Computational physics

Heiko Rieger and Adam Wysocki

Theoretical Physics
Saarland University

SS2017

Contents

- Stochastic simulation of coupled chemical reactions

Stochastic chemical kinetics

- Consider a well-stirred system with N chemically active species $\{S_1, \ldots, S_N\}$ each with a population X_i (number of molecules) in volume V.
- Species can interact via M types of unidirectional chemical reactions $\{R_1, \ldots, R_M\}$ (reversible reactions can be modeled as two opposite and parallel running processes), for example:

$$
\begin{aligned}
\emptyset &\rightarrow \text{products} \\
S_j &\rightarrow \text{products} \\
S_j + S_k &\rightarrow \text{products} \quad (j \neq k) \\
2S_j &\rightarrow \text{products} \\
S_i + S_j + S_k &\rightarrow \text{products} \quad (i \neq j \neq k \neq i) \\
S_j + 2S_k &\rightarrow \text{products} \quad (j \neq k) \\
3S_j &\rightarrow \text{products},
\end{aligned}
$$

where products are again combinations of $\{S_1, \ldots, S_N\}$.

Stochastic chemical kinetics

- The reaction velocity of each reaction type $\{R_1, \ldots, R_M\}$ is characterized by a constant parameter $\{c_1, \ldots, c_M\}$.
- **Fundamental hypothesis of stochastic chemical kinetics:**

$$c_\mu \delta t \equiv \text{the average probability that a particular combination of reactant molecules will react according to } R_\mu \text{ in } V \text{ in the next infinitesimal time interval } \delta t. \quad (1)$$

- Valid for well stirred systems dominated by elastic collisions such that positions of molecules are always uniformly randomized in V and their velocities are Maxwell-Boltzmann distributed.
- Remark: It is assumed that the occurrence of multiple reactions is of order $o(\delta t)$ and thus vanishes for $\delta t \to 0$.

Chemical master equation

- The aim is, under the above assumptions, to construct an algorithm for the time evolution of $\{X_i\}(t)$ given the initial conditions $\{X_i\}(0)$ and reaction parameters $\{c_\mu\}$.
- Chemical master equation for the probability, $\mathcal{P}(\{X_i\}; t)$, that $\{X_i\}$ molecules of species $\{S_i\}$ are present in V at time t is:

$$\frac{\partial}{\partial t}\mathcal{P}(\{X_i\}; t) = - \sum_{\{X_i^*\}} w(\{X_i\} \to \{X_i^*\})\mathcal{P}(\{X_i\}; t) + \sum_{\{X_i^*\}} w(\{X_i^*\} \to \{X_i\})\mathcal{P}(\{X_i^*\}; t) \quad (2)$$

- Eq.(2) completely determines $\mathcal{P}(\{X_i\}; t)$.
- Particle number moments $\langle X_i^k \rangle = \sum_{\{X_i\}} X_i^k \cdot \mathcal{P}(\{X_i\}; t)$, like $\langle X_i \rangle$ and $\langle X_i^2 \rangle$, could be obtained.
- However, Eq.(2) is not easily solvable neither analytically nor numerically.

Stochastic simulation algorithm

Instead of computing probability density function $P(\{X_i\}; t)$ for $\{X_i\}(t)$ simulate trajectories of $\{X_i\}(t)$ (compare random walk).

- Define new probability:

$$P(\tau, \mu)\, d\tau \equiv \text{the probability at time } t \text{ that the next reaction in the system will occur in the infinitesimal time interval } [t+\tau, t+\tau+d\tau) \text{ and will be an } R_\mu \text{ reaction.} \quad (3)$$

- Further, define

$$h_\mu \equiv \text{the number of distinct molecular reactant combinations for } R_\mu \text{ reaction at time } t \quad (4)$$

Stochastic simulation algorithm

Reaction type R_μ and distinct reactant combinations h_μ.

