Numerical approximation methods for stochastically modeled biochemical reaction networks

David F. Anderson*

*anderson@math.wisc.edu

Department of Mathematics
University of Wisconsin - Madison

University of Florida: Conference on Computational and Systems Biology

March 17th - 19th, 2011

Grant support: NSF-DMS-1009275
1. Discuss Poisson process and time changes.


3. Numerical methods and error approximation.

4. Mathematical question for today: how do we effectively quantify how well different methods approximate the continuous time Markov chain model.
   - Directly lowers the computational complexity of Monte Carlo methods.
   - Used in variance reduction techniques (lowering computational complexity).

5. Discuss/use multi-scale nature of biochemical reaction networks.
A Poisson process, $Y(\cdot)$, is a model for a series of random observations occurring in time.

(a) Let $\{\xi_i\}$ be i.i.d. exponential random variables with parameter one.

(b) Now, put points down on a line with spacing equal to the $\xi_i$.

Let $Y(t)$ denote the number of points hit by time $t$.

In the figure above, $Y(t) = 6$. 
The Poisson process

Let

- $Y$ be a unit rate Poisson process.

- Define $Y_\lambda(t) \equiv Y(\lambda t)$,

Then $Y_\lambda$ is a Poisson process with parameter $\lambda$.

Intuition: The Poisson process with rate $\lambda$ is simply the number of points hit (of the unit-rate point process) when we run along the time frame at rate $\lambda$.  

\[
\begin{array}{cccccc}
\xi_1 & \xi_2 & \xi_3 & \cdots & & \\
\hline
x & x & x & x & x & x \\
\hline
\end{array}
\]
The Poisson process

There is no reason $\lambda$ needs to be constant in time, in which case

$$Y_{\lambda}(t) \equiv Y \left( \int_0^t \lambda(s)ds \right)$$

is an inhomogeneous Poisson process with intensity $\lambda(t)$. 

Points:

1. We have "changed time" to convert a unit-rate Poisson point process to one which has rate or intensity or propensity $\lambda(t)$.

2. Will use more complicated time changes of unit-rate processes to build models of interest.
The Poisson process

There is no reason $\lambda$ needs to be constant in time, in which case

$$Y_\lambda(t) \equiv Y \left( \int_0^t \lambda(s)ds \right)$$

is an inhomogeneous Poisson process with intensity $\lambda(t)$.

Points:

1. We have “changed time” to convert a unit-rate Poisson point process to one which has rate or intensity or propensity $\lambda(t)$.

2. Will use more complicated time changes of unit-rate processes to build models of interest.
Outline

1. Discuss Poisson process and time changes.


3. Numerical methods and error approximation.

4. Mathematical question for today: how do we effectively quantify how well different methods approximate the continuous time Markov chain model.

5. Discuss/use multi-scale nature of biochemical reaction networks.
Models of interest

Consider the simple system

\[ A + B \rightarrow C \]

where one molecule each of \( A \) and \( B \) is being converted to one of \( C \).

Intuition for standard stochastic model:

\[
P\{\text{reaction occurs in } (t, t + \Delta t) | F_t \} \approx \kappa X_A(t) X_B(t) \Delta t
\]

where

- \( F_t \) represents the information about the system available at time \( t \) and
- \( \kappa \) is a positive constant, the reaction rate constant.
Models of interest

\[
A + B \rightarrow C
\]

Simple book-keeping: if \( X(t) = (X_A(t), X_B(t), X_C(t)) \) gives the state at time \( t \),

\[
X(t) = X(0) + R(t) \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix},
\]

where

- \( R(t) \) is the # of times the reaction has occurred by time \( t \) and
- \( X(0) \) is the initial condition.

