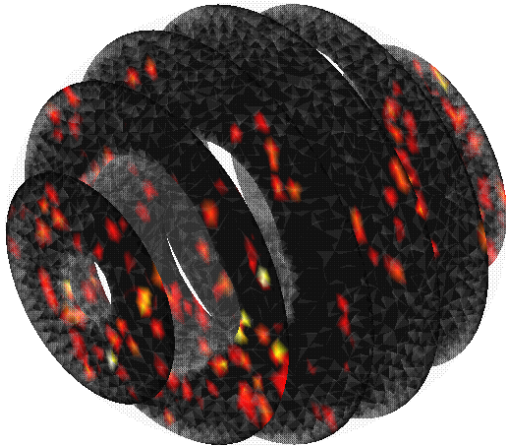
Mesoscopic Stochastic Modeling of Reaction-Diffusion Processes



Stefan Engblom

CSC/NA

Royal Institute of Technology (KTH)

KCSE Seminar, Stockholm, February 10, 2010

Joint work with

Andreas Hellander, Lars Ferm, Per Lötstedt.

- Motivation for stochastic chemical kinetics
- Well-stirred chemical kinetics
- Spatially inhomogeneous kinetics
- Unstructured meshes
- Hybrid simulation
- Examples
- Conclusions

Modelling chemical reactions

System size Ω (# molecules)	Model	Idea
$\lesssim 10^2$	Micro	Movement of individual atoms/molecules
		Collisions $\rightarrow$ (Possible) reactions
$\sim 10^1$–10^6	Meso	Non-individual, assuming well-stirred mixture
		A *stochastic model* is used for reactions
$\gtrsim 10^6$	Macro	"Average"; —in the limit of many molecules

Diffusion-controlled kinetics

Model	Assumption
BD (Smoluchowski)	Brownian motion of individual molecules
CTMC (Master equation)	Non-individual, (locally) well-stirred
SDE (Langevin)	Continuous *approximation*
ODE (Reaction rate)	Continuous, deterministic

-With a mesoscopic Continuous-Time Markov Chain, an accurate but still manageable *non-individual* model is possible thanks to stochasticity.

-There are many examples of when stochastic kinetics more easily captures actual behavior...

Multistability (*Gardner/Cantor/Collins*)

