Reduced Stochastic Simulation Times

James P. Keener and Kyle Gaffney
Department of Mathematics,
University of Utah, Salt Lake City, UT 84112 USA

April 19, 2012

Abstract We present an analysis of Stochastic simulation (via the Gillespie algorithm) that shows how to use time scale differences to reduce the complexity and computational cost of stochastic simulation. We apply this analysis to several important chemical reactions, including the Michaelis Menten enzyme kinetics reaction and find an algorithm that improves on previously published reduction algorithms.

1 Introduction

The Gillespie algorithm for simulation of chemical reaction networks **refs** has become the gold standard for stochastic simulation. We describe the algorithm in more detail below. What has become apparent is that in large chemical reaction networks there may be "loops", i.e., a collection of states among which there is a lot of cycling, with the result that the much of the computational time is devoted to this cycling. An important challenge to developing efficient computational schemes is to determine how to correctly reduce the system so that the computation in these loops is minimized or eliminated.

The classical example of this problem is the well-known Michaelis-Menten enzymatic reaction

$$S + E \underset{k_2}{\overset{k_1}{\rightleftarrows}} C \xrightarrow{k_3} P + E \tag{1.1}$$

where S is substrate, P is product, E is enzyme and C is the enzyme-substrate complex. If $k_2 >> k_3$ then most of the time that enzyme combines with substrate to form complex C, the complex degrades back to enzyme and substrate rather than forming product. If this is the case then there are many cycles of complex formation and degradation before a product molecule is produced, and a simulation that computes all these reactions is performing a lot of "unnecessary" reactions.

This problem with the Gillespie algorithm is well known, and a number of previous works have described reduction schemes.**add ref's** The purpose of this paper is to provide a new formal analysis of this system reduction procedure. We find two reductions of the Gillespie algorithm. The first reduction results in a system whose expected travel times are identical to those of the original system, although the travel time distributions are not the same. The second reduction is a "fast-slow" reduction, closely related to those previously reported (add refs), but here we give a formal mathematical analysis of the validity of this reduction.

2 The Setting

Finite-state, continuous-time jump processes are a class of Markov processes that arise in a variety of applications, especially as stochastic models of chemical reactions [9, 1] and ion channel gating kinetics [7, 6]. Such processes describe the stochastic evolution of a state variable $S : \mathbb{R}_+ \to \mathcal{S}$, where $\mathcal{S} = \{s_1, \ldots, s_n\}$ is a finite state space and the *transition probabilities*

$$p(s_i, t|s_j, s) = \text{Prob}\{X(t) = s_i \mid X(s) = s_j\} \quad (t \geq s \geq 0) \tag{2.1}$$

satisfy the *Chapman-Kolmogorov equation*

$$p(s_i, t|s_j, s) = \sum_{k=1}^{n} p(s_i, t|s_k, u)p(s_k, u|s_j, s) \quad (t \geq u \geq s). \tag{2.2}$$

For a time-autonomous process, the transition probabilities (2.1) are of the form

$$p(s_i, t + \Delta t|s_j, t) = \alpha_{i,j}\Delta t + o(\Delta t), \tag{2.3}$$

for $i \neq j$ where each *transition rate* $\alpha_{i,j}$ is nonnegative. Of necessity,

$$p(s_i, t + \Delta t|s_i, t) = 1 - \sum_{i} \alpha_{i,j}\Delta t + o(\Delta t). \tag{2.4}$$

It is an immediate consequence of this that the probability distribution $\mathbf{p}(t)$ of $X(t)$ satisfies the linear ordinary differential equation, called the *master equation* [2, 3, 9],

$$\frac{d\mathbf{p}}{dt} = W\mathbf{p}. \tag{2.5}$$

Here W is the $n \times n$ matrix whose off-diagonal entries are the transition rates $\alpha_{i,j}$ of (2.3) and whose diagonal entries satisfy

$$\alpha_{i,i}(t) = -\sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_{j,i}(t).$$

Shortened notation for this is

$$W = A - D, \tag{2.6}$$

where $D = \Sigma_\alpha = \text{diag}(\mathbf{1}^T A) = \text{diag}(\sum_{i \neq j} \alpha_{i,j})$. Notice that $\mathbf{1}^T \tilde{W} = -\beta^T$. That is, the *transition rate matrix* W is a $\mathbb{W}$-matrix as defined by van Kampen [9]. As a consequence, the sum of each column of W is zero; i.e.,

$$\mathbf{1}^T W = \mathbf{0}^T,$$

where $\mathbf{1}^T = (1, \ldots, 1)$ and $\mathbf{0}^T = (0, \ldots, 0)$.

