Deterministic and stochastic modelling of biochemical processes

An introduction and open problems

Tomáš Vejchodský

vejchod@math.cas.cz

Wolfson Centre
for Mathematical Biology
Mathematical Institute

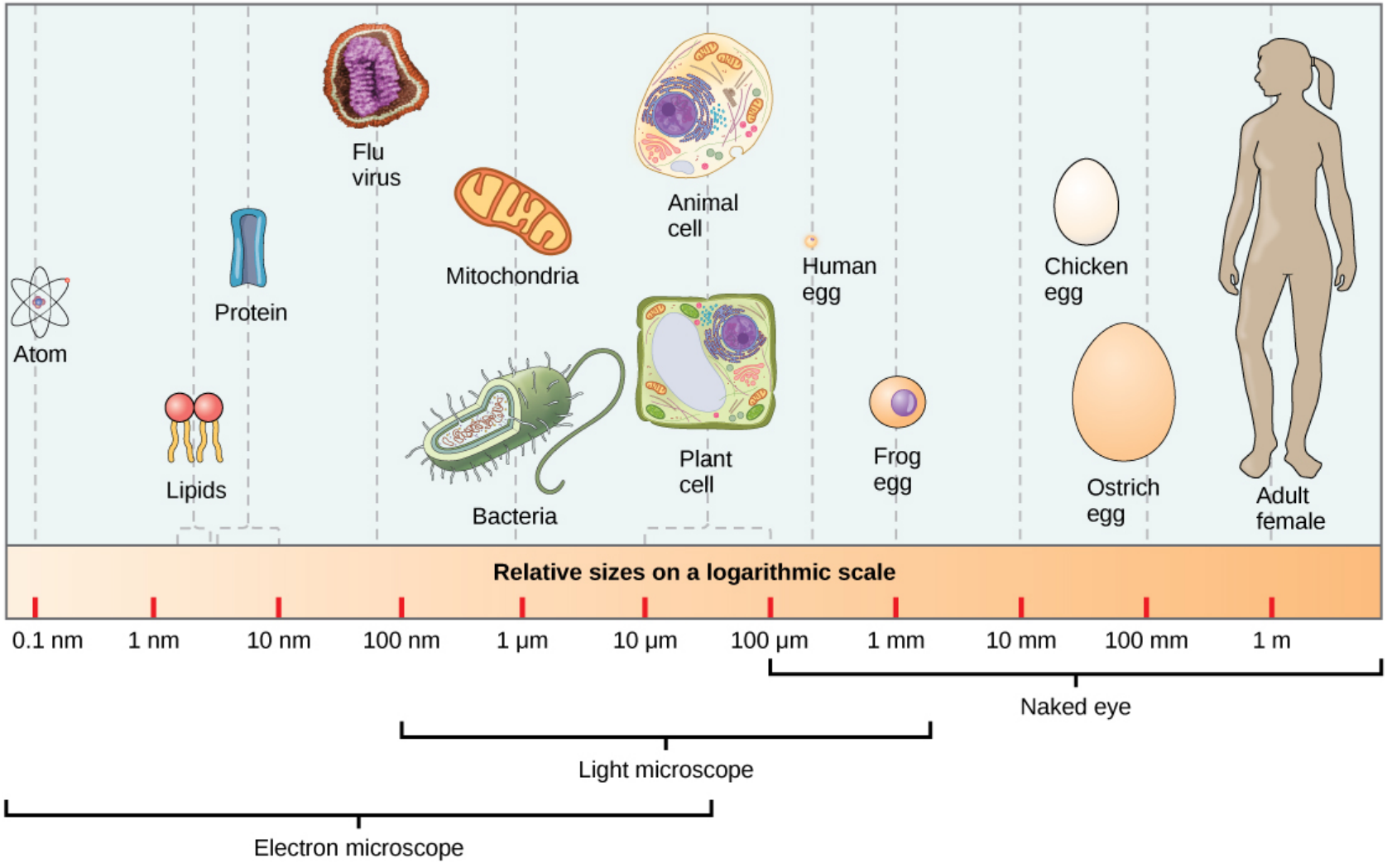
Institute of Mathematics
Academy of Sciences
Czech Republic

AIME@CZ, Prague, 11–12 March 2014

How does the life function?

How does the life function?

Levels and scales

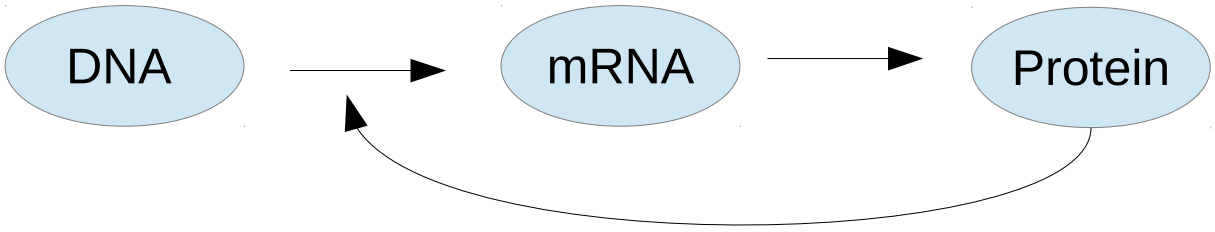


Biochemistry

Protein production

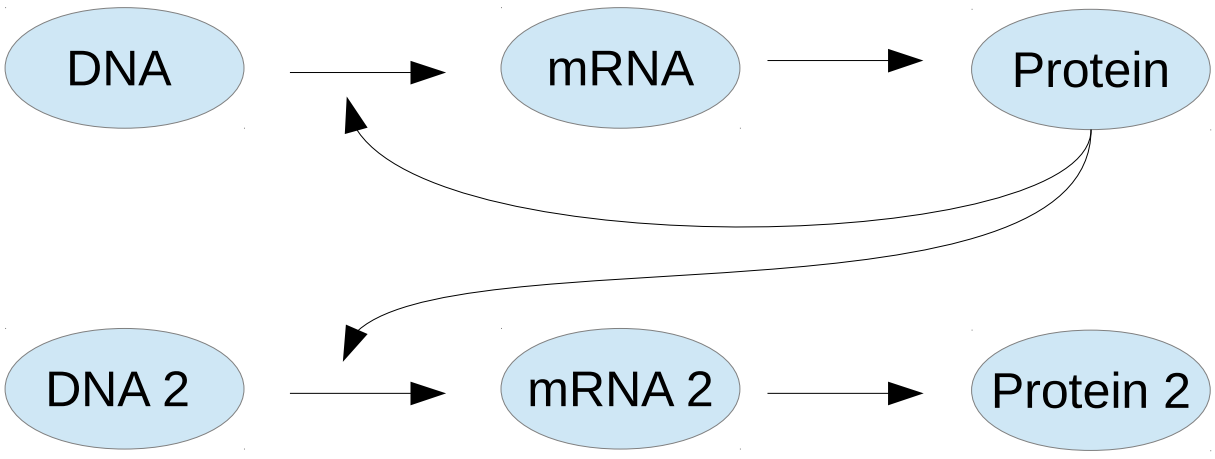
Biochemistry

Feedback loops (transcription factors)



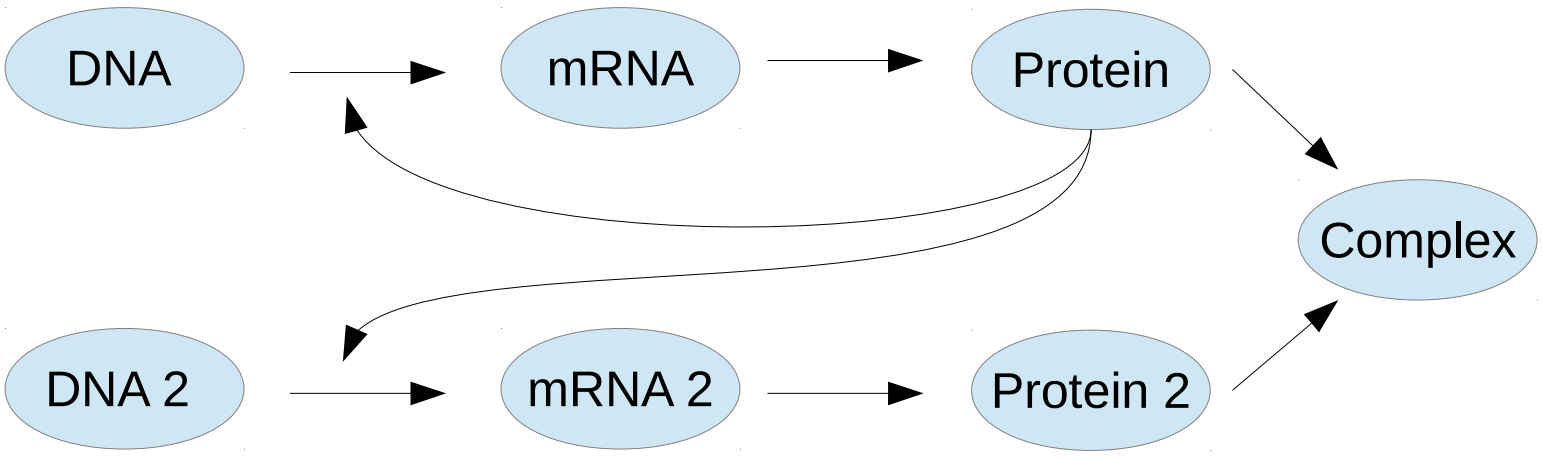
Biochemistry

Feedback loops (transcription factors)

