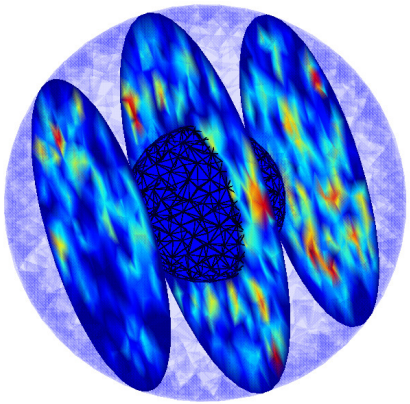
Stochastic Modeling in Reaction-Transport Processes: Forward and Backward considerations



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

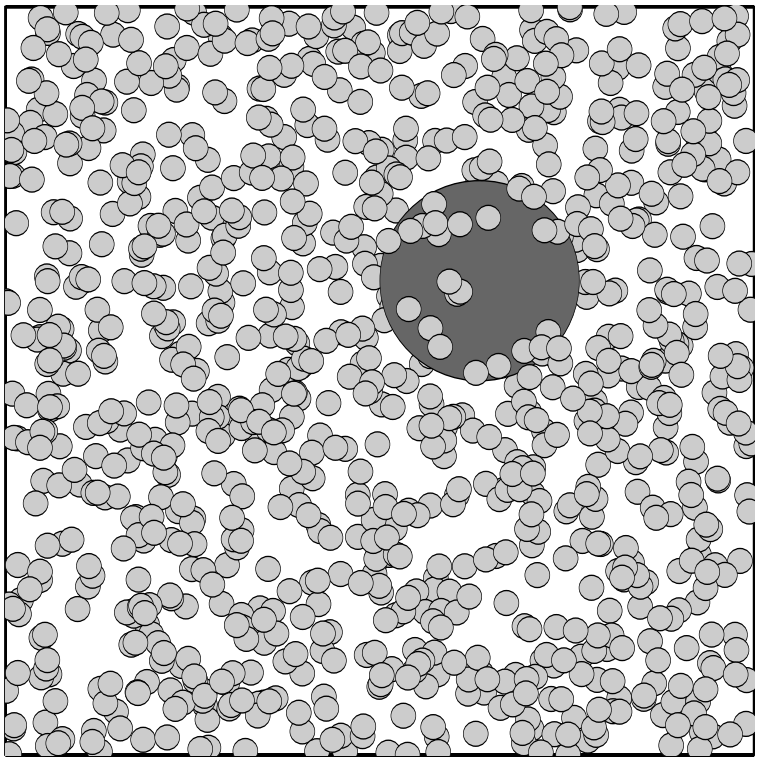
3. Inverse formulations
   - Reaction rates from observations
   - Molecular movements from observations
   - Optimal rates

Conclusions

Brownian motion

Einstein 1905, & some others...

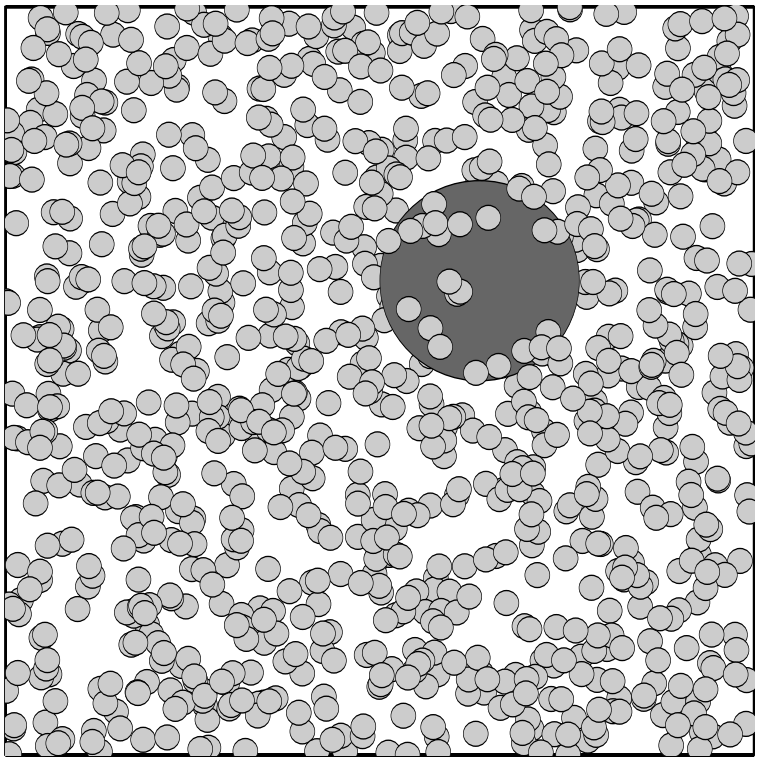
Example: Particle in a fluid.

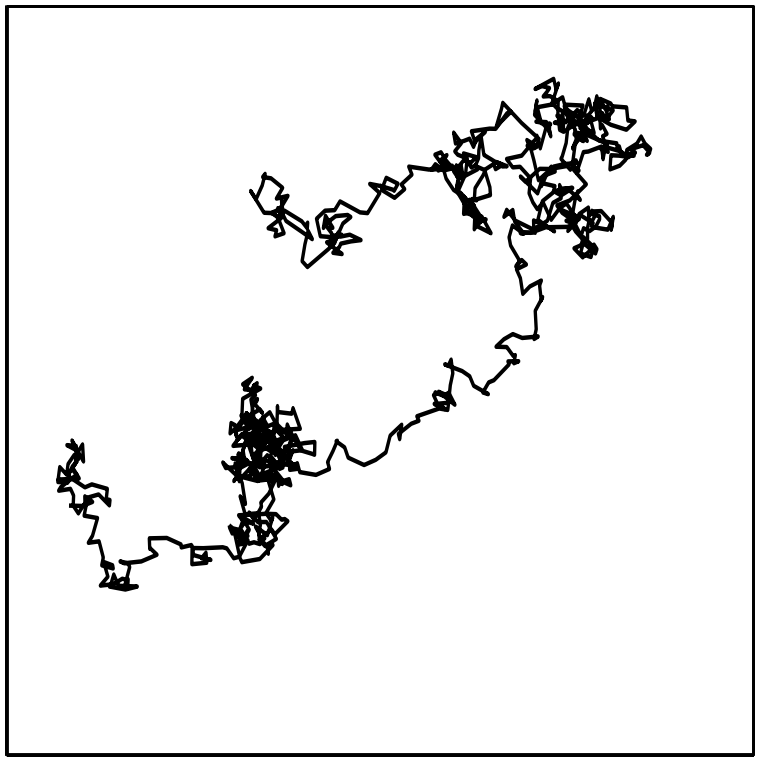


Brownian motion

Einstein 1905, & some others...

Example: Particle in a fluid.



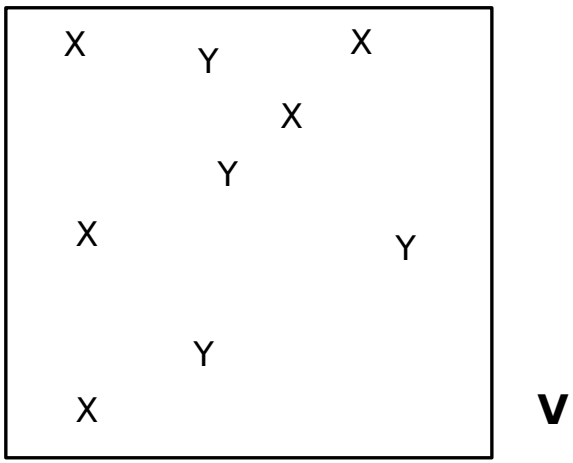


A stochastic model is simpler but depends on randomness.

Stochastic modeling of biochemical reactions

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

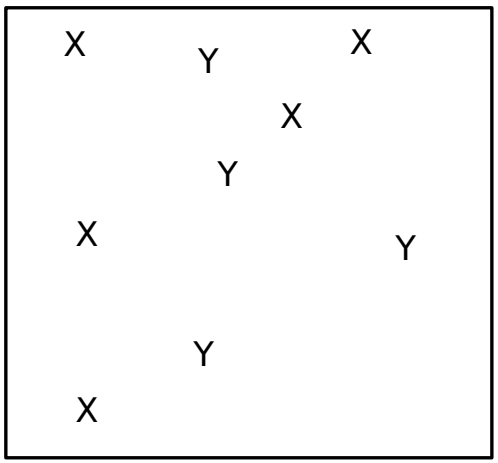
-What is the probability $P(1X \text{ and } 1Y \text{ reacts in the interval } [0, \Delta t])$?



