Stochastic Chemical Kinetics

Dan T. Gillespie
GillespieDT@mailaps.org

Abstract: The time evolution of a well-stirred chemically reacting system is traditionally modeled by a set of coupled ode's called the reaction rate equations (RRE). The resulting picture of continuous deterministic evolution is, however, valid only for infinitely large systems. That condition is adequately approximated in most "macroscopic" chemical systems. But in biological systems formed by single living cells, the small population numbers of some reactant species can result in dynamical behavior that is noticeably discrete rather than continuous, and noticeably stochastic rather than deterministic. In that case, a more accurate mathematical modeling is obtained by using the machinery of Markov process theory, specifically, the chemical master equation (CME) and the stochastic simulation algorithm (SSA). This talk will review the theoretical foundations of stochastic chemical kinetics, and then discuss some recent efforts to (1) approximate the SSA by a faster simulation procedure, and (2) establish the formal connection between the CME/SSA description and the RRE description.
STOCHASTIC CHEMICAL KINETICS

Dan Gillespie
GillespieDT@mailaps.org

Current Support: AFOSR (#F30602-01-2-0558) and Caltech

Past Support: ONR

Collaborators: John Doyle (Caltech)
Linda Petzold (UCSB)
Muruhan Rathinam (UMBC)
Adam Arkin (UCB, LBNL)
Harley McAdams (Stanford)
Yang Cao (UCSB)
Andrew Hall (UCSB)
CHEMICAL KINETICS

- \( N \) chemical species \( S_1, \ldots, S_N \). \( M \) reaction channels \( R_1, \ldots, R_M \).
  
  E.g., the \( N = M = 2 \) system: \( S_1 + S_2 \overset{c_1}{\underset{c_2}{\rightarrow}} 2S_1 \)

- Assume system has constant volume \( \Omega \), and constant temperature \( T \).
- Also assume system is \textit{well-stirred} (spatially homogeneous).
- Let \( X_i(t) \overset{\hat{\Delta}}{=} \) the \textit{number} of \( S_i \) molecules in the system at time \( t \).
- \( (X_1(t), \ldots, X_N(t)) \overset{\hat{\Delta}}{=} X(t) \), the \textit{state} of the system at time \( t \).

\[ \Rightarrow \textbf{The Problem:} \] Given \( X(t_0) = x_0 \), find \( X(t) \) for \( t > t_0 \).
The Traditional Approach

- Asserts that \( X(t) \) evolves in time according to a set of coupled, first-order, *ordinary differential equations* of the form

\[
\frac{dX_i}{dt} = f_i(X_1, \ldots, X_N) \quad (i = 1, \ldots, N),
\]

where the \( f_i \) are determined by the forms of the \( M \) reaction channels.

- Called the *reaction rate equations* (RRE).
- Usually written in terms of the concentrations, \( Z_i \equiv X_i / \Omega \).
- RRE \( \Rightarrow X(t) \) is a *continuous, deterministic* process.
But in fact . . .

- **X(t)** is *not continuous*; it’s *discrete*:
  - Molecules come in *whole numbers*.
  - Molecular populations change only by *integer* amounts.

- **X(t)** is *not deterministic*; it’s *stochastic*:
  - Only if we were to define the system’s state as the positions and the velocities of all the molecules (and assume Newtonian mechanics) could we regard the system as being “deterministic”.
  - But even then, the extreme *sensitivity to initial conditions* will render the system *effectively* stochastic – like a tossed coin.
  - Chemical reactions occur as discrete events, as a result of molecular collisions that cannot be precisely predicted.
  - At best, we can predict only the *probability* that a reaction event will occur.
\[ R_j: \quad \text{products} \]

\[
\begin{align*}
(\overline{V}_{12} dt)(\pi r_{12}^2) \times \exp\left(-E_j/k_B T\right) & \times x_1 x_2 = \\
\left(\frac{\overline{V}_{12} \pi r_{12}^2}{\Omega} \exp\left(-E_j/k_B T\right)\right) x_1 x_2 \ dt
\end{align*}
\]