Goal: represent \( R(t) \) in terms of Poisson process.
Models of interest

Recall that for $A + B \rightarrow C$ our intuition was

$$P\{\text{reaction occurs in } (t, t + \Delta t) | \mathcal{F}_t\} \approx \kappa X_A(t) X_B(t) \Delta t,$$

where $Y$ is a unit-rate Poisson process.
Models of interest

Recall that for $A + B \rightarrow C$ our intuition was

$$P\{\text{reaction occurs in } (t, t + \Delta t)|\mathcal{F}_t\} \approx \kappa X_A(t) X_B(t) \Delta t,$$

and that for an inhomogeneous Poisson process with rate $\lambda(t)$ we have

$$P\{Y_\lambda(t + \Delta t) - Y_\lambda(t) > 0|\mathcal{F}_t\} \approx \lambda(t) \Delta t.$$
Models of interest

Recall that for $A + B \rightarrow C$ our intuition was

$$P\{\text{reaction occurs in } (t, t + \Delta t) | \mathcal{F}_t \} \approx \kappa X_A(t)X_B(t)\Delta t,$$

and that for an inhomogeneous Poisson process with rate $\lambda(t)$ we have

$$P\{Y_\lambda(t + \Delta t) - Y_\lambda(t) > 0 | \mathcal{F}_t \} \approx \lambda(t)\Delta t.$$

This suggests we can model

$$R(t) = Y \left( \int_0^t \kappa X_A(s)X_B(s)ds \right)$$

where $Y$ is a unit-rate Poisson process.
Models of interest

Recall that for $A + B \rightarrow C$ our intuition was

$$P\{\text{reaction occurs in } (t, t + \Delta t) | \mathcal{F}_t\} \approx \kappa X_A(t) X_B(t) \Delta t,$$

and that for an inhomogeneous Poisson process with rate $\lambda(t)$ we have

$$P\{Y_\lambda(t + \Delta t) - Y_\lambda(t) > 0 | \mathcal{F}_t\} \approx \lambda(t) \Delta t.$$

This suggests we can model

$$R(t) = Y \left( \int_0^t \kappa X_A(s) X_B(s) ds \right)$$

where $Y$ is a unit-rate Poisson process.

Hence

$$\begin{pmatrix} X_A(t) \\ X_B(t) \\ X_C(t) \end{pmatrix} \equiv X(t) = X(0) + \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} Y \left( \int_0^t \kappa X_A(s) X_B(s) ds \right).$$

This equation uniquely determines $X(t)$ for all $t > 0$. 
General stochastic models of biochemical reactions

- We consider a network of reactions involving \( d \) chemical species, \( S_1, \ldots, S_d \):

\[
\sum_{i=1}^{d} \nu_{ik} S_i \rightarrow \sum_{i=1}^{d} \nu'_{ik} S_i
\]

- \( \nu_k \in \mathbb{Z}^d_{\geq 0} \): number of molecules of each chemical species consumed in the \( k \)th reaction.

- \( \nu'_k \in \mathbb{Z}^d_{\geq 0} \): number of molecules of each chemical species created in the \( k \)th reaction.

Denote reaction vector as

\[
\zeta_k = \nu'_k - \nu_k.
\]
General stochastic models of biochemical reactions

- We consider a network of reactions involving \(d\) chemical species, \(S_1, \ldots, S_d\):
  \[
  \sum_{i=1}^{d} \nu_{ik} S_i \longrightarrow \sum_{i=1}^{d} \nu'_{ik} S_i
  \]

- \(\nu_k \in \mathbb{Z}_{\geq 0}^d\): number of molecules of each chemical species consumed in the \(k\)th reaction.

- \(\nu'_k \in \mathbb{Z}_{\geq 0}^d\): number of molecules of each chemical species created in the \(k\)th reaction.

Denote reaction vector as
\[
\zeta_k = \nu'_k - \nu_k.
\]

- The intensity of \(k\)th reaction is \(\lambda_k : \mathbb{Z}_{\geq 0}^d \rightarrow \mathbb{R}\) (Mass action kinetics).

- By analogy with before
\[
X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s)) \, ds \right) \zeta_k,
\]

