus the average values of the corresponding propensity functions over time during this new longer step can be approximated by the expectations of those propensity functions over the distributions.

For the abridgment of reaction set (3) to reaction set (4), the conditions for validity of the abridgment are given in Eqs. (8) or (9). To examine the extent of speed-up, we compare the time step τ_f of the full model (3) with the time step τ_a of the abridged model (4) under different conditions in the Appendix. Several interesting conclusions can be summarized from the analysis:

1. The abridgment always speeds up tau-leaping, as shown in (A9) and (A18).
2. The conditions corresponding to a substantial speed-up from the abridgment for both tau-leaping and SSA are compared in Table I. The condition $\varepsilon x_{12} \gg 1$ is assumed in all the conditions for tau-leaping, as this

TABLE I. Conditions for a substantial speed-up from the abridgment, for tau-leaping and SSA, applied to reaction set (3).

		$c_2 \gg c_3$			
	$c_2 \gg c_1$	$c_3 \gg c_1$	$c_1 \gg c_3$		$c_2 \gg c_3$
Tau-leaping	$x_{12} \ll \dfrac{2(c_2+c_3)^3}{\varepsilon c_1^2 c_3}$		$\begin{cases} x_{12} \gg \frac{1}{\varepsilon}, & \text{if } c_2 \gg c_1 \\ \frac{c_3}{2\varepsilon c_2} \ll x_{12} \ll \frac{2c_1^2}{\varepsilon c_2 c_3}, & \text{if } c_2 \ll c_3 \\ x_{12} \ll \frac{2c_1^2}{\varepsilon c_2 c_3}, & \text{otherwise} \end{cases}$	$c_2 \gg c_1$ or $x_{12} \ll \dfrac{2c_1^2}{\varepsilon c_2 c_3}$	
SSA			$\emptyset$		$x_{12} > 0$

is the situation where tau-leaping is advantageous over SSA. It is easy to see that for most situations (except for when $c_2 \gg c_3$), tau-leaping simulation benefits from the abridgment under a wider range of conditions than is the case for SSA simulation.

C. Abridgment of enzyme-substrate system

For the enzyme–substrate system (1), Sanft *et al.*[10] showed that the condition for validity of the stochastic M–M approximation is the same as that of the deterministic case,[3] namely,

$$E_T \ll S_0 + K_m, \tag{11}$$

where $E_T = E(t) + ES(t)$, $S_0 = S(0)$, and $K_m = (c_2 + c_3)/c_1$. The propensity function of the abridged system is also the same as the deterministic M–M rate:

$$a(\mathbf{x}) \approx \frac{V_{\max} S}{K_m + S}, \tag{12}$$

where $V_{\max} = c_3 E_T$. And, under only one condition will the abridgment speed up the SSA simulation significantly: $c_2 \gg c_3$.

Applying the same arguments in Sec. III A, the condition for validity and the stochastic M–M rate are the same for tau-leaping as for SSA.

On the other hand, for tau-leaping, the abridged system is no longer a mass-action system. However, it is easy to see that the propensity function (12) depends on the population of S, and

$$\frac{\Delta_\tau a}{a} = \frac{K_m \Delta_\tau S}{(K_m + S - \Delta_\tau S)S} < \frac{\Delta_\tau S}{S}. \tag{13}$$

Thus the τ selection strategy in Eq. (7) with $g = 1$ can be used to calculate the tau-leaping step size of the abridged system:

$$\tau_a = \frac{\varepsilon S}{V_{\max} S/(K_m + S)} = \frac{\varepsilon(K_m + S)}{V_{\max}}. \tag{14}$$

To calculate the characteristic tau-leaping step size of the original system, we can make use of the corresponding ODE results as estimates of the expected populations:

$$\overline{E} \approx \frac{E_T K_m}{K_m + S}, \tag{15a}$$

$$\overline{ES} \approx \frac{E_T S}{K_m + S}. \tag{15b}$$

Applying the τ selection strategy (7), we can easily see that $\tau_f < \tau_a$ by examining the contribution to τ of the change in S. Then by examining the contribution to τ of the changes in E and ES, we have

$$\tau_f < \frac{\max\left\{\varepsilon \min\left\{\overline{E}, \overline{ES}\right\}, 1\right\}^2}{c_1 S\overline{E} + (c_2 + c_3)\overline{ES}}. \tag{16}$$

