Stochastic Modeling in Reaction-Transport Processes: Forward and Backward considerations

Stefan Engblom

Division of Scientific Computing
Department of Information Technology
Uppsala University

BSSE Seminar, ETH, March 26, 2013
Today

**Agenda:** give an overview of computational stochastic modeling in (bio-)chemical kinetics, specifically targeting cell biology. I also like to discuss some different possibilities for inverse formulations (“given observations, find the model”).
Program

1. Stochastic modeling
   Brownian motion
   (Bio-)Chemical kinetics
   Spatial chemical kinetics

2. Computations by examples

3. Inverse formulations
   Reaction rates from observations
   Molecular movements from observations
   Optimal rates

Conclusions
Brownian motion
Einstein 1905, & some others...

*Example*: Particle in a fluid.
Brownian motion
Einstein 1905, & some others...

*Example*: Particle in a fluid.

A stochastic model is simpler but depends on randomness.
Stochastic modeling of biochemical reactions

*Example:* Bimolecular reaction $X + Y \rightarrow Z$.

- What is the probability $P(1X$ and $1Y$ reacts in the interval $[0, \Delta t])$?

$$P \propto n_X \quad \text{("number of } X\text{-molecules")},$$

$$P \propto n_Y,$$

$$P \propto \frac{1}{V},$$

$$P \propto \Delta t = \Rightarrow P(X + Y \rightarrow Z\text{ in the interval } [0, \Delta t]) = \text{const} \cdot n_X n_Y \Delta t / V.$$
Stochastic modeling of biochemical reactions

*Example:* Bimolecular reaction $X + Y \rightarrow Z$.

- What is the probability $P(1X$ and $1Y$ reacts in the interval $[0, \Delta t])$?

- $P \propto n_X$ ("number of $X$-molecules")

- $P \propto n_Y$
Stochastic modeling of biochemical reactions

*Example:* Bimolecular reaction $X + Y \rightarrow Z$.

- What is the probability $P(1X$ and $1Y$ react in the interval $[0, \Delta t])$?

$$
\begin{array}{c|c|c|c}
X & Y & X \\
X & & \\
& X & \\
X & & Y \\
& & Y \\
X & & \\
\end{array}
$$

- $P \propto n_X$ ("number of $X$-molecules")
- $P \propto n_Y$
- $P \propto 1/V$
- $P \propto \Delta t$
Stochastic modeling of biochemical reactions

Example: Bimolecular reaction $X + Y \rightarrow Z$.

- What is the probability $P(1X$ and $1Y$ reacts in the interval $[0, \Delta t])$?

\[
P \propto n_X \text{ ("number of } X\text{-molecules" )}
\]

\[
P \propto n_Y
\]

\[
P \propto 1/V
\]

\[
P \propto \Delta t
\]

$\implies P(X + Y \rightarrow Z \text{ in the interval } [0, \Delta t]) = \text{const} \cdot n_X n_Y \Delta t/V$.

It so happens that this receipt describes a continuous-time Markov chain.
Well-stirred kinetics

**Assumption #1**: the chance of finding a molecule is equal throughout the volume (*homogeneous*).

**Assumption #2**: the energy of a molecule does not depend on its position in the volume (*thermal equilibrium*).
Well-stirred kinetics

Assumption #1: the chance of finding a molecule is equal throughout the volume (homogeneous).
Assumption #2: the energy of a molecule does not depend on its position in the volume (thermal equilibrium).

- State vector $x \in \mathbb{Z}_+^D$ counting the number of molecules of each of $D$ species.
- $R$ specified reactions defined as transitions between these states,

\[
\begin{align*}
x & \xrightarrow{w_r(x)} x - N_r, \\
N & \in \mathbb{Z}^{D \times R} \text{ (stoichiometric matrix)}
\end{align*}
\]

where each transition intensity or propensity $w_r : \mathbb{Z}_+^D \to \mathbb{R}_+$ is the probability of reacting per unit of time.
Simulating the chain

(Doob ’45, Gillespie ’76)

Simulate a single stochastic trajectory $X(t)$ “an outcome”:

0. Let $t = 0$ and set the state $x$ to the initial number of molecules.

1. Compute the total reaction intensity $W := \sum_r w_r(x)$. Generate the time to the next reaction $\tau := -W^{-1} \log u_1$ where $u_1 \in (0, 1)$ is a uniform random number. Determine also the next reaction $r$ by the requirement that

$$\sum_{s=1}^{r-1} w_s(x) < Wu_2 \leq \sum_{s=1}^{r} w_s(x),$$

where $u_2$ is again a uniform random deviate in $(0, 1)$.