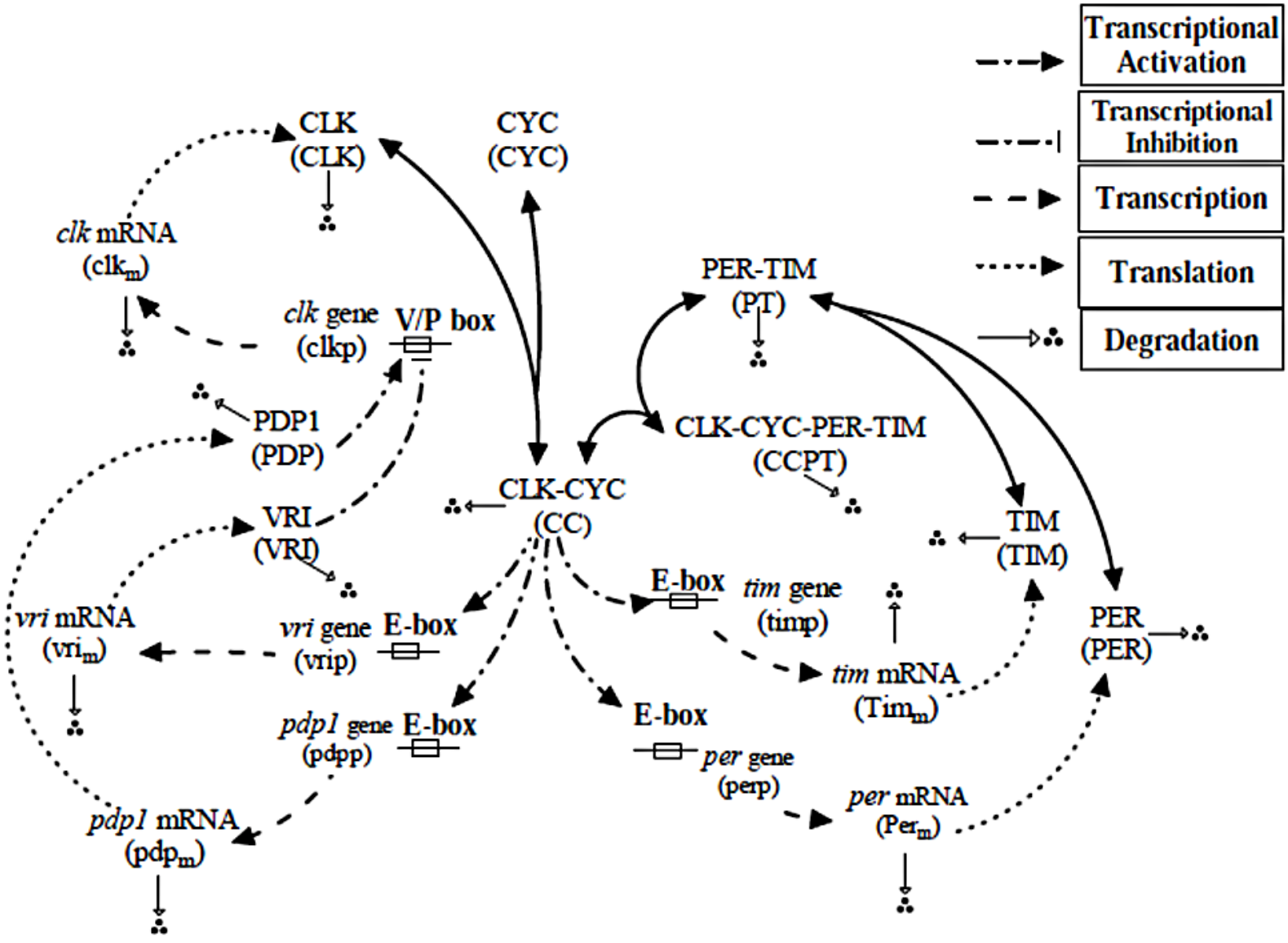


Biochemistry

Feedback loops (transcription factors)

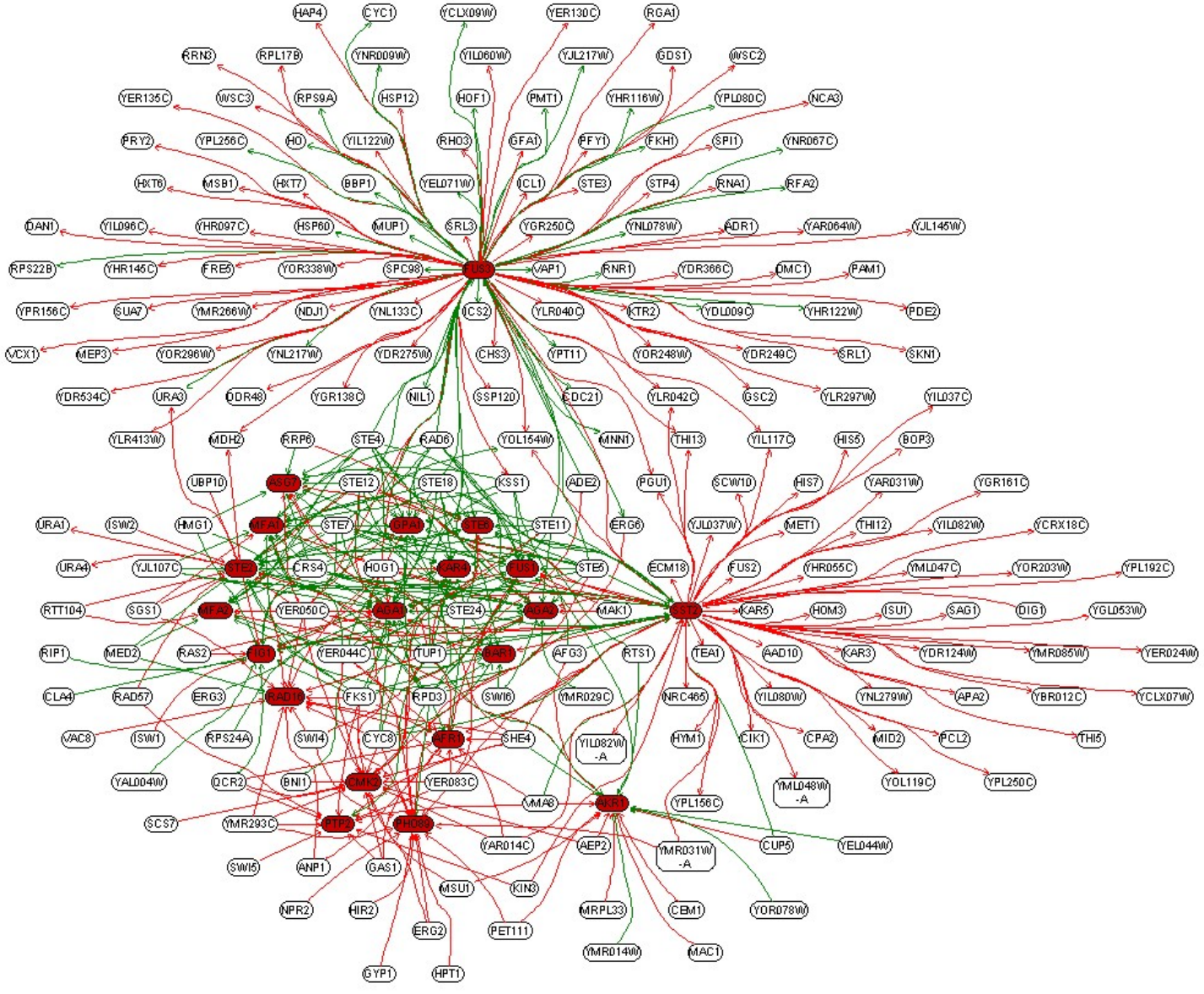


Gene regulatory networks



Circadian rhythms in Drosophila [Xie, Kulasiri, 2007]