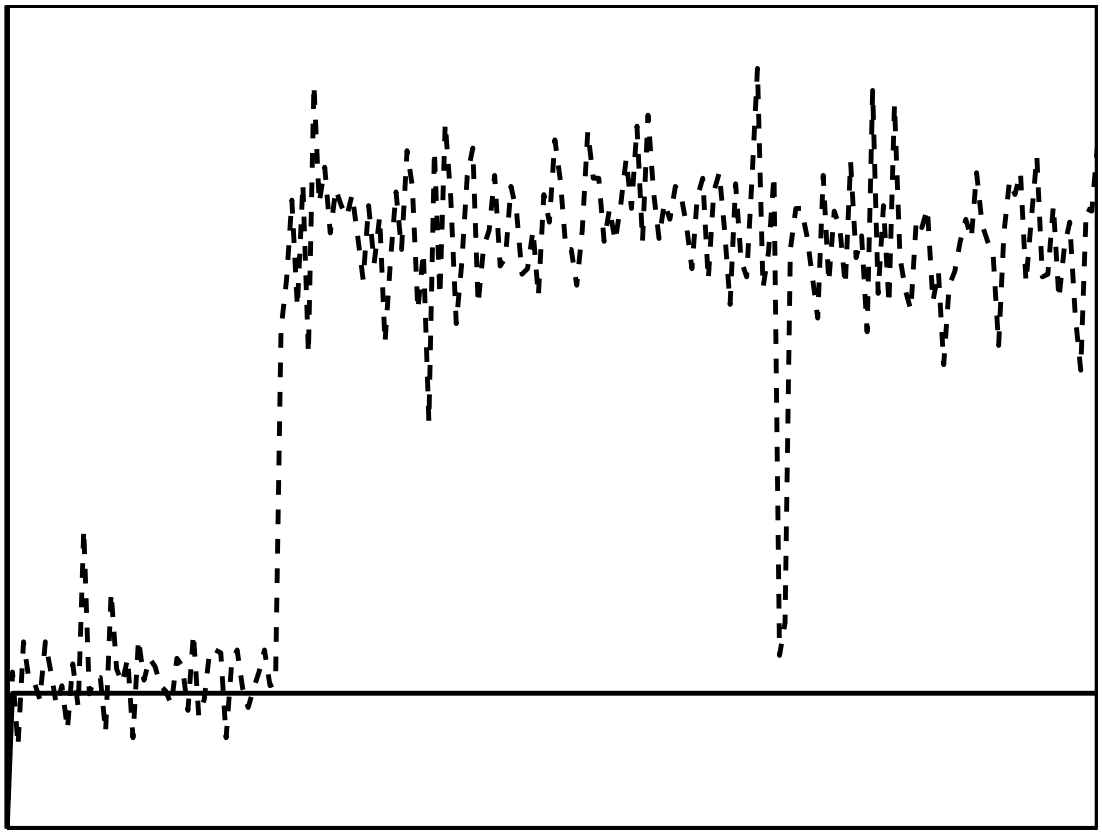


Figure 1: Solid: deterministic, dashed: stochastic.

Stochastic resonance (*Barkai/Leibler*)

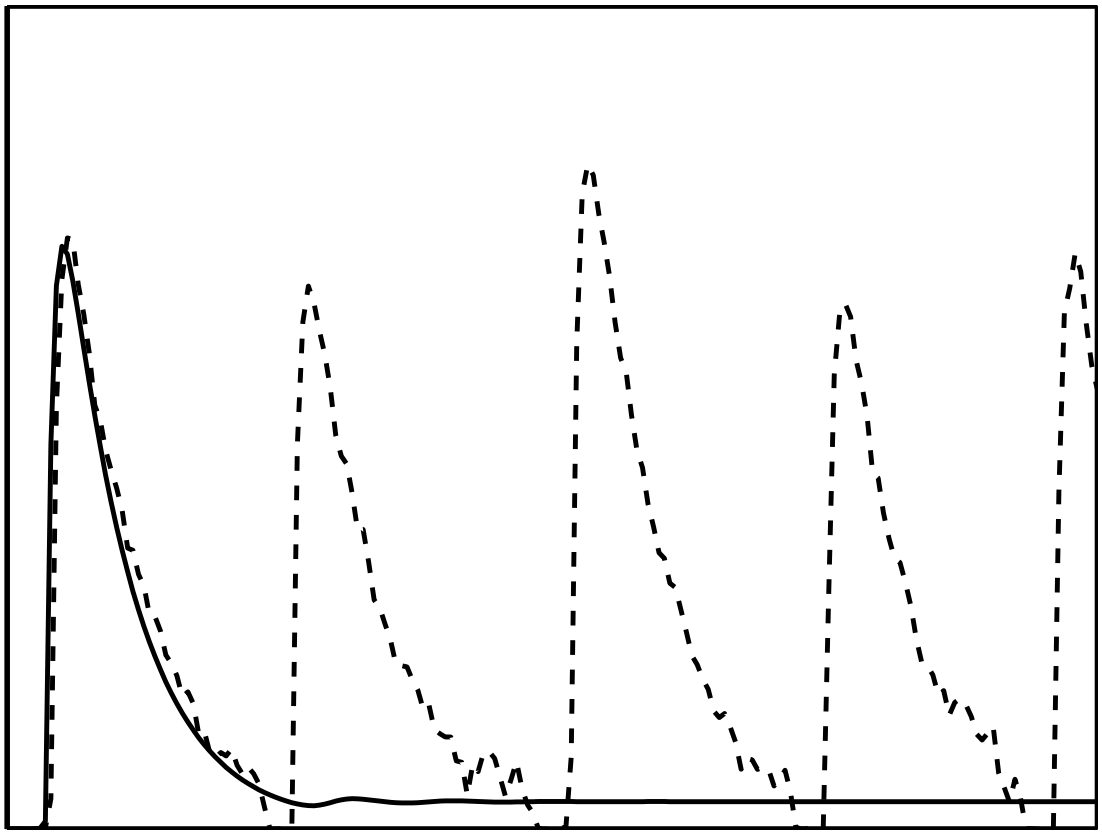


Figure 2: Solid: deterministic, dashed: stochastic.

