Mesoscopic Stochastic Modeling: Diffusion Operators, Multiphysics Couplings, and Convergence

Stefan Engblom

Div of Scientific Computing, Dept of Information Technology, Uppsala University

BIRS Workshop, Banff, Alberta, Canada, November 10, 2014
Outline

Session: Mesoscopic Methods and Modeling I

1A. Motivation and notation
   Simplicity & Accuracy
   Well-stirred kinetics
   Personal observations

1B. Spatial modeling
   Unstructured meshes: finite elements vs. volumes
   Convergence (of diffusion)

2. Multiphysics couplings

Summary
Brownian motion

*Example:* Particle in a fluid (Einstein 1905, & others...).

A stochastic model is *simpler* but depends on randomness.
Chemical reactions

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

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

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

\[ P \propto n_Y, \]

\[ P \propto 1/V, \]

\[ P \propto \Delta t \]

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

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

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

  $\begin{array}{ccc}
  X & Y & X \\
  X & & X \\
  & Y & \\
  X & & Y \\
  & Y & \\
  X & & \\
  \end{array}$

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

Let $\Delta t \rightarrow 0$. Then it so happens that this receipt describes a continuous-time Markov chain.

“Simpler, but random.”
Multistability

Figure: Solid: deterministic, dashed: stochastic.
Stochastic resonance

Figure: Solid: deterministic, dashed: stochastic.
Stochastic focusing

Figure: Nonlinear response to twofold signal increase; solid: partially deterministic, dashed: fully stochastic.
Well-stirred kinetics

**Assuming** a homogeneous probability of finding a molecule throughout the volume, and an energy which is independent on position.

- State \( X \in \mathbb{Z}_+^D \), counting the number of molecules of each of \( D \) species.
- Reactions are transitions between these states,

\[
X \xrightarrow{w_r(X)} X - N_r, \quad N \in \mathbb{Z}^{D \times R} \text{ (stoichiometric matrix)}
\]

where the *propensity* \( w_r : \mathbb{Z}_+^D \rightarrow \mathbb{R}_+ \), \( r = 1 \ldots R \), is the probability of reacting per unit of time.

\[\implies \text{Jump SDE formulation: } dX_t = -N \mu(dt)\]
Well-stirred kinetics

Assuming a homogeneous probability of finding a molecule throughout the volume, and an energy which is independent on position.

- State $X \in \mathbb{Z}_+^D$, counting the number of molecules of each of $D$ species.
- Reactions are transitions between these states,

$$
X \xrightarrow{w_r(X)} X - N_r, \quad N \in \mathbb{Z}^{D \times R} \text{ (stoichiometric matrix)}
$$

where the *propensity* $w_r : \mathbb{Z}_+^D \rightarrow \mathbb{R}_+$, $r = 1 \ldots R$, is the probability of reacting per unit of time.

$$
\Longrightarrow \text{ Jump SDE formulation: } dX_t = -N \mu(dt) = -N \mu(w(X_{t-}); dt) \text{ such that } E[\mu(w(X_{t-}); dt)] = w(X_{t-}) dt.
$$
More on notation

Where did the CME go?

Kolmogorov’s forward differential system/Master equation, (Kolmogorov ’31, Nordsieck/Lamb/Uhlenbeck ’40), with $p(x, t) := P(X(t) = x | X(0))$.

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

$$=: \mathcal{M}p.$$
More on notation

Where did the CME go?

Kolmogorov’s forward differential system/Master equation, (Kolmogorov '31, Nordsieck/Lamb/Uhlenbeck ’40), with \( p(x, t) := P(X(t) = x | X(0)) \).

\[
\frac{\partial p(x, t)}{\partial t} = \sum_{r=1}^{R} w_r(x + N_r)p(x + N_r, t) - \sum_{r=1}^{R} w_r(x)p(x, t) =: \mathcal{M}p.
\]

Jump SDE, competing counting measures:

\[
dX_t = -N \mu(w(X_{t-}); dt) \quad \text{("First reaction method")}
\]

or, via a thinning of a single marked counting process,

\[
dX_t = -N \int \hat{w}(X_{t-}; z) \mu(dt \times dz) \quad \text{("Direct method")}
\]
Some personal observations
Mesoscopic Methods and Modeling

- By now, many *methods*, many *acronyms*, a few good ideas, various trade-offs, combined modeling...

- **Who is the customer?**
- **What exactly are the customer’s demands?**
  - accuracy vs. speed trade-off
  - actual point of use
Some personal observations
Mesoscopic Methods and Modeling

-By now, many *methods*, many *acronyms*, a few good ideas, various trade-offs, combined modeling...