Gene regulatory networks



Neighbourhood of mating response genes in yeast

[Rung, Schlitt, et al, 2002]

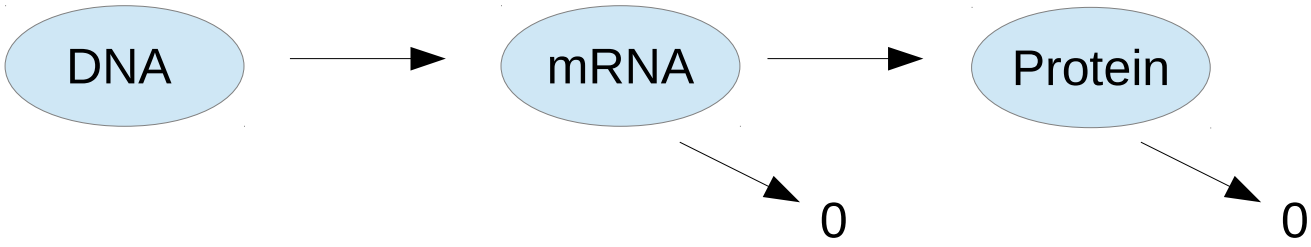
Outline

1. Mathematical models of biochemical systems
2. Deterministic vs. stochastic models
3. Modelling and numerics

1. Mathematical models of biochemical systems

Deterministic model – law of mass action

Protein production



Chemical system

$$D \xrightarrow{\alpha_R} D + R \qquad R \xrightarrow{\delta_R} \emptyset$$
$$R \xrightarrow{\alpha_P} R + P \qquad P \xrightarrow{\delta_P} \emptyset$$

Mass action ODE

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \alpha_R D - \delta_R R$$
$$\frac{\mathrm{d}P}{\mathrm{d}t} = \alpha_P R - \delta_P P$$

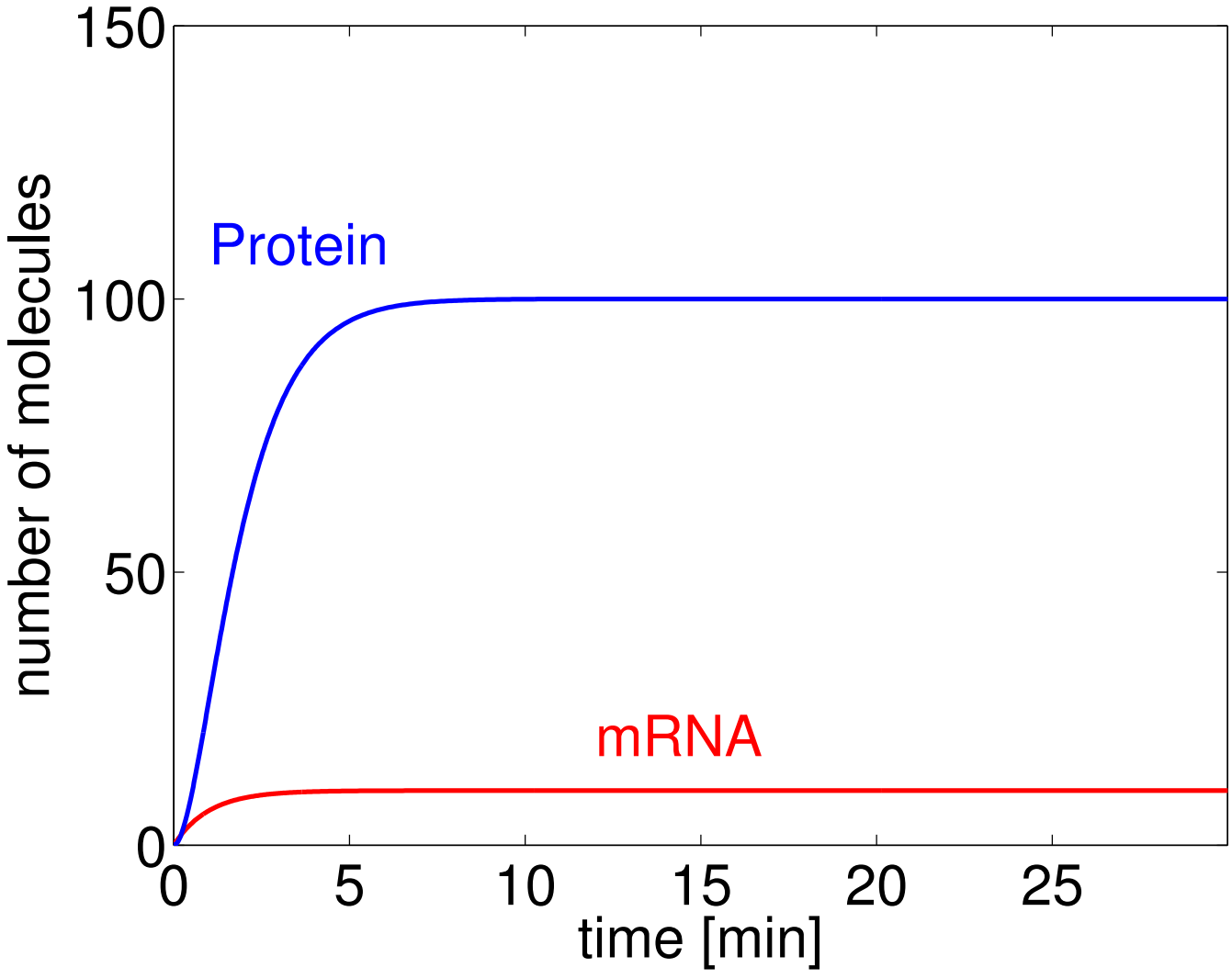
Notation

$D = D(t)$ ... number of DNA molecules
$R = R(t)$ ... number of mRNA molecules
$P = P(t)$ ... number of Protein molecules

Initial condition

$D(0) = 1,\ R(0) = 0,\ P(0) = 0$

Deterministic model – law of mass action



$$\alpha_R = \alpha_P = 10, \quad \delta_R = \delta_P = 1 \quad [\text{min}^{-1}]$$

Stochastic model

Chemical system

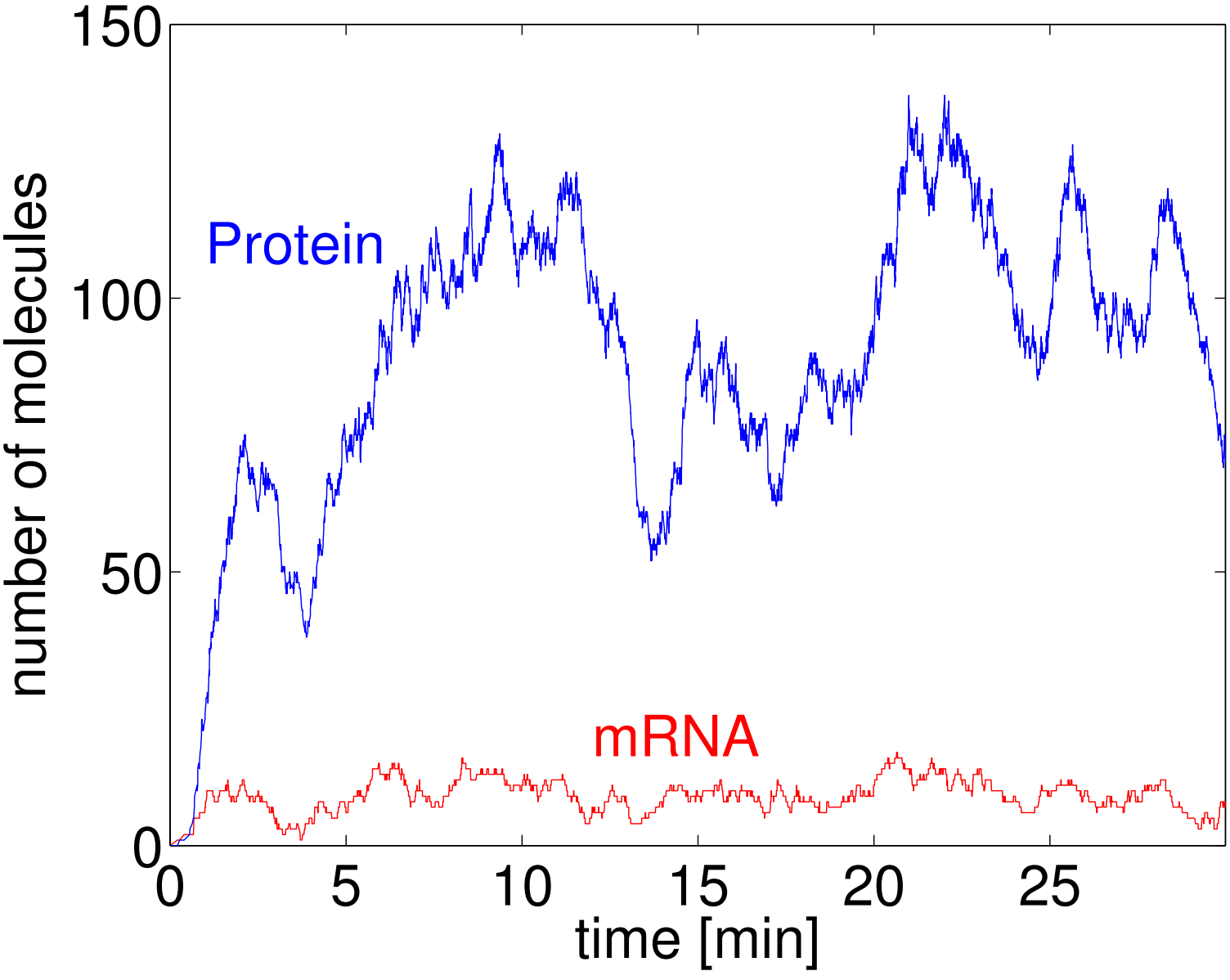
$$D \xrightarrow{\alpha_R} D + R \qquad R \xrightarrow{\delta_R} \emptyset$$
$$R \xrightarrow{\alpha_P} R + P \qquad P \xrightarrow{\delta_P} \emptyset$$