When $\varepsilon \overline{E} \leq 1$ or $\varepsilon \overline{ES} \leq 1$, which is equivalent to

$$\min\{K_m, S\} \leq \frac{K_m + S}{\varepsilon E_T}, \tag{17}$$

we obtain

$$\tau_f < \frac{1}{c_1 S\overline{E} + (c_2 + c_3)\overline{ES}} = \frac{K_m + S}{2(c_2 + c_3)E_T S}, \tag{18}$$

$$\frac{\tau_f}{\tau_a} < \frac{c_3}{2(c_2 + c_3)\varepsilon S}. \tag{19}$$

Again, we are only interested in the case $\varepsilon S \gg 1$, where tau-leaping is advantageous over SSA. In this case, $\tau_f \ll \tau_a$, and the speed gain grows as c_2/c_3 increases. Note that since $K_m + S \gg E_T$ and $\varepsilon \ll 1$, condition (17) is actually a very loose condition.

When $\varepsilon \overline{E} > 1$ and $\varepsilon \overline{ES} > 1$, which is equivalent to

$$\min\{K_m, S\} > \frac{K_m + S}{\varepsilon E_T}, \tag{20}$$

TABLE II. Conditions for a substantial speed-up from the M–M abridgment, for tau-leaping and SSA, applied to the enzyme–substrate system (1).

	$c_2 \gg c_3$	$c_2 \gg c_3$
Tau-leaping	$S \gg \varepsilon E_T^2$ or $\left(\dfrac{K_m + S}{\min\{K_m, S\}}\right)^2 \gg \dfrac{\varepsilon E_T^2}{S}$	$S \ll \varepsilon E_T^2$ or $\left(\dfrac{K_m + S}{\min\{K_m, S\}}\right)^2 \gg \dfrac{c_3 \varepsilon E_T^2}{c_2 S}$
SSA	$\emptyset$	$S > 0$

TABLE III. Simulation time and speed-up of SSA and tau-leaping under different initial conditions, for parameter set I, applied to reaction set (3).

$x_1(0)$	t_{end}	SSA			Tau-leaping		
		Original(s)	Abridged(s)	Speed-up	Original (s)	Abridged (s)	Speed-up
1×10^5	1	58	9	6	50	1	50
1×10^6	1	650	110	6	5	1	5
1×10^7	1	5800	900	6	3	1	3
1×10^6	10	1660	290	6	220	2	110

we obtain

$$\tau_f < \frac{\varepsilon^2 E_T \min\{K_m, S\}^2}{2(c_2 + c_3)S(K_m + S)}, \quad (21)$$

$$\frac{\tau_f}{\tau_a} < \frac{\varepsilon E_T^2 c_3 \min\{K_m, S\}^2}{2(c_2 + c_3)S(K_m + S)^2}$$

$$= \frac{c_3}{2(c_2 + c_3)} \frac{1}{\varepsilon S} \left(\frac{\varepsilon E_T \min\{K_m, S\}}{K_m + S} \right)^2. \quad (22)$$

The condition for the abridgment to gain a substantial speed-up in this case is given by

$$\frac{c_3}{2(c_2 + c_3)} \frac{1}{\varepsilon S} \ll \left(\frac{K_m + S}{\varepsilon E_T \min\{K_m, S\}} \right)^2. \quad (23)$$

Either $S \gg \varepsilon E_T^2$, or $c_2 \gg c_3$ and $S \ll \varepsilon E_T^2$, or if the values of K_m and S are widely separated will there be a large speed-up.

The conditions corresponding to a substantial speed-up from the M–M abridgment for both tau-leaping and SSA are compared in Table II. It is easily seen that the M–M abridgment yields a significant speed-up for tau-leaping under a different range of conditions than it does for SSA, for the enzyme–substrate system.

IV. NUMERICAL EXAMPLES

Based on the analysis in Sec. III, we tested and timed the abridgment to different reaction systems under different conditions using the adaptive tau-leaping code in STOCHKIT 2.0.[14] The tests were carried on a personal computer with an Intel Core 2 Quad Q9300 2.5 GHz central processing unit and 4 GB RAM.

A. Speed-up of reaction set (3)

For the abridgment of reaction set (3), we choose two sets of parameters to show the different speed-up behavior of SSA and tau-leaping. Parameter set I was chosen as follows: $c_1 = 1$, $c_2 = 100$, $c_3 = 100$. This corresponds to the condition $c_2 + c_3 \gg c_1$. To show how condition (A16) factors into the speed-up, we set $x_2(0) = x_3(0) = 0$, and varied $x_1(0) = x_{12}(0)$. The simulation time and speed-up of SSA and tau-leaping under different initial conditions or to the different system end time t_{end}, all with 10 000 runs, is shown in Table III. ε was set to 0.01 in the τ selection. We can see from the table that: (a) the speed-up gain of SSA from the abridgment does not change as the initial condition changes.