Stochastic modeling of biochemical reactions

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

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

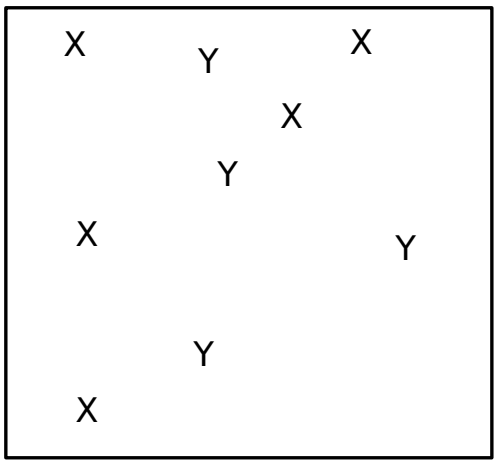


- $P \propto n_X$ ("number of X-molecules")
- $P \propto n_Y$

Stochastic modeling of biochemical reactions

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

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

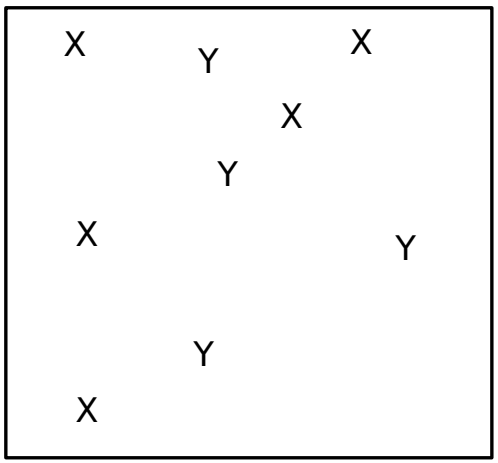


- $P \propto n_X$ ("number of X-molecules")
- $P \propto n_Y$
- $P \propto 1/V$
- $P \propto \Delta t$

Stochastic modeling of biochemical reactions

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

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



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

$\Longrightarrow P(X + Y \to Z$ 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 $x \in \mathbf{Z}_+^D$ counting the number of molecules of each of D species.
-R specified reactions defined as *transitions* between these states,

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

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

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

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

   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.

Kolmogorov's forward differential system/Master equation

(Kolmogorov '31, Nordsieck/Lamb/Uhlenbeck '40)

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

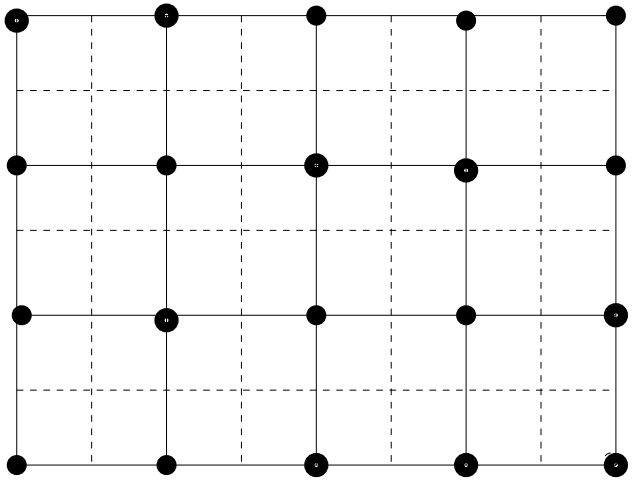


Figure: Primal mesh (solid), dual mesh (dashed). The nodal dofs are the # of molecules in each dual cell.

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*,

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

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.

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.

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

"RDME"

Combining reactions with diffusions,

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

The reaction-diffusion master equation

"RDME"

Combining reactions with diffusions,

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

-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 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

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

Unstructured meshes

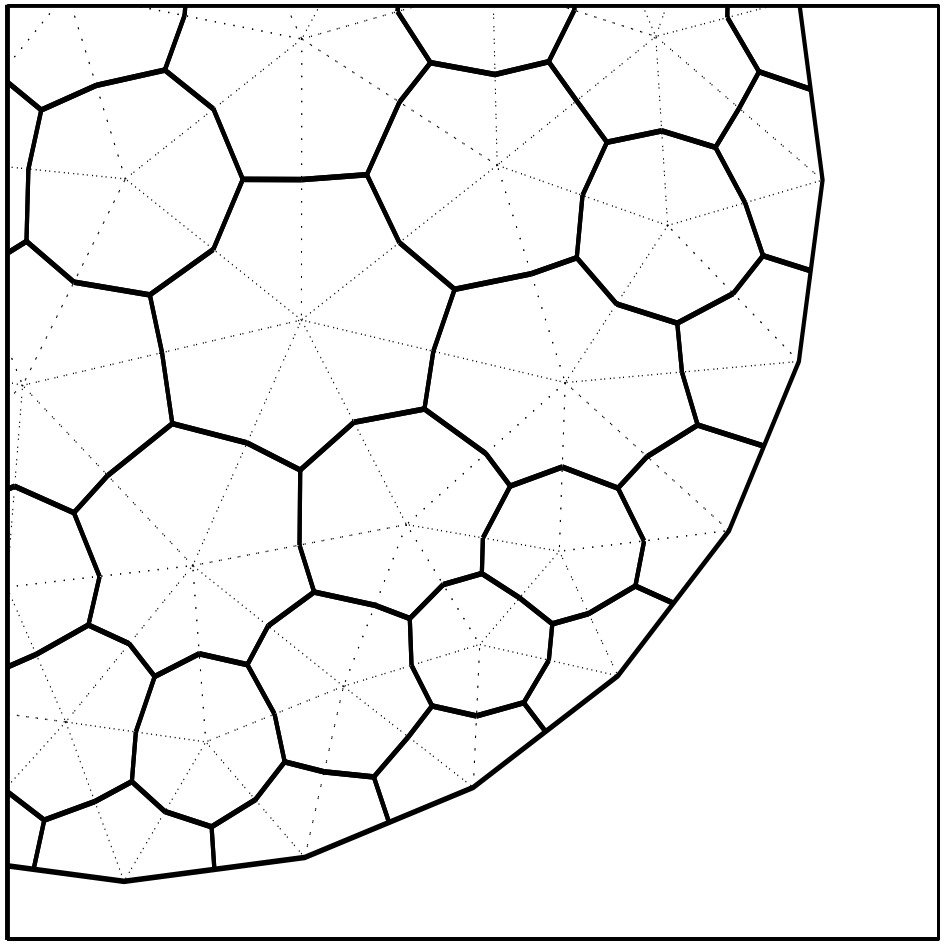
-Mean first exit time only known for very simple geometries (e.g. circles).
-*How to handle complicated geometries?* Attempt to converge in expectation to the macroscopic diffusion equation. Briefly, a numerical method applied to $u_t = \sigma^2/2\,\Delta u$ yields the *discretized* form

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

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

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

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



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

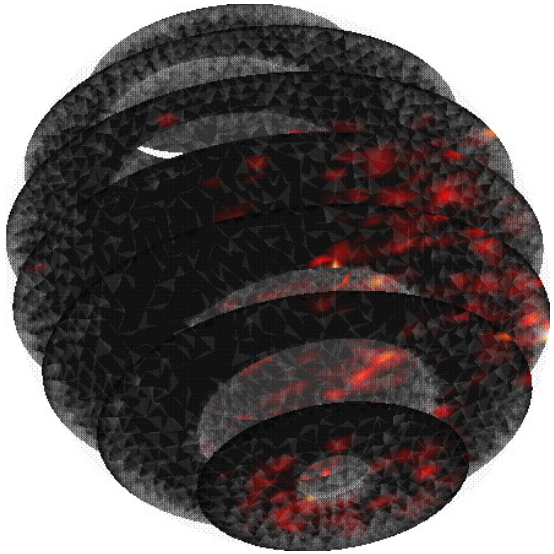
(Forward) Computations by examples

- Bistable model
- Spatial oscillations in *E. coli*.

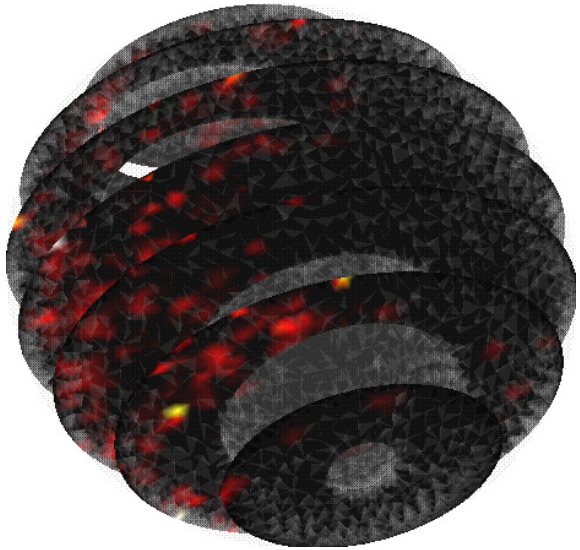
Bistable double-negative feedback system

$$
\begin{aligned}
E_A &\xrightarrow{k_1} E_A + A & \qquad E_B &\xrightarrow{k_1} E_B + B \\
E_A + B &\underset{k_d}{\overset{k_a}{\rightleftharpoons}} E_A B & \qquad E_B + A &\underset{k_d}{\overset{k_a}{\rightleftharpoons}} E_B A \\
E_A B + B &\underset{k_d}{\overset{k_a}{\rightleftharpoons}} E_A B_2 & \qquad E_B A + A &\underset{k_d}{\overset{k_a}{\rightleftharpoons}} E_B A_2 \\
A &\xrightarrow{k_4} \emptyset & \qquad B &\xrightarrow{k_4} \emptyset
\end{aligned}
$$

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.