Discrete space continuous time Markov process

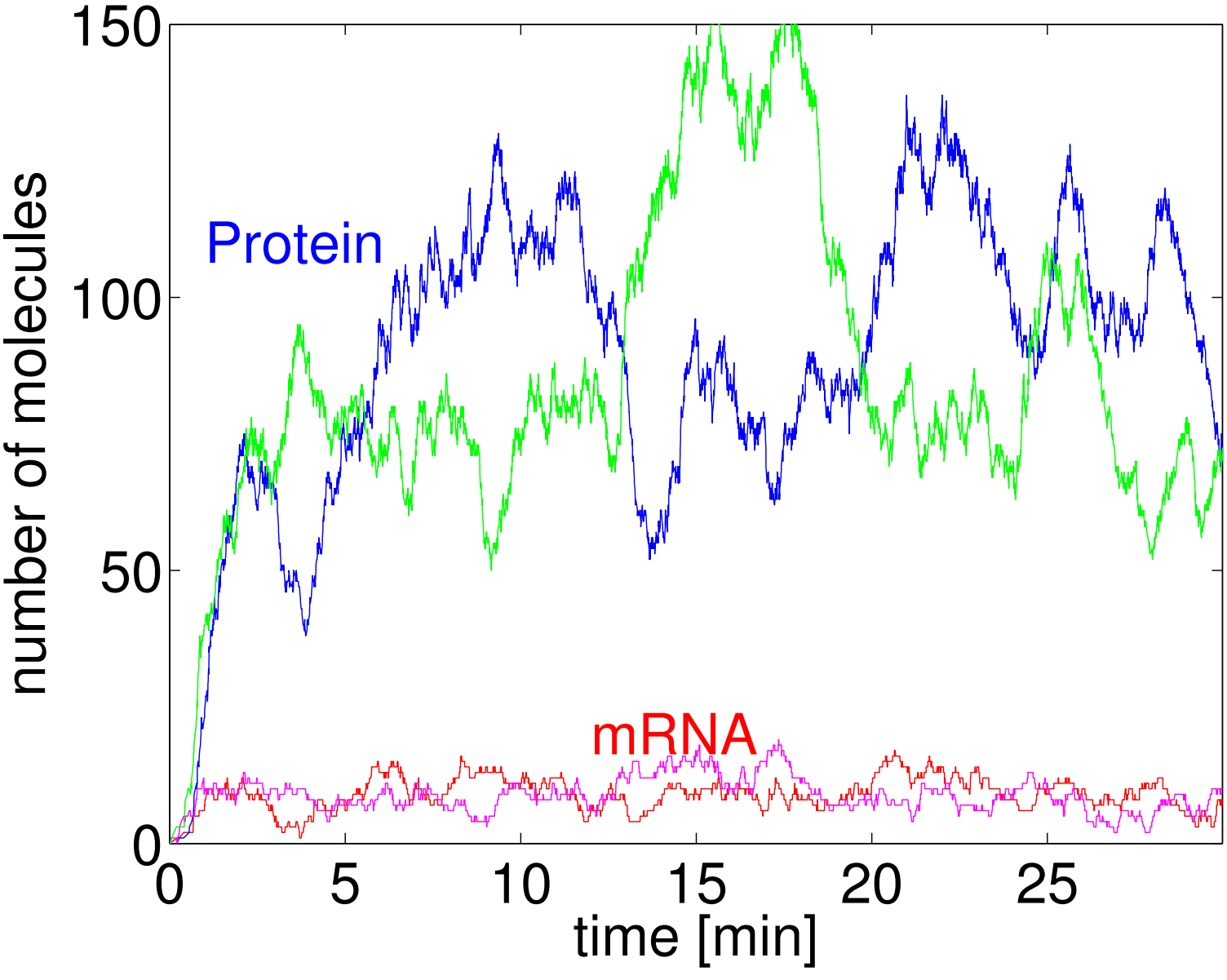
- State at time t: $[R, P](t)$
- Change of state:
  $[R, P](t) \mapsto [R+1, P](t+\mathrm{d}t)$ with probability $\alpha_R D\,\mathrm{d}t$
  $[R, P](t) \mapsto [R-1, P](t+\mathrm{d}t)$ with probability $\delta_R R\,\mathrm{d}t$
  $[R, P](t) \mapsto [R, P+1](t+\mathrm{d}t)$ with probability $\alpha_P R\,\mathrm{d}t$
  $[R, P](t) \mapsto [R, P-1](t+\mathrm{d}t)$ with probability $\delta_P P\,\mathrm{d}t$