Probability that a randomly chosen \( S_1 - S_2 \) pair will react according to \( R_j \) in next \( dt \).

\[ \Delta (c_j x_1 x_2) dt \triangleq a_j(x) \ dt \]

\[ a_j(x) \]
STOCHASTIC CHEMICAL KINETICS

Each elemental reaction channel $R_j$ is defined by two quantities:

- a **propensity function** $a_j(x)$, where $a_j(x)dt \triangleq$ the probability, given $X(t) = x$, that one $R_j$ reaction event will occur in $[t, t+dt]$;

- a **state change vector** $\nu_j \equiv (\nu_{1j}, \ldots, \nu_{Nj})$, where $\nu_{ij} \triangleq$ the change in the $S_i$ population caused by one $R_j$ reaction event.

E.g.,

$$S_1 + S_2 \xrightarrow{c_1 \, c_2} 2S_1 \quad \Rightarrow \quad \begin{cases} a_1(x) = c_1 x_1 x_2, & \nu_1 = (+1, -1, 0, \ldots, 0) \\ a_2(x) = c_2 \frac{x_1(x_1-1)}{2}, & \nu_2 = (-1, +1, 0, \ldots, 0) \end{cases}$$

- **Implication:** $X(t)$ is a **jump Markov process** (a continuous-time, discrete-state, past-forgetting, stochastic process).
Two Approaches to Stochastic Chemical Kinetics

- An *analytical* approach, and a *simulation* approach.
- They are *logically equivalent* – both follow rigorously from the same stochastic premise.
- They are *exact* for *well-stirred* (or *self-stirring*) systems.
THE ANALYTICAL APPROACH

Focuses on the *probability density function* of the random variable $X(t)$, namely

$$P(x,t|x_0,t_0) \equiv \text{Prob}\{X(t) = x, \text{ given that } X(t_0) = x_0\}.$$  

Can prove that $P$ obeys the time-evolution equation

$$\frac{\partial P(x,t|x_0,t_0)}{\partial t} = \sum_{j=1}^{M} \left[ a_j(x - \nu_j)P(x - \nu_j,t|x_0,t_0) - a_j(x)P(x,t|x_0,t_0) \right]$$

- Called the *chemical master equation* (CME).
- In principle, it completely determines $P(x,t|x_0,t_0)$, and hence $X(t)$.
- In practice, it’s impossible to solve for all but the simplest of systems.
**Derivation of the Chemical Master Equation**

Write $\Pr \{ \mathbf{X}(t+dt) = \mathbf{x} \}$ as the sum of the probabilities of all (mutually exclusive) paths leading to state $\mathbf{x}$ at time $t+dt$ via possible states at time $t$:

$$
P(\mathbf{x}, t+dt | \mathbf{x}_0, t_0) = P(\mathbf{x}, t | \mathbf{x}_0, t_0) \times \left[ 1 - \sum_{j=1}^{M} \left( a_j(\mathbf{x}) dt \right) \right]$$

$$
+ \sum_{j=1}^{M} P(\mathbf{x}-\mathbf{v}_j, t | \mathbf{x}_0, t_0) \times (a_j(\mathbf{x}-\mathbf{v}_j) dt)
$$

Subtract $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$, divide through by $dt$, then let $dt \to 0 \Rightarrow$ CME.