\(Y_k\) are independent, unit-rate Poisson processes.
Some examples

**E. coli Heat Shock Response Model.** 9 species, 18 reactions.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Intensity</th>
<th>Reaction</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset \rightarrow A_8$</td>
<td>$4.00 \times 10^0$</td>
<td>$A_6 + A_8 \rightarrow A_9$</td>
<td>$3.62 \times 10^{-4}X_{A_6}X_{A_8}$</td>
</tr>
<tr>
<td>$A_2 \rightarrow A_3$</td>
<td>$7.00 \times 10^{-1}X_{A_2}$</td>
<td>$A_8 \rightarrow \emptyset$</td>
<td>$9.99 \times 10^{-5}X_{A_8}$</td>
</tr>
<tr>
<td>$A_3 \rightarrow A_2$</td>
<td>$1.30 \times 10^{-1}X_{A_3}$</td>
<td>$A_9 \rightarrow A_6 + A_8$</td>
<td>$4.40 \times 10^{-5}X_{A_9}$</td>
</tr>
<tr>
<td>$\emptyset \xrightarrow{A_1} A_2$</td>
<td>$7.00 \times 10^{-3}X_{A_1}$</td>
<td>$\emptyset \rightarrow A_1$</td>
<td>$1.40 \times 10^{-5}$</td>
</tr>
<tr>
<td>stuff + $A_3 \rightarrow A_5 + A_2$</td>
<td>$6.30 \times 10^{-3}X_{A_3}$</td>
<td>$A_1 \rightarrow \emptyset$</td>
<td>$1.40 \times 10^{-6}X_{A_1}$</td>
</tr>
<tr>
<td>stuff + $A_3 \rightarrow A_4 + A_2$</td>
<td>$4.88 \times 10^{-3}X_{A_3}$</td>
<td>$A_4 \xrightarrow{A_7} A_6$</td>
<td>$1.42 \times 10^{-6}X_{A_4}X_{A_7}$</td>
</tr>
<tr>
<td>stuff + $A_3 \rightarrow A_6 + A_2$</td>
<td>$4.88 \times 10^{-3}X_{A_3}$</td>
<td>$A_5 \rightarrow \emptyset$</td>
<td>$1.80 \times 10^{-8}X_{A_5}$</td>
</tr>
<tr>
<td>$A_7 \rightarrow A_2 + A_6$</td>
<td>$4.40 \times 10^{-4}X_{A_7}$</td>
<td>$A_6 \rightarrow \emptyset$</td>
<td>$6.40 \times 10^{-10}X_{A_6}$</td>
</tr>
<tr>
<td>$A_2 + A_6 \rightarrow A_7$</td>
<td>$3.62 \times 10^{-4}X_{A_2}X_{A_6}$</td>
<td>$A_4 \rightarrow \emptyset$</td>
<td>$7.40 \times 10^{-11}X_{A_4}$</td>
</tr>
</tbody>
</table>

---

\(^1\) Hye Won Kang, presentation at SPA in 2007.
Some examples

**E. coli Heat Shock Response Model.** 9 species, 18 reactions.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Intensity</th>
<th>Reaction</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset \rightarrow A_8$</td>
<td>$4.00 \times 10^0$</td>
<td>$A_6 + A_8 \rightarrow A_9$</td>
<td>$3.62 \times 10^{-4} X_{A_6} X_{A_8}$</td>
</tr>
<tr>
<td>$A_2 \rightarrow A_3$</td>
<td>$7.00 \times 10^{-1} X_{A_2}$</td>
<td>$A_8 \rightarrow \emptyset$</td>
<td>$9.99 \times 10^{-5} X_{A_8}$</td>
</tr>
<tr>
<td>$A_3 \rightarrow A_2$</td>
<td>$1.30 \times 10^{-1} X_{A_3}$</td>
<td>$A_9 \rightarrow A_6 + A_8$</td>
<td>$4.40 \times 10^{-5} X_{A_9}$</td>
</tr>
<tr>
<td>$\emptyset \xrightarrow{A_1} A_2$</td>
<td>$7.00 \times 10^{-3} X_{A_1}$</td>
<td>$\emptyset \rightarrow A_1$</td>
<td>$1.40 \times 10^{-5}$</td>
</tr>
<tr>
<td>stuff $+ A_3 \rightarrow A_5 + A_2$</td>
<td>$6.30 \times 10^{-3} X_{A_3}$</td>
<td>$A_1 \rightarrow \emptyset$</td>
<td>$1.40 \times 10^{-6} X_{A_1}$</td>
</tr>
<tr>
<td>stuff $+ A_3 \rightarrow A_4 + A_2$</td>
<td>$4.88 \times 10^{-3} X_{A_3}$</td>
<td>$A_7 \xrightarrow{A_4} A_6$</td>
<td>$1.42 \times 10^{-6} X_{A_4} X_{A_7}$</td>
</tr>
<tr>
<td>stuff $+ A_3 \rightarrow A_6 + A_2$</td>
<td>$4.88 \times 10^{-3} X_{A_3}$</td>
<td>$A_5 \rightarrow \emptyset$</td>
<td>$1.80 \times 10^{-8} X_{A_5}$</td>
</tr>
<tr>
<td>$A_7 \rightarrow A_2 + A_6$</td>
<td>$4.40 \times 10^{-4} X_{A_7}$</td>
<td>$A_6 \rightarrow \emptyset$</td>
<td>$6.40 \times 10^{-10} X_{A_6}$</td>
</tr>
<tr>
<td>$A_2 + A_6 \rightarrow A_7$</td>
<td>$3.62 \times 10^{-4} X_{A_2} X_{A_6}$</td>
<td>$A_4 \rightarrow \emptyset$</td>
<td>$7.40 \times 10^{-11} X_{A_4}$</td>
</tr>
</tbody>
</table>