An immediate consequence of this is that exiting any particular state, say the j^{th} state, is a Poisson process with rate $\lambda_j = \sum_{i \neq j} \alpha_{i,j}$ and exponential exit time distribution

$$P(t_{exit} < t) = 1 - \exp(-\lambda_j t), \tag{2.7}$$

and expected exit time (or mean first exit time)

$$T_e = \frac{1}{\lambda_j}. \tag{2.8}$$

More generally, consider exit from a collection of states $\tilde{\mathcal{S}} \in \mathcal{S}$. We split the matrix A into two parts, namely, $\tilde{A} = (\tilde{\alpha}_{i,j})$, where $\tilde{\alpha}_{i,j} = \alpha_{i,j}$ for $s_i, s_j \in \tilde{\mathcal{S}}$ and set $B = (\beta_{i,j})$ with $\beta_{i,j} = \alpha_{i,j}$ for $s_i \notin \tilde{\mathcal{S}}$, $s_j \in \tilde{\mathcal{S}}$. Notice that the master equation (2.5) can be split into two equations, namely,

$$\frac{d\tilde{\mathbf{p}}}{dt} = \tilde{W}\tilde{\mathbf{p}}, \qquad \frac{d\mathbf{q}}{dt} = B\tilde{\mathbf{p}}, \tag{2.9}$$

where $\tilde{\mathbf{p}} = \text{Prob}(X \in \tilde{\mathcal{S}})$ and $\mathbf{q} = \text{Prob}(X \notin \tilde{\mathcal{S}})$, and $\tilde{W} = \tilde{A} - \tilde{D}$, where $\tilde{D} = \Sigma_{\tilde{\alpha}} + \Sigma_\beta$.

We assume that all the eigenvalues of $\tilde{W}$ are negative, so that the nullspace of $\tilde{W}$ is trivial, in other words, there is no nontrivial stationary distribution for the set $\tilde{\mathcal{S}}$. As a result, it must be that $\mathbf{1^T q} \to \mathbf{1}$ as $t \to \infty$, and $\mathbf{1^T \frac{dq}{dt}}$ is the exit time distribution. In general, this is not a simple exponential distribution. Nonetheless, we can calculate several statistics for this process. For example,

$$\mathbf{q}(\infty) = \int_0^\infty \frac{d\mathbf{q}}{dt} dt = B\tilde{W}^{-1} \int_0^\infty \frac{d\tilde{\mathbf{p}}}{dt} dt = -B\tilde{W}^{-1}\tilde{\mathbf{p}}_0, \tag{2.10}$$

where $\tilde{\mathbf{p}}_0$ is the initial distribution for $\tilde{\mathbf{p}}$, $\mathbf{1}^T\tilde{\mathbf{p}}_0 = 1$. The quantity $\mathbf{q}(\infty)$ is known as the *splitting probability*.

The expected exit time is

$$\begin{aligned} T_e &= \int_0^\infty t\mathbf{1}^T \frac{d\mathbf{q}}{dt} dt \\ &= \beta^T \int_0^\infty t\tilde{\mathbf{p}} dt = \beta^T \tilde{W}^{-1} \int_0^\infty t \frac{d\tilde{\mathbf{p}}}{dt} dt \\ &= \mathbf{1}^T \int_0^\infty \tilde{\mathbf{p}} dt = -\mathbf{1}^T \tilde{W}^{-1} \tilde{\mathbf{p}}_0, \end{aligned} \tag{2.11}$$

where $\beta^T = \mathbf{1}^T B$.

Exit times are additive. To see this, suppose there are two groups of states $\tilde{\mathcal{S}}_1$ and $\tilde{\mathcal{S}}_2$ and that the exit from states $\tilde{\mathcal{S}}_2$ into states $\tilde{\mathcal{S}}_1$ is into a single state of $\tilde{\mathcal{S}}_1$ (say, state s_1) with rates β_2, and exit from states $\tilde{\mathcal{S}}_1$ is at rates β_1. In this case the master equation reduces to having

$$\tilde{W} = \begin{pmatrix} \tilde{W}_1 & \mathbf{e}_1 \beta_2^T \\ 0 & \tilde{W}_2 \end{pmatrix}, \tag{2.12}$$

from which we calculate directly that

$$T_e = -\mathbf{1}^T {\tilde{W}_1}^{-1} \mathbf{e}_1 - \mathbf{1}^T \tilde{W}_2^{-1} \mathbf{p}_0 = T_e^1 + T_e^2. \tag{2.13}$$

3 Stochastic Simulation

The classical way to simulate this stochastic process is via Monte Carlo simulation. This process works as follows. We define the state $S(t)$ to be some integer $1 \leq j \leq N$ and then track S as a function of time by discretizing time into small steps of size dt. Then, the probability that the state j switches to the state i in a time step dt is $\alpha_{i,j} dt$ and the probability that it does not switch is $1 - \sum_i \alpha_{i,j} dt$. In a Monte Carlo simulation, one makes decisions about changing states on the basis of these switching probabilities compared with a random number chosen from a uniform distribution on $[0, 1]$. Using many such simulation trajectories, one gathers statistics about the process; it is well-known that these statistics converge to the solution of the Kolmogorov equation as $dt \to 0$.

The Gillespie algorithm takes a different view of a stochastic simulation. Rather than following the process as a function of time, it follows a process as a function of transition. That is, we define the state s_n to be one of the integers $1, 2, \cdots, N$ after the n^{th} transition. Then, a transition from state j to state i takes place with probability

$$m_{i,j} = \frac{\alpha_{i,j}}{\sum_{k \neq j} \alpha_{k,j}}, \tag{3.1}$$

and the time at which the n^{th} transition takes place is t_n, where $t_n = t_{n-1} + \Delta t_n$, and Δt_n is an exponentially distributed random variable with rate $\lambda_j = \sum_{k \neq j} \alpha_{k,j}$, i.e.

$$\Delta t_n = -\frac{1}{\lambda_j} \ln(R), \tag{3.2}$$

where R is a uniformly distributed random number $R \in [0, 1]$. Thus, to simulate this process, we draw two uniformly distributed numbers, one used in conjunction with (3.1) to determine which reaction takes place and the second used in conjunction with (3.2) to determine when the next reaction takes place.