This agrees with the results of Gillispie *et al.*[9] because the speed-up gain of SSA is determined solely by the c_2/c_3 ratio; (b) tau-leaping gains a considerable speed-up when condition (A16) is satisfied; and (c) the speed-up of tau-leaping is not large when condition (A16) is not satisfied, but if t_{end} is large enough the system will go into the region of condition (A16) eventually and tau-leaping will benefit substantially from the abridgment. All of these results agree with the analysis.

Parameter set II was chosen so that $c_1 \gg c_3 \gg c_2$ to illustrate the only case for which tau-leaping would not gain a substantial speed-up, regardless of the system end time t_{end}. The speed-up of tau-leaping under different initial conditions or to different t_{end} is shown in Table IV. The other parameters and values are set to: $c_1 = 100$, $c_2 = 0.01$, $c_3 = 1$, $x_2(0) = x_3(0) = 0$, $\varepsilon = 0.01$. Noting that $c_3/(2\varepsilon c_2) = 5000$ and $2c_1^2/(\varepsilon c_2 c_3) = 2 \times 10^8$, we can see that the results are consistent with condition (A26): tau-leaping will benefit substantially from the abridgment when $c_3/(2\varepsilon c_2) \ll x_{12} \ll 2c_1^2/(\varepsilon c_2 c_3)$. A quick test of SSA shows that the speed-up for SSA under this set of parameters is between 3 and 4. Tau-leaping benefits from the abridgment more significantly than SSA when condition (A26) is satisfied.

B. Speed-up of enzyme–substrate system

For the abridgment of enzyme–substrate system (1), we chose three sets of parameters and initial conditions to show the different speed-up behavior of SSA and tau-leaping:

1. $c_1 = 1$, $c_2 = 10$, $c_3 = 10$, $S(0) = 1 \times 10^5$, $E(0) = 100$, $ES(0) = P(0) = 0$, $t_{end} = 10$. This set corresponds to the condition $c_2 \gg c_3$ and $S \gg \varepsilon E_T^2$. According to the analysis, only tau-leaping should be able to gain a substantial speed-up from the abridgment;

2. $c_1 = 1$, $c_2 = 100$, $c_3 = 1$, $S(0) = 1 \times 10^5$, $E(0) = 100$, $ES(0) = P(0) = 0$, $t_{end} = 10$. This set corresponds to the condition $c_2 \gg c_3$ and $S \gg \varepsilon E_T^2$. Un-

TABLE IV. Speed-up of tau-leaping (abridged system vs original system) for different initial conditions or system end times, for parameter set II, applied to reaction set (3).

$x_1(0)$	1×10^4	1×10^6	1×10^8
$t_{end} = 1$	4	20	3
$t_{end} = 10$	3	17	34

TABLE V. Simulation time and speed-up of SSA and tau-leaping under different parameters and initial conditions, applied to enzyme–substrate system (1).

Parameters set	SSA			Tau-leaping		
	Original (s)	Abridged (s)	Speed-up	Original (s)	Abridged (s)	Speed-up
I	28	8	4	39	0.08	488
II	139	0.8	174	221	0.02	1105
III	695	3	232	0.03	0.02	2

der this condition, both SSA and tau-leaping should be able to gain a substantial speed-up; and

3. $c_1 = 1$, $c_2 = 1 \times 10^{10}$, $c_3 = 1 \times 10^8$, $S(0) = 1 \times 10^{10}$, $E(0) = 1 \times 10^8$, $ES(0) = P(0) = 0$, $t_{end} = 1 \times 10^{-12}$. This set corresponds to the condition that $c_2 \gg c_3$ holds but neither Eq. (17) nor Eq. (23) holds. The analysis suggests that only SSA should be able to gain a substantial speed-up from the abridgment.

The simulation time and speed-up of SSA and tau-leaping under different parameters and initial conditions, all with 10 000 runs, are shown in Table V. It is easy to see that the results agree with the analysis.