Stochastic focusing (*Paulsson/Berg/Ehrenberg*)

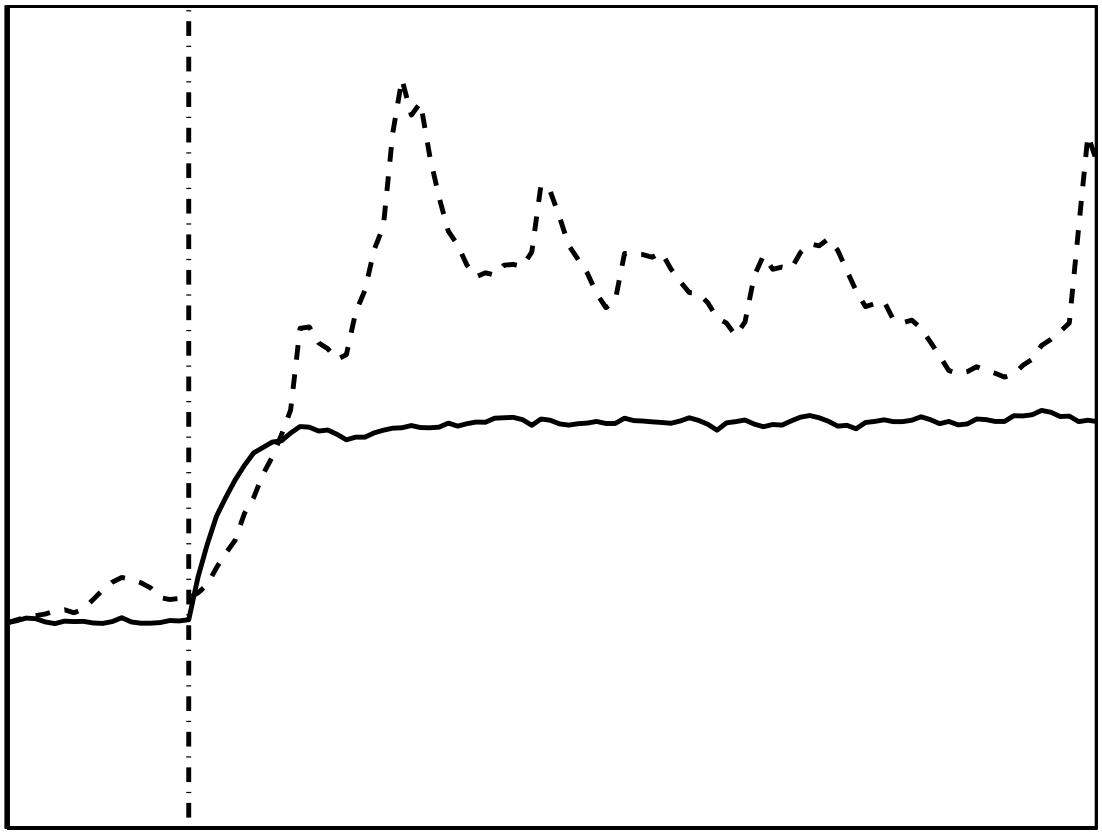


Figure 3: Nonlinear response to twofold signal increase; solid: partially deterministic, dashed: fully stochastic.

Well-stirred kinetics (*Gillespie '76, '92, Gardiner, van Kampen*)

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)*.

-Let the state vector $x \in \mathbf{Z}_+^D$ count the number of molecules of each of D species.

-Let R specified reactions be defined as *transitions* between these states,

$$x \xrightarrow{w_r(x)} x - \mathbb{N}_r, \qquad \mathbb{N} \in \mathbf{Z}^{D \times R} \ (\textit{stoichiometric matrix})$$

where each transition intensity or *propensity* $w_r : \mathbf{Z}_+^D \to \mathbf{R}_+$ is the probability of reacting per unit of time. *This probability can be shown to exist provided that the system is well-stirred!*

The *chemical master equation* is given by

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

-A gain-loss discrete PDE in D dimensions for the probability density.

-Several *exact* simulation algorithms exist ("SSA", "NRM", ...); determine *what* event and *when*.