The evolution of the statistics for the Gillespie algorithm are described in a different way as well. Let $\mathbf{p}_n$ be the probability distribution of s_n, i.e, the state vector after n transitions. In this way of thinking, the Markov process can be represented by

$$\mathbf{p}_{n+1} = M\mathbf{p}_n, \tag{3.3}$$

where the transition probability matrix is $M = A\Sigma_\alpha^{-1} = W\Sigma_\alpha^{-1} + I$. Notice that $\mathbf{1}^T M = \mathbf{1}^T$. From the equation (3.3), it is obvious that $\mathbf{p}_n = M^n\mathbf{p}_0$, where $\mathbf{p}_0$ is the initial probability distribution, $\mathbf{1}^T\mathbf{p}_0 = 1$, and of course, $\mathbf{1}^T\mathbf{p}_n = 1$.

Suppose we wish to know the number of times a particular state is visited after n transitions. Let $z_i|n$ be the number of times state i is visited after n transitions. Then,

$$P(z_i|n) = P(z_i - 1|n-1)\sum_j m_{i,j}p_{n-1_j} + P(z_i|n-1)(1 - \sum_j m_{i,j}p_{n-1_j}), \tag{3.4}$$

where p_{n_j} is j^{th} component of the vector $\mathbf{p}_n$. The expected number of visits to site i is

$$\begin{aligned} E(Z_i|n) &= \sum_{z_i} z_i P(z_i|n) \\ &= \sum_{z_i} z_i P(z_i - 1|n-1)\sum_j m_{i,j}p_{n-1_j} + \sum_{z_i} z_i P(z_i|n-1)(1 - \sum_j m_{i,j}p_{n-1_j}) \\ &= E(Z_i|n-1) \ + \sum_j m_{i,j}p_{n-1_j}, \end{aligned} \tag{3.5}$$

or in matrix notation

$$\mathbf{Z}_{n+1} = \mathbf{Z}_n + M\mathbf{p}_n, \tag{3.6}$$

where $\mathbf{Z_n}$ is the vector of expected number of times states are visited. Necessarily, $\mathbf{1}^T\mathbf{Z}_n = n$.

Other interesting statistics for this process can be determined. For example, let Π_n be the vector with entries $\pi_{i,j}$ which is the number of times the transition from state j to state i occurred in n total transitions. Necessarily, $\mathbf{1}^T\Pi_n = n$. Setting $P(\Pi_n|i)$ to be the probability of Π_n given that the process is in state i after the n^{th} transition, we have

$$P(\Pi_n|i) = \sum_j m_{i,j}P(\Pi_n - \mathbf{e}_{i,j}|j), \tag{3.7}$$

where $\mathbf{e}_{i,j}$ is a vector with $N^2 - 1$ zeroes and the entry whose index is (i,j) is one. Now suppose that $\delta t_{i,j}$ is the expected time for the transition from state j to state i. Then, the expected time to be in state i after n transitions is

$$E(T_n|i) = \sum_{\mathbf{1}^T\Pi_n = n} P(\Pi_n|i)\Pi_n \cdot \Delta T, \tag{3.8}$$

where ΔT is the vector with components $\delta t_{i,j}$ of expected transition times. Using (3.7), we find that

$$\begin{aligned} E(T_n|i) &= \sum_{\mathbf{1}^T\Pi_n=n} \sum_j m_{i,j} P(\Pi_n - \mathbf{e}_{i,j}|j)\Pi_n \cdot \Delta T && (3.9)\\ &= \sum_{\mathbf{1}^T\Pi_n=n} \sum_j m_{i,j} P(\Pi_n - \mathbf{e}_{i,j}|j)(\Pi_n - \mathbf{e}_{i,j}) \cdot \Delta T && (3.10)\\ &\quad + \sum_{\mathbf{1}^T\Pi_n=n} \sum_j m_{i,j} P(\Pi_n - \mathbf{e}_{i,j}|j)\mathbf{e}_{i,j} \cdot \Delta T && (3.11)\\ &= \sum_j m_{i,j} E(T_{n-1}|i) + \sum_j m_{i,j}\delta t_{i,j} p_{n_j}. && (3.12) \end{aligned}$$

Let $\mathbf{T}_n$ be the vector with components $E(T_n|i)$. It follows from (3.12) that

$$\mathbf{T}_{n+1} = M\mathbf{T}_n + \Delta\mathbf{T}\mathbf{p}_n, \tag{3.13}$$

where $\Delta\mathbf{T} = M.*\Delta T$ is the matrix with components $m_{i,j}\delta t_{i,j}$ (i.e., component-wise multiplication).