Lotka-Volterra predator-prey model: think of $A$ as a prey and $B$ as a predator.

$$A \xrightarrow{\kappa_1} 2A, \quad A + B \xrightarrow{\kappa_2} 2B, \quad B \xrightarrow{\kappa_3} \emptyset,$$

\[\text{1 Hye Won Kang, presentation at SPA in 2007.}\]
Markov chain models

Path-wise representation

\[ X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s)) ds \right) \zeta_k. \]

is equivalent to Kolmogorov’s forward equation (chemical master equation)

\[ \frac{d}{dt} P_x(t) = \sum_k \lambda_k(x - \zeta_k) P_{x-\zeta_k}(t) - \sum_k \lambda_k(x) P_x(t), \]

where \( P_x(t) \) is probability \( X(t) = x \).
Outline

1. Discuss Poisson process and time changes.


3. → Numerical methods and error approximation.

4. Mathematical question for today: how do we effectively quantify how well different methods approximate the continuous time Markov chain model.

5. Discuss/use multi-scale nature of biochemical reaction networks.
Numerical methods

\[ X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s)) \, ds \right) \zeta_k, \]

**Good news.** There are a number of numerical methods that produce statistically exact sample paths:

→ Gillespie’s algorithm, first reaction method, next reaction method.
Numerical methods

\[ X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s)) ds \right) \zeta_k, \]

**Good news.** There are a number of numerical methods that produce statistically exact sample paths:

→ Gillespie’s algorithm, first reaction method, next reaction method.

For each step of these methods one must find:

(i) the amount of time that passes until the next reaction takes place:

\[ \Delta_n \sim \exp \left( \sum_k \lambda_k(X(t)) \right) \]

(the minimum of exponential RVs)

(ii) which reaction takes place at that time.
Good news. There are a number of numerical methods that produce statistically exact sample paths:
→ Gillespie’s algorithm, first reaction method, next reaction method.

For each step of these methods one must find:
(i) the amount of time that passes until the next reaction takes place:
\[ \Delta_n \sim \exp \left( \sum_k \lambda_k(X(t)) \right) \]
(the minimum of exponential RVs)
(ii) which reaction takes place at that time.

Bad news. If \( \sum_k \lambda_k(X(t)) \gg 1 \), then \[ \mathbb{E} \Delta_n = \frac{1}{\sum_k \lambda_k(X(t))} \ll 1 \]
* time to produce a single path over an interval \([0, T]\) can be prohibitive.
Explicit “$\tau$-leaping” \(^2\) was developed by Dan Gillespie in an effort to overcome the problem that $\Delta_n$ may be prohibitively small.

**Tau-leaping** is essentially an **Euler approximation** of \(\int_0^t \lambda_k(X(s)) \, ds\):

\[
Z(\tau) = Z(0) + \sum_k Y_k \left( \int_0^\tau \lambda_k(Z(s)) \, ds \right) \zeta_k
\]

\[
\approx Z(0) + \sum_k Y_k \left( \lambda_k(Z(0)) \, \tau \right) \zeta_k
\]

\[
\overset{d}{=} Z(0) + \sum_k \text{Poisson} \left( \lambda_k(Z(0)) \, \tau \right) \zeta_k.
\]
One “intuitive” representation for $Z(t)$ generated by $\tau$-leaping is

$$Z(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(Z \circ \eta(s)) \, ds \right) \zeta_k,$$

where

$$\eta(s) = t_n, \quad \text{if} \quad t_n \leq s < t_{n+1} = t_n + \tau$$

is a step function giving left endpoints of time discretization.
Another algorithm: A midpoint method