R_μ	h_μ
$\emptyset$	1
S_j	X_j
$S_j + S_k \quad (j \neq k)$	$X_j X_k$
$2S_j$	$\frac{1}{2} X_j (X_j - 1)$
$S_i + S_j + S_k \quad (i \neq j \neq k \neq i)$	$X_i X_j X_k$
$S_j + 2S_k \quad (j \neq k)$	$\frac{1}{2} X_j X_k (X_k - 1)$
$3S_j$	$\frac{1}{6} X_j (X_j - 1)(X_j - 2)$

$$h_\mu c_\mu \delta t = \text{probability of a reaction of type } R_\mu \text{ occurring in the next time interval } \delta t \qquad (5)$$

$P(\tau, \mu) d\tau = P_0(\tau) \cdot h_\mu c_\mu d\tau$, where $P_0(\tau)$ is the probability at time t that no reaction will occur in $[t, t + \tau)$ and $h_\mu c_\mu d\tau$ is the probability that R_μ will occur subsequent in $[t + \tau, t + \tau + d\tau)$.

Reaction probability density function $P(\tau, \mu)$

Estimate of $P_0(\tau)$:

- Divide $[t, t+\tau)$ into K equal subintervals of length $\epsilon = \frac{\tau}{K}$.
- The probability that none of the reactions $\{R_1, \ldots, R_M\}$ occurred in $[t, t+\epsilon)$ is

$$\prod_{\nu=1}^{M}[1 - h_\nu c_\nu \epsilon + o(\epsilon)] \approx 1 - \sum_{\nu=1}^{M} h_\nu c_\nu \epsilon + o(\epsilon). \qquad (6)$$

- Same expressions result for the remaining $K-1$ subintervals

$$\begin{aligned} P_0(\tau) &= \left[1 - \sum_{\nu=1}^{M} h_\nu c_\nu \epsilon + o(\epsilon)\right]^K && (7) \\ &= \left[1 - \sum_{\nu=1}^{M} h_\nu c_\nu \tfrac{\tau}{K} + o(K^{-1})\right]^K && (8) \\ &= \left[1 - \left(\sum_{\nu=1}^{M} h_\nu c_\nu \tau + \tfrac{o(K^{-1})}{K^{-1}}\right)\frac{1}{K}\right]^K && (9) \end{aligned}$$

Reaction probability density function

- In the limit $K \to \infty$ we obtain:

$$P_0(\tau) = \exp\left[-\sum_{\nu=1}^{M} h_\nu c_\nu \tau\right] \tag{10}$$

- Finally, the reaction probability density function reads as

$$P(\tau, \mu) = P_0(\tau) h_\mu c_\mu = h_\mu c_\mu \exp\left[-\sum_{\nu=1}^{M} h_\nu c_\nu \tau\right] \tag{11}$$

with normalization

$$\int_0^\infty d\tau \sum_{\mu=1}^{M} P(\tau, \mu) = \sum_{\mu=1}^{M} h_\mu c_\mu \int_0^\infty d\tau \exp\left[-\sum_{\nu=1}^{M} h_\nu c_\nu \tau\right] = 1 \tag{12}$$

- Eq.(11) is the mathematical basis for the stochastic simulation approach, since it contains all the information needed to treat stochastic chemical kinetics via Monte-Carlo method.

Overview of Gillespie's stochastic simulation algorithm

- Step 0 (Initialization): set initial molecular populations $\{X_i\}(0)$ and reaction parameters $\{c_\mu\}$, calculate $\{h_\mu\}$.
- Step 1 (Monte Carlo): generate a random reaction time and type (τ, μ) according to $P(\tau, \mu)$.
- Step 2 (Update): advance $t \to t + \tau$, update the populations $\{X_i\}(t)$ of species $\{S_i\}$ involved in the reaction R_μ, update $\{h_\mu\}$ accordingly.
- Step 3 (Terminate): If $t > t_{max}$ or no reaction type is possible, i.e., $\{h_1, \ldots, h_M\} = \{0, \ldots, 0\}$, then terminate, else go to Step 1.

Gillespie's direct method

How to generate (τ, μ) from $P(\tau, \mu)$?

Apply chain rule, i.e., write joint probability distribution

$$P(\tau, \mu) = P_1(\tau) \cdot P_2(\mu|\tau) \tag{13}$$

as product of

$P_1(\tau)d\tau \equiv$ probability that the next reaction (irrespective of type) will occur in $[t + \tau, t + \tau + d\tau)$

and

$P_2(\mu|\tau) \equiv$ conditional probability that the next reaction will be an R_μ reaction given that it occurs at time $t + \tau$

Using Eq.(11) we get

$$P_1(\tau) = \sum_{\mu=1}^{M} P(\tau, \mu) = a_0 \exp(-a_0\tau) \tag{14}$$

$$P_2(\mu|\tau) = P(\tau, \mu)/P_1(\tau) = a_\mu/a_0, \tag{15}$$

with abbreviations $a_\mu \equiv h_\mu c_\mu$ and $a_0 \equiv \sum_\mu h_\mu c_\mu \equiv \sum_\mu a_\mu$.