Stochastic simulation algorithm — direct method (*Gillespie '76*)

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

$$\sum_{s=1}^{r-1} w_s(x) < W u_2 \leq \sum_{s=1}^{r} w_s(x),$$

   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 $x := x - \mathbb{N}_r$.
3. Repeat from step 1 until some final time T is reached.

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 *localised* — e.g. depend on an enzyme molecule situated at a precise position.

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

Microscopic kinetics

- Molecular dynamics...
  -Many different algorithms, usually *very* expensive simulations.
- Smoluchowski kinetics (diffusion-controlled limit): individual coordinates of molecules, Brownian motion in space. The *Smoluchowski PDE* evolves the spatial probability density in time and the reactions are to be incorporated as boundary conditions.
  -One exact algorithm: Green's function reaction dynamics (GFRD).
  -Various software for approximations: "MCell", "SmolDyn", "ChemCell"...

Mesoscopic spatial kinetics

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

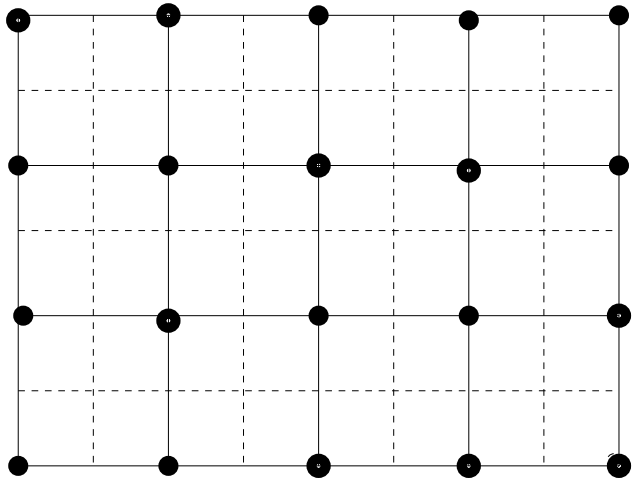


Figure 4: Primal mesh (solid), dual mesh (dashed). The nodal dofs are the # of molecules in each dual cell.

- 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 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 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*,

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

Diffusion

A natural model of diffusion from one cell Ω_k to another cell Ω_j is

$$X_{ik} \xrightarrow{q_{kj}\mathbf{x}_{ik}} X_{ij},$$

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 Ω_k to Ω_j. $\Longrightarrow q_{kj} \propto \sigma^2/h^2$, where $\sigma^2/2$ is the macroscopic diffusion, h the local length.

The diffusion master equation can therefore be written

$$\frac{\partial p(\mathbf{x},t)}{\partial t} = \sum_{i=1}^{D}\sum_{k=1}^{K}\sum_{j=1}^{K} q_{kj}(\mathbf{x}_{ik} + \mathbb{M}_{kj,k})p(\mathbf{x}_{1\cdot},\ldots,\mathbf{x}_{i\cdot}+\mathbb{M}_{kj},\ldots,\mathbf{x}_{D\cdot},t)$$
$$-q_{kj}\mathbf{x}_{ik}p(\mathbf{x},t) =: \mathcal{D}p(\mathbf{x},t).$$

The transition vector $\mathbb{M}_{kj}$ is zero except for $\mathbb{M}_{kj,k} = -\mathbb{M}_{kj,j} = 1$.

The reaction-diffusion master equation (*Gardiner, van Kampen*)

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

-An *approximation*! Valid when

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

ρ the molecular radius, τ_Δ average molecular survival time.

-Once formulated, any well-stirred algorithm can simulate the RDME. For a spatially resolved model, most of the simulation time is spent on diffusion events.