**Note:** Invocation of the Addition Law here relies on the fact that $dt$ is so small that *no more than one* reaction event will ever occur in any $dt$. 
THE SIMULATION APPROACH

Focuses on the function \( p(\tau, j|\mathbf{x}, t) \), defined by

\[
p(\tau, j|\mathbf{x}, t) \, d\tau \triangleq \text{the probability, given } \mathbf{X}(t) = \mathbf{x}, \text{ that the next reaction in the system will occur in the infinitesimal time interval } [t+\tau, t+\tau+d\tau), \text{ and will be an } R_j \text{ reaction.}
\]

Can prove that \( p \) is given by

\[
p(\tau, j|\mathbf{x}, t) = a_j(\mathbf{x}) \exp(-a_0(\mathbf{x}) \tau), \quad \text{where } a_0(\mathbf{x}) \triangleq \sum_{j=1}^{M} a_j(\mathbf{x}).
\]

- Implies: The time \( \tau \) to the next reaction event is an exponentially distributed random variable with mean \( 1/a_0(\mathbf{x}) \), and the channel index \( j \) of that reaction is an integer random variable with prob \( a_j(\mathbf{x})/a_0(\mathbf{x}) \).

- It’s easy to generate on a computer (pseudo)random samples of \( \tau \) and \( j \) according to this prescription.
Derivation of the Next-Reaction Probability Density Function

Let \( P_0(\tau | x, t) \) be the probability, given \( X(t) = x \), that no reactions will occur in time interval \([t, t+\tau)\). Then

\[
p(\tau, j | x, t) d\tau = P_0(\tau | x, t) \times \left(a_j(x) d\tau\right),
\]

so

\[
p(\tau, j | x, t) = a_j(x) P_0(\tau | x, t).
\]

To calculate \( P_0(\tau | x, t) \), observe that it must satisfy

\[
P_0(\tau + d\tau | x, t) = P_0(\tau | x, t) \times \left[1 - \sum_{j' = 1}^{M} \left(a_{j'}(x) d\tau\right)\right] = P_0(\tau | x, t) \left[1 - a_0(x) d\tau\right],
\]

whence

\[
\frac{dP_0(\tau | x, t)}{d\tau} = -a_0(x) P_0(\tau | x, t).
\]

The solution to this ODE for the initial condition \( P_0(\tau=0 | x, t) = 1 \) is

\[
P_0(\tau | x, t) = \exp\left(-a_0(x) \tau\right).
\]

QED
The Stochastic Simulation Algorithm (SSA)

An explicit, exact procedure for constructing a numerical realization of the stochastic process $X(t)$. The “direct” version of the SSA is:

1. With the system in state $\mathbf{x}$ at time $t$, compute $a_0(\mathbf{x}) \triangleq \sum_{j'=1}^{M} a_{j'}(\mathbf{x})$.

2. Draw two unit-interval uniform random numbers $r_1$ and $r_2$, and compute $\tau$ and $j$ according to
   - $\tau = \frac{1}{a_0(\mathbf{x})} \ln \left( \frac{1}{r_1} \right)$,
   - $j =$ the smallest integer satisfying $\sum_{j'=1}^{j} a_{j'}(\mathbf{x}) > r_2 a_0(\mathbf{x})$.

3. Effect the next reaction: Replace $t \leftarrow t + \tau$ and $\mathbf{x} \leftarrow \mathbf{x} + \mathbf{v}_j$.

4. Record $(\mathbf{x},t)$. Then return to Step 1, or else end the simulation.
The Stochastic Simulation Algorithm . . .

- Numerically simulates the time evolution of a *well-stirred* chemically reacting system in a way that accurately reproduces all effects of naturally occurring randomness.
- Does *not* entail approximating a “\( dt \)” by a “\( \Delta t \)”.
- Is *procedurally simple*, even for systems whose CME is intractable.
- Has been redesigned to be faster and more efficient (though more complicated to code) by M. Gibson and J. Bruck [ *J. Phys. Chem. A* **104**, 1876 (2000)]
- *Remains too slow for most practical problems*: Since it simulates every reaction event, one at a time, it can take a very long time to execute if any reactant is present in very large numbers – as is usually the case.
WHY BOTHER?