Gillespie stochastic simulation algorithm

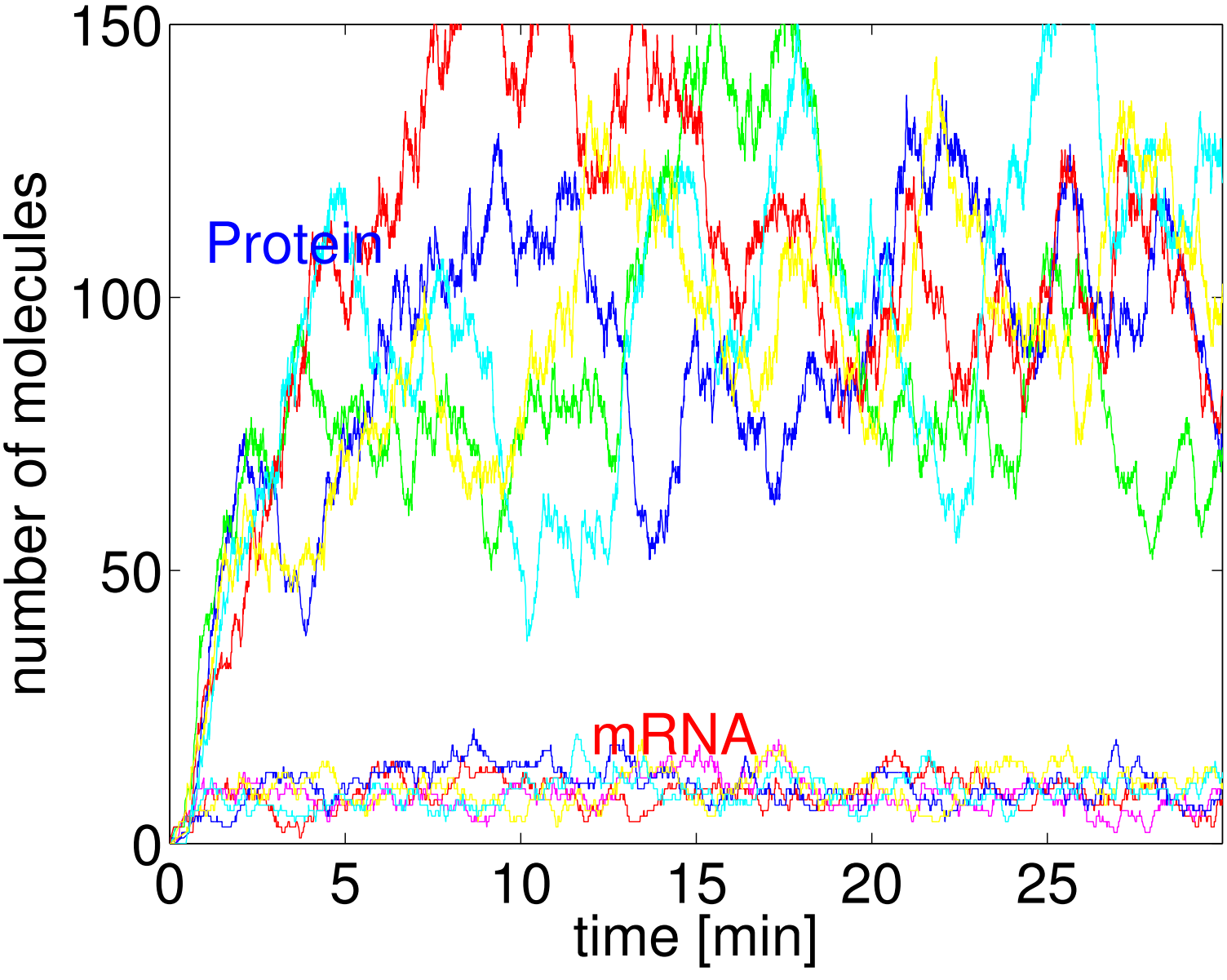
Stochastic model – Gillespie algorithm



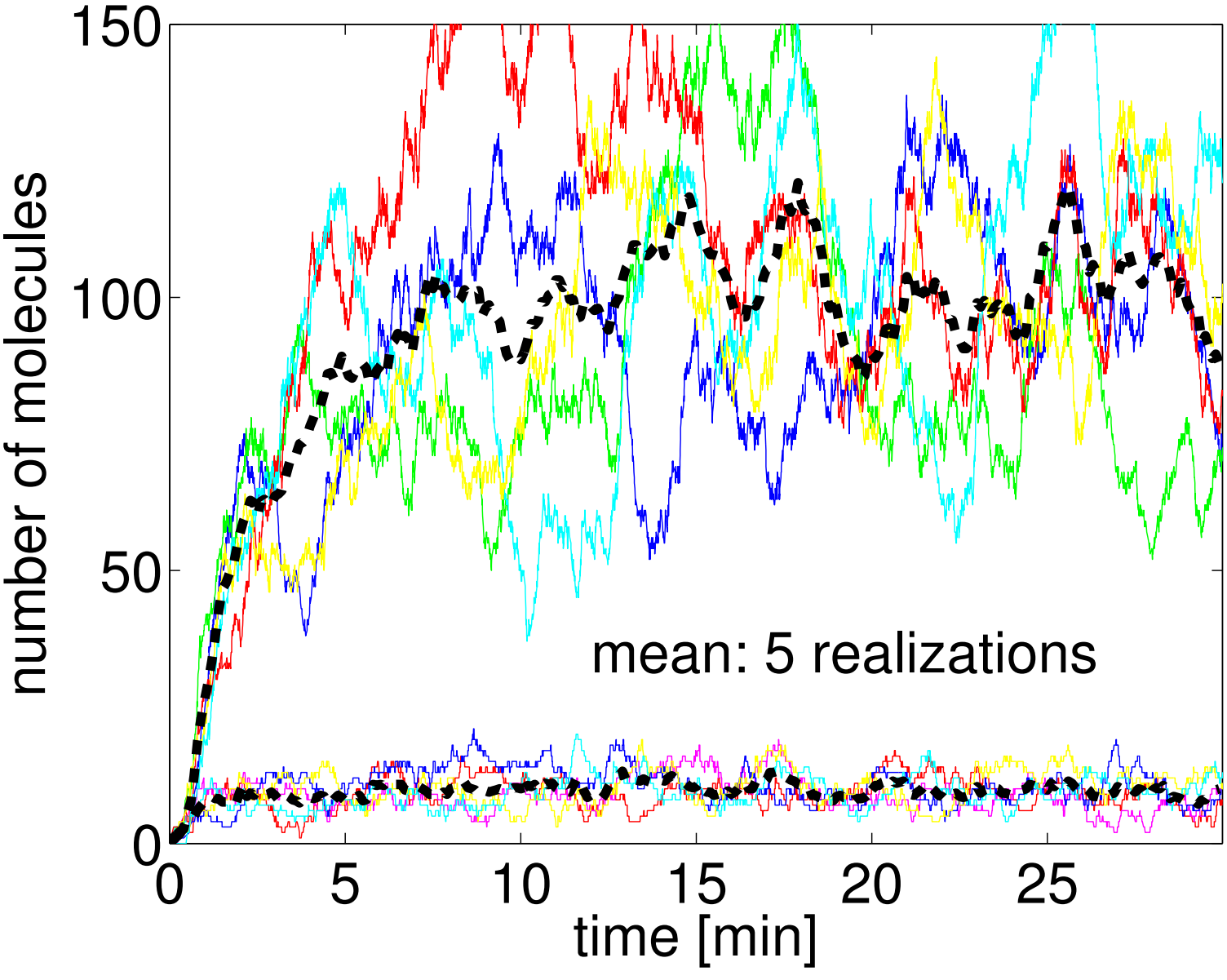
Stochastic model – Gillespie algorithm



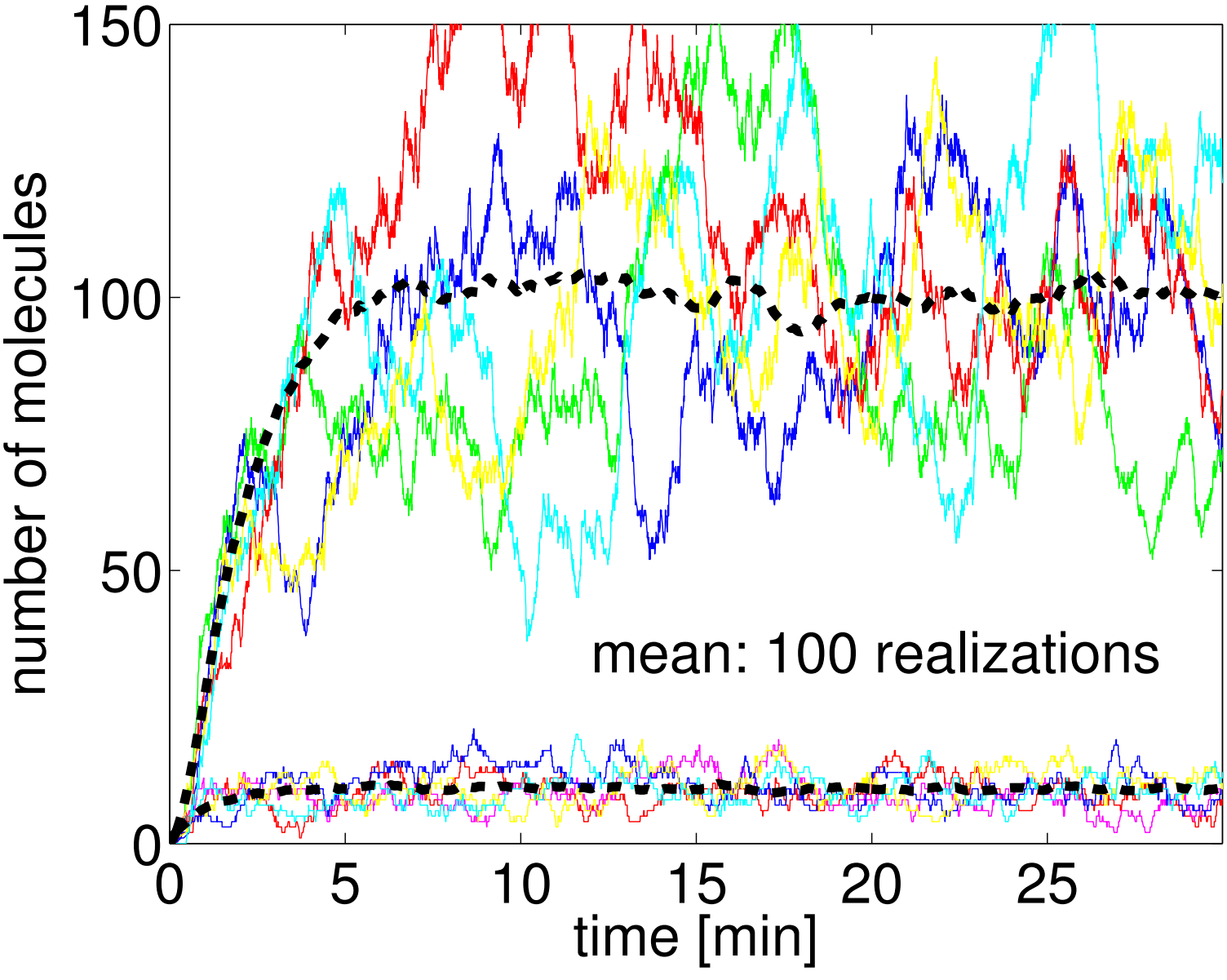
Stochastic model – Gillespie algorithm



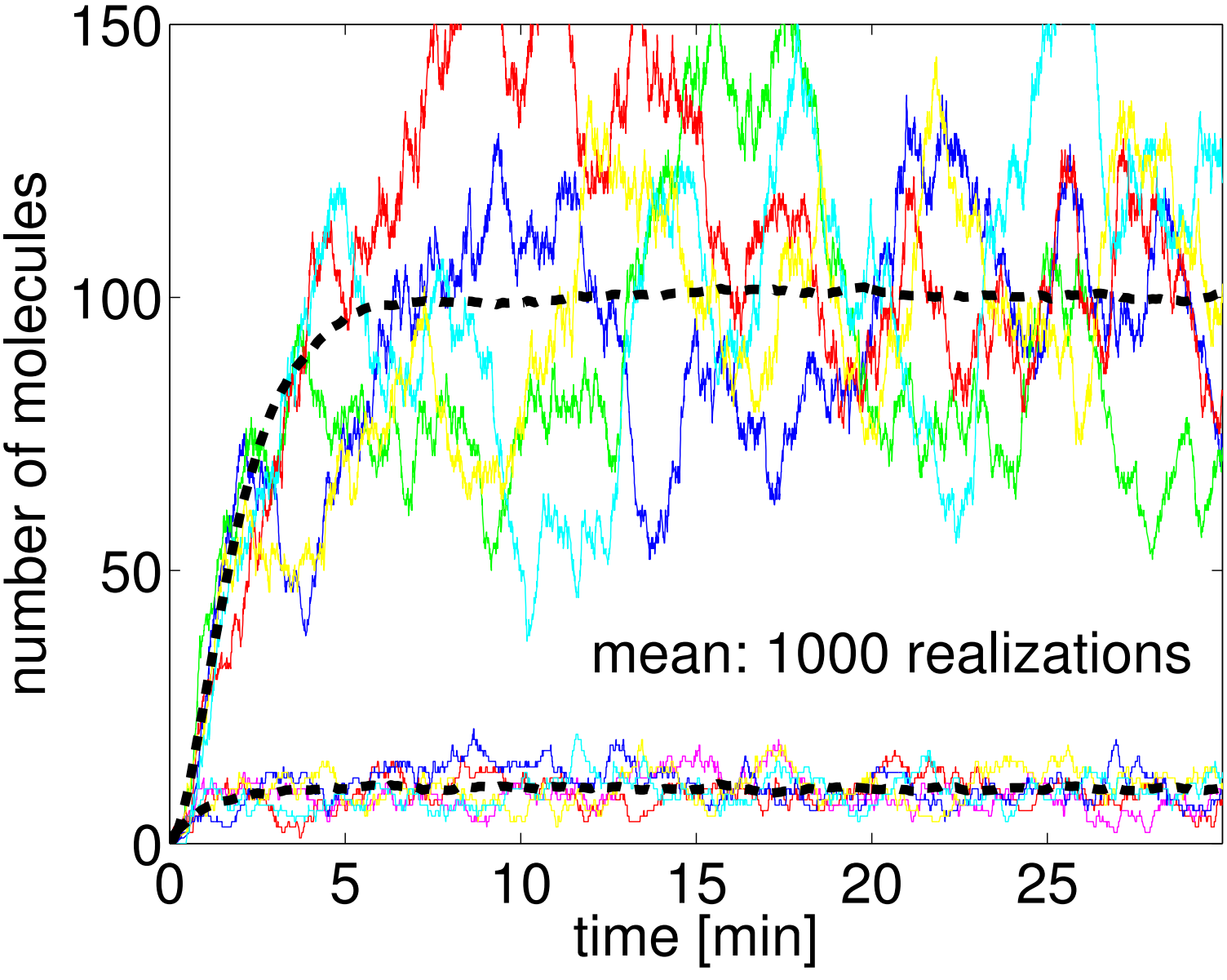