For a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$, with $\tau = t_n - t_{n-1}$, let

$$\rho(z) = z + \frac{1}{2} \tau \sum_k \lambda_k(z) \zeta_k,$$

be a “deterministic” midpoint approximation

and let $\mathcal{Z}(t)$ solve:

$$\mathcal{Z}(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k \circ \rho(\mathcal{Z} \circ \eta(s)) ds \right) \zeta_k,$$

where $\eta(s) = t_n$, if $t_n \leq s < t_{n+1} = t_n + \tau$. 
A new algorithm: Weak trapezoidal method

Fixing a $\theta \in (0, 1)$, we define

$$\xi_1 = \frac{1}{2} \frac{1}{\theta(1-\theta)}$$ \quad and \quad $$\xi_2 = \frac{1}{2} \frac{(1-\theta)^2 + \theta^2}{\theta(1-\theta)} .$$

Repeat the following:

1. Set $y^* = Z(t_n) + \sum_k y_{k,n,1}(\lambda_k(Z(t_n))\theta h)\zeta_k$.
2. Set $Z(t_{n+1}) = y^* + \sum_k y_{k,n,2}([\xi_1 \lambda_k(y^*) - \xi_2 \lambda_k(t_n)]^+(1-\theta)h)\zeta_k$.
3. Set $t_{n+1} = t_n + h$.

Notes:
1. If $\theta = 1/2$, then $\xi_1 = 2$ and $\xi_2 = 1$.
2. First step is just Euler's method on partial interval.
Weak trapezoidal method: intuition

Alternative representation for

\[ X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s)) ds \right) \zeta_k, \]

is

\[ X(t) = X(0) + \sum_k \zeta_k \int_0^t \int_{\infty} \mathbf{1}_{[0,\lambda_k(X(s))]}(u) \mu_k(du \times ds), \]

where \( \mu_k \) are independent Poisson random measures with Lebesgue mean measure (area):
Weak trapezoidal method: intuition

Alternative representation for

\[ X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda_k(X(s)) ds \right) \zeta_k, \]

is

\[ X(t) = X(0) + \sum_k \zeta_k \int_0^t \int_0^\infty 1_{[0,\lambda_k(X(s))]}(u) \mu_k(du \times ds), \]

where \( \mu_k \) are independent Poisson random measures with Lebesgue mean measure (area):

1. For disjoint \( A, B \subset \mathbb{R}^2 \),

2. \( \mu_k(A) \) and \( \mu_k(B) \) are Poisson with parameters \( \text{Area}(A) \) and \( \text{Area}(B) \), respectively.
Why this works: $\theta = 1/2$
Why this works: $\theta = 1/2$
Why this works: $\theta = 1/2$
Error analysis

Under the scaling $h \to 0$:

1. Li$^3$ and also Rathinam, Petzold, Cao, and Gillespie$^4$ showed, among other things (implicit methods, etc.), that explicit Euler tau-leaping has a strong error (in the $L^2$ norm) of order $1/2$ and a weak error of order one:

$$\sup_{n \leq N} \sqrt{\mathbb{E}_{x_0} |Z(t_n) - X(t_n)|^2} \leq Ch^{1/2}$$

$$|\mathbb{E}_{x_0} f(Z(T)) - \mathbb{E}_{x_0} f(X(T))| \leq Ch,$$

where $0 = t_0 < t_1 < \cdots < t_N = T$ is a partition of $[0, T]$.

2. The midpoint method has the same order accurate as explicit Euler tau-leaping as $h \to 0$.

---


Error analysis

Under the scaling $h \to 0$:

1. Li$^3$ and also Rathinam, Petzold, Cao, and Gillespie$^4$ showed, among other things (implicit methods, etc.), that explicit Euler tau-leaping has a strong error (in the $L^2$ norm) of order $1/2$ and a weak error of order one:

$$\sup_{n \leq N} \sqrt{\mathbb{E}_X |Z(t_n) - X(t_n)|^2} \leq Ch^{1/2}$$

$$|\mathbb{E}_X f(Z(T)) - \mathbb{E}_X f(X(T))| \leq Ch,$$

where $0 = t_0 < t_1 < \cdots < t_N = T$ is a partition of $[0, T]$.