C. Accuracy of the M–M abridgment

Next, we tested the accuracy of the M–M abridgment for tau-leaping. We chose an enyzme substrate reaction set with the following set of parameters: $c_1 = 1, c_2 = 10, c_3 = 10$. Initial conditions were set to $S(0) = 10^5$, $E(0) = 100$, $ES(0) = P(0) = 0$. The system end time was set to 10. Thus, the validity condition for the M–M approximation (11) will be satisfied all the time. The baseline is the histogram of product P at the system end time of 10 000 SSA simulations of the original enzyme substrate model. 10 000 tau-leaping simulations were performed for both the original model and the abridged M–M model. The comparison of the histograms of the product P for both SSA simulation and tau-leaping simulation of the original model is shown in Fig. 1, while the comparison of the histograms of P for SSA simulation and tau-leaping simulation of the abridged model is shown in Fig. 2. The Euclidian distance and Manhattan distance in the figures are respectively L^2 norm and L^1 norm of the histogram distance:[15] suppose X and Y are two groups of samples with N samples in X and M samples in Y, and all the sample values are bounded in the interval $I = [x_{min}, x_{max})$. Let $L = x_{max} - x_{min}$. Divide the interval I into K subintervals
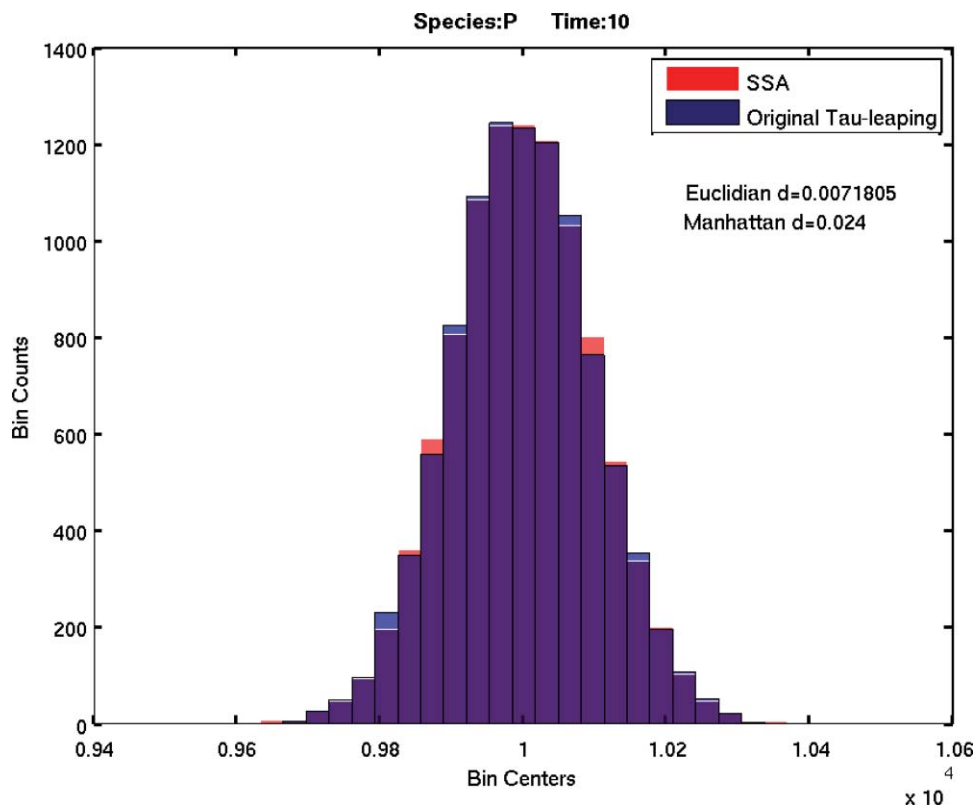
FIG. 1. The comparison of histograms of product P at the system end time of SSA simulation and tau-leaping simulation of the original model.
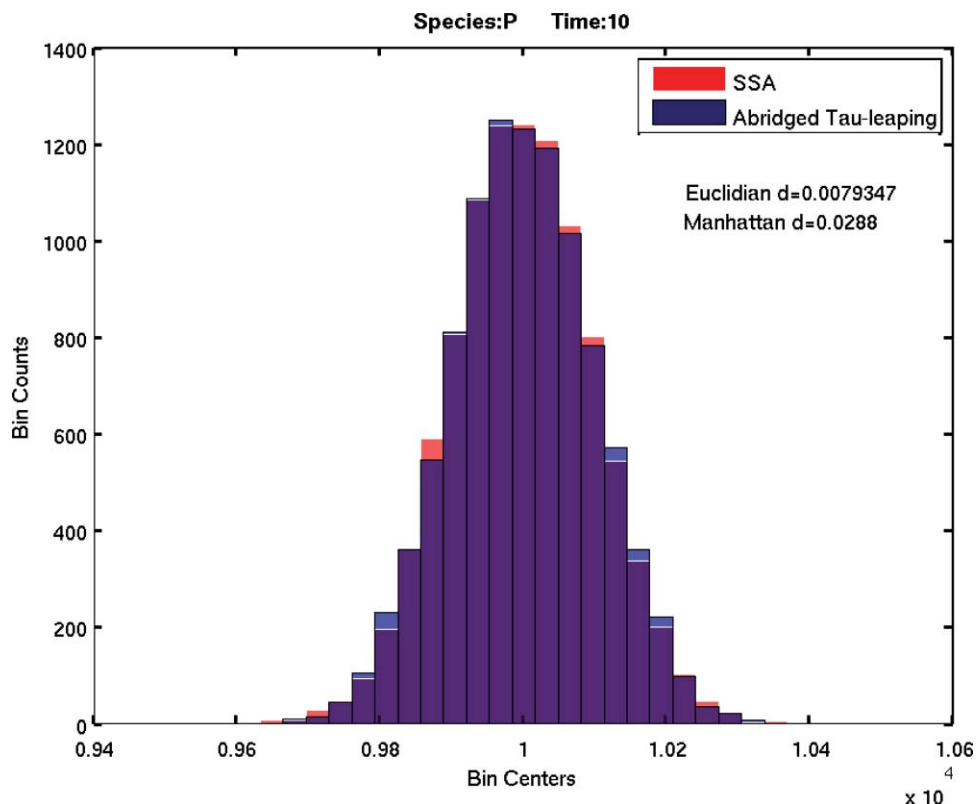
FIG. 2. The comparison of histograms of product P at the system end time of SSA simulation and tau-leaping simulation of the abridged model.