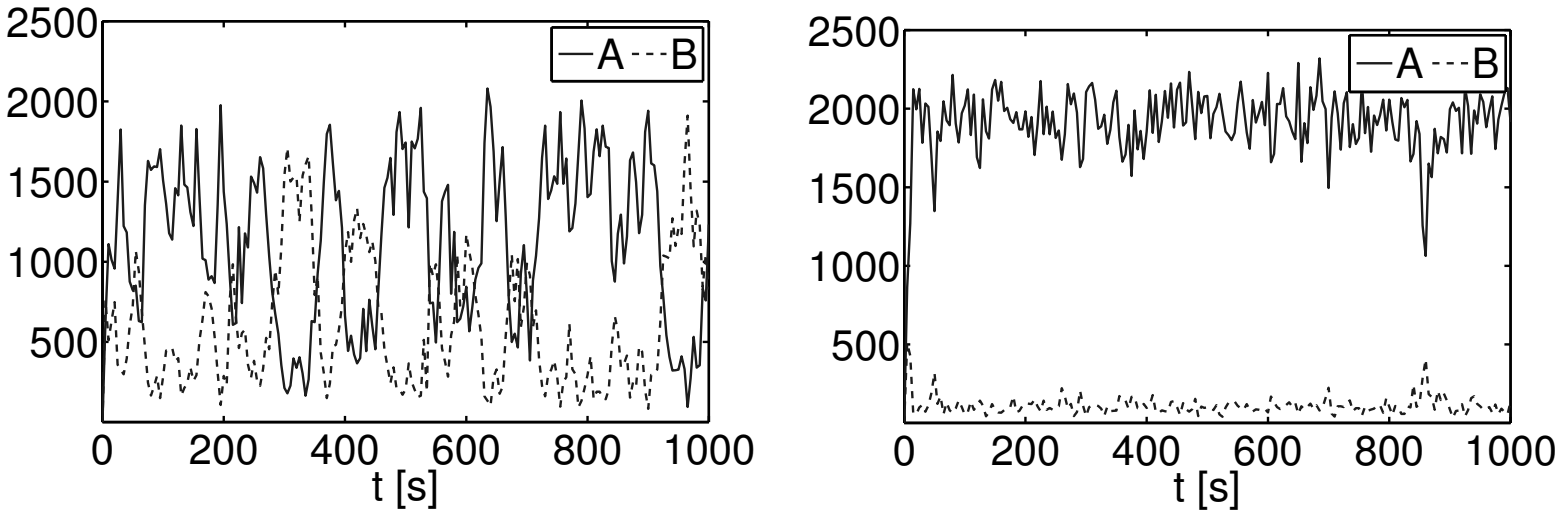


(a) Species A.



(b) Species B.

"URDME" software www.urdme.org.



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

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

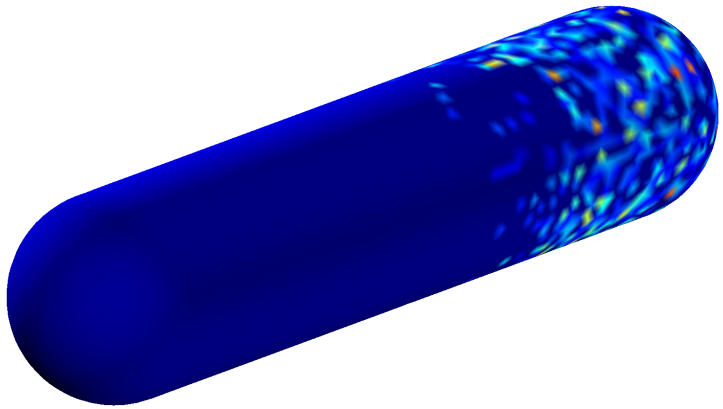
MinD oscillations

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

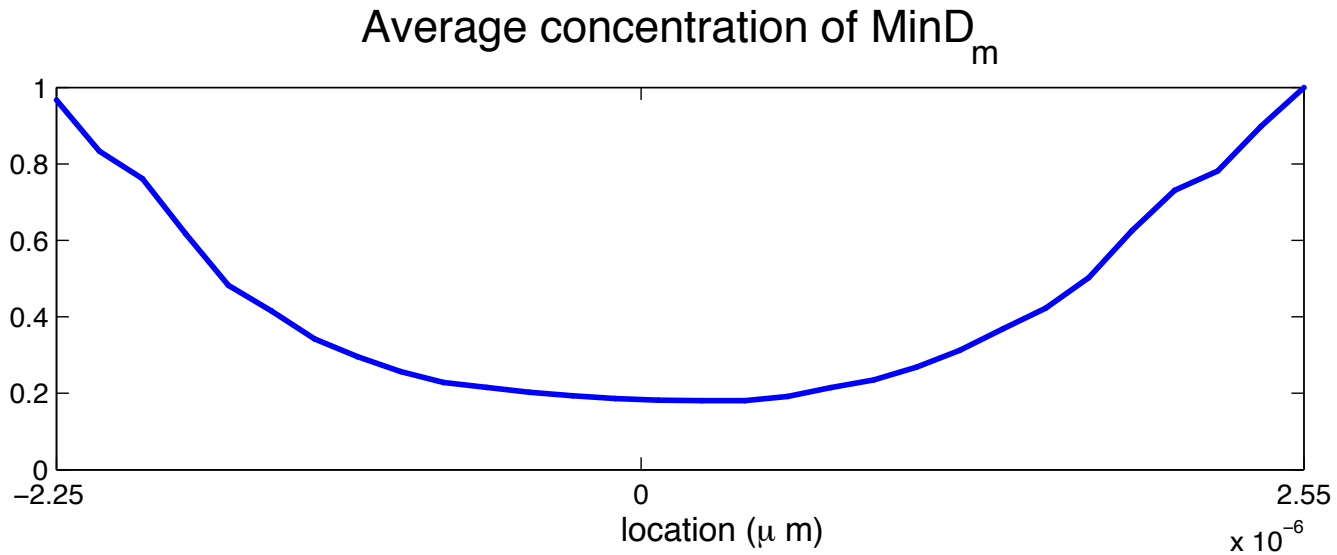
$$\text{MinD\_c\_atp} \xrightarrow{k_d} \text{MinD\_m} \qquad \text{MinD\_c\_atp} + \text{MinD\_m} \xrightarrow{k_{dD}} 2\text{MinD\_m}$$

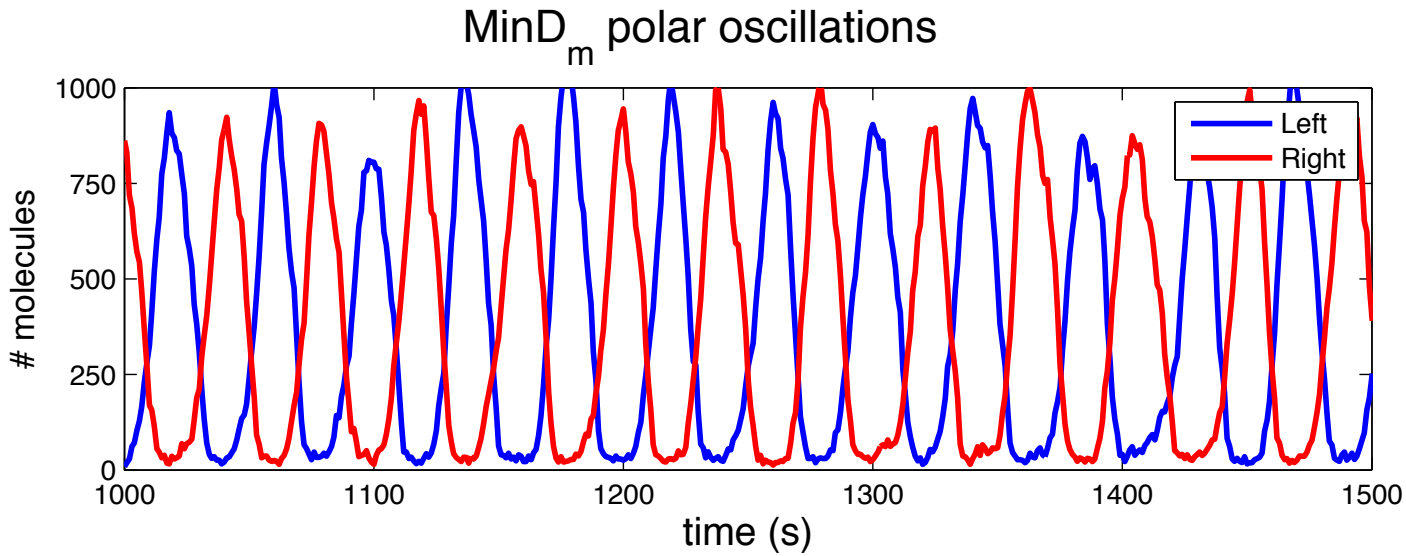
$$\text{Min\_e} + \text{MinD\_m} \xrightarrow{k_{de}} \text{MinDE} \qquad \text{MinDE} \xrightarrow{k_e} \text{MinD\_c\_adp} + \text{Min\_e}$$