2. Update the state of the system by setting $t := t + \tau$ and $x := x - N_r$.

3. Repeat from step 1 until some final time $T$ is reached.
Kolmogorov’s forward differential system/Master equation
(Kolmogorov ’31, Nordsieck/Lamb/Uhlenbeck ’40)

With states $x \in \mathbb{Z}_+^D$, let $p(x, t) := P(X(t) = x|X(0))$. Then the chemical master equation (CME) is given by

$$
\frac{\partial p(x, t)}{\partial t} = \sum_{r=1}^{R} w_r(x + \mathbb{N}_r)p(x + \mathbb{N}_r, t) - \sum_{r=1}^{R} w_r(x)p(x, t) =: \mathcal{M}p.
$$

-A gain-loss discrete PDE in $D$ dimensions for the probability density conditioned upon an initial state.
Inhomogeneous kinetics

*Not* well-stirred:

- When the molecular movement (diffusion) is slow compared to the reaction intensity — large *local* concentrations may easily build up.
- When some reactions are *localized* — e.g. depend on an enzyme emitted from a precise position, or are located to, say, a membrane.

These conditions are not unusual for reactions taking place inside living cells!
Mesoscopic spatial kinetics

-Not well-stirred in the whole volume, but if the domain $\Omega$ is subdivided into smaller computational cells $\Omega_j$ such that their individual volume $|\Omega_j|$ is small, then diffusion suffices to make each cell well-stirred.

**Figure:** Primal mesh (solid), dual mesh (dashed). The nodal dofs are the # of molecules in each dual cell.
Mesoscopic spatial kinetics (cont)

- $D$ chemically active species $X_{ij}$ for $i = 1, \ldots, D$ but now counted separately in $K$ cells, $j = 1, \ldots, K$.
- The state of the system is now an array $\mathbf{x}$ with $D \times K$ elements.
- This state is changed by chemical reactions occurring between the molecules in the same cell (vertically in $\mathbf{x}$) and by diffusion/transport where molecules move to adjacent cells (horizontally in $\mathbf{x}$).
Reactions

By assumption, each cell is well-stirred and consequently the master equation is valid as a description of reactions,

\[ \frac{\partial p(x, t)}{\partial t} = \mathcal{M} p(x, t) := \]

\[ \sum_{j=1}^{K} \sum_{r=1}^{R} w_r(x_j + \mathbb{N}_r)p(x_1, \ldots, x_j + \mathbb{N}_r, \ldots, x_K, t) \]

\[ - \sum_{j=1}^{K} \sum_{r=1}^{R} w_r(x_j)p(x, t). \]
Diffusion

A natural model of diffusion from one cell $\Omega_k$ to another cell $\Omega_j$ is

$$X_{ik} \xrightarrow{q_{kj}x_{ik}} X_{ij},$$

where $q_{kj}$ is non-zero only for connected cells.
Diffusion

A natural model of diffusion from one cell $\Omega_k$ to another cell $\Omega_j$ is

$$X_{ik} \xrightarrow{q_{kj} x_{ik}} X_{ij},$$

where $q_{kj}$ is non-zero only for connected cells.

Ideally, $q_{kj}$ should be taken as the inverse of the \textit{mean first exit time} for a single molecule of species $i$ from cell $\Omega_k$ to $\Omega_j$. $\implies q_{kj} \propto \sigma^2/h^2$, where $\sigma^2/2$ is the macroscopic diffusion, $h$ the local length.
Diffusion

A natural model of diffusion from one cell $\Omega_k$ to another cell $\Omega_j$ is

$$X_{ik} \xrightarrow{q_{kj} \cdot x_{ik}} X_{ij},$$

where $q_{kj}$ is non-zero only for connected cells.

Ideally, $q_{kj}$ should be taken as the inverse of the mean first exit time for a single molecule of species $i$ from cell $\Omega_k$ to $\Omega_j$. $\implies q_{kj} \propto \sigma^2 / h^2$, where $\sigma^2 / 2$ is the macroscopic diffusion, $h$ the local length.

The diffusion master equation can therefore be written

$$\frac{\partial p(x, t)}{\partial t} = \sum_{i=1}^{D} \sum_{k=1}^{K} \sum_{j=1}^{K} q_{kj} (x_{ik} + M_{kj,k}) p(x_1, \ldots, x_i + M_{kj}, \ldots, x_D, t) - q_{kj} x_{ik} p(x, t) =: D p(x, t).$$

The transition vector $M_{kj}$ is zero except for $M_{kj,k} = -M_{kj,j} = 1$. 
The reaction-diffusion master equation
“RDME”

Combining reactions with diffusions,

\[ \frac{\partial p(x, t)}{\partial t} = (\mathcal{M} + \mathcal{D})p(x, t). \]
The reaction-diffusion master equation
“RDME”

Combining reactions with diffusions,

$$\frac{\partial p(x, t)}{\partial t} = (\mathcal{M} + \mathcal{D})p(x, t).$$

-An approximation! Valid when

$$\rho^2 \ll h^2 \ll \sigma^2 \tau_{\Delta},$$

$\rho$ the molecular radius, $\tau_{\Delta}$ average molecular survival time.