- For most *practical* chemical systems the molecular populations are *very large*, and the phenomenological RRE,

\[
\frac{dX(t)}{dt} = \sum_{j=1}^{M} \nu_j a_j(X(t)),
\]

is *accurate* and *fast*.

- But in *small biochemical systems* (like the inside of a living cell) it can happen that *some* critical reactant species are present in *low numbers*. Then the inherent stochasticity of the system becomes noticeable, and the deterministic RRE can be *inaccurate* and *misleading*.

➢ For such systems, we must instead use either the CME or the SSA.
TWO QUESTIONS:

1. Is there a way to *speed up* the SSA for large systems, even at the cost of some exactness?

2. How does it happen that the *rigorous* but computationally intractable CME,

\[
\frac{\partial P(x,t|x_0,t_0)}{\partial t} = \sum_{j=1}^{M} \left[ a_j(x - v_j)P(x - v_j,t|x_0,t_0) - a_j(x)P(x,t|x_0,t_0) \right],
\]

segues for “large” systems to the *heuristic* but computationally efficient RRE,

\[
\frac{dX(t)}{dt} = \sum_{j=1}^{M} v_j a_j(X(t))?
\]
Tau-Leaping

An approximate, accelerated stochastic simulation procedure.

• Advances the process by a pre-selected time $\tau$, which may encompass more than one reaction event.

• The size of $\tau$ is limited by the Leap Condition: The changes in the propensity function values during the leap must be “small”.

• Accelerated simulation occurs whenever $\tau$ can satisfy the Leap Condition, yet encompass many reaction events.

• But must use with caution! Leap over only the “unimportant” reactions, not the “important” ones.
Key Ideas Underlying Tau-Leaping

- For positive constants $a$ and $\tau$, the Poisson random variable $P(a, \tau)$ is defined as the number of “events” that will occur in time $\tau$, given that the probability of an event occurring in any infinitesimal $dt$ is $adt$.

- Therefore, if $X(t) = x$, and $\tau$ is such that $a_j(x) \approx$ constant in $[t, t+\tau]$, then the number of $R_j$ reactions that will occur in $[t, t+\tau]$ is (approximately) $P(a_j(x), \tau)$.

- Numerical procedures exist for generating sample values of $P(a, \tau)$ [Press, et al., Numerical Recipes: The Art of Scientific Computing].

- A way has been developed to estimate in advance the largest $\tau$ that is consistent with the Leap Condition.
A Method for Choosing $\tau$

If $X(t) = x$, then the \textit{state change} induced by a time leap $\tau$ is

$$X(t + \tau) - x \triangleq \Lambda(\tau, x) = \sum_{j=1}^{M} K_j(x, \tau) \nu_j \approx \sum_{j=1}^{M} \mathcal{P}_j(a_j(x), \tau) \nu_j.$$ 

The resultant change in propensity function $a_j$ is

$$\Delta a_j(\tau, x) \triangleq a_j(x + \Lambda(\tau, x)) - a_j(x) \approx \sum_{i=1}^{N} \frac{\partial a_j(x)}{\partial x_i} \Lambda_i(\tau, x) \approx \sum_{i=1}^{N} \frac{\partial a_j(x)}{\partial x_i} \sum_{j'=1}^{M} \mathcal{P}_{j'}(a_{j'}(x), \tau) \nu_{ij'}.$$ 

$$\Delta a_j(\tau, x) \approx \sum_{j'=1}^{M} f_{jj'}(x) \mathcal{P}_{j'}(a_{j'}(x), \tau), \text{ where } f_{jj'}(x) \triangleq \sum_{i=1}^{N} \frac{\partial a_j(x)}{\partial x_i} \nu_{ij'}.$$ 