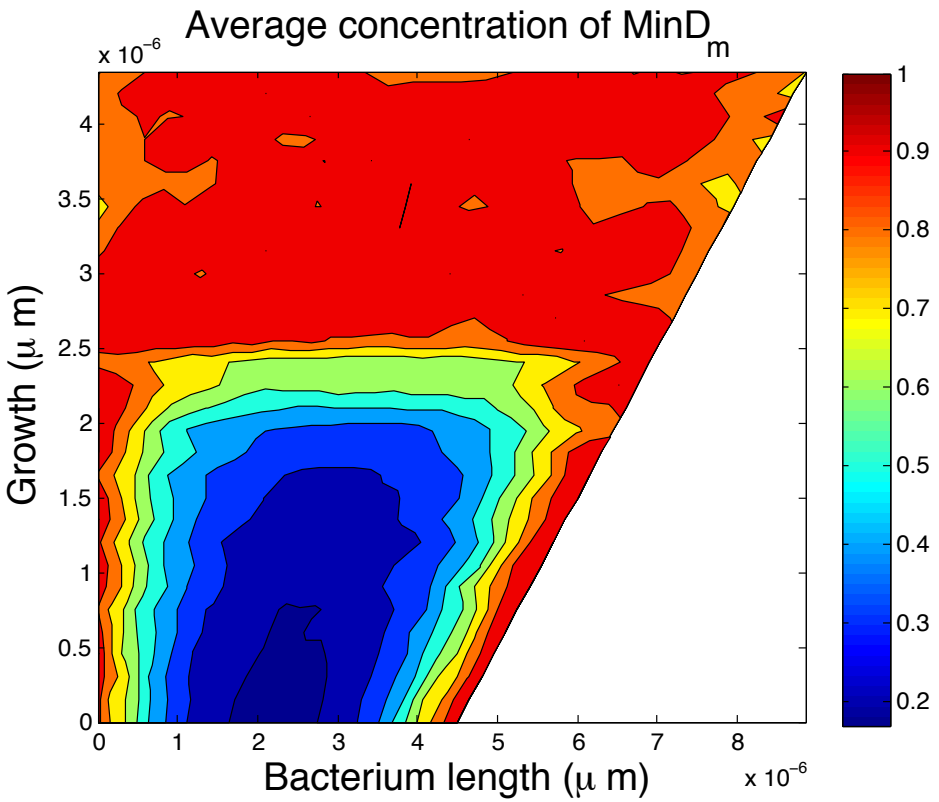
$$\text{MinD\_c\_adp} \xrightarrow{k_p} \text{MinD\_c\_atp}$$

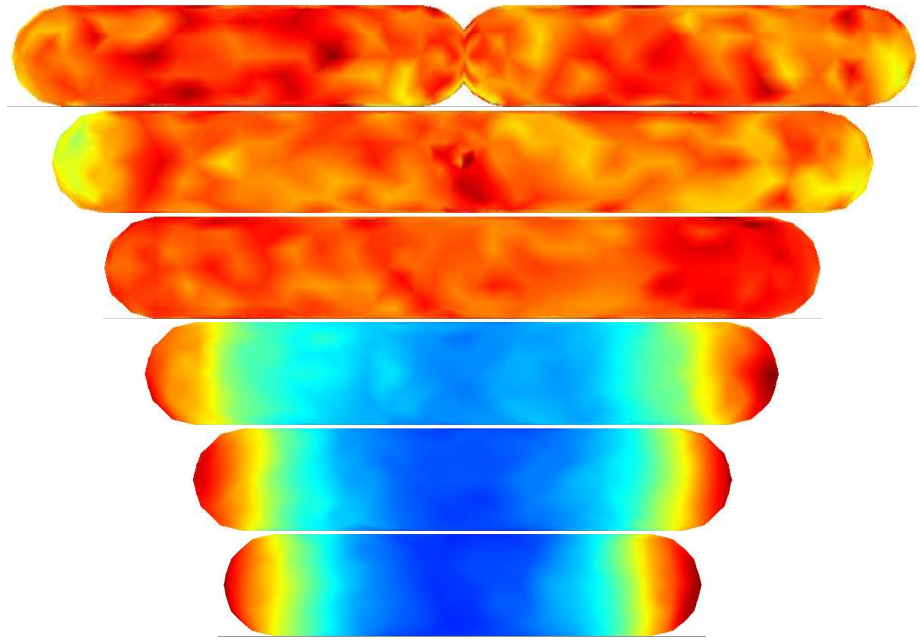


"URDME" software www.urdme.org.


Average concentration of MinD$_m$
1
0.8
0.6
0.4
0.2
0
−2.25
0
2.55
location (μ m)
x 10^{-6}


MinD$_m$ polar oscillations
1000
750
500
250
0
molecules
1000
1100
1200
1300
1400
1500
time (s)
Left
Right


Average concentration of MinD$_m$
x 10^{-6}
Growth (μ m)
Bacterium length (μ m)
x 10^{-6}
0
0.5
1
1.5
2
2.5
3
3.5
4
0
1
2
3
4
5
6
7
8
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2



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

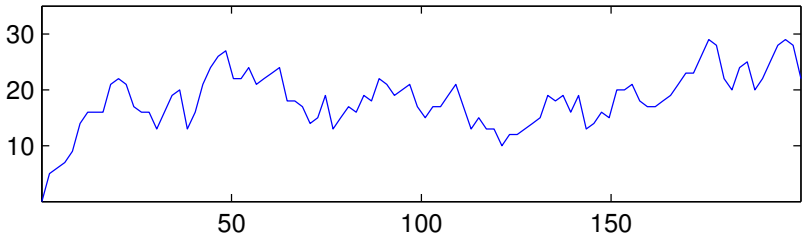
Physics: linear birth-death process with hidden parameters (k, μ):

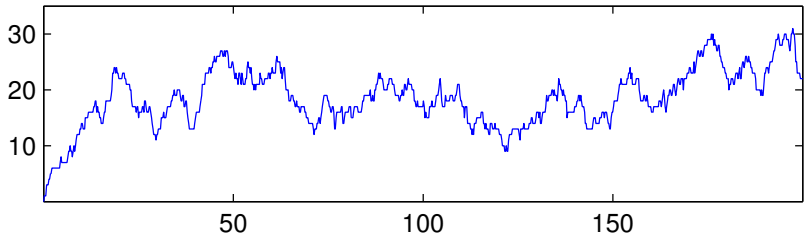
$$\left.\begin{array}{lcl} \emptyset & \xrightarrow{k} & X \\ X & \xrightarrow{\mu x} & \emptyset \end{array}\right\}$$

Task: find the rates (k, μ).

Convergence: increasing the temporal resolution:

Data:





Maximum Likelihood

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

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

$$P(X(t_1) = x_1 | X(0) = x_0) \propto \exp\left(-[x_1 - x(t_1|x_0, t_0)]^2/2\sigma^2\right)$$
$$P(X(t_2) = x_2 | X(t_1) = x_1) \propto \exp\left(-[x_2 - x(t_2|x_1, t_1)]^2/2\sigma^2\right)$$
$$\ldots$$

Maximum Likelihood

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

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

$$P(X(t_1) = x_1 | X(0) = x_0) \propto \exp\left(-[x_1 - x(t_1|x_0, t_0)]^2/2\sigma^2\right)$$
$$P(X(t_2) = x_2 | X(t_1) = x_1) \propto \exp\left(-[x_2 - x(t_2|x_1, t_1)]^2/2\sigma^2\right)$$
$$\ldots$$

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

Maximum Likelihood

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

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

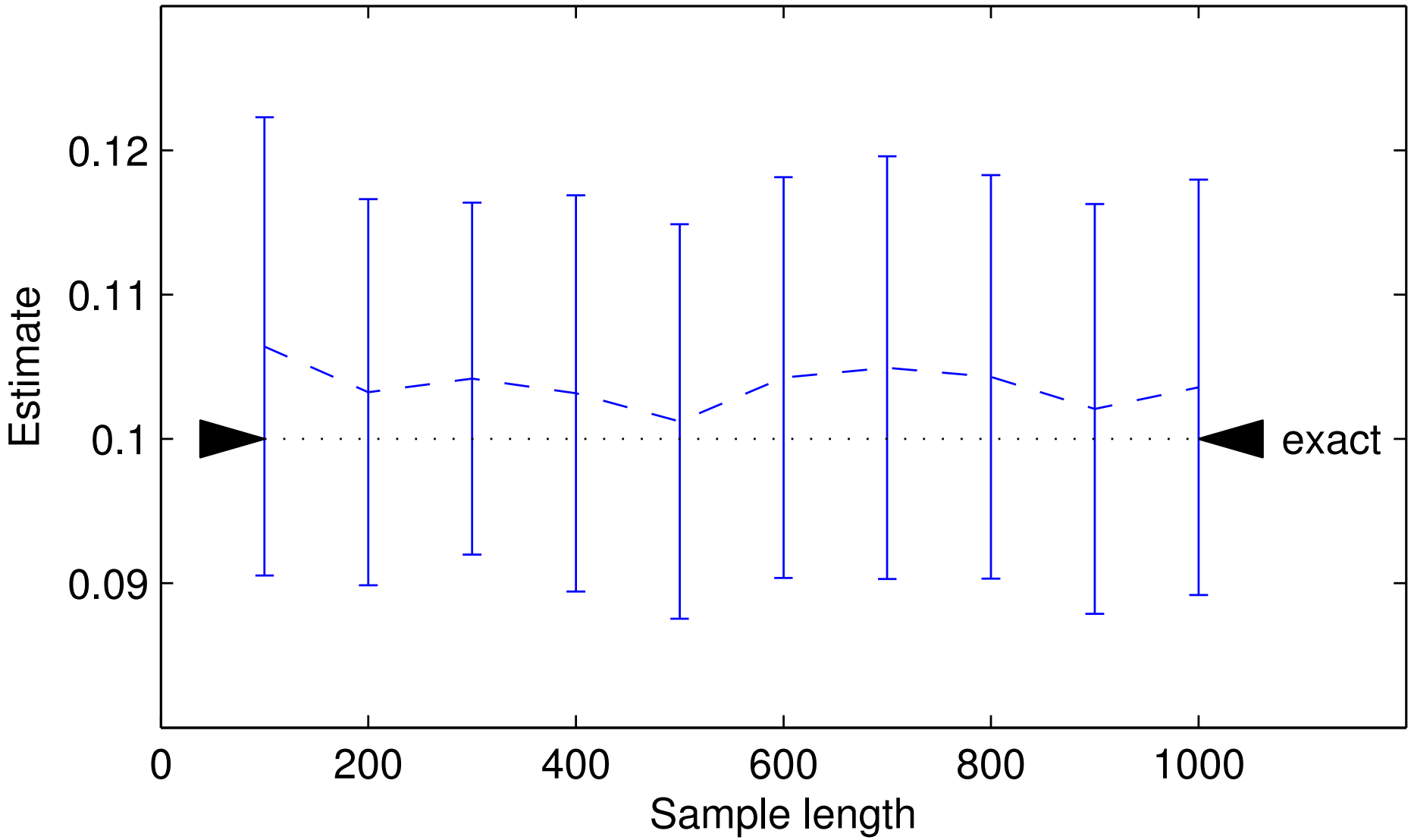
$$P(X(t_1) = x_1 | X(0) = x_0) \propto \exp\left(-[x_1 - x(t_1|x_0, t_0)]^2/2\sigma^2\right)$$
$$P(X(t_2) = x_2 | X(t_1) = x_1) \propto \exp\left(-[x_2 - x(t_2|x_1, t_1)]^2/2\sigma^2\right)$$
$$\dots$$