Stochastic model – Gillespie algorithm



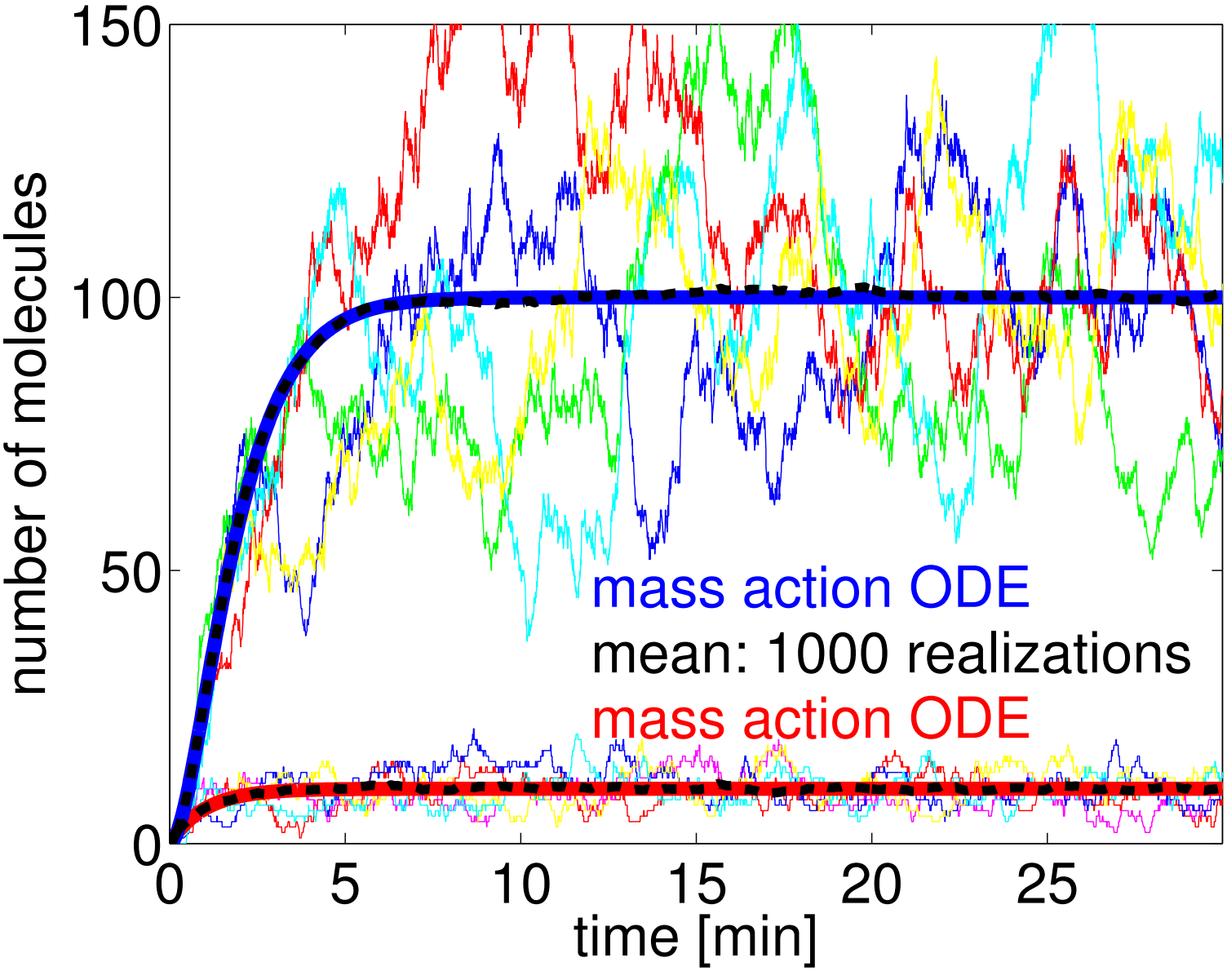
Stochastic model – Gillespie algorithm



Stochastic model – Gillespie algorithm



Stochastic model – Gillespie algorithm



Stochastic model – analysis

Definition: $p_{n,m}(t) = \Pr[R(t) = n, P(t) = m]$

Chemical master equation (CME):