$I_i = [x_{\min} + (i-1)L/K, x_{\min} + iL/K)$. Then the histogram distance is given by

$$D_K(X, Y) = \sum_{i=1}^{K} \left| \frac{\sum_{j=1}^{N} \chi(x_j, I_i)}{N} - \frac{\sum_{j=1}^{M} \chi(y_j, I_i)}{M} \right|,$$

$$(24)$$

where the characteristic function $\chi(x, I_i)$ is defined as

$$\chi(x, I_i) = \begin{cases} 1, & \text{if } x \in I_i, \\ 0, & \text{otherwise.} \end{cases} \quad (25)$$

The histogram distance results show that the M–M abridgment is accurate and valid under this condition. Different conditions are also tested and all of them show that the M–M abridgment is valid when the validity condition (11) is satisfied.

D. Circadian oscillation model

Next, we consider a more complicated model: a circadian oscillation model in the Drosophila period protein (PER).[16] The scheme of the model is shown in Fig. 3. In this model, per mRNA (M) is synthesized in the nucleus and transfers to the cytosol at a maximum rate v_s. It is also degraded by an enzyme there with a maximum rate V_m and
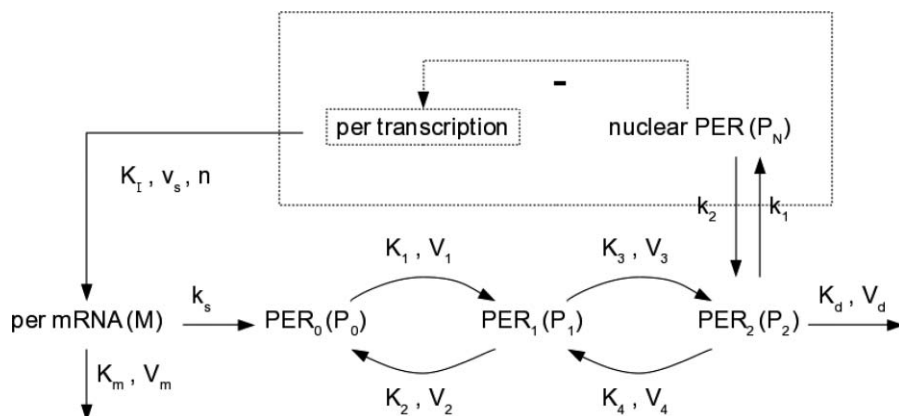

FIG. 3. Scheme of the model for circadian oscillations in PER. Six out of the ten reactions are Michaelis–Menten reactions.