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

-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

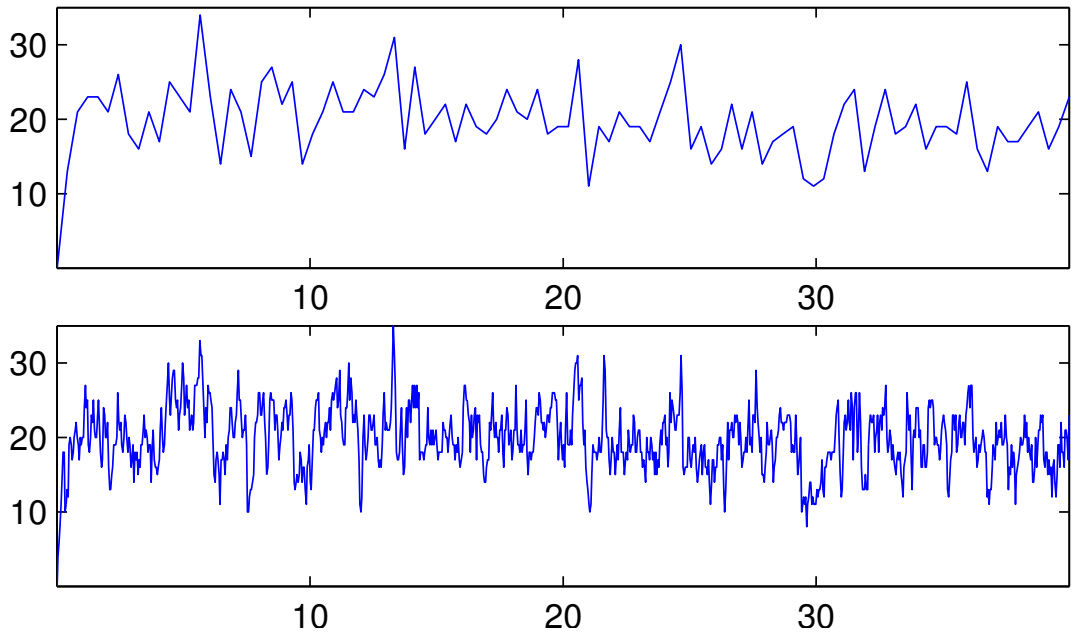
Results (μ)



Dimerization

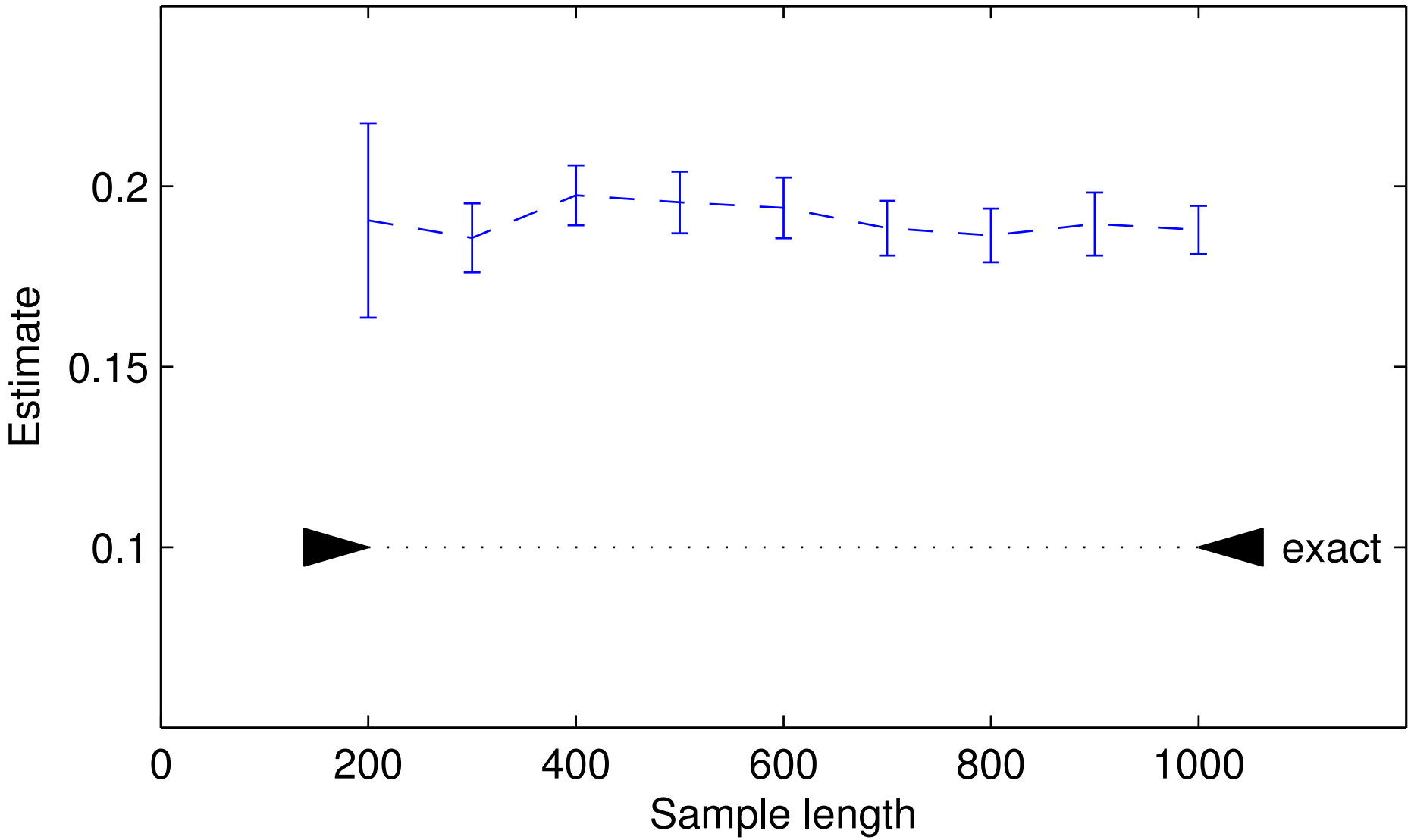
Slightly more difficult (nonlinear)

$$\left.\begin{array}{rcl} \emptyset & \xrightarrow{k} & X \\ X+X & \xrightarrow{\nu x(x-1)} & \emptyset \end{array}\right\}$$



Dimerization

Results (ν)