-Once formulated, any algorithm for sampling from the CME can also simulate the RDME. For a spatially resolved model, most of the simulation time is spent on diffusion events.
Unstructured meshes

-Mean first exit time only known for very simple geometries (e.g. circles).

-How to handle complicated geometries?
Unstructured meshes

- Mean first exit time only known for very simple geometries (e.g. circles).
- *How to handle complicated geometries?* Attempt to converge in expectation to the *macroscopic diffusion equation*. Briefly, a numerical method applied to $u_t = \sigma^2 / 2 \Delta u$ yields the *discretized* form

$$
\frac{du}{dt} = \frac{\sigma^2}{2} Du.
$$
Unstructured meshes

-Mean first exit time only known for very simple geometries (e.g. circles).

-How to handle complicated geometries? Attempt to converge in expectation to the macroscopic diffusion equation. Briefly, a numerical method applied to \( u_t = \sigma^2/2 \Delta u \) yields the discretized form

\[
\frac{d\mathbf{u}}{dt} = \frac{\sigma^2}{2} \mathbf{D}\mathbf{u}.
\]

-Define \( \varphi_{ij} = E \Omega_j^{-1} x_{ij} \). By linearity of the diffusion intensities, the diffusion master equation implies

\[
\frac{d\varphi_{ij}}{dt} = \sum_{k=1}^{K} \frac{\Omega_k}{|\Omega_j|} q_{kj} \varphi_{ik} - \left( \sum_{k=1}^{K} q_{jk} \right) \varphi_{ij},
\]

\[\iff\]

\[
\frac{d\varphi_{i.}}{dt} = \mathbf{Q}\varphi_{i.}.
\]
Assuming point-wise convergence of the numerical discretization → diffusion PDE, the consistency in this interpretation ensures convergence in distribution to the correct Brownian motion as the mesh-size $h \to 0$. 
(Forward) Computations by examples

- Bistable model
- Spatial oscillations in *E. coli*.
Bistable double-negative feedback system

\[ E_A \xrightarrow{k_1} E_A + A \quad E_B \xrightarrow{k_1} E_B + B \]

\[ E_A + B \xrightarrow[k_a]{k_d} E_A B \quad E_B + A \xrightarrow[k_a]{k_d} E_B A \]

\[ E_A B + B \xrightarrow[k_a]{k_d} E_A B_2 \quad E_B A + A \xrightarrow[k_a]{k_d} E_B A_2 \]

\[ A \xrightarrow{k_4} \emptyset \quad B \xrightarrow{k_4} \emptyset \]

Slow/intermediate/fast diffusion in a simple model of an *S. cerevisiae* cell with internal structures in the form of a nucleus and a large vacuole. Molecules are not allowed to diffuse across the membranes and enter the organelles.
2. Computations by examples

(a) Species A.

(b) Species B.

Figure: The total number of $A$ and $B$ molecules as the diffusion constant is varied. *Right:* local bistability is lost.
MinD oscillations

Oscillations of proteins involved in the cell division of *E. coli*:

\[
\begin{align*}
\text{MinD\_c\_atp} & \xrightarrow{k_d} \text{MinD\_m} & \text{MinD\_c\_atp} + \text{MinD\_m} & \xrightarrow{k_{dD}} 2\text{MinD\_m} \\
\text{Min\_e} + \text{MinD\_m} & \xrightarrow{k_{de}} \text{MinDE} & \text{MinDE} & \xrightarrow{k_e} \text{MinD\_c\_adp} + \text{Min\_e} \\
\text{MinD\_c\_adp} & \xrightarrow{k_p} \text{MinD\_c\_atp}
\end{align*}
\]

“URDME” software [www.urdme.org](http://www.urdme.org).
2. Computations by examples

Average concentration of MinD$_m$

MinD$_m$ polar oscillations

S. Engblom (Uppsala University)
2. Computations by examples

Bacterium length (µ m)

Growth (µ m)

Average concentration of MinD

<table>
<thead>
<tr>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>x 10^{-6}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
</tr>
<tr>
<td>0.4</td>
</tr>
<tr>
<td>0.6</td>
</tr>
<tr>
<td>0.8</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