TABLE VI. Simulation time and speed-up of SSA and tau-leaping with different c_2/c_3 ratios, applied to the circadian oscillation model.

c_2/c_3	SSA			Tau-leaping		
	Original (s)	Abridged (s)	Speed-up	Original (s)	Abridged (s)	Speed-up
0.01	1042	674	2	1271	3	424
1	1701	674	3	2103	3	701
100	78 200	674	116	83 800	3	2.8×10^4

Michaelis constant K_m. The PER protein (PER_0) is synthesized at a rate proportional to M by a first-order rate constant k_s. The reversible phosphorylations of the PER proteins, between P_0 and P_1 and between P_1 and P_2, are governed by a set of Michaelis–Menten reactions at maximum rates V_i, with Michaelis constants K_i ($i = 1, 2, 3, 4$). The bisphosphorylated form P_2 is degraded by an enzyme with a maximum rate V_d and Michaelis constant K_d, and transported into the nucleus with a first-order rate constant k_1. The nuclear bisphosphorylated form of PER P_N is transported into the cytosol with a first-order rate constant k_2, while it also exerts a negative feedback on per transcription described by a Hill equation with repression threshold constant K_I and Hill coefficient n. The detailed kinetic laws, parameter values, and initial conditions can be found in Goldbeter.[16]

To compare the results with the original M–M model to a comparable mass-action model, we converted this 5 species, 10 reaction model into a full 17 species, 22 reaction stochastic model by replacing all the 6 Michaelis–Menten reactions with enzyme–substrate reactions. The system volume was chosen to be the characteristic size of a cell nucleus, 1000 μm^3. Then we varied the ratios of backward dissociation rates [c_2 in system (1)] to forward dissociation rates [c_3 in system (1)] of the enzyme–substrate compounds, to see the different speed-up behavior of the Michaelis–Menten abridgment applied to SSA and tau-leaping. To simplify the problem, we took all of the six ratios (corresponding to c_2/c_3) to be the same. The simulation time and speed-up of SSA and tau-leaping with different c_2/c_3 ratios, all with 1000 runs and $t_{end} = 1$, is shown in Table VI. The result shows that for this model, SSA will gain a large speed-up from the M–M abridgment only when $c_2 \gg c_3$ is satisfied, while tau-leaping can benefit substantially from the M–M abridgment under a wider range of conditions.

APPENDIX: ANALYSIS OF SPEED-UP OF THE ABRIDGMENT OF REACTION SET (3)

For this analysis, we assume that $\varepsilon x_{12} \gg 1$, as this is the situation where tau-leaping is advantageous over SSA.

For the abridged model (4), applying the τ selection strategy (7), we get

$$\tau_a = \frac{\varepsilon x_{12}}{c_1 c_3 x_{12}/(c_1 + c_2 + c_3)} = \frac{\varepsilon(c_1 + c_2 + c_3)}{c_1 c_3}. \quad (A1)$$

On the other hand, the step size τ_f of the original model (3) is restricted by the minimum of

$$\tau_{11} = \frac{\max\{\varepsilon x_1, 1\}}{|c_1 x_1 - c_2 x_2|}, \quad (A2a)$$

$$\tau_{12} = \frac{\max\{\varepsilon x_1, 1\}^2}{|c_1 x_1 + c_2 x_2|}, \quad (A2b)$$

$$\tau_{21} = \frac{\max\{\varepsilon x_2, 1\}}{|c_1 x_1 - (c_2 + c_3)x_2|}, \quad (A2c)$$

$$\tau_{22} = \frac{\max\{\varepsilon x_2, 1\}^2}{|c_1 x_1 + (c_2 + c_3)x_2|}. \quad (A2d)$$

Since τ_f varies with x_1 and x_2, the expectations of x_1 and x_2 are used to calculate the characteristic time step of the original system. Because it is a linear system, the expectations of x_1 and x_2 are the same as the results of the corresponding ODE solutions. Let $\overline{x_1}$ and $\overline{x_2}$ denote the expectation of x_1 and x_2, and $x' = dx/dt$. Then

$$\overline{x_1}' = -c_1\overline{x_1} + c_2\overline{x_2}, \quad (A3a)$$

$$\overline{x_2}' = c_1\overline{x_1} - (c_2 + c_3)\overline{x_2}. \quad (A3b)$$

To evaluate the τ values in (A2), we are concerned with the ratio r between $\overline{x_2}$ and $\overline{x_1}$:

$$\overline{x_2} = r\overline{x_1}. \quad (A4)$$

Substituting (A4) into (A3), we have

$$r' = -c_2 r^2 + (c_1 - c_2 - c_3)r + c_1. \quad (A5)$$

Applying the QSSA ($r' = 0$) and requiring that $r > 0$, we obtain

$$r = \frac{(c_1 - c_2 - c_3) + \sqrt{(c_1 - c_2 - c_3)^2 + 4c_1 c_2}}{2c_2}. \quad (A6)$$

Now consider the different conditions under which the abridgment is valid:

1. Conditions (9a) or (9b): If either of these conditions holds, then $c_2 + c_3 \gg c_1$, hence

$$r = \frac{c_1 - c_2 - c_3}{2c_2} \left(1 - \sqrt{1 + \frac{4c_1 c_2}{(c_1 - c_2 - c_3)^2}}\right)$$

$$\approx \frac{c_1}{c_2 + c_3} \ll 1. \tag{A7}$$

In this case,

$$x_1 \approx \frac{c_2 + c_3}{c_1 + c_2 + c_3} x_{12} \gg x_2 \approx \frac{c_1}{c_1 + c_2 + c_3} x_{12}, \tag{A8}$$

and $\varepsilon x_1 \approx \varepsilon x_{12} \gg 1$ as assumed. Thus the τ_f is at most

$$\tau_{11} \approx \frac{\varepsilon(c_2 + c_3)}{c_1 c_3} < \tau_a. \tag{A9}$$

Further calculation reveals that $\tau_{21} = \infty$ and $\tau_{12} > \tau_{22}$. Thus $\tau_f = \min\{\tau_{11}, \tau_{22}\}$.
When $\varepsilon x_2 \leq 1$, we have

$$\tau_{22} \approx \frac{1}{2c_1 x_1} \approx \frac{c_1 + c_2 + c_3}{2c_1(c_2 + c_3)x_{12}}, \tag{A10}$$

thus (recalling that $\varepsilon x_{12} \gg 1$)

$$\frac{\tau_{22}}{\tau_a} \approx \frac{c_3}{2\varepsilon x_{12}(c_2 + c_3)} \ll 1. \tag{A11}$$

In this case a significant speed up is ensured. Condition $\varepsilon x_2 \leq 1$ is equivalent to

$$x_{12} \leq \frac{c_1 + c_2 + c_3}{\varepsilon c_1}. \tag{A12}$$

When $\varepsilon x_2 > 1$, we have

$$\tau_{22} \approx \frac{\varepsilon^2 x_{12} c_1}{2(c_2 + c_3)(c_1 + c_2 + c_3)}, \tag{A13}$$

$$\frac{\tau_{22}}{\tau_a} \approx \frac{\varepsilon x_{12} c_1^2 c_3}{2(c_2 + c_3)(c_1 + c_2 + c_3)^2} \approx \frac{\varepsilon x_{12} c_1^2 c_3}{2(c_2 + c_3)^3}. \tag{A14}$$

To yield a substantial speed-up in this case, the following condition must be satisfied:

$$\frac{c_1 + c_2 + c_3}{\varepsilon c_1} < x_{12} \ll \frac{2(c_2 + c_3)^3}{\varepsilon c_1^2 c_3}. \tag{A15}$$

As a consequence, when $c_2 + c_3 \gg c_1$, the condition for the abridgment to gain a significant speed up in tau-leaping is (A12) or (A15):

$$x_{12} \ll \frac{2(c_2 + c_3)^3}{\varepsilon c_1^2 c_3}. \tag{A16}$$

Since $c_2 + c_3 \gg c_1$ and $\varepsilon \ll 1$, this is actually a very loose condition. If the end time of a simulation is long enough, the simulation will always gain a large speed up when there are enough S_1 and S_2 consumed

and the system goes into the region of condition (A16).

2. Conditions (9c) or (9d): If either of these conditions holds, then $c_1 + c_2 \gg c_3$, hence

$$r = \frac{c_1 - c_2 - c_3}{2c_2} + \frac{c_1 + c_2 + c_3}{2c_2}$$

$$\times \sqrt{1 - \frac{4c_2 c_3}{(c_1 + c_2 + c_3)^2}} \approx \frac{c_1}{c_2}. \tag{A17}$$

In this case, $x_1 \approx c_2 x_{12}/(c_1 + c_2)$, $x_2 \approx c_1 x_{12}/(c_1 + c_2)$, and the τ_f is at most

$$\tau_{21} \approx \frac{\varepsilon}{c_3} < \tau_a. \tag{A18}$$

Further calculation reveals that $\tau_{11} = \infty$. Thus $\tau_f = \min\{\tau_{21}, \tau_{12}, \tau_{22}\}$.
If $c_2 \gg c_1$, then $\tau_{21} \ll \tau_a$, thus $\tau_f < \tau_{12} \ll \tau_a$.
If $c_2 \sim c_1$, then $\overline{x_2} \sim \overline{x_1}$, both $\varepsilon x_1 \gg 1$ and $\varepsilon x_2 \gg 1$ hold as assumed, thus