Maximum Likelihood using the CME

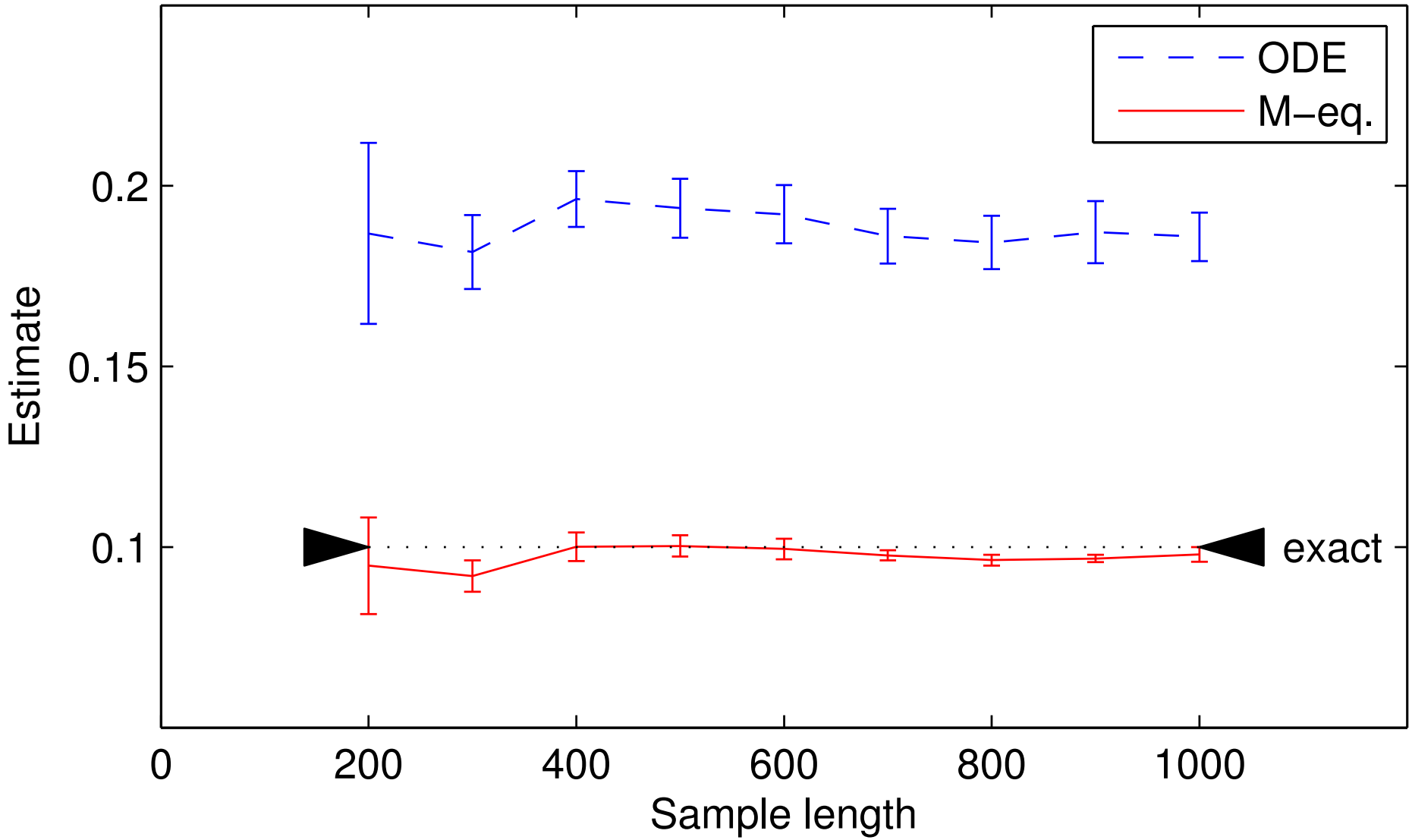
-Previously we used a Gaussian probability model. A better model is the chemical master equation.

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

-Observations are still independent *(Markov property)*.

Dimerization

Results (ν)

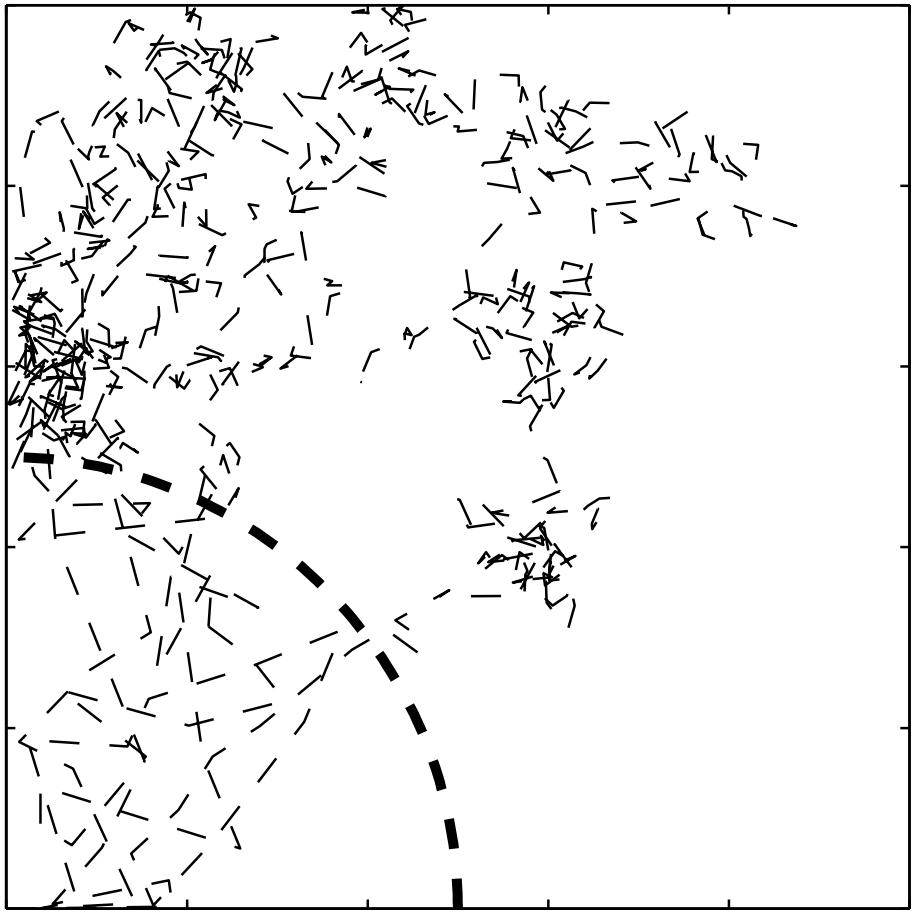


Diffusion rates from observations

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 nth observation had diffusion constant σ_k.

Expectation-Maximization algorithm

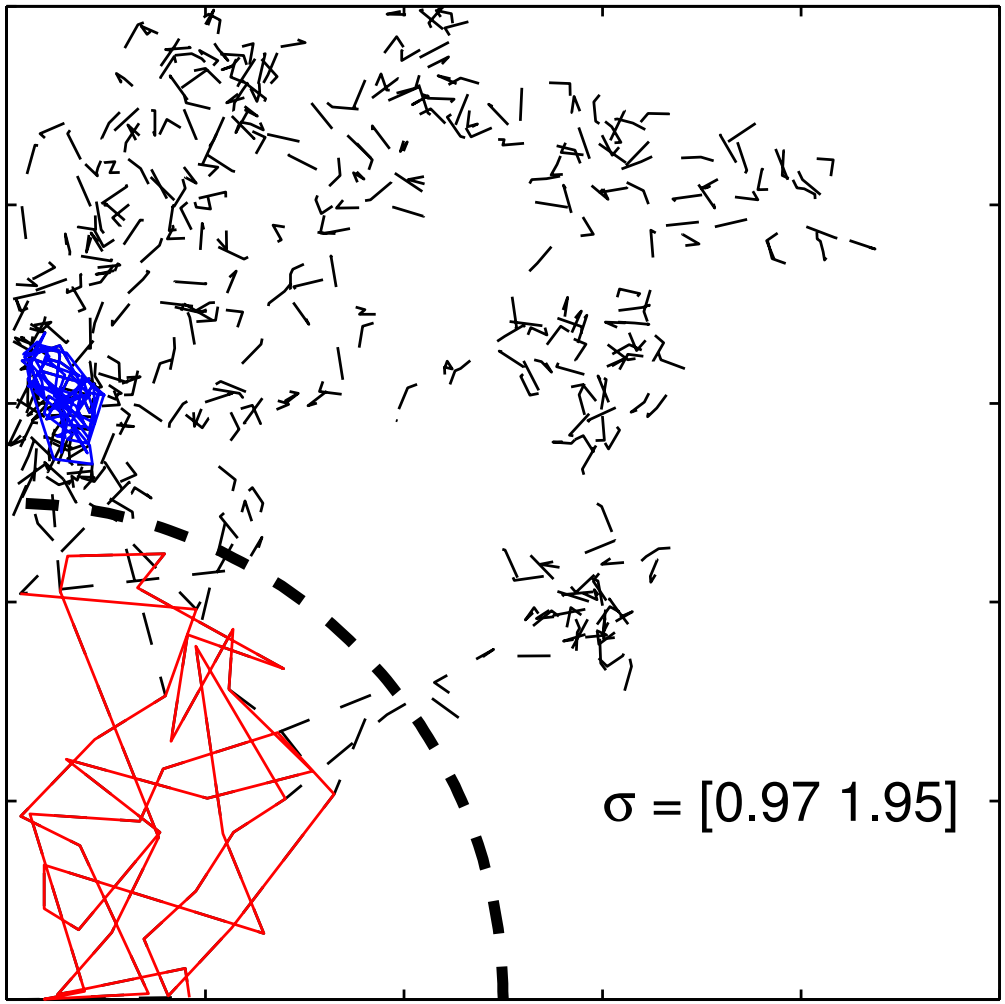
Briefly, an iterative algorithm which upon convergence produces a Maximum-Likelihood estimator of *(i)* $\sigma = [\sigma_1\ \sigma_2]$ as well as of *(ii)* p_{nk}, the probability that the nth observation had diffusion constant σ_k.

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.