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{n,m} = \alpha_R D p_{n-1,m} - \alpha_R D p_{n,m} + \delta_R (n+1) p_{n+1,m} - \delta_R n p_{n,m}$$
$$+ \alpha_P n p_{n,m-1} - \alpha_P n p_{n,m} + \delta_P (m+1) p_{n,m+1} - \delta_P m p_{n,m}$$
$$n, m = 0, 1, 2, \ldots$$

Stochastic model – analysis

Definition: $p_{n,m}(t) = \Pr[R(t) = n, P(t) = m]$

Chemical master equation (CME):

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{n,m} = \alpha_R D p_{n-1,m} - \alpha_R D p_{n,m} + \delta_R (n+1) p_{n+1,m} - \delta_R n p_{n,m} + \alpha_P n p_{n,m-1} - \alpha_P n p_{n,m} + \delta_P (m+1) p_{n,m+1} - \delta_P m p_{n,m}$$

$$n, m = 0, 1, 2, \ldots$$

Definition: $p(x, y, t) \approx \Pr[R(t) = x, P(t) = y]$

Chemical Fokker-Planck equation (CFPE):

$$\frac{\partial p}{\partial t} = \mathrm{div}(\mathcal{A} \nabla p - \mathbf{b} p), \qquad (x, y) \in (0, \infty)^2$$

where

$$\mathcal{A} = \frac{1}{2} \begin{bmatrix} \alpha_R D + \delta_R x & 0 \\ 0 & \alpha_P x + \delta_P y \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \alpha_R D - \delta_R x + \delta_R/2 \\ \alpha_P x - \delta_P y + \delta_P/2 \end{bmatrix}$$

Stochastic model – analysis

Definition: $p_{n,m}(t) = \Pr[R(t) = n, P(t) = m]$

Chemical master equation (CME):

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{n,m} = \alpha_R D p_{n-1,m} - \alpha_R D p_{n,m} + \delta_R (n+1) p_{n+1,m} - \delta_R n p_{n,m} + \alpha_P n p_{n,m-1} - \alpha_P n p_{n,m} + \delta_P (m+1) p_{n,m+1} - \delta_P m p_{n,m}$$

$$n, m = 0, 1, 2, \ldots$$

Definition: $p(x, y, t) \approx \Pr[R(t) = x, P(t) = y]$

Chemical Fokker-Planck equation (CFPE):

$$\frac{\partial p}{\partial t} = \mathrm{div}(\mathcal{A} \nabla p - \mathbf{b} p), \qquad (x, y) \in (0, \infty)^2$$

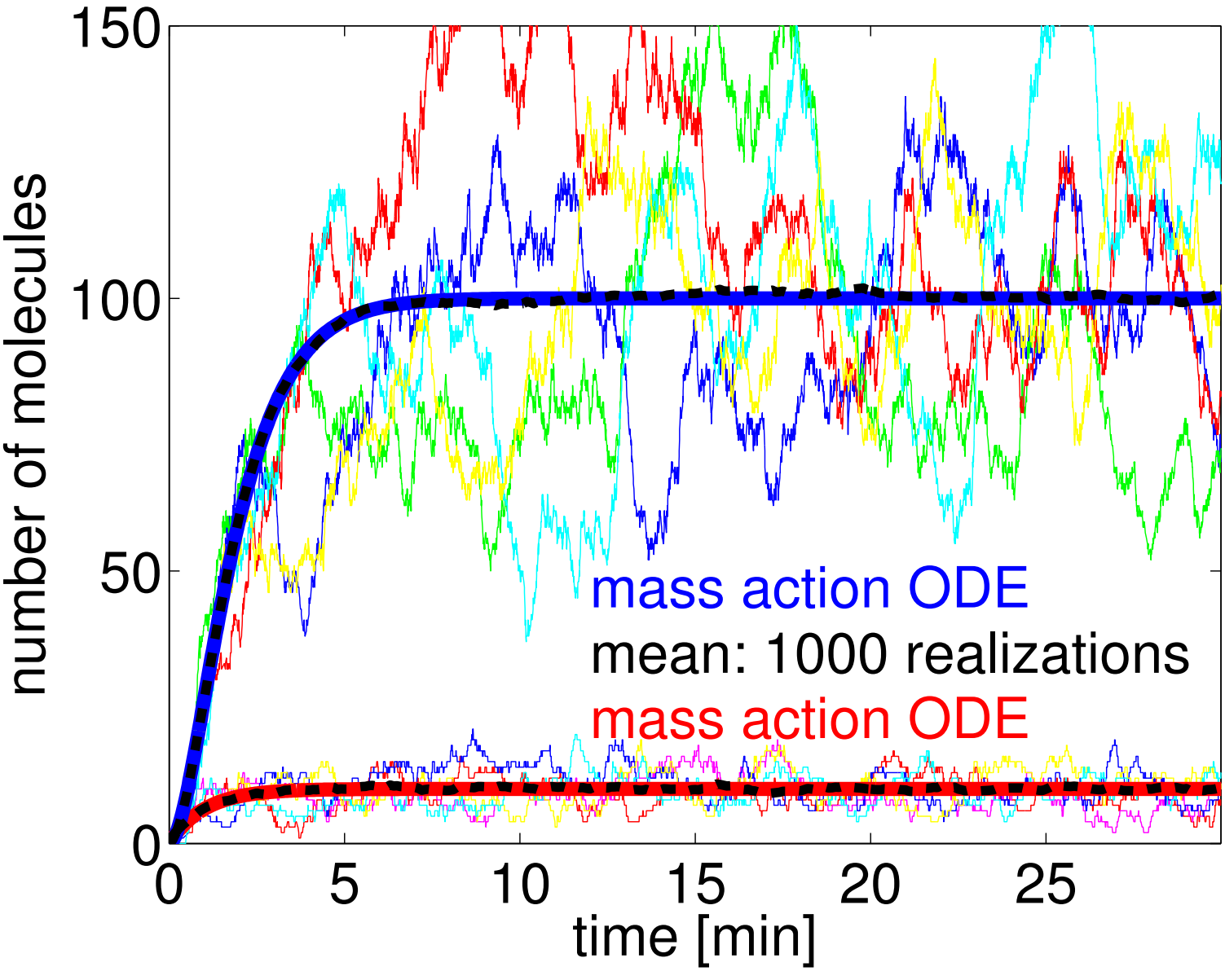
where

$$\mathcal{A} = \frac{1}{2} \begin{bmatrix} \alpha_R D + \delta_R x & 0 \\ 0 & \alpha_P x + \delta_P y \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \alpha_R D - \delta_R x + \delta_R/2 \\ \alpha_P x - \delta_P y + \delta_P/2 \end{bmatrix}$$