- **Who is the customer?**
- What exactly are the customer’s demands?
  - accuracy vs. speed trade-off
  - actual point of use

- Sometimes the customer is another researcher (my *only* experience so far)
  - results of simulations will be reported in a paper
  - well characterized accuracy is highly desirable (eg. backward analysis: method solves exactly a certain perturbed problem)
  - heuristics/unknown accuracy acceptable iff very serious speed-ups
Mesoscopic spatial kinetics

The conditions for well-stirred kinetics are often violated, particularly so for reactions taking place inside living cells.

- 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.
The state of the system is now an array $\mathbf{X}$ with $D \times K$ elements; $D$ chemically active species $X_{ij}$, $i = 1, \ldots, D$, counted separately in $K$ cells, $j = 1, \ldots, K$.  

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 $\mathbf{X}$).
By assumption, each cell is well-stirred and consequently the jump SDE is valid as a description of reactions,

\[ dX_t = -N \mu(dt), \]

where \( \mu \) is now \( R \)-by-\( K \); \( E[\mu_{rj}]dt^{-1} \) = propensity of the \( r \)th reaction in the \( j \)th cell.
Diffusion

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

\[
\mathbf{X}_{ik} \xrightarrow{q_{kj}\mathbf{X}_{ik}} \mathbf{X}_{ij},
\]

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 \( \Omega_j \). \( q_{kj} \propto \frac{\sigma^2}{h^2} \), 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

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

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 $\Omega_j$. $\Rightarrow q_{kj} \propto \sigma^2/h^2$, where $\sigma^2/2$ is the macroscopic diffusion, $h$ the local length.

Assuming that the diffusion constants are the same for all species,

$$dX_t = E(-\nu^T + \nu)(dt),$$

where $E$ is $D$-by-$K$ of all 1's, and $\nu$ is $K$-by-$K$; $E[\nu_{kj}] = q_{kj}X_{ik} dt$. 

The reaction-diffusion jump SDE

“RDME”

Combining reactions with diffusions we arrive at

\[ dX_t = -N\mu(dt) + \mathbb{E}(-\nu^T + \nu)(dt). \]

For example, (“Next Subvolume method”)

\[ dX_t = -N \int_I \hat{\nu}(X_t^-; z) \otimes \bar{\mu}(dt \times dz) + \]
\[ \mathbb{E} \int_I (-\hat{\nu}^T + \hat{\nu})(X_t^-; z) \otimes \bar{\nu}(dt \times dz). \]

-An approximation, valid when

\[ \rho^2 \ll h^2 \ll \sigma^2 \tau_\Delta, \]

\( \rho \) the molecular radius, \( \tau_\Delta \) average molecular survival time.
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. A numerical method applied to \( u_t = \sigma^2 / 2 \Delta u \) yields the discretized form

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

-Define \( \varphi_{ij} = E[\Omega_j^{-1} X_{ij}] \). By linearity of the diffusion intensities,

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

\( \iff \)

\[
\frac{d\varphi_{i.}}{dt} = Q \varphi_{i.}.
\]
FEM vs. FVM

An insane summary

Consider the strong formulation $u_t = \Delta u$ in $\Omega$,

1. Variational form (Green’s theorem): find $u \in V$
   s.t. $(v, u_t) = - (\nabla v, \nabla u)$ for
   $\forall v \in V$, where
   $(f, g) \equiv \int_\Omega fg \, dx$.

2. A FEM is obtained by approximating
   $V \approx V_h = \text{span}_i \phi_i \subset V$.

3. With $u_h = \sum_i u_i(t) \phi_i$ we get
   $Mu_t = -Au$; $M_{ij} = (\phi_i, \phi_j)$,
   $A_{ij} = (\nabla \phi_i, \nabla \phi_j)$.
FEM vs. FVM

An insane summary

Consider the strong formulation \( u_t = \Delta u \) in \( \Omega \),

1. Variational form (Green’s theorem): find \( u \in V \)
   s.t. \( (v, u_t) = - (\nabla v, \nabla u) \) for \( \forall v \in V \), where
   \( (f, g) \equiv \int_\Omega fg \, dx \).

2. A FEM is obtained by approximating
   \[ V \approx V_h = \text{span}_i \varphi_i \subset V. \]

3. With \( u_h = \sum_i u_i(t) \varphi_i \) we get
   \[ Mu_t = -Au; \quad M_{ij} = (\varphi_i, \varphi_j), \]
   \[ A_{ij} = (\nabla \varphi_i, \nabla \varphi_j). \]

1. Integrating over the \( j \)th finite volume and invoking the divergence theorem we get
   \[ \int_{\omega_j} u_t \, dx = \int_{\partial\omega_j} \mathbf{n} \cdot \nabla u \, da. \]

2. Approximating \( \nabla \) with a difference and defining \( u_j \) as a volume average gives
   \[ |\omega_j| d\,/\,dt \ u_j = \sum_k |\partial\omega_{jk}| \ e_{jk}^{-1} (u_k - u_j), \]
   \( e_{jk} \) the distance between nodes \( j \) and \( k \).
Weak convergence

<Key observation>: by linearity, the diffusion CTMC on the unstructured grid has an expected value which coincides with the exact solution to the deterministic numerical method.
FEM convergence

\[ M u_t = -A u \text{ or } u_t = -M^{-1} A u \approx -\tilde{M}^{-1} A u =: D u. \]

1) Converges in \( L^2 \), \( \| u_h - u \| = O(h^2) \) as \( h \to 0 \), under very mild assumptions on the mesh.