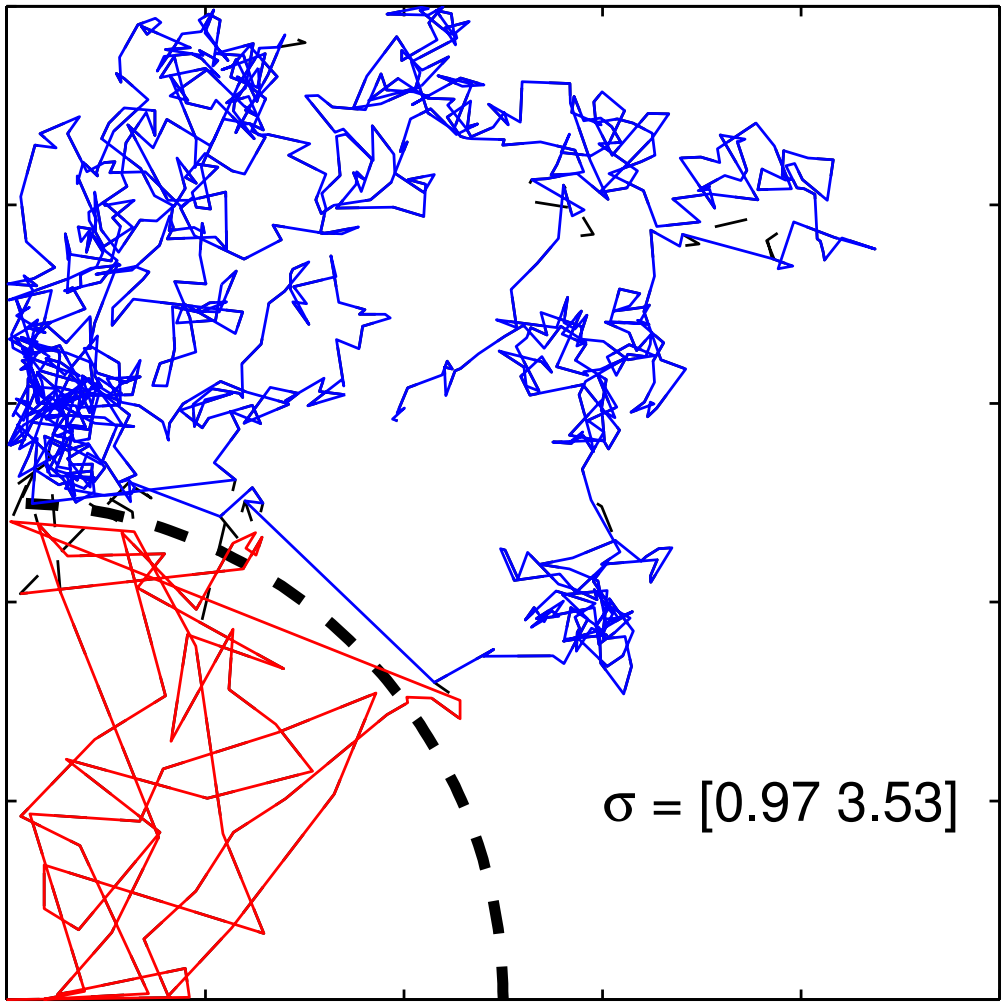
Diffusion rates

Results iteration #2



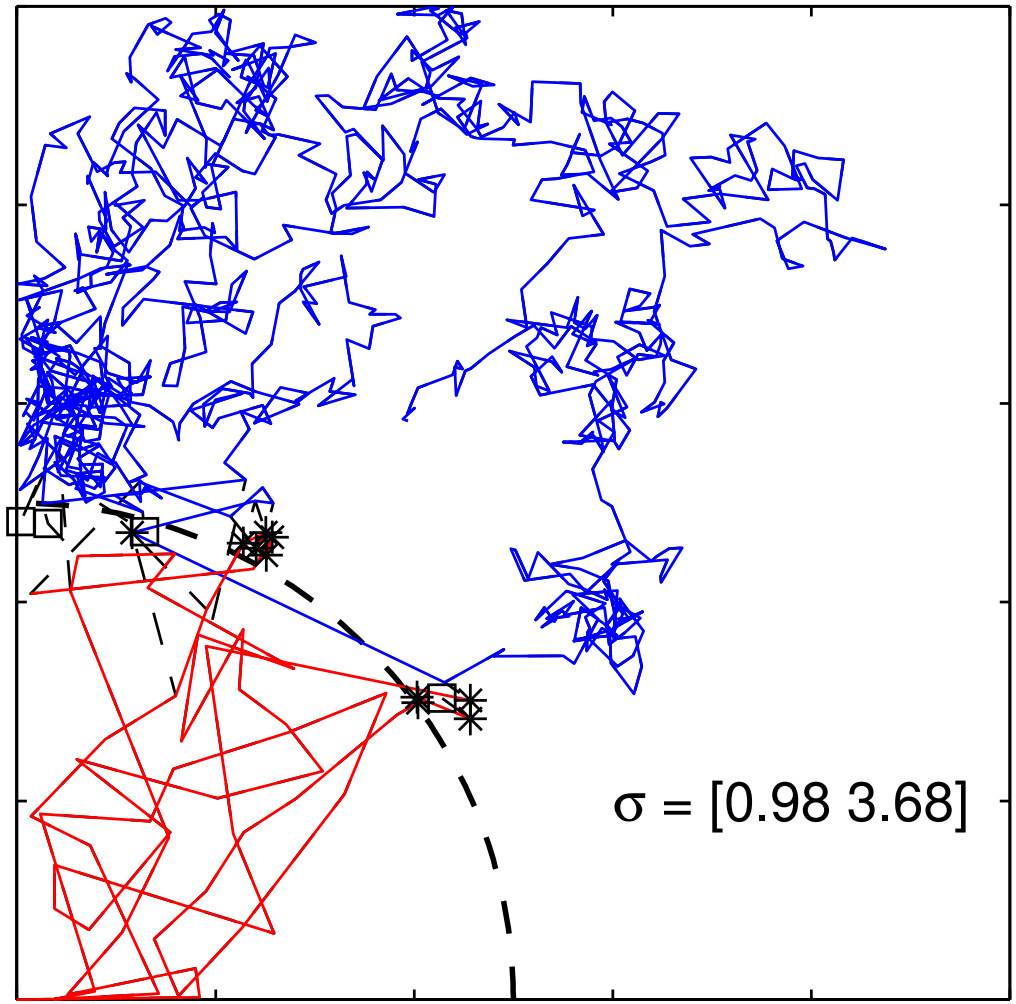
Diffusion rates

Results iteration #3



Diffusion rates

Results iteration #6



Optimal control of rates

Enzymatic reaction of a complex into a product,

$$C + E \xrightarrow{\nu\, C\cdot E} P + E.$$

Combine with

$$\emptyset \underset{\mu_E E}{\overset{s(t)}{\rightleftharpoons}} E, \quad \emptyset \underset{\mu_C C}{\overset{k_C}{\rightleftharpoons}} C, \quad P \xrightarrow{\mu_P P} \emptyset$$

such that E is under *control* through the signal $s(t)$.

Optimal control of rates (cont)

Maximize

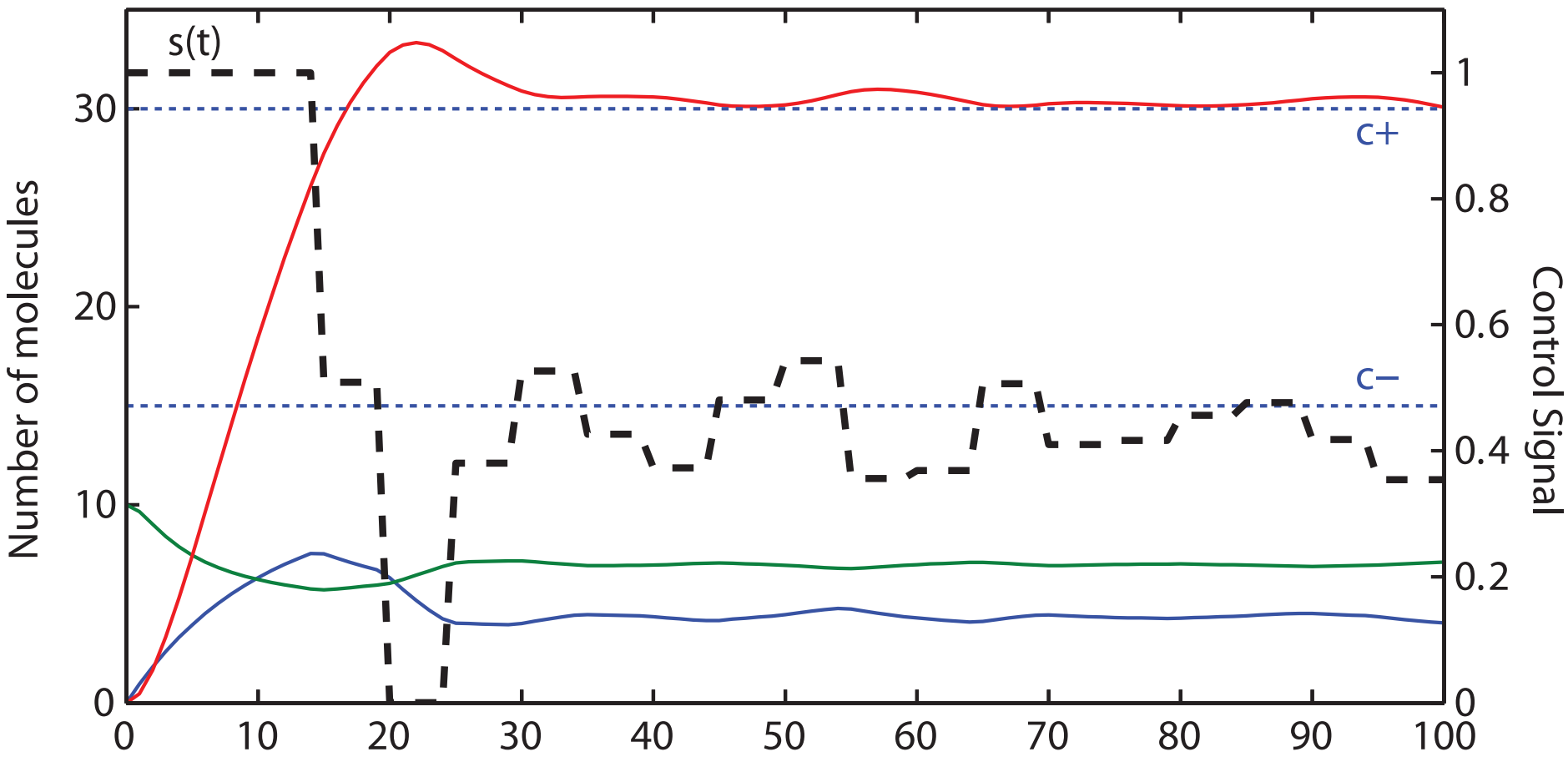
$$\mathcal{M}[P] := \int_0^T \varphi(P_t)\, dt,$$