So, 

$$\langle \Delta a_j(\tau, x) \rangle \approx \sum_{j'=1}^{M} f_{jj'}(x) \langle \mathcal{P}_{j'}(a_{j'}(x), \tau) \rangle = \sum_{j'=1}^{M} f_{jj'}(x) \left( a_{j'}(x) \tau \right),$$

$$\text{var}\{\Delta a_j(\tau, x)\} \approx \sum_{j'=1}^{M} f_{jj'}^2(x) \text{var}\{\mathcal{P}_{j'}(a_{j'}(x), \tau)\} = \sum_{j'=1}^{M} f_{jj'}^2(x) \left( a_{j'}(x) \tau \right).$$
Satisfy the **Leap Condition** by requiring the *absolute mean* and the *standard deviation* of $\Delta a_j(\tau, \mathbf{x})$ to be $\leq \epsilon a_0(\mathbf{x}) \quad \forall j$, where $\epsilon \ll 1$ is an *accuracy control parameter*. Find that the largest $\tau$ satisfying this requirement is

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

where, with $f_{j,j'}(\mathbf{x}) \triangleq \sum_{i=1}^{N} \frac{\partial a_j(\mathbf{x})}{\partial x_i} \nu_{ij'}$,

$$
\mu_j(\mathbf{x}) \triangleq \sum_{j'=1}^{M} f_{j,j'}(\mathbf{x}) a_{j'}(\mathbf{x}), \quad \sigma_j^2(\mathbf{x}) \triangleq \sum_{j'=1}^{M} f_{j,j'}^2(\mathbf{x}) a_{j'}^2(\mathbf{x}).
$$

The functions $f_{j,j'}$, $\mu_j$, and $\sigma_j^2$ can be explicitly computed *prior* to simulation, and they will usually be easy to evaluate for a given $\mathbf{x}$. 
The Tau-Leap Simulation Algorithm

- **1.** In state \( \mathbf{x} \) at time \( t \), choose \( \tau \) so that the expected change in every propensity function in \([t, t+\tau]\) is \( \leq \varepsilon a_0(\mathbf{x}) \).

- **2.** Generate the number of firings \( k_j \) of channel \( R_j \) in \([t, t+\tau]\) as
  \[
  k_j = \mathcal{P} \left( a_j(\mathbf{x}), \tau \right) \quad (j = 1, \ldots, M).
  \]

- **3.** *Leap*: Replace \( t \leftarrow t + \tau \) and \( \mathbf{x} \leftarrow \mathbf{x} + \sum_{j=1}^{M} k_j \nu_j \).

- **4.** Record \((\mathbf{x}, t)\). Then return to Step 1, or else end the simulation.
Decaying-Dimerizing Reaction Set

\begin{align*}
S_1 & \rightarrow 0 \quad c_1 = 1 \\
S_1 + S_1 & \rightarrow S_2 \quad c_2 = 0.002 \\
S_2 & \rightarrow S_1 + S_1 \quad c_3 = 0.5 \\
S_2 & \rightarrow S_3 \quad c_4 = 0.04
\end{align*}

- Exact SSA Run
- 500 steps (reactions) per plotted dot.
- Initially: \(X_1=100,000;\ X_2=X_3=0\).
- Last reaction at \(T=48.07\).
- Final \(X(3)=17,106\).
- 524,112 steps (reactions) total.
Decaying-Dimerizing Reaction Set

\[ \text{S1} \rightarrow 0 \quad c_1 = 1 \]
\[ \text{S1} + \text{S1} \rightarrow \text{S2} \quad c_2 = 0.002 \]
\[ \text{S2} \rightarrow \text{S1} + \text{S1} \quad c_3 = 0.5 \]
\[ \text{S2} \rightarrow \text{S3} \quad c_4 = 0.04 \]

- Explicit Tau Leaping Run (Epsilon=0.03)
- 1 leap per plotted dot.
- Initially: \( \text{X1}=100,000; \quad \text{X2}=\text{X3}=0 \).
- Last reaction at \( T=44.52 \).
- Final \( X(3)=17,033 \).
- 592 leaps total.
Tau-Leaping usually speeds things up, but...