Formulation and consistency

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

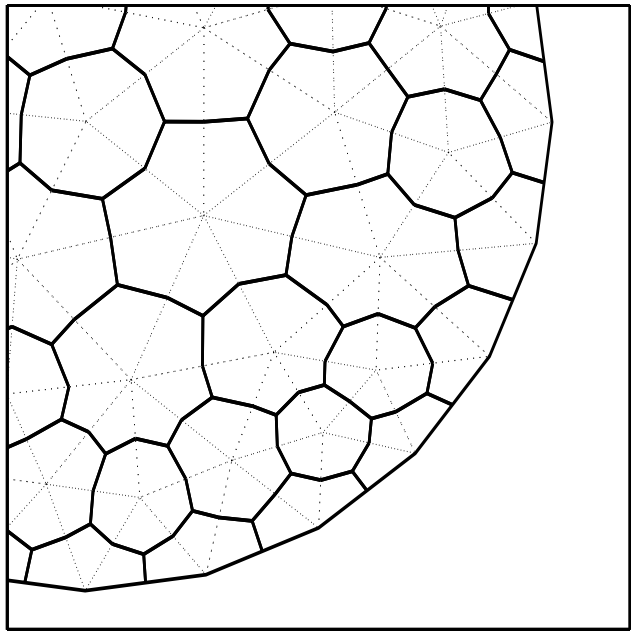
-A solution in the Cartesian case: ensure that the expected value limits to the macroscopic diffusion equation.

Define $\varphi_{ij} = E\,\Omega_j^{-1}\mathbf{x}_{ij}$. By linearity of the diffusion intensities, the diffusion master equation implies

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

or simply

$$\frac{d\varphi_{i\cdot}^T}{dt} = Q\varphi_{i\cdot}^T.$$

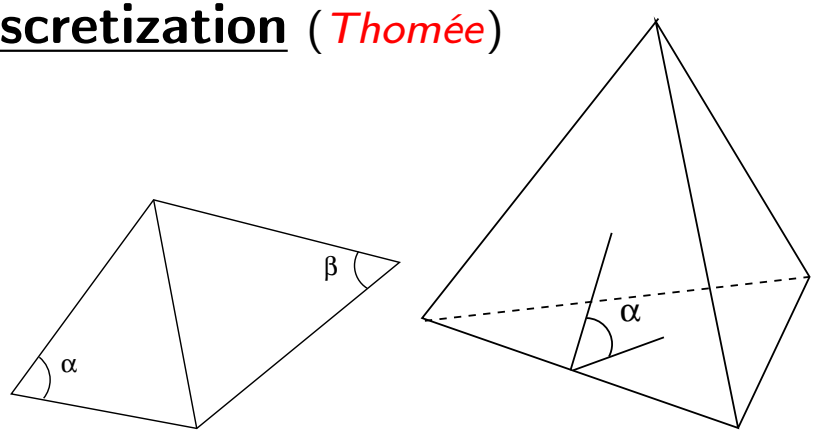


-FEM applied to the macroscopic equation $u_t = \sigma^2/2\,\Delta u$ with piecewise linear basis functions and inversion of the lumped mass-matrix yields

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

Assuming point-wise convergence FEM $\to$ diffusion PDE, the consistency of this interpretation ensures convergence in distribution to the correct Brownian motion as $h \to 0$.

Finite element discretization (*Thomée*)



(a) $\alpha + \beta < \pi$ (b) $\alpha < \pi/2$

Figure 5: The critical angles for positive off-diagonal elements.

With Neumann boundary conditions,

$$D_{jk} \geq 0,\ D_{jj} < 0,\ \sum_{k=1}^{K} D_{jk} = 0.$$

The same sufficient conditions implies the maximum principle for parabolic equations.

Diffusion moments...

Using the exact equation for the covariance matrix C of the diffusion process one can show:

-Standard deviation $\sim \sqrt{\|E\,\mathbf{x}_{i\cdot}\|}$.

-Quotient between standard deviation and expected values is $\sim 1/\sqrt{\|E\,\mathbf{x}_{i\cdot}\|}$ and is small for species i with large copy numbers $\Longrightarrow$ the expected value is a good approximation of the copy number.

The diffusion of such species can be evolved efficiently with mean field equations.

Time integration

Order the species X_i such that

$X_i,\ i = 1, \ldots, D_L$, have low copy numbers and

$X_i,\ i = D_L + 1, \ldots, D$, have high copy numbers.