Implementing the Monte Carlo Step

- Step 1a (pick the reaction time): generate a uniformly distributed random number $r_1 \in [0, 1]$ and set

$$\tau = \frac{1}{a_0} \ln\left(\frac{1}{r_1}\right), \tag{16}$$

  see inverse transform sampling.

- Step 1b (pick the reaction type): generate a uniformly distributed random number $r_2 \in [0, 1]$ and set μ to be the integer for which

$$\sum_{\nu=1}^{\mu-1} a_\nu < r_2 a_0 < \sum_{\nu=1}^{\mu} a_\nu, \tag{17}$$

  see rejection sampling.

First-reaction method

A reminder: $a_\mu \delta t$ is the probability that a reaction R_μ occurs in δt.

$$\lim_{K\to\infty} \left(1 - a_\mu \frac{\tau}{K}\right)^K a_\mu \delta t = \text{probability that no reaction } R_\mu \text{ takes place in } [0,\tau] \text{ but occurs later in } [\tau, \tau+\delta t]$$

and thus

$$P_\mu(\tau)d\tau = e^{-a_\mu \tau} a_\mu d\tau = \text{probability at time } t \text{ that a reaction } R_\mu \text{ takes place } [t+\tau, t+\tau+d\tau] \text{ provided that the population involved in } R_\mu \text{ does not change in } [t, t+\tau]$$

Implementing the Monte Carlo Step:

- Generate M uniform random numbers $\{r_1, ..., r_M\} \in [0,1]$.
- Compute tentative reaction times $\tau_\nu = \frac{1}{a_\nu} \ln\left(\frac{1}{r_\nu}\right)$ for $\nu \in \{1, ..., M\}$.
- Choose as the actual next reaction the one which occurs first:

$$\tau = \text{the smallest of the } \{\tau_\nu\} \quad (18)$$

$$\mu = \text{the index of the smallest } \{\tau_\nu\} \quad (19)$$

First-reaction method

Proof that this method generates $P(\tau, \mu)$ from Eq.(11).

Reaction probability corresponding to procedure described above:

$$\widetilde{P}(\tau, \mu)d\tau = \Pr(\tau < \tau_\mu < \tau + d\tau) \cdot \Pr(\tau_\nu > \tau,\ \forall \nu \neq \mu) \qquad (20)$$

Since

$$\Pr(\tau < \tau_\mu < \tau + d\tau) = \exp(-a_\mu \tau) a_\mu d\tau \qquad (21)$$

and

$$\begin{aligned} \Pr(\tau_\nu > \tau,\ \forall \nu \neq \mu) &= \Pr\{(1/a_\nu)\ln(1/r_\nu) > \tau,\ \forall \nu \neq \mu\} \\ &= \Pr\{r_\nu < \exp(-a_\nu \tau),\ \forall \nu \neq \mu\} \\ &= \prod_{\nu \neq \mu} \Pr\{r_\nu < \exp(-a_\nu \tau)\} \\ &= \prod_{\nu \neq \mu} \exp(-a_\nu \tau) \qquad (22) \end{aligned}$$

we obtain

$$\begin{aligned} \widetilde{P}(\tau, \mu)d\tau &= e^{-a_\mu \tau} a_\mu d\tau \prod_{\nu \neq \mu} e^{-a_\nu \tau} \\ &= a_\mu e^{-a_0 \tau} d\tau = P(\tau, \mu)d\tau \qquad (23) \end{aligned}$$

Example

System of four chemical species W, X, Y and Z subject to six coupled chemical reactions:

$$X \underset{c_2}{\overset{c_1}{\rightleftharpoons}} Y \tag{24}$$

$$2X \underset{c_4}{\overset{c_3}{\rightleftharpoons}} Z \tag{25}$$

$$W + X \underset{c_6}{\overset{c_5}{\rightleftharpoons}} 2X \tag{26}$$

In the deterministic approach the following system of coupled nonlinear ODES must be solved:

$$\frac{dW}{dt} = -c_5 WX + \frac{1}{2} c_6 X^2 \tag{27}$$

$$\frac{dX}{dt} = -c_1 X + c_2 Y - c_3 X^2 + 2c_4 Z + c_5 WX - \frac{1}{2} c_6 X^2 \tag{28}$$