- Selecting the largest $\tau$ consistent with the Leap Condition entails some computational overhead.
- Generating Poisson random numbers is more time consuming than generating the exponential random numbers used by the SSA.
- Care must be taken not to leap over “important” reactions.
- Additional strategies are needed to cope with stiffness (widely varying time-scales).

- By itself, it’s not the final answer. But it provides a context for such.
How does the CME/SSA segue to the RRE for “large” systems?

• We saw in tau-leaping that, if \( X(t) = x \) and \( \tau \) is macroscopically infinitesimal (in that none of the \( a_j \)'s change “noticeably” during \( \tau \)), then to a good approximation,

\[
X(t + \tau) = x + \sum_{j=1}^{M} \mathcal{P}_j \left( a_j(x), \tau \right) \nu_j,
\]

where the \( \mathcal{P}_j \)'s are statistically independent Poisson random variables.

• Suppose it also happens that every reaction channel \( R_j \) fires many more times than once in the next \( \tau \). (This nearly always happens whenever all the reactant molecular populations are large enough.)
• This would mean, since \( \langle \mathcal{P}_j \left( a_j(x), \tau \right) \rangle = a_j(x) \tau \), that \( a_j(x) \tau \gg 1 \ \forall j \).

• But when \( a \tau \gg 1 \), can approximate \( \mathcal{P}(a, \tau) \approx \mathcal{N}(a \tau, a \tau) \); hence,

\[
X(t + \tau) = x + \sum_{j=1}^{M} \mathcal{N}_j \left( a_j(x) \tau, a_j(x) \tau \right) v_j
\]

\[
= x + \sum_{j=1}^{M} \left\{ a_j(x) \tau + \left[ a_j(x) \tau \right]^{1/2} \mathcal{N}_j(0,1) \right\} v_j
\]

• Collecting terms, and remembering that \( x = X(t) \), this is

\[
X(t + \tau) = X(t) + \sum_{j=1}^{M} v_j a_j \left( X(t) \right) \tau + \sum_{j=1}^{M} v_j \sqrt{a_j \left( X(t) \right)} \mathcal{N}_j(0,1) \sqrt{\tau}
\]

We have just proved the following . . .
**Theorem:** If $\Delta t$ is a *macroscopic infinitesimal*, in that during $\Delta t$, 
- no propensity function changes its value significantly, yet 
- every reaction channel fires many more times that once, 
then we can approximate the $t$ to $t + \Delta t$ system update by 

$$
X(t + \Delta t) = X(t) + \sum_{j=1}^{M} \nu_j a_j(X(t)) \Delta t + \sum_{j=1}^{M} \nu_j \sqrt{a_j(X(t))} N_j(t) \sqrt{\Delta t} .
$$

Here, the $N_j(t)$ are statistically independent, temporally uncorrelated, *normal* random variables with means 0 and variances 1.

- This is the **Chemical Langevin Equation (CLE)**.
- It approximates $X(t)$ as a *continuous* (versus a *jump*) Markov process.
- It is mathematically the same as the SDE 

$$
\frac{dX(t)}{dt} = \sum_{j=1}^{M} \nu_j a_j(X(t)) + \sum_{j=1}^{M} \nu_j \sqrt{a_j(X(t))} \Gamma_j(t) ,
$$

where $\left< \Gamma_j(t) \Gamma_{j'}(t') \right> = \delta_{jj'} \delta(t - t') \quad (j, j' = 1, \ldots, M)$.
In the theory of continuous Markov processes, every Langevin equation for $X(t)$ implies a Fokker-Planck equation for $P(x,t|x_0,t_0)$.