2. The midpoint method has the same order accurate as explicit Euler tau-leaping as $h \to 0$.

3. Many examples show midpoint method more accurate than Euler: led to work in “classical scaling.”$^5$


---

Example

Again think of $A$ as a prey and $B$ as a predator.

\[
A \xrightarrow{\kappa_1} 2A, \quad A + B \xrightarrow{\kappa_2} 2B, \quad B \xrightarrow{\kappa_3} \emptyset,
\]

with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = .002$, $\kappa_3 = 2$.

Letting $\tau = 1/20$ and simulating 30,000 sample paths with each method yields the following approximate distributions for $B(10)$:
Recall, tau-leaping methods are used when \( h \gg E\Delta_n \), for otherwise an exact method would be performed. Therefore, we should require that

\[
h \gg \frac{1}{\sum_k \lambda_k(X(t))} = E\Delta_n\quad \text{while} \quad \sum_k \lambda_k(X(t)) \gg 1.
\]

So \( h \to 0 \) does not tell the whole story \cdots what to do?
Recall, tau-leaping methods are used when $h \gg \mathbb{E} \Delta_n$, for otherwise an exact method would be performed. Therefore, we should require that

$$h \gg \frac{1}{\sum_k \lambda_k(X(t))} = \mathbb{E} \Delta_n$$

while

$$\sum_k \lambda_k(X(t)) \gg 1.$$ 

So $h \to 0$ does not tell the whole story ⋯ what to do?

Explicitly take the multi-scale nature of the systems into account.
Let \( N \gg 1 \). Assume that we are given a model of the form

\[
X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda'_k(X(s)) ds \right) \zeta_k,
\]

where the \( \lambda'_k \) are of the form

\[
\lambda'_k(x) = \kappa'_k \prod_i \nu_{ik}! \prod_i \binom{x_i}{\nu_{ik}}.
\]
Let \( N \gg 1 \). Assume that we are given a model of the form

\[
X(t) = X(0) + \sum_k Y_k \left( \int_0^t \lambda'_k(X(s)) ds \right) \zeta_k,
\]

where the \( \lambda'_k \) are of the form

\[
\lambda'_k(x) = \kappa'_k \prod_i \nu_{ik}! \prod_i \left( \frac{x_i}{\nu_{ik}} \right).
\]

For each species \( i \), define the normalized abundance

\[
X_i^N(t) = \frac{X_i(t)}{N^{\alpha_i}},
\]

where \( \alpha_i \geq 0 \) should be selected so that \( X_i^N = O(1) \).
Rate constants, $\kappa'_k$, may also vary over several orders of magnitude. We write

$$\kappa'_k = \kappa_k N^{\beta_k}$$

where the $\beta_k$ are selected so that $\kappa_k = O(1)$. 

Rate constants, $\kappa'_k$, may also vary over several orders of magnitude. We write

$$\kappa'_k = \kappa_k N^{\beta_k}$$

where the $\beta_k$ are selected so that $\kappa_k = O(1)$.

Mass action kinetics implies

$$\lambda'_k(x) = N^{\beta_k + \nu_k \cdot \alpha} \lambda_k(z)$$

where

$$z_i = \frac{x_i}{N^{\alpha_i}}.$$
Our model has become

\[ X_i^N(t) = X_i^N(0) + \sum_k N^{-\alpha_i} Y_k \left( N^{\beta_k+\nu_k \cdot \alpha} \int_0^t \lambda_k(X^N(s)) ds \right) \zeta_k, \quad i \in \{1, \ldots, d\}. \]
Mulit-scale nature of biochemical reaction networks

Our model has become

\[ X_i^N(t) = X_i^N(0) + \sum_k N^{-\alpha_i} Y_k \left( N^\beta_k + \nu_k \cdot \alpha \int_0^t \lambda_k(X^N(s)) ds \right) \zeta_k, \quad i \in \{1, \ldots, d\}. \]

Remark:

If

\[ \beta_k + \nu_k \cdot \alpha = \alpha_i = 1 \]

for all \( i, k \), then model is

\[ X^N(t) = X^N(0) + \sum_k \frac{1}{N} Y_k \left( N \int_0^t \lambda_k(X^N(s)) ds \right) \zeta_k. \]

which is typically called the classical scaling.