Split the diffusion operator accordingly,

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = [\mathcal{M} + \mathcal{D}_L]p(\mathbf{x}, t) + \mathcal{D}_H p(\mathbf{x}, t).$$

Strang splitting:

1. Advance $p_t = \mathcal{D}_H p$ from t to $t + \Delta t/2$
2. Advance $p_t = [\mathcal{M} + \mathcal{D}_L]p$ by Δt (stochastic algorithm)
3. Advance $p_t = \mathcal{D}_H p$ from $t + \Delta t/2$ to $t + \Delta t$

Numerical simulations

-Deterministic diffusion is solved by the trapezoidal method:

$$\left(I - \frac{\Delta t}{2}\frac{\sigma^2}{2}D^T\right)\left(\mathbf{x}_{i\cdot}^{n+1}\right)^T = \left(I + \frac{\Delta t}{2}\frac{\sigma^2}{2}D^T\right)\left(\mathbf{x}_{i\cdot}^{n}\right)^T .$$

-Mesoscopic diffusion and reactions are simulated by NSM (*Fange/Elf*).

Properties:

- Non-negativity of $\mathbf{x}_{ij}$ is preserved with a bound on Δt.
- Total number of molecules of each species is conserved by the diffusion.

Bistable double-negative feedback system (*Elf/Ehrenberg*)

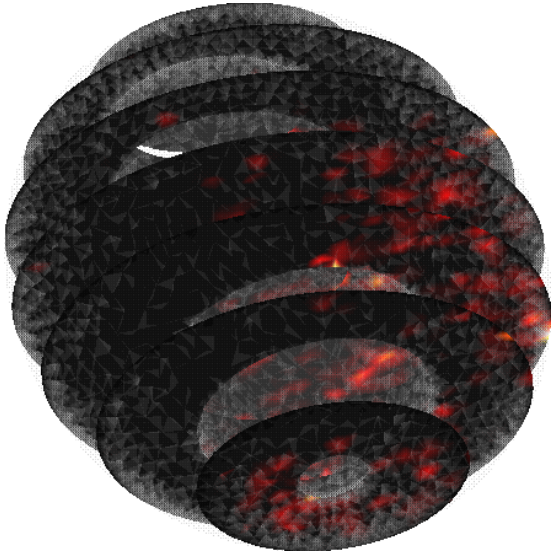
$$E_A \xrightarrow{k_1} E_A + A \qquad E_B \xrightarrow{k_1} E_B + B$$

$$E_A + B \underset{k_d}{\overset{k_a}{\rightleftharpoons}} E_A B \qquad E_B + A \underset{k_d}{\overset{k_a}{\rightleftharpoons}} E_B A$$

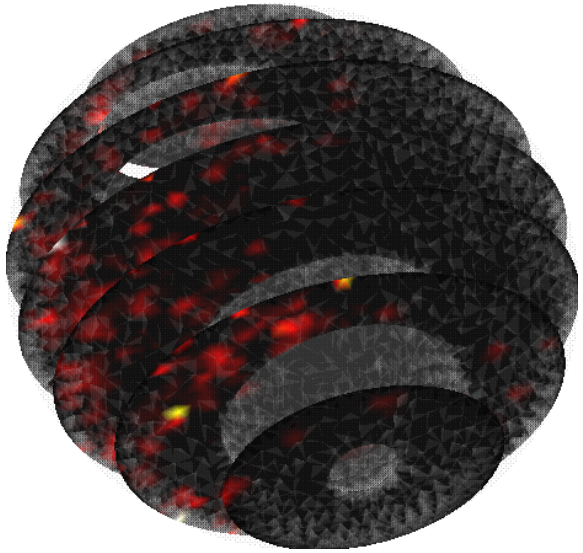
$$E_A B + B \underset{k_d}{\overset{k_a}{\rightleftharpoons}} E_A B_2 \qquad E_B A + A \underset{k_d}{\overset{k_a}{\rightleftharpoons}} E_B A_2$$

