Kinetic Investigation of Small Systems Using Different Algorithms

F. Taherkhani\textsuperscript{a}, G.A. Parsafara\textsuperscript{a,}\textsuperscript{*} and M.R. Rahimitabar\textsuperscript{b}

\textsuperscript{a}Department of Chemistry and Nanotechnology Research Center, Sharif University of Technology, Tehran, Iran
\textsuperscript{b}Department of Physics, Sharif University of Technology, Tehran, Iran

(Received 10 June 2006, Accepted 8 August 2006)

We have investigated different algorithms for the simulation of kinetics of small systems. We have simulated the first order reversible reaction with the Gillespie, Gibson and Bruck time simulation as a function of Poisson distribution and compared the results of three algorithms. We have also simulated intracellular viral kinetics for a genome with Gillespie and Poisson distribution algorithms.

Keywords: Chemical kinetics, Small systems, Stochastic approach, Algorithms

INTRODUCTION

To investigate the kinetics of small systems, the use of a stochastic approach is necessary. The classical approach to chemical kinetics is called a deterministic approach, since once the state of the system is known at time \( t_1 \), its state at any other time will be known and no fluctuation about this value is observed \([1,6]\).

The stochastic approach uses the inherent random nature of microscopic molecular collision to build a probabilistic model for a chemical reaction. This approach is useful in studying the kinetics of small systems. For small systems the validity of the deterministic approach becomes worse \([1-7]\). The calculated average concentration \( \text{vs.} \) time given by a stochastic approach and deterministic approach for a linear system are found to be equal. However, the results of these two approaches for nonlinear small systems are completely different \([4]\).

The stochastic formulation proceeds by considering the grand probability function \( P(X;t) \), the density probability of particles in volume \( V \) at time \( t \), where \( X \) is the number of \( S_i \) species and \( X = (X_1, X_2, \ldots, X_N) \) is a vector for the molecular species populations. Evidently, knowledge of this function provides a complete understanding of the probability distribution of all possible states at any time. By considering a discrete infinitesimal time interval \( (t, t+dt) \) in which either 0 or 1 reaction occurs, we see that there exists only \( M+1 \) distinct configurations at time \( t \) that can lead to the state \( X \) at time \( t+dt \). We can write our grand probability function at time \( t+dt \) as a function of all possible states at time \( t \) as follows \([2,9,12]\):

\[
P(X; t+dt) = P(X; t)*P(\text{no state change over } dt) + \sum_{\mu} P(X - \nu_\mu; t)*P(\text{state changes to } X \text{ over } dt)
\]

where \( \nu_\mu \) is a stoichiometric vector defining the result of reaction \( \mu \) on a state vector \( X \), which means \( X \rightarrow X + \nu_\mu \) after an occurrence of reaction \( \mu \). \( P(\text{no state changes over } dt) = 1 - \sum_{\mu=1}^{M} a_\mu (X)dt \) \( P(\text{state changes to } X \text{ over } dt) = \sum_{\mu=1}^{M} P(X - \nu_\mu; t) \).

By using this formula, we may derive the chemical master equation (CME) \([2,9,12]\) that describes the stochastic dynamics of the system as
Due to the complexity of the CME, an analytical solution is rarely possible. Hence, for such a problem, simulation is an appropriate approach to solve the CME.

**STOCHASTIC SIMULATION ALGORITHMS**

Essentially, there are three modeling regimes, namely the discrete and stochastic, continuous and stochastic, and the continuous and deterministic, which depend on the nature of the reaction and the number of molecules in the system under study. A key simulation technique is the stochastic simulation approach to chemical reactions, which was developed by Gillespie \textit{via} the stochastic simulation algorithm [8]. This is an exact procedure for numerically simulating the time evolution of a well-stirred reacting system that takes proper account of the randomness inherent in such a system [2]. It is rigorously based on the same microphysical premise that underlies the CME described above and gives a more realistic representation of a system evolution than the deterministic reaction rate equation represented mathematically by ordinary differential equations that are to be solved simultaneously.

Recently, considerable attention has been paid to reduce the computational time of simulation algorithms for stochastic chemical kinetics. Gibson and Bruck [9] refined the first stochastic reaction simulation algorithm of Gillespie by reducing the number of random variables needed to be simulated. This algorithm can be effective for systems in which some reactions occur much more frequently than others. This algorithm is more efficient than Gillespie’s direct method in the sense that only one new random number must be simulated for each reaction event that takes place, unlike Gillespie’s method in which two random numbers are required. Note, however, that although selective recalculation of the hazards, \( h_i(x,c) \) is also possible for the Gillespie algorithm [8,10,11], it could speed up the algorithm enormously for large systems, which will be introduced later.

**Gillespie’s Direct Method**

For a system in a given state, Gillespie’s direct algorithm [9] asks two equations: Which reaction occurs next and when does it occur? Clearly, both of these equations must be answered probabilistically by specifying the probability density \( P(\mu, \tau) \) that the next reaction is \( \mu \) and it occurs at time \( \tau \). It can be shown that

\[
P(\mu, \tau) d\tau = a_\mu \exp(-\tau \sum_i a_i)d\tau
\]

This equation leads directly to the answers of the two aforementioned questions. First, what is the probability distribution for reaction? Integrating \( P(\mu, \tau) \) over all \( \tau \) from 0 to \( \infty \) results in

\[
P(\mu) = \frac{a_\mu}{\sum_i a_i}
\]

Second, what is the probability distribution as function of time? Summing \( P(\mu, \tau) \) over all \( \mu \) results in

\[
P(\tau) d\tau = (\sum_i a_i) \exp(-\tau \sum_i a_i)d\tau
\]

These two distributions lead to Gillespie’s direct algorithm as:

1. Initialize (i.e., set initial numbers of molecules, set \( 0 \rightarrow t \)) [9].
2. Calculate the propensity function, \( a_i \), for all \( i \).
3. Choose \( \mu \) according to the distribution in equation 3.
4. Choose \( \tau \) according to an exponential with parameter \( \sum_i a_i \) (as in equation 4).
5. Change the number of molecules to reflect execution of reaction \( \mu \). Set \( t + \tau \rightarrow t \)
6. Go to step 2.

**Gillespie’s First Reaction Method**

Algorithm I is direct in the sense that it generates \( \mu \) and \( \tau \) directly. Gillespie also developed the first reaction method, which generates a putative time \( \tau_i \) for each reaction; the reaction with the smallest putative time is allowed to occur, which will be shown by \( \tau_\mu \). Formally, the algorithm for the first reaction method is as follows [9]:

\[
\frac{\partial P(X;t)}{\partial t} = \sum_{\mu=1}^N a_\mu (X - \mu_\mu) P(X - \mu_\mu; t) - a_\mu P(X; t)
\]