- In this case it is natural to consider \( X^N \) as a vector whose \( i \)th component gives the concentration, in moles per unit volume, of the \( i \)th species.
- For large \( N \), can use Langevin/diffusion approximation.
- As \( N \to \infty \), converge to ODE model with mass action kinetics.
Importance of simulation

Our model has become

\[ X_i^N(t) = X_i^N(0) + \sum_k N^{-\alpha_i} Y_k \left( N^{\beta_k + \nu_k \cdot \alpha} \int_0^t \lambda_k(X^N(s)) ds \right) \zeta_k, \quad i \in \{1, \ldots, d\}. \]

In this general setting, methods of model simplification/reduction:

1. Langevin/Diffusion approximations, or
2. linear noise/Van Kampen approximations, or
3. Averaging ideas to reduce system, or
4. ODE/Law of large numbers types of approximations

are exceedingly difficult.

Very active line of current research: Kurtz, Popovic, Kang, etc.
Weak error analysis on normalized processes

\[ X_i^N(t) = X_i^N(0) + \sum_k N^{-\alpha_i} Y_k \left( N^{\beta_k+\nu_k \cdot \alpha} \int_0^t \lambda_k(X^N(s)) \, ds \right) \zeta_k, \quad i \in \{1, \ldots, d\}. \]

- Denote exact process as \( X^N \) and approximate by \( Z^N \).
- For \( f \in C_0(\mathbb{R}^d, \mathbb{R}) \), we define the operators \( \mathcal{P}_t \) and \( P_t \) by
  \[
  \mathcal{P}_t f(x) \overset{\text{def}}{=} \mathbb{E}_x f(X^N(t)) \]
  \[
  P_t f(x) \overset{\text{def}}{=} \mathbb{E}_x f(Z^N(t)).
  \]

Want to understand

\[
\mathcal{P}_t f(x) - P_t f(x) = \mathbb{E}_x f(X^N(t)) - \mathbb{E}_x f(Z^N(t)).
\]
Weak error analysis

\[ \mathcal{P}_t f(x) \overset{\text{def}}{=} \mathbb{E}_x f(X^N(t)), \quad P_t f(x) \overset{\text{def}}{=} \mathbb{E}_x f(Z^N(t)). \]

\[ |\mathbb{E}_x f(X^N(t)) - \mathbb{E}_x f(Z^N(t))| = |\mathcal{P}_t f(x) - P_t f(x)| \leq \|\mathcal{P}_t - P_t\|_{m \to 0} \|f\|_{\nabla^N} \]

Theorem (A., Koyama – 2011) For any \( n, m \geq 0, \) and \( h > 0 \) (think: \( T = nh \))

\[ \|\mathcal{P}_n h - \mathcal{P}_n h\|_{\nabla^N} \rightarrow 0 = O(n \|\mathcal{P}_h - \mathcal{P}_h\|_{\nabla^N} \max_{\ell \in \{1, \ldots, n\}} \|\mathcal{P}_\ell h\|_{\nabla^N}). \]

Points:
1. Last part is independent of the method. Just depends on behavior of exact process.
2. Rest says: one step (local error) is \( O(h^q + 1) \) \( \Rightarrow \) global error is \( O(h^q) \).
Weak error analysis

\[ P_t f(x) \overset{\text{def}}{=} \mathbb{E}_x f(X^N(t)), \quad P_t f(x) \overset{\text{def}}{=} \mathbb{E}_x f(Z^N(t)). \]

\[ |\mathbb{E}_x f(X^N(t)) - \mathbb{E}_x f(Z^N(t))| = |P_t f(x) - P_t f(x)| \leq \|P_t - P_t\|_{\nabla^N m \to 0} \|f\|_{\nabla^N m} \]

**Theorem** (A., Koyama – 2011) For any \( n, m \geq 0 \), and \( h > 0 \) (think: \( T = nh \))

\[ \|P^n_h - P^m_{nh}\|_{\nabla^N m \to 0} = O(n \|P_h - P_h\|_{\nabla^N m \to 0} \max_{\ell \in \{1, \ldots, n\}} \{\|P_\ell h\|_{\nabla^N m \to m}\}) \]

**Points:**

1. Last part is independent of the method.
   Just depends on behavior of exact process.

2. Rest says:

   one step (local error) is \( O(h^{q+1}) \) \( \implies \) global error is \( O(h^q) \).