For the chemical Langevin equation (CLE)

$$X(t + dt) = X(t) + \sum_{j=1}^{M} \nu_j a_j(X(t)) \bar{d}t + \sum_{j=1}^{M} \nu_j \sqrt{a_j(X(t))} N_j(t) \sqrt{d}t,$$

the corresponding chemical Fokker-Planck equation (CFPE) can be shown to be

$$\frac{\partial P(x,t|x_0,t_0)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i} \left[ \left( \sum_{j=1}^{M} \nu_{ij} a_j(x) \right) P(x,t|x_0,t_0) \right]$$

$$+ \frac{1}{2} \sum_{i=1}^{N} \frac{\partial^2}{\partial x_i^2} \left[ \left( \sum_{j=1}^{M} \nu_{ij}^2 a_j(x) \right) P(x,t|x_0,t_0) \right]$$

$$+ \sum_{i,i'=1}^{N} \frac{\partial^2}{\partial x_i \partial x_{i'}} \left[ \left( \sum_{j=1}^{M} \nu_{ij} \nu_{ij'} a_j(x) \right) P(x,t|x_0,t_0) \right],$$

for $i < i'$. 
Now consider . . . The Thermodynamic Limit.

**Definition:** All $X_i \to \infty$, and $\Omega \to \infty$, with $X_i/\Omega \to \text{constant}$.

- Can prove that, in this limit, *all* propensity functions grow *linearly* with the system size.
- Therefore, in the CLE

$$X(t + \bar{dt}) = X(t) + \sum_{j=1}^{M} \nu_j a_j (X(t)) \bar{dt} + \sum_{j=1}^{M} \nu_j \sqrt{a_j (X(t))} N_j(t) \sqrt{\bar{dt}}$$

the *deterministic* term grows like $X(t)$, while the *stochastic* term grows like $\sqrt{X(t)}$. So in the thermodynamic limit, the CLE reduces to

$$X(t + \bar{dt}) = X(t) + \sum_{j=1}^{M} \nu_j a_j (X(t)) \bar{dt}$$

- This is how the RRE arises in stochastic chemical kinetics.
A Hierarchy of Schemes for Modeling Chemical Kinetics

1. **Molecular Dynamics** (MD)
   - Tracks the position and velocity of every molecule.
   - Simulates every collision, non-reactive as well as reactive.
   - Shows changes in species populations and spatial concentrations.
   - Is essentially exact.
   - Is extremely slow for realistic systems.

2. **Stochastic Simulation Algorithm** (SSA)
   - Assumes that the non-reactive collisions merely keep the molecules well-stirred; hence, assumes the system is spatial homogeneous.
   - Simulates only the reactive collisions, i.e., every chemical reaction event.
   - Tracks only the species populations.
   - Is exactly equivalent to the Chemical Master Equation (CME).
   - Is much faster than MD.
3. **Tau-Leaping**

- *An approximation to the SSA.*
- Advances time by a *pre-selected* $\tau$ during which no propensity function changes its value “noticeably” and more than one reaction may occur.
- Number of $R_j$ firings in $\tau$ is approximated as $\mathcal{P}(a_j(x), \tau)$.
- Is faster than the SSA if *many* reactions occur in $\tau$.

4. **Chemical Langevin Equation (CLE)**

- *A special case of Tau-Leaping.*
- Applies whenever $\tau$ is such that $a_j(x)\tau \gg 1$ for every $R_j$.
- Number of $R_j$ firings in $\tau$ is approximated as $\mathcal{N}(a_j(x)\tau, a_j(x)\tau^2)$.
- Is an SDE that approximates the *jump* Markov process $X(t)$ of the CME/SSA by a *continuous* Markov process.
- Is faster than ordinary Tau-Leaping.
5. **Reaction Rate Equation** (RRE)
   - *The thermodynamic limit of the CLE.*
   - The (random) diffusion term in the CLE becomes negligibly small compared to the (deterministic) drift term.
   - The CLE then reduces from an SDE to an ODE.
   - Defines a continuous *deterministic* process.
   - Computationally the fastest.

**The Challenge:**

To build a seamless, easy-to-use software package that automatically selects the optimal appropriate method to advance the state in time, and then executes that method correctly and efficiently.