$$\frac{dY}{dt} = c_1 X - c_2 Y \tag{29}$$

$$\frac{dZ}{dt} = \frac{1}{2} c_3 X^2 - c_4 Z \tag{30}$$

Example

Chemical reaction system

$$X \underset{c_2}{\overset{c_1}{\rightleftharpoons}} Y \tag{31}$$

$$2X \underset{c_4}{\overset{c_3}{\rightleftharpoons}} Z \tag{32}$$

$$W + X \underset{c_6}{\overset{c_5}{\rightleftharpoons}} 2X \tag{33}$$

and the corresponding master equation

$$\begin{aligned}
\frac{d\mathcal{P}(W,X,Y,Z;t)}{dt} = \\
c_1\{(X+1)\mathcal{P}(W,X+1,Y-1,Z;t) - X\mathcal{P}(W,X,Y,Z;t)\} \\
+ c_2\{(Y+1)\mathcal{P}(W,X-1,Y+1,Z;t) - Y\mathcal{P}(W,X,Y,Z;t)\} \\
+ c_3\{\tfrac{1}{2}(X+2)(X+1)\mathcal{P}(W,X+2,Y,Z-1;t) \\
- \tfrac{1}{2}X(X-1)\mathcal{P}(W,X,Y,Z;t)\} \\
+ c_4\{(Z+1)\mathcal{P}(W,X-2,Y,Z+1;t) - Z\mathcal{P}(W,X,Y,Z;t)\} \\
+ \ldots
\end{aligned}$$

Next-subvolume method

- So far simulation methods apply to well stirred systems, i.e., diffusion is so fast that all concentrations are homogenous in space.
- If system size is to large to be homogenized by diffusion on the timescale of the chemical reactions the system becomes spatially heterogeneous and a method with spatial resolution is needed.
- Elf *et al.* proposed an extension of Gillespie's direct method to simulate spatially resolved reaction-diffusion kinetics on mesoscopic level.
- The total system is divided into N subvolumes (SVs), chosen so small that the concentrations of reactants in a SV are near-homogeneous in space.
- The molecules in a SV can either undergo chemical reactions or diffuse to a neighboring SV.

Sketch of the next-subvolume method

- Calculate first the next event time (reaction or diffusion) in each SV and identify the SV with the smallest next event time, very similar to the first-reaction method.
- Apply Gillespie's direct method to the SV with the smallest next event time in order to decide if the next event is a chemical reaction or a diffusion jump and which species are involved in this event.
- The time for the next event in each SV is ordered in an event queue, which makes the computation time linear in $\log N$, rather then in N.