Notice that the diagonal elements of D represent the net rate of moving out of a particular state. Hence, the elements of D^{-1} represent the expected waiting time to leave a particular state. Thus, $\Delta\mathbf{T} = MD^{-1}$.

There are other features of the Markov process that may be of interest. For example, changes of state may lead to changes of some physical characteristic x, characterized by the numbers $\delta x_{i,j}$ where $\delta x_{i,j}$ is the change in x when there is a transition from state j to state i. In this case, the expected value of x after $n+1$ transitions is

$$\mathbf{X}_{n+1} = M\mathbf{X}_n + \Delta\mathbf{X}\mathbf{p}_n, \tag{3.14}$$

where $\Delta\mathbf{X} = M.*\delta X$ has entries $m_{i,j}\delta x_{i,j}$. This equation is identical in form to (3.13).

Finally, we may wish to know the number of times each transition occurs. It follows from (3.7) that

$$E(\Pi_n|i) = \sum_{\mathbf{1}^T\Pi_n=n} P(\Pi_n|i)\Pi_n. \tag{3.15}$$

Using (3.7), we find that

$$\begin{aligned} E(\Pi_n|i) &= \sum_{\mathbf{1}^T\Pi_n=n} \sum_j m_{i,j} P(\Pi_n - \mathbf{e}_{i,j}|j)\Pi_n && (3.16)\\ &= \sum_{\mathbf{1}^T\Pi_n=n} \sum_j m_{i,j} P(\Pi_n - \mathbf{e}_{i,j}|j)(\Pi_n - \mathbf{e}_{i,j}) && (3.17)\\ &\quad + \sum_{\mathbf{1}^T\Pi_n=n} \sum_j m_{i,j} P(\Pi_n - \mathbf{e}_{i,j}|j)\mathbf{e}_{i,j} && (3.18)\\ &= \sum_j m_{i,j} E(T_{n-1}|i) + \sum_j m_{i,j}\mathbf{e}_{i,j} p_{n_j}. && (3.19) \end{aligned}$$

In matrix notation, with $\mathbf{E}_n$ the $N \times N^2$ matrix of expected number of transitions $E(\Pi_n|i)$, we have

$$\mathbf{E}_{n+1} = M\mathbf{E}_n + \Delta\mathbf{E}\mathbf{p}_n, \tag{3.20}$$

where ΔE is an $N \times N^2 \times N$ matrix operator (tensor) with components $m_{i,j}\mathbf{e}_{i,j}$.

To summarize, we have the following iterative formulas for statistics of this process

$$\begin{aligned}
\mathbf{p}_{n+1} &= M\mathbf{p}_n, & (3.21)\\
\mathbf{Z}_{n+1} &= \mathbf{Z}_n + M\mathbf{p}_n, & (3.22)\\
\mathbf{T}_{n+1} &= M\mathbf{T}_n + \Delta\mathbf{T}\mathbf{p}_n, & (3.23)\\
\mathbf{X}_{n+1} &= M\mathbf{X}_n + \Delta\mathbf{X}\mathbf{p}_n, & (3.24)\\
\mathbf{E}_{n+1} &= M\mathbf{E}_n + \Delta\mathbf{E}\mathbf{p}_n, & (3.25)
\end{aligned}$$

with initial values $\mathbf{Z}_0 = \mathbf{T}_0 = \mathbf{X}_0 = 0$, $\mathbf{E}_0 = 0$, and $\mathbf{1}^T\mathbf{p}_0 = 1$.

Now, we apply these formulas to the case discussed above, in which there is exit from the group of states $\tilde{\mathcal{S}}$. In this case,

$$M = \begin{pmatrix} I & B\tilde{D}^{-1} \\ 0 & \\ \vdots & \tilde{M} \\ 0 & \end{pmatrix},$$

with $\tilde{M} = \tilde{A}\tilde{D}^{-1} = \tilde{W}\tilde{D}^{-1} + I$. Then, (3.3) reduces to

$$\mathbf{q}_{n+1} = \mathbf{q}_n + B\tilde{D}^{-1}\tilde{\mathbf{p}}_n, \qquad \tilde{\mathbf{p}}_{n+1} = \tilde{M}\tilde{\mathbf{p}}_n. \tag{3.26}$$

It follows that

$$\mathbf{q}_{n+1} = B\tilde{D}^{-1}\sum_{k=0}^{n}\tilde{M}^n\mathbf{p}_0 = B\tilde{D}^{-1}(\tilde{M} - I)^{-1}(\tilde{M}^{n+1} - I)\mathbf{p}_0, \tag{3.27}$$

and, since the eigenvalues of $\tilde{M}$ are all strictly less than one in magnitude, we find that in the limit $n \to \infty$

$$\mathbf{q}_\infty = -B\tilde{D}^{-1}(\tilde{M} - I)^{-1}\mathbf{p}_0 = -B\tilde{W}^{-1}\mathbf{p}_0, \tag{3.28}$$

which is exactly the splitting probability found above in (2.10).