$$A \xrightarrow{k_4} \emptyset \qquad B \xrightarrow{k_4} \emptyset$$

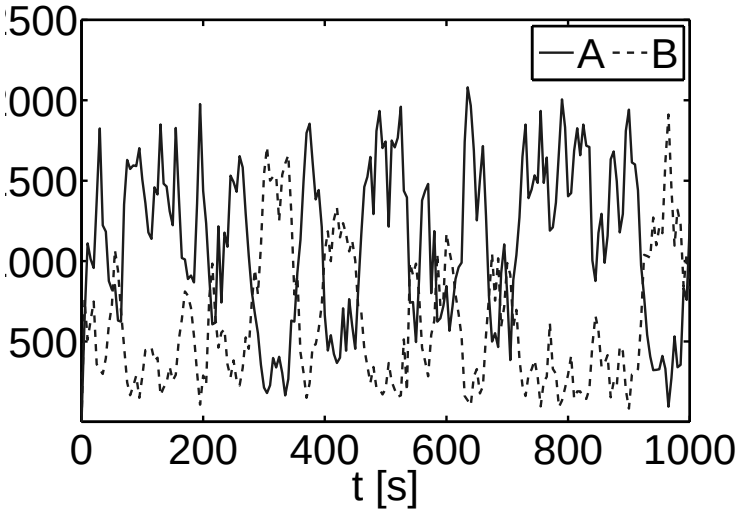
Slow/intermediate/fast diffusion in a simple model of an *S. cerevisiae* cell with internal structures in the form of a nucleus and a large vacuole. Molecules are not allowed to diffuse across the membranes and enter the organelles.



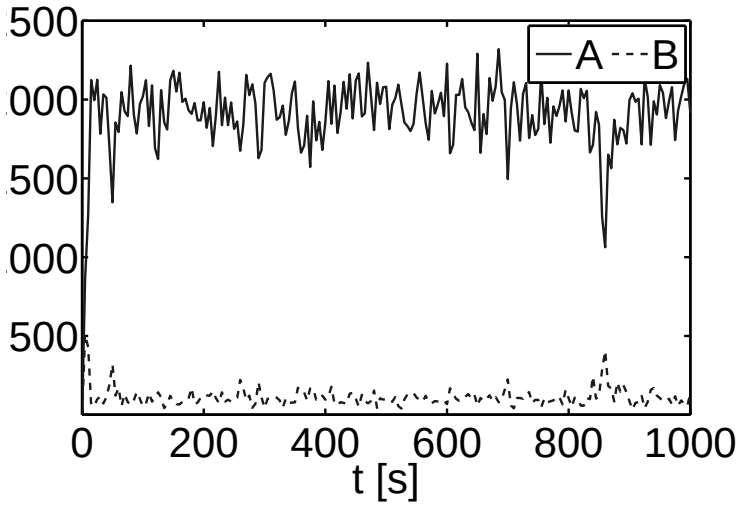
(a) Species A.



(b) Species B.



(c) $\sigma^2 = 2 \times 10^{-13}$



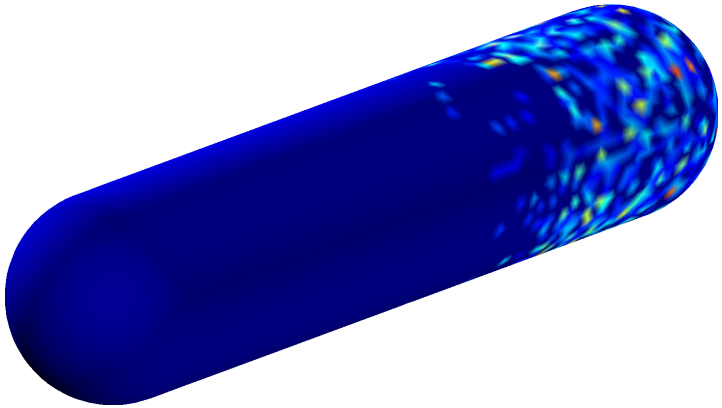
(d) $\sigma^2 = 4 \times 10^{-13}$

Figure 6: The total number of A and B molecules as the diffusion constant is varied. Right: local bi-stability is lost.

Oscillations of proteins involved in the cell division of *Escherichia coli* bacterium:

-Five species, five reactions (*Fange/Elf*).

-"URDME" software (*Cullhed/Engblom/Hellander*).



- Mesoscopic stochastic kinetics (CTMC/master equation)
  -well stirred chemical reactions
- Spatially inhomogeneous case:
  -microscopic kinetics usually very expensive
  -local well-stirredness implies the reaction-diffusion master equation
  -the RDME is computationally simpler
- Unstructured meshes: consistency with macroscopic equations
- Expensive but structurally simple diffusion $\Longrightarrow$ hybrid method
- Publicly available software ANSI-C99/Matlab/Comsol "URDME".