Average concentration of MinD

<table>
<thead>
<tr>
<th>location (µ m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
</tr>
<tr>
<td>1100</td>
</tr>
<tr>
<td>1200</td>
</tr>
<tr>
<td>1300</td>
</tr>
<tr>
<td>1400</td>
</tr>
<tr>
<td>1500</td>
</tr>
</tbody>
</table>

MinD polar oscillations

time (s)

<table>
<thead>
<tr>
<th># molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
</tr>
<tr>
<td>Right</td>
</tr>
</tbody>
</table>

Figure 3: (A) Geometry and mesh modeling of an E. Coli cell. (B) Temporal average concentration of MinD protein as a function of position along the long axis of the E. Coli cell (top), and the time series plot of the oscillations. (C) Six E. Coli cells of increasing lengths, as specified in the parameter sweep described in Table 1. The color intensity shows the temporal average concentration of MinD protein along the membrane. (D) Parameter sweep shows how the relative concentration of MinD changes as the bacterium grows.
Inverse or ‘Backwards’ formulations

- Reaction rates from observations...
- ...diffusion rates from observations
- “Evolutionary” optimal control setup

None of these formulations are in a ‘final’ state.
Rate coefficients from observations

**Physics:** linear birth-death process with hidden parameters \((k, \mu)\):

\[
\emptyset \xrightarrow{k} X \quad X \xrightarrow{\mu X} \emptyset
\]

**Task:** find the rates \((k, \mu)\).

**Convergence:** increasing the temporal resolution:

![Graph](image1)

![Graph](image2)
Maximum Likelihood

*In a nutshell:* find the parameters \((k, \mu)\) that maximizes the probability of obtaining the data we did observe (…)

*Model:* let us assume independent observations, say, Gaussians around a certain predicted value \(x(t)\),

\[
P(X(t_1) = x_1 | X(0) = x_0) \propto \exp \left( -\frac{[x_1 - x(t_1 | x_0, t_0)]^2}{2\sigma^2} \right)
\]

\[
P(X(t_2) = x_2 | X(t_1) = x_1) \propto \exp \left( -\frac{[x_2 - x(t_2 | x_1, t_1)]^2}{2\sigma^2} \right)
\]

…
Maximum Likelihood

*In a nutshell:* find the parameters \((k, \mu)\) that maximizes the probability of obtaining the data we did observe (…)

- *Model:* let us assume independent observations, say, Gaussians around a certain predicted value \(x(t)\),

\[
P(X(t_1) = x_1 | X(0) = x_0) \propto \exp \left( - \left[ x_1 - x(t_1 | x_0, t_0) \right]^2 / 2\sigma^2 \right)
\]

\[
P(X(t_2) = x_2 | X(t_1) = x_1) \propto \exp \left( - \left[ x_2 - x(t_2 | x_1, t_1) \right]^2 / 2\sigma^2 \right)
\]

\[
\ldots
\]

indep. \(\implies\) \(P \propto \exp \left( - \frac{1}{2\sigma^2} \sum_i [x_i - x(t_i | x_{i-1}, t_{i-1})]^2 \right)

\[
\text{minimize}
\]
3. Inverse formulations

Reaction rates from observations

**Maximum Likelihood**

*In a nutshell:* find the parameters \((k, \mu)\) that *maximizes* the probability of obtaining the data we did observe (...)

**Model:** let us assume independent observations, say, Gaussians around a certain predicted value \(x(t)\),

\[
P(X(t_1) = x_1 | X(0) = x_0) \propto \exp\left( - \frac{[x_1 - x(t_1 | x_0, t_0)]^2}{2\sigma^2} \right)
\]

\[
P(X(t_2) = x_2 | X(t_1) = x_1) \propto \exp\left( - \frac{[x_2 - x(t_2 | x_1, t_1)]^2}{2\sigma^2} \right)
\]

... 

\[
\text{indep. } \implies P \propto \exp\left( - \frac{1}{2\sigma^2} \sum_i [x_i - x(t_i | x_{i-1}, t_{i-1})]^2 \right)
\]

minimize

-Linear birth-death ODE is \(x'(t) = k - \mu x(t)\). Use for the predictor \(x(t | x_i, t_i)\) the solution to the ODE at time \(t\) given initial data \((x_i, t_i)\).
Linear birth-death

Results ($\mu$)

![Plot of linear birth-death results](image)
Dimerization

Slightly more difficult (nonlinear)

\[
\begin{align*}
\emptyset & \xrightarrow{k} X \\
X + X & \xrightarrow{\nu x(x-1)} \emptyset
\end{align*}
\]
3. Inverse formulations