The expected number of visits to the states $\tilde{\mathcal{S}}$ satisfy

$$\begin{aligned}
\tilde{\mathbf{Z}}_{n+1} &= \tilde{\mathbf{Z}}_n + \tilde{M}^{n+1}\mathbf{p}_0 & (3.29)\\
&= \sum_{k=1}^{n}\tilde{M}^k\mathbf{p}_0 & (3.30)\\
&= (\tilde{M} - I)^{-1}(\tilde{M}^{n+1} - I)\mathbf{p}_0, & (3.31)
\end{aligned}$$

so that

$$\tilde{\mathbf{Z}}_\infty = \lim_{n\to\infty} = -(\tilde{M} - I)^{-1}\mathbf{p}_0 = -\tilde{D}\tilde{W}^{-1}\mathbf{p}_0. \tag{3.32}$$

This implies that the expected number of transitions to exit is

$$Z_e = 1 - \mathbf{1}^T\tilde{D}\tilde{W}^{-1}\mathbf{p}_0. \tag{3.33}$$

To calculate exit statistics, we observe that the three equations (3.23)-(3.25) are of the form

$$\mathbf{Y}_{n+1} = M\mathbf{Y}_n + M_Y\mathbf{p}_n, \tag{3.34}$$

with M_Y of the form

$$M_Y = \begin{pmatrix} 0 & B_Y \\ 0 & \tilde{M}_Y \end{pmatrix}. \tag{3.35}$$

It follows that

$$\mathbf{Y}_n = \sum_{k=1}^{n} M^{k-1}M_Y M^{n-k}\mathbf{p}_0. \tag{3.36}$$

Using that

$$M^k = \begin{pmatrix} I & B\tilde{W}^{-1}(\tilde{M}^k - I) \\ 0 & \tilde{M}^k \end{pmatrix}, \tag{3.37}$$

we find that the first component of $\mathbf{Y}_n$, denoted Y_{e_n} (as it corresponds to the exit states), is

$$\begin{aligned} Y_{e_n} &= \sum_{k=1}^{n}\left(B_Y + B\tilde{W}^{-1}(\tilde{M}^{k-1} - I)\tilde{M}_Y\right)\tilde{M}^{n-k}\tilde{\mathbf{p}}_0 \\ &= (B_Y - B\tilde{W}^{-1}\tilde{M}_Y)\tilde{D}\tilde{W}^{-1}(\tilde{M}^n - I)\tilde{\mathbf{p}}_0 + B\tilde{W}^{-1}\sum_{k=1}^{n}\tilde{M}^{k-1}\tilde{M}_Y\tilde{M}^{n-k}\tilde{\mathbf{p}}_0. \end{aligned} \tag{3.38}$$

In the limit $n \to \infty$ we find

$$Y_e = (-B_Y + B\tilde{W}^{-1}M_Y)\tilde{D}\tilde{W}^{-1}\tilde{\mathbf{p}}_0. \tag{3.39}$$

Using that $\Delta\mathbf{T} = MD^{-1}$, we determine the expected exit time to be

$$\begin{aligned} T_e &= (-\beta^T\tilde{D}^{-2} + \beta^T\tilde{W}^{-1}\tilde{M}\tilde{D}^{-1})\tilde{D}\tilde{W}^{-1}\tilde{\mathbf{p}}_0 &(3.40)\\ &= -\mathbf{1}^T\tilde{W}^{-1}\tilde{\mathbf{p}}_0, &(3.41)\end{aligned}$$

which agrees with our earlier calculation (2.11).

A similar calculation shows that the expected value of x at absorption is

$$X_e = -\mathbf{1}^T(\tilde{M}.*\tilde{\Delta X})\tilde{D}\tilde{W}^{-1}\tilde{\mathbf{p}}_0 - \beta^T.*\Delta X_e\tilde{\mathbf{p}}_0. \tag{3.42}$$

The number X_e is the expected change in the variable x before absorption, e.g., the distance traveled.

The point of this is that the iteration (3.13) is an equivalent way to calculate T_e. Furthermore, this iteration views the stochastic process in the same way that the Gillespie algorithm views the simulation of the process, namely by keeping track of transitions. Thus, T_e as calculated using (3.13) is *exactly* the statistic that is calculated when simulating a stochastic process using the Gillespie algorithm.

What we have learned from these calculations is that the Gillespie algorithm produces exactly the same exit time statistics as does solving the master equation. However, it also demonstrates the validity of the following reduction principle:

Reduction Principle 1: *If a collection of states has a single entry point (say s_1), then that collection of states can be reduced into a single state with exit probabilities $-B\tilde{W}^{-1}\mathbf{e}_1$ and expected exit time $-\mathbf{1}^T\tilde{W}^{-1}\mathbf{e}_1$. The Gillespie simulation of the reduced Markov process has exactly the same expected travel times and travel distances as the original (uncollapsed) process. This reduction results in the elimination of $Z_e = -\mathbf{1}^T\tilde{D}\tilde{W}^{-1}\mathbf{p}_0$ transitions.*

4 Fast/Slow Dynamics

Perhaps the biggest source of inefficiency with the Gillespie algorithm is in the situation where there is a lot of cycling within a group of states before exiting. This typically happens because some intra-group transition rates are much larger than those that result in exiting. To examine this situation we suppose that the intragroup dynamics are fast compared to the exiting dynamics, that is, we assume that

$$\tilde{W} = \frac{1}{\epsilon}W_f - \Sigma_\beta \tag{4.1}$$

for some small parameter $0 < \epsilon << 1$. Notice that $\mathbf{1}^T W_f = 0$. We denote the negative diagonal part of W_f as $D = \frac{1}{\epsilon}\Sigma_f - \Sigma_\beta$. Necessarily, there is a null vector of W_f, say ρ_0, with $\mathbf{1^T}\boldsymbol{\rho_0} = \mathbf{1}$ and $W_f\rho_0 = 0$.

To solve the problem

$$\tilde{W}x = \tilde{\mathbf{p}}_0, \tag{4.2}$$

we try a power series solution $x = a\rho_0 + \epsilon\mathbf{p}_1 + \cdots$ with $\mathbf{1}^T\mathbf{p}_1 = 0$. It must be that

$$W_f\mathbf{p}_1 - a\Sigma_\beta\rho_0 = \tilde{\mathbf{p}}_0, \tag{4.3}$$

which, from the Fredholm alternative, requires $a = -\frac{\mathbf{1}^T\tilde{\mathbf{p}}_0}{\beta^T\rho_0}$, from which it follows that

$$x = -\frac{\rho_0}{\beta^T\rho_0} + O(\epsilon). \tag{4.4}$$

The splitting probability is

$$q(\infty) = \frac{B\rho_0}{\beta^T\rho_0} + O(\epsilon), \tag{4.5}$$

and the expected escape time from the group is

$$T_e = -\mathbf{1}^T x = \frac{1}{\beta^T \rho_0} + O(\epsilon). \tag{4.6}$$

The exit time distribution can be calculated using standard slow-manifold analysis. That is, after a short initial time,

$$\mathbf{p}(t) = a(t)\rho_0 + \epsilon \mathbf{p}_1. \tag{4.7}$$

Substituting this into the differential equation for $\mathbf{p}$ we find

$$\frac{da}{dt}\rho_0 + \epsilon \frac{d\mathbf{p}_1}{dt} = W_f \mathbf{p}_1 - a\Sigma_\beta \rho_0 + O(\epsilon) \tag{4.8}$$

from which it follows that

$$\frac{da}{dt} = -a\beta^T \rho_0 + O(\epsilon), \tag{4.9}$$

implying that the exit time distribution is approximately exponential with rate $\beta^T \rho_0$.

The number of transitions before exiting is

$$Z_e = -\mathbf{1}^T \tilde{D} \tilde{W}^{-1} \mathbf{p}_0 = \frac{1}{\epsilon} \frac{\mathbf{1}^T \Sigma_f \rho_0}{\beta^T \rho_0} + O(1). \tag{4.10}$$

In other words, the number of transitions before exiting is of the order of $\frac{1}{\epsilon}$, obviously representing a significant computational cost in any simulation, if ϵ is small.

This leads us to our second reduction possibility:

Reduction Principle 2: *If the nonzero eigenvalues of the matrix $\tilde{A} - \Sigma_{\tilde{\alpha}}$ are large compared to the eigenvalues of Σ_β, then the collection of states $\tilde{\mathcal{S}}$ can be reduced to a single state with an exponential exit time distribution with rate $\beta^T \rho_0$ and splitting probabilities $\frac{B\rho_0}{\beta^T \rho_0}$. This reduction results in an improvement in simulation time of order $\frac{\mathbf{1}^T \Sigma_{\tilde{\alpha}} \rho_0}{\beta^T \rho_0}$.*

Notice that this reduction does not require information about the entry into the group of states $\tilde{\mathcal{S}}$.

5 Examples

Consider the Michaelis-Menten enzyme kinetic reaction scheme

$$S + E \underset{k_2}{\overset{k_1}{\rightleftarrows}} C \xrightarrow{k_3} P + E \tag{5.1}$$

where S is substrate, P is product, E is enzyme and C is the enzyme-substrate complex. Suppose that there is a single enzyme molecule and s substrate molecules. We wish to know how long will

the conversion of substrate to product take. Notice, also, that if $k_2 >> k_3$ then there will be a lot of cycling of the first reaction before product is produced, and this is a problem for the Gillespie algorithm, so we would like to apply our analysis to reduce this problem.

We would like to use this analysis to reduce the Michaelis-Menten equations. Previous analysis of this problem (Rao and Arkin, Kuwahara and Myers) [**?**] [**?**][**?**]have given some analyses of this problem, but they fail in important parameter ranges. I think we can fix this.

We consider the problem of producing one molecule of product starting with S_0 substrate molecules and E_0 enzyme molecules. Because of conservation, it must be that if there are j complex molecules C, there there are $E_0 - j$ free enzyme molecules and $S_0 - j$ substrate molecules. For this group of states, the matrix $\tilde{A}$ is the matrix with upper diagonal elements $\tilde{\alpha}_{j,j-1} \equiv a_j = \tilde{k}_1(E_0 - j_(S_0 - j)$ and $\tilde{\alpha}_{j-1,j} \equiv b_j = k_2 j$, $j = 0, 1, \cdots, N$, where $N = \min(E_0, S_0)$, and where $\tilde{k}_1 = \frac{k_1}{\nu}$, with ν the characteristic volume of the system (necessary to convert units of concentration to units of numbers of molecules).

For this problem we can readily compute that the nullspace of A has

$$p_j = a \left(\frac{\tilde{k}_1}{k_2} \right)^j \frac{E_0! S_0!}{j!(E_0 - j)!(S_0 - j)!}, \tag{5.2}$$

where a is a normalization constant (which I don't know how to calculate). Here j is the number of complex molecules. Thus, we propose the rate of production of product to be

$$k_3 \frac{\sum_j j p_j}{\sum_j p_j}. \tag{5.3}$$

There is no simple analytical expression for this. However, numerical calculations show that this is extremely well approximated by

$$\frac{\sum_j j p_j}{\sum_j p_j} \approx \frac{1}{2}(C_e - \sqrt{C_e^2 - 4 S_0 E_0}). \tag{5.4}$$

for all values of E_0, S_0. where $C_e = \frac{k_2}{k_1} + S_0 + E_0$. Furthermore, it is an easy

$$C \approx \frac{S_0 E_0}{C_e} + O\left(\frac{S_0^2 E_0^2}{C_e^3} \right) \tag{5.5}$$

The expression

First, the continuous (ode) problem implies that in quasisteady state

$$k_1(S_0 - C)(E_0 - C) = k_2 C. \tag{5.6}$$

We solve this for C finding

$$C = \frac{1}{2}(C_e - \sqrt{C_e^2 - 4 S_0 E_0}). \tag{5.7}$$

where $C_e = \frac{k_2}{k_1} + S_0 + E_0$. It is an easy matter to verify that $\frac{4S_0E_0}{C_e} < 1$ so that C is real. Furthermore, it follows that

$$C \approx \frac{S_0E_0}{C_e} + O\left(\frac{S_0^2E_0^2}{C_e^3}\right) \tag{5.8}$$

and it is significant that C is symmetric in S_0 and E_0.

Our proposal is that the degradation rate to product should be k_3C.

There are two things to do now. First, prove that this rate is well approximated by the one above, and second, find the spectral gap for this matrix, and verify that the reduction is legitimate.

5.1 Spectral gap analysis

Our final objective is to justify the claims made in the previous sections. That is, we want to show that an initial distribution will move sufficiently quickly onto the slow manifold to consider the slow mnifold reductions valid. This is done by examining the spectrum of the operator A. In particular, we wish to show that the non-zero eigenvalue with least magnitude is bounded some sufficient distance from the origin to consider ϵ and $\tilde{\epsilon}$ comparatively small parameters. First, we consider the spectrum of the matrix $\tilde{A}$.

To this end we take advantage of the work by Granovsky and Zeifman [4],[5] in bounding the spectral gap β of birth-death processes, which is the minimum of all moduli of the non-zero eigenvalues of the matrix. In particular, they defined $\mathscr{C}$ to be the set of all real matrices $C = \{c_{i,j}, i, j = 1, \ldots, N\}$ with non-negative off-diagonal entries,

$$c_{i,j} \geq 0, \quad i \neq j, \quad i, j = 1, \ldots, N \tag{5.9}$$

and non-positive column sums

$$c_j = \sum_{i=1}^{N} c_{i,j} \leq 0, \quad j = 1, \ldots, N. \tag{5.10}$$

Then, they proved that for any matrix $C \in \mathscr{C}$, the magnitude of the smallest eigenvalue

$$|z_0| = \min\{|z_i|, z_i \in \sigma(C)\} \tag{5.11}$$

satisfies

$$-\bar{c} \leq |z_0| \leq -\underline{c} \tag{5.12}$$

where

$$\begin{aligned} \underline{c} &= \min\{c_j, \ \ j = 1, \ldots, N\} \\ \bar{c} &= \max\{c_j, \ \ j = 1, \ldots, N\}. \end{aligned} \tag{5.13}$$

The matrix $W = \tilde{A} - \Sigma_{\tilde{\alpha}}$ is a $\mathbb{W}$-matrix as defined by Van Kampen [8], so every column sum is 0. Hence, the bounds given in (5.12) are not immediately useful for bounding the spectral gap of

W. We let p_j be the probability that there are j complex molecules at time t. Then, the differential equation

$$\frac{d\mathbf{p}}{dt} = W\mathbf{p} \tag{5.14}$$

can be transformed via

$$p_0(t) = 1 - \sum_{i=1}^{N} p_i(t) \tag{5.15}$$

into the equivalent equation

$$\frac{dy}{dt} = \bar{B}y(t) + f, \tag{5.16}$$

where **notation here is wrong**

$$\begin{aligned} \bar{B} &= \{\bar{a}_{i,j} - \bar{a}_{i1}, \;\; i,j = 1,\ldots,N\}, \\ y(t) &= \left[\; p_2(t), \;\; \ldots, \;\; p_N(t) \; \right]^{\mathrm{T}} \\ f &= \left[\; \bar{a}_{11} \;\; \ldots, \;\; \bar{a}_{N1} \; \right]^{\mathrm{T}}. \end{aligned} \tag{5.17}$$

This is useful for determining β because $\sigma(\bar{B}) = \sigma(\bar{A})/\{0\}$. Hence, the smallest eigenvalue in $\sigma(\bar{B})$ is the spectral gap for our original matrix $\tilde{A}$. They used the similarity transformation $\bar{C} = T\bar{B}T^{-1}$ where

$$T = \begin{bmatrix} d_1 & d_1 & \ldots & d_1 \\ 0 & d_2 & \ldots & d_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & d_N \end{bmatrix}, \tag{5.18}$$

$d_i > 0$, $i = 1,\ldots,N$, and

$$T^{-1} = \begin{bmatrix} d_1^{-1} & -d_2^{-1} & 0 & \ldots & 0 \\ 0 & d_2^{-1} & -d_3^{-1} & \ldots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & -d_N^{-1} \\ 0 & \ldots & \ldots & 0 & d_N^{-1} \end{bmatrix}. \tag{5.19}$$

For any choice of the sequence $d_1,\ldots,d_{N-1}$, $\bar{C} \in \mathscr{C}$, and column sums of $\bar{C}$ take the form

$$c_j = -a_{j-1} - b_j + \delta_j a_j + \delta_{j-1}^{-1} b_{j-1} \tag{5.20}$$

where

$$\begin{aligned} a_j &= \tilde{k}_1(E_0 - j)(S_0 - j), \\ b_j &= k_2 j, \\ \delta_j &= \frac{d_{j+1}}{d_j}, \qquad j = 1\ldots N. \end{aligned} \tag{5.21}$$

Clearly, $a_N = b_0 = 0$. Picking $d_j = 1$, yields

$$c_j = -\tilde{k}_1(E_0 + S_0 - 2j + 1) - k_2 \tag{5.22}$$

which is negative for all $j = 1 \cdots, N$. The smallest $|c_j|$ is when $j = N$, in whch case

$$|c_N| = \tilde{k}_1(|E_0 - S_0| + 1) + k_2 \tag{5.23}$$

which is therefore our estimate of the spectral gap.

We conclude that the approximate analysis described above is appropriate whenever

$$\tilde{k}_1(|E_0 - S_0| + 1) + k_2 \gg k_3. \tag{5.24}$$

6 Discussion

It is well-appreciated that the Gillespie algorithm has many features that make it the preferred method for simulation of discrete state Markov processes. What we have shown here is that the formulation of the Markov process as a transition process via what we identify as the Gillespie equation (3.3) gives access to statistics about exit or absorption times that are not readily available from the Kolmogorov equation (**??**). Furthermore, this formulation gives alternate ways to calculate these statistics.

The traditional way to calculate a statistic is to simulate the process hundreds or thousands of times and calculate the appropriate statistic from the collected simulation data. Here we find two additional ways to calculate these statistics. First, if the state space is not too large, a direct method may be feasible, in which one directly inverts the matrix $\tilde{M} - I$ and uses the formulas (**??**), (**??**), or (3.42). However, the equations (3.13),(**??**), and (3.14) provide iterative methods to calculate these statistics.

It is a straightforward matter to generalize these results to processes with multiple absorbing states and to examine splitting probabilities and their associated statistics.

These formulas do nothing to alleviate the difficulties associated with studying stochastic processes where there are fast reactions (so that $\tilde{M}$ has eigenvalues very close to one), or when the state space is very large. However, just as there are ways to speed up the Gillespie algorithm (τ-leaping), there are also ways to speed up the convergence of the matrix iterative statistics described here.

7 Questions

1. What is the Gillespie formulation for time-dependent processes?
2. What is the analogue of an invariant manifold?
3. Can invariant manifolds be simulated efficiently with a Gilespie algorithm?

4. What are the implications of this for model reduction? (This is probably the easiest question to deal with.)

References

[1] C. Gadgil, C. H. Lee, and H. G. Othmer, *A stochastic analysis of first-order reaction networks*, Bull. Math. Biol., 67 (2005), pp. 901–946.

[2] C. W. Gardiner, *Handbook of stochastic methods for physics, chemistry and the natural sciences*, vol. 13 of Springer Series in Synergetics, Springer-Verlag, Berlin, third ed., 2004.

[3] Ī. Gihman and A. V. Skorohod, *The theory of stochastic processes. II*, Springer-Verlag, New York, 1975. Translated from the Russian by Samuel Kotz, Die Grundlehren der Mathematischen Wissenschaften, Band 218.

[4] B. L. Granovsky and A. I. Zeifman, *The decay function of nonhomogeneous birth-death processes, with application to mean-field models*, Stochastic Processes and their Applications, 72 (1997), pp. 105–120.

[5] ——, *The N-limit of spectral gap of a class of birth-death Markov chains*, Applied Stochastic Models in Business and Industry, 16 (2000), pp. 235–248.

[6] J. P. Keener and J. Sneyd, *Mathematical Physiology*, Springer-Verlag, New York, NY, 2nd ed., 2008.

[7] C. Koch, *Biophysics of Computation*, Oxford University Press, New York, NY, 2004.

[8] N. G. van Kampen, *Stochastic processes in physics and chemistry*, vol. 888 of Lecture Notes in Mathematics, North-Holland Publishing Co., Amsterdam, 1981.

[9] ——, *Stochastic Processes in Physics and Chemistry (Revised and enlarged edition)*, Elsevier Science, Amsterdam, 1992.