\( \implies \) Can focus on local, one step error.
One step error analyses

**Theorem** (A., Koyama – 2011) For forward Euler, if \( h < N^{-\gamma} \)

\[
\| P_h - P_h \|_{2 \to 0}^{\nabla N} \in O(h^2).
\]

**Corollary** If \( Z^N \) is generated via forward Euler, then

\[
\sup_{t \leq T} | \mathbb{E}_x f(X^N(t)) - \mathbb{E}_x f(Z^N(t)) | = O(h).
\]
One step error analyses

Define $\eta_k$ via

$$|\zeta_k^N| = O(N^{-\eta_k}).$$

Theorem (A., Koyama – 2011) For midpoint method, if $h < N^{-\gamma}$

$$\|P_h - \mathcal{P}_h\|_{3 \to 0} = O(h^3 + N^{-\min\{\eta_k\}} h^2).$$

Corollay If $Z^N$ is generated via midpoint method, then

$$\sup_{t \leq T} |\mathbb{E}_x f(X^N(t)) - \mathbb{E}_x f(Z^N(t))| = O(h^2 + N^{-\min\{\eta_k\}} h).$$
One step error analyses

**Theorem** (A., Koyama – 2011) For weak trapezoidal, if $h < N^{-\gamma}$

$$\| P_h - \mathcal{P}_h \|_{3 \rightarrow 0}^{\nabla N} \in O(h^3).$$

**Corollary** If $Z^N$ is generated via weak trapezoidal, then

$$\sup_{t \leq T} | \mathbb{E}_x f(X^N(t)) - \mathbb{E}_x f(Z^N(t)) | = O(h^2).$$
Example

Consider

$A \xleftarrow{1} B \xrightarrow{1} C,$

with $X(0) = [1000, 1000, 4000]$. Log-Log plot of $|EZ_C(1) - EX_C(1)|$:

Slopes:

Euler = .9957, Midpoint = 2.7, Weak Trap = 2.32.
Example

Consider

\[
A \xrightleftharpoons[1]{1} B \xrightleftharpoons[1]{1} C,
\]

with \(X(0) = [1000, 1000, 4000]\). Log-Log plot of second moments:
Careful!

Consider

\[
A \overset{1}{\leftrightarrow} B \overset{1}{\leftrightarrow} C,
\]

with \( X(0) = [1000, 1000, 4000] \). Log-Log plot of \( |\text{Var}(Z_C(1)) - \text{Var}(X_C(1))| \):

Slopes:

Euler = .9831, \quad \text{Midpoint} = 1.0525, \quad \text{Weak Trap} = 2.2291.
Example

Consider

\[
\begin{align*}
A & \overset{3/100}{\leftrightarrow} B \overset{1/10}{\leftrightarrow} C, \\
1 & 1
\end{align*}
\]

with \( X(0) = [13000, 100, 20] \). Log-Log of \(|\mathbb{E}Z_C(1)^2 - \mathbb{E}X_C(1)^2|\):

![Log-Log Graph]

Slopes: Euler = 1.366, Midpt = 2.03

Slopes: Euler = 1.091, Midpt = 1.12
Example

Consider

\[ \begin{align*}
A^{3/100} & \iff B^{1/10} \iff C, \\
& \quad 1 \quad 1
\end{align*} \]

with \( X(0) = [13000, 100, 20] \). Log-Log of \( |E Z_C(1)^2 - E X_C(1)^2| \):

Slopes: Euler = 1.366, Mdpt = 2.03

Slopes: Euler = 1.091, Mdpt = 1.12
Take home messages

1. Simulation of continuous time Markov chains is easy.....unless it’s not.

2. Calculating error estimates must be done with care.

3. Incorporating the natural multiple scales of the process plays an important role.
Take home messages

1. Simulation of continuous time Markov chains is easy.....unless it’s not.

2. Calculating error estimates must be done with care.

3. Incorporating the natural multiple scales of the process plays an important role.

Thank you!