2) Under stringent conditions on the mesh, the maximum principle holds.

These conditions are needed to ensure that

\[ D_{jk} \geq 0, \quad D_{jj} < 0, \quad \sum_{k=1}^{K} D_{jk} = 0. \]
FVM convergence

\[ |\omega_j| \frac{d}{dt} u_j = \sum_k |\partial \omega_{jk}| |e_{jk}|^{-1} (u_k - u_j) \]

1) The **maximum principle** always holds.
2) If the mesh is a Delaunay triangulation, the method **converges** as 
\[ \|u_h - u\| = O(h^2) \]. Unfortunately (in 3D) such meshes have a very poor quality except for very simple geometries. Then the “C” in \( O(h^2) \) is very large.
On balance...

- With a (very) good mesh both methods converge as $h \to 0$ and satisfy the maximum principle. With an “average” mesh, (truncated) FEM seems to have an accuracy edge to FVM.

- Importantly, truncated FEM is amenable to backward analysis: the solution satisfies exactly a perturbed equation $u_t = \nabla \cdot \left( \tilde{\sigma}^2(x)/2 \times \nabla u \right)$ where $\tilde{\sigma}$ can be explicitly obtained, and where $\|\tilde{\sigma} - \sigma\|$ is small and localized.

- Key challenges: (i) convergence in distribution — retrieving the correct Brownian motion, (ii) convergence with reactions, (iii) getting to grip of when it actually matters...
Application: multiscale neuronal model

Joint work with Pavol Bauer and Emil Berwald.
Bottom level
Ion channel gating

Figure: Gating process: sodium channels.
Bottom level
Ion channel gating

The gating process of ion channels can be mesoscopically described as

\[ N_0 \overset{3\alpha_m(V_m)N_0}{\iff} N_1 \overset{2\alpha_m(V_m)N_1}{\iff} N_2 \overset{\alpha_m(V_m)N_2}{\iff} N_3, \]

again a continuous-time Markov chain. Output: \( N_3 \), the number of open gates.

For efficient model coupling we freeze the voltage dependency for a short time-step \( \tau \) ("Euler method/1st order Strang split"):

\[ X_{t+\tau} = X_t - \int_t^{t+\tau} \mathbb{N}\mu(V_m(t), w(X_{s-})); \, ds. \]
Middle level
Membrane dynamics

Figure: Cable equation circuit.
Middle level
Membrane dynamics

\[ I_m = c_m \frac{dV_m}{dt} + \sum_{i \in C_v} \gamma_i N_3^i(t)[V_m(t) - E_i] \]

- Morphological information extracted using the Trees toolbox
- System of current-balance and cable equations is solved for each time step \( \tau \)
Electric field intensity $\mathbf{E}$ in terms of the electric scalar potential $V$, 

$$\mathbf{E} = -\nabla V.$$ 

Trans-membrane current $I_m$ is scaled with the compartment surface area and coupled as a current source,

$$-\nabla \cdot \left( \sigma \nabla V + \varepsilon_0 \varepsilon_r \frac{\partial}{\partial t} \nabla V \right) = \frac{1}{\Omega_c} I_m,$$

with conductivity $\sigma$ and permittivity $\varepsilon$. Finally, the time dependent potential $V$ is solved via finite element methods.
2. Multiphysics couplings

Top level

Geometry coupling

- Bottom and middle level: compartments (cylindrical volumes)
- Coupling with PDE requires a mesh
- Approximation with curves much more efficient than volumetric elements
Coupled solution
Summary

- *Simplicity* and *Accuracy*: Who is the customer?
- Accurately capturing a stochastic nonlinear phenomenon is a very hard constraint for method’s development!

- Spatially inhomogeneous case, consistency with macroscopic equations, FEM vs. FV
- The numerical method’s convergence to the macroscopic equation implies weak convergence of the corresponding stochastic model, **backward analysis**

- Sample multiscale neuronal application solved in URDME ([www.urdme.org](http://www.urdme.org)): coupling different types of models
In case somebody asks: more on notation

Compare with ODEs...

For states $x \in \mathbb{R}^D$; either one may consider

$$\frac{\partial}{\partial t} p(t, x) = -\Delta \cdot (f(x)p(t, x)),$$

or,

$$x'(t) = f(x).$$

-In method’s development and modeling, “$f \rightarrow f_h$”, one almost exclusively deals with the latter representation.