$$\tau_{22} \sim \tau_{12} \approx \frac{\varepsilon^2 x_{12} c_2}{2c_1(c_1 + c_2)} \sim \frac{\varepsilon^2 x_{12}}{4c_1}$$

$$= \frac{\varepsilon x_{12} c_3}{4(c_1 + c_2 + c_3)} \tau_a. \tag{A19}$$

The condition to yield a substantial speed gain in this case is

$$x_{12} \ll \frac{4(c_1 + c_2 + c_3)}{\varepsilon c_3}; \tag{A20}$$

If $c_2 \ll c_1$, then $\overline{x_2} \gg \overline{x_1}$,

$$\tau_{22} \approx \frac{\varepsilon^2 x_{12}}{2c_2 + c_3}, \tag{A21}$$

$$\tau_{12} \approx \max\left\{\frac{\varepsilon^2 x_{12} c_2}{2c_1(c_1 + c_2)}, \frac{c_1 + c_2}{2c_1 c_2 x_{12}}\right\}, \tag{A22}$$

whence

$$\frac{\tau_{22}}{\tau_a} \approx \frac{\varepsilon x_{12} c_3}{2c_2 + c_3}, \tag{A23}$$

$$\frac{\tau_{12}}{\tau_a} \approx \max\left\{\frac{\varepsilon x_{12} c_2 c_3}{2c_1^2}, \frac{c_3}{2\varepsilon x_{12} c_2}\right\}. \tag{A24}$$

To yield a major speed-up, at least one of (A23) or (A24) should be much less than 1, which is equivalent to

$$\begin{cases} x_{12} \ll \dfrac{2c_1^2}{\varepsilon c_2 c_3}, & \text{if } c_2 \gg c_3 \text{ or } c_2 \sim c_3; \\[2mm] \dfrac{c_3}{2\varepsilon c_2} \ll x_{12} \ll \dfrac{2c_1^2}{\varepsilon c_2 c_3}, & \text{if } c_2 \ll c_3. \end{cases} \tag{A25}$$

Consequently, when $c_1 + c_2 \gg c_3$, the condition of a substantial speed up in tau-leaping is

$$
\begin{cases}
x_{12} \gg \dfrac{1}{\varepsilon}, & \text{if } c_2 \gg c_1; \\[2mm]
\dfrac{c_3}{2\varepsilon c_2} \ll x_{12} \ll \dfrac{2c_1^2}{\varepsilon c_2 c_3}, & \text{if } c_2 \ll c_3; \\[2mm]
x_{12} \ll \dfrac{2c_1^2}{\varepsilon c_2 c_3}, & \text{otherwise.}
\end{cases} \qquad \text{(A26)}
$$

Notice $c_1 + c_2 \gg c_3$, this is also a very loose condition. If the end time of a simulation is long enough, the simulation can always gain a large speed up except for the case $c_2 \ll c_3$ and $x_{12} \gg c_3/(2\varepsilon c_2)$.

[1] L. Michaelis and M. L. Menten, Biochem. Z. **49**, 333 (1913).
[2] D. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 4th ed. (Freeman, New York, 2005).
[3] L. A. Segel and M. Slemrod, SIAM Rev. **31**, 446 (1989).
[4] A. R. Tzafriri and E. R. Edelman, J. Theor. Biol. **226**, 303 (2004).
[5] M. G. Pedersen, A. M. Bersani, and E. Bersani, Bull. Math. Biol. **69**, 433 (2005).
[6] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[7] C. Rao and A. P. Arkin, J. Chem. Phys. **118**, 4999 (2003).
[8] E. A. Mastny, E. L. Haseltine, and J. B. Rawlings, J. Chem. Phys. **127**, 094106 (2007).
[9] D. T. Gillespie, Y. Cao, K. R. Sanft, and L. R. Petzold, J. Chem. Phys. **130**, 064103 (2009).
[10] K. R. Sanft, D. T. Gillespie, and L. R. Petzold, IET Syst. Biol. **5**, 58 (2011).
[11] D. T. Gillespie, J. Chem. Phys. **115**, 1716 (2001).
[12] D. T. Gillespie, Annu. Rev. Phys. Chem. **58**, 35 (2007).
[13] Y. Cao, D. T. Gillespie, and L. R. Petzold, J. Chem. Phys. **124**, 044109 (2006).
[14] See http://engineering.ucsb.edu/~cse/StochKit for details of the STOCHKIT 2.0 software.
[15] Y. Cao and L. R. Petzold, J. Comp. Phys. **212**, 6 (2006).
[16] A. Goldbeter, Proc. Biol. Sci. **261**, 319 (1995).