Stochastic differential equations

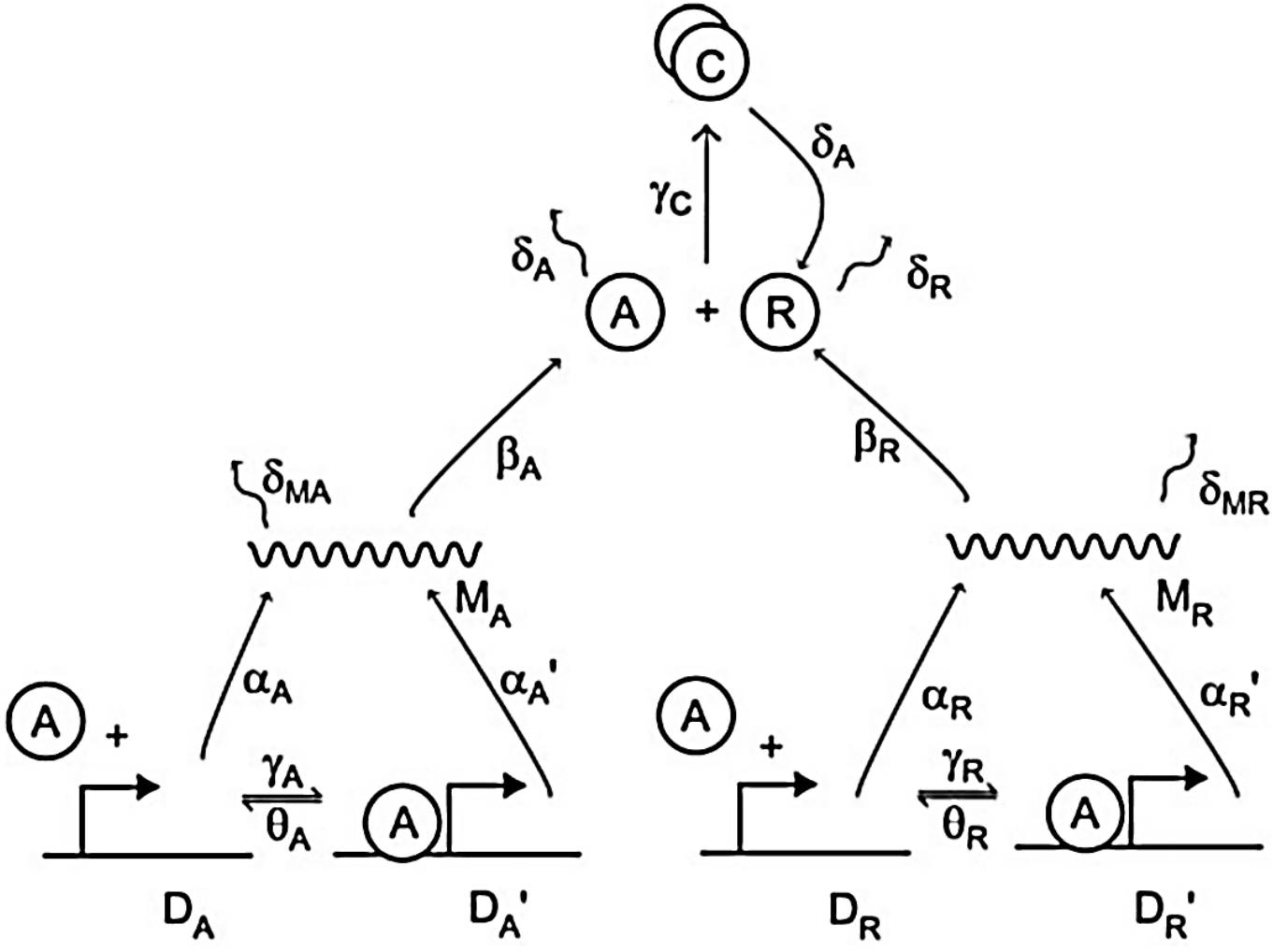
2. Deterministic vs. Stochastic

Protein production



Stochastic bifurcations

VKBL model of circadian rhythms



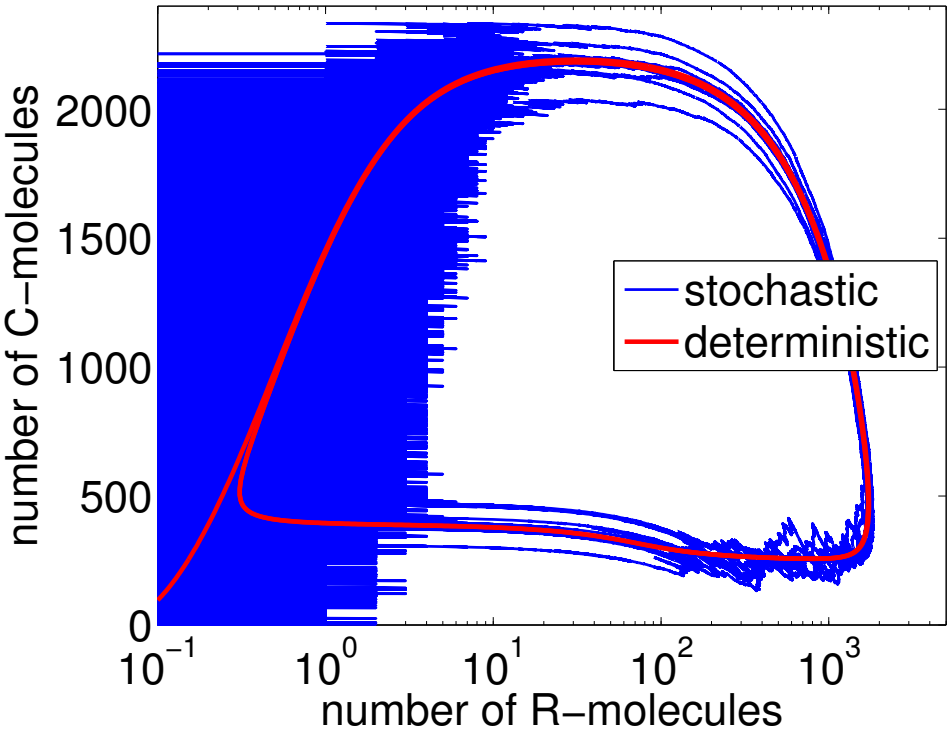
$\alpha_A = 50\,\mathrm{h}^{-1}$
$\alpha'_A = 500\,\mathrm{h}^{-1}$
$\alpha_R = 0.01\,\mathrm{h}^{-1}$
$\alpha'_R = 50\,\mathrm{h}^{-1}$
$\beta_A = 50\,\mathrm{h}^{-1}$
$\beta_R = 5\,\mathrm{h}^{-1}$
$\gamma_A = 1\,\mathrm{mol}^{-1}\ \mathrm{h}^{-1}$
$\gamma_R = 1\,\mathrm{mol}^{-1}\ \mathrm{h}^{-1}$
$\gamma_C = 2\,\mathrm{mol}^{-1}\ \mathrm{h}^{-1}$
$\delta_A = 1\,\mathrm{h}^{-1}$
$\delta_R = 0.2\,\mathrm{h}^{-1}$
$\delta_{M_A} = 10\,\mathrm{h}^{-1}$
$\delta_{M_R} = 0.5\,\mathrm{h}^{-1}$
$\theta_A = 50\,\mathrm{h}^{-1}$
$\theta_R = 100\,\mathrm{h}^{-1}$

[Vilar, Kueh, Barkai, Leibler, 2002]

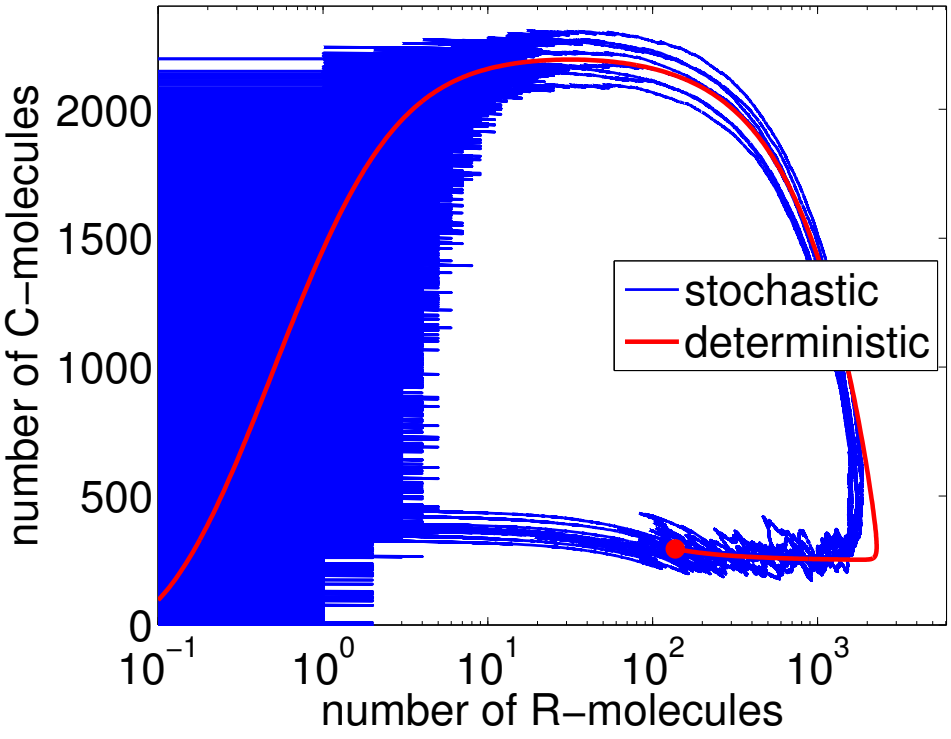
Stochastic bifurcations

VKBL model: phase diagram

$\delta_R = 0.2$