Next-subvolume method

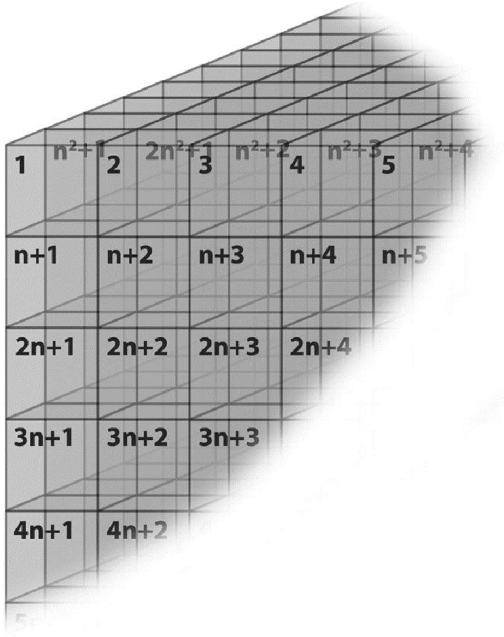

Figure 1: An example of indexing n^3 cells. From Elf *et al.*

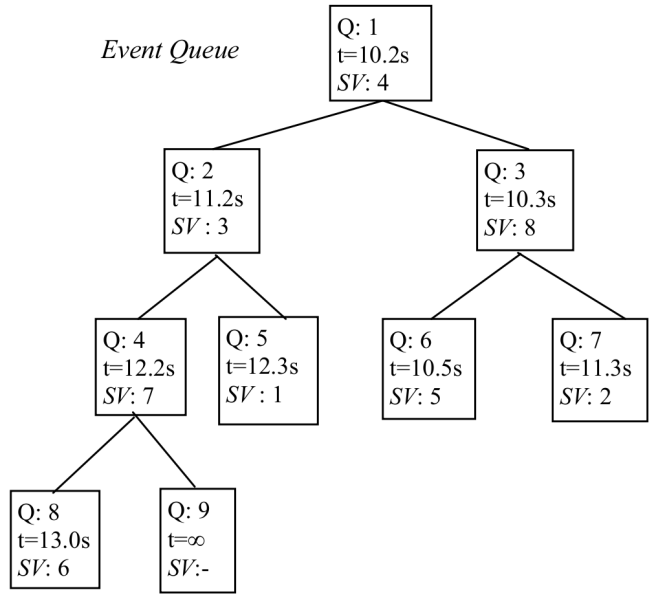

Figure 2: In the event queue, the elements are ordered such that, in each branch of the binary tree, a SV with an earlier event time t is higher up. The Q array keeps a reference to the SVs position in the event queue. From Elf *et al.*

Application of the next-subvolume method

- Simulation of the spatial oscillation patterns that are displayed by the Min system of *Escherichia coli*.
- In wild-type *E. coli*, the Min proteins oscillate back and forth between the cell poles to help the bacterium find its middle before cell division.

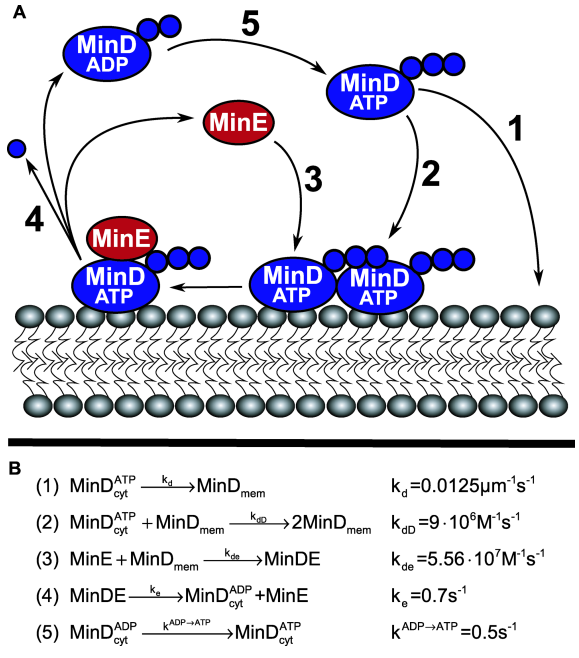

Figure 3: Min system reaction scheme and rate constants. From Fange *et al.*

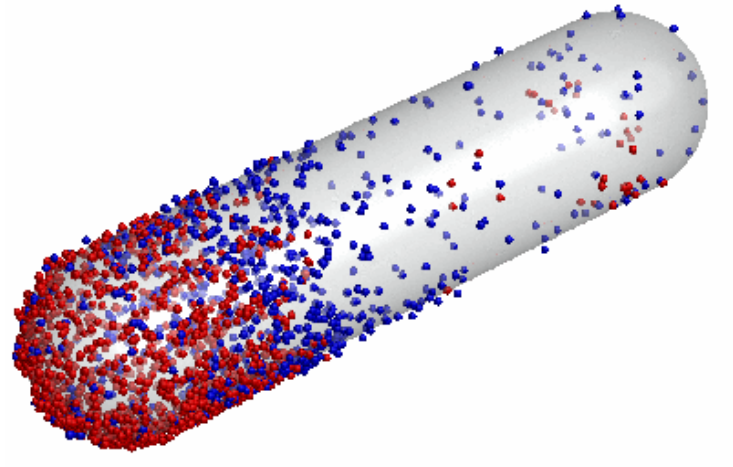

Figure 4: Membrane-bound MinD is shown in blue, and MinE in complex with MinD on the membrane is shown in red, see movie https://doi.org/10.1371/journal.pcbi.0020080

Literature

- Daniel T. Gillespie: *Stochastic Simulation of Chemical Kinetics.* Annu. Rev. Phys. Chem. 58 (2007): 35-55
- Elf, Johan *et al.*: *Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases.* Systems biology 1.2 (2004): 230-236
- Fange, David, and Johan Elf: *Noise-induced Min phenotypes in E. coli.* PLoS Comput Biol 2.6 (2006): e80.