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

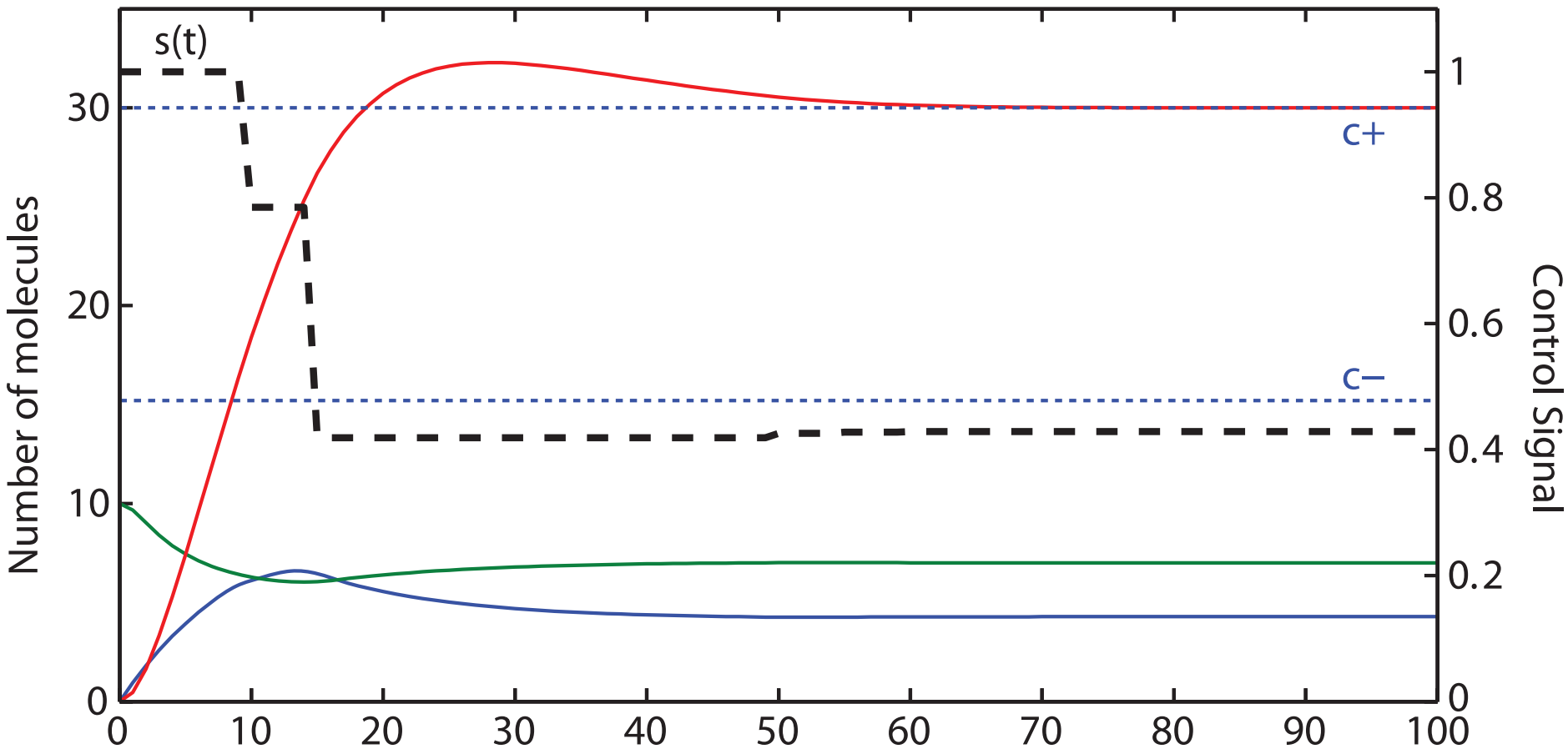
$$\left.\begin{array}{lcll} \varphi(P) & = & 0, & P \le c_- \\ \varphi(P) & = & \tau(P - c_-), & c_- < P \le C_+ \\ \varphi(P) & = & \tau(C_+ - c_-), & C_+ < P \end{array}\right\}$$

Constraints on the production signal s

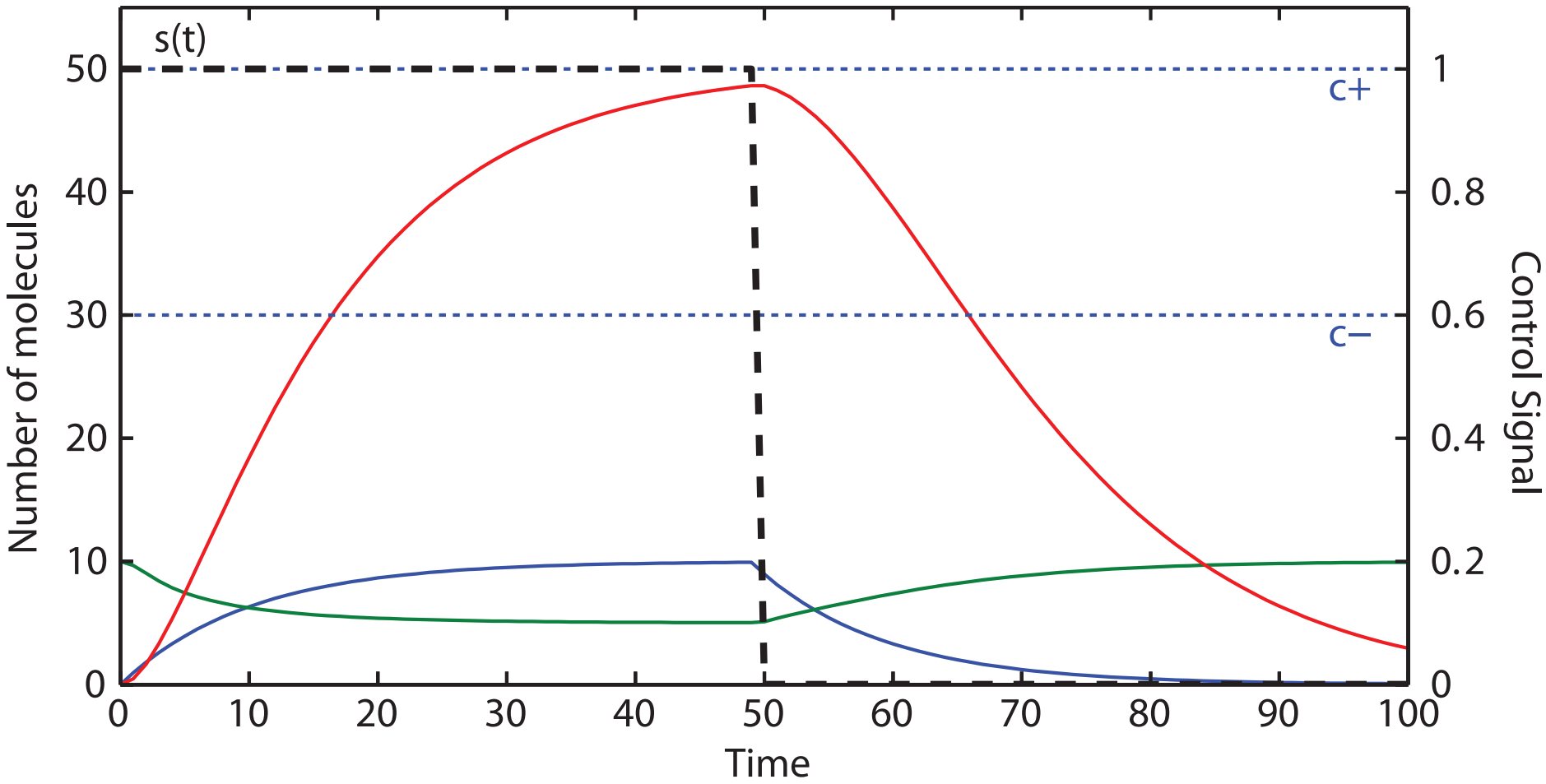
$$\left.\begin{array}{l} \max_{t\in[0,T]} s(t) \le S_\infty, \\ \int_0^T s(t)\, dt \le S_1, \end{array}\right\}$$



-Results from non-spatial deterministic ODE.



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


s(t)
c+
c−
Number of molecules
Control Signal
Time
0
10
20
30
40
50
60
70
80
90
100
0
0.2
0.4
0.6
0.8
1

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`)
- 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!