Reaction rates from observations

Dimerization

Results (\(\nu\))

![Graph showing dimerization results](https://example.com/graph.png)
Maximum Likelihood using the CME

- Previously we used a Gaussian probability model. A better model is the chemical master equation.

\[ P(X(t_i) = x_i | X(t_{i-1}) = x_{i-1}) = p(x_i, t_i) \text{ with } p \text{ a solution to the CME with initial data } p(x_{i-1}, t_{i-1}) = 1. \]

- Observations are still independent \((Markov \ property)\).
3. Inverse formulations

Reaction rates from observations

Dimerization

Results ($\nu$)

![Graph showing dimerization results](image)

- **ODE**
- **M−eq.**

Estimate

- **Exact**
Diffusion rates from observations

Physics: a single particle at position $Z(t)$ undergoing 2D Brownian motion with hidden diffusion constant: $dZ_t = \sigma \, dW_t$, where $Z_t = [X_t \ Y_t]$, $W_t = [W_{t}^{(x)} \ W_{t}^{(y)}]$.

Data: $N = 1000$ observations, $\sigma \in \{1, 4\}$. Only $\sim 50$ observations from within the quarter circle where $\sigma = 4$.

Task: determine $\sigma_{1,2}$ and classify the observations accordingly (hence determine $\sigma = \sigma(x, y)$).
Expectation-Maximization algorithm

Briefly, an iterative algorithm which upon convergence produces a Maximum-Likelihood estimator of (i) $\sigma = [\sigma_1 \, \sigma_2]$ as well as of (ii) $p_{nk}$, the probability that the $n$th observation had diffusion constant $\sigma_k$. 
3. Inverse formulations

Molecular movements from observations

**Expectation-Maximization algorithm**

Briefly, an iterative algorithm which upon convergence produces a Maximum-Likelihood estimator of $(i)$ $\sigma = [\sigma_1 \sigma_2]$ as well as of $(ii)$ $p_{nk}$, the probability that the $n$th observation had diffusion constant $\sigma_k$.

1. *Given* values of $\sigma_{1,2}$, we can estimate $p_{nk}$. (Gaussian increments)
2. *Given* values of $p_{nk}$, we can estimate $\sigma_{1,2}$. (Sample means)

The iteration defined by iterating step #1 and 2 is (a version of) the *Expectation-Maximization* algorithm.
Diffusion rates
Results iteration #2

\[ \sigma = [0.97 \ 1.95] \]
Diffusion rates
Results iteration #3

\[ \sigma = [0.97 \ 3.53] \]
Diffusion rates

Results iteration #6

\[ \sigma = [0.98 \ 3.68] \]
Optimal control of rates

Enzymatic reaction of a complex into a product,

\[ C + E \xrightarrow{\nu C \cdot E} P + E. \]

Combine with

\[
\emptyset \underset{\mu E E}{\xrightleftharpoons{s(t)}} E, \quad \emptyset \underset{\mu C C}{\xrightleftharpoons{k_C}} C, \quad P \xrightarrow{\mu P P} \emptyset
\]

such that \( E \) is under \textit{control} through the signal \( s(t) \).
Optimal control of rates (cont)

Maximize

\[ M[P] := \int_0^T \varphi(P_t) \, dt, \]

with a nonlinear payoff function \( \varphi(P) \),

\[
\begin{align*}
\varphi(P) &= 0, & P \leq c_- \\
\varphi(P) &= \tau(P - c_-), & c_- < P \leq C_+ \\
\varphi(P) &= \tau(C_+ - c_-), & C_+ < P
\end{align*}
\]

Constraints on the production signal \( s \)

\[
\begin{align*}
\max_{t \in [0, T]} s(t) &\leq S_\infty, \\
\int_0^T s(t) \, dt &\leq S_1,
\end{align*}
\]
-Results from non-spatial deterministic ODE.
Inverse formulations

Optimal rates

\[ \mathcal{M}_2[P] := \int_0^T \varphi(P_t) + \varepsilon |s'(t)| \, dt \]
3. Inverse formulations

Optimal rates

---

![Graph showing time series analysis](image-url)
Summary & Conclusions

- Stochastic mesoscopic modeling in chemical kinetics can combine *simplicity* with *accuracy*
- Spatial modeling is also possible and often necessary, computational issues arise due to high temporal resolution
- Free software URDME ([www.urdme.org](http://www.urdme.org))
- Examples of inverse formulations, many possibilities; I like to think that it is important to be *data- and question driven*
- Input is welcome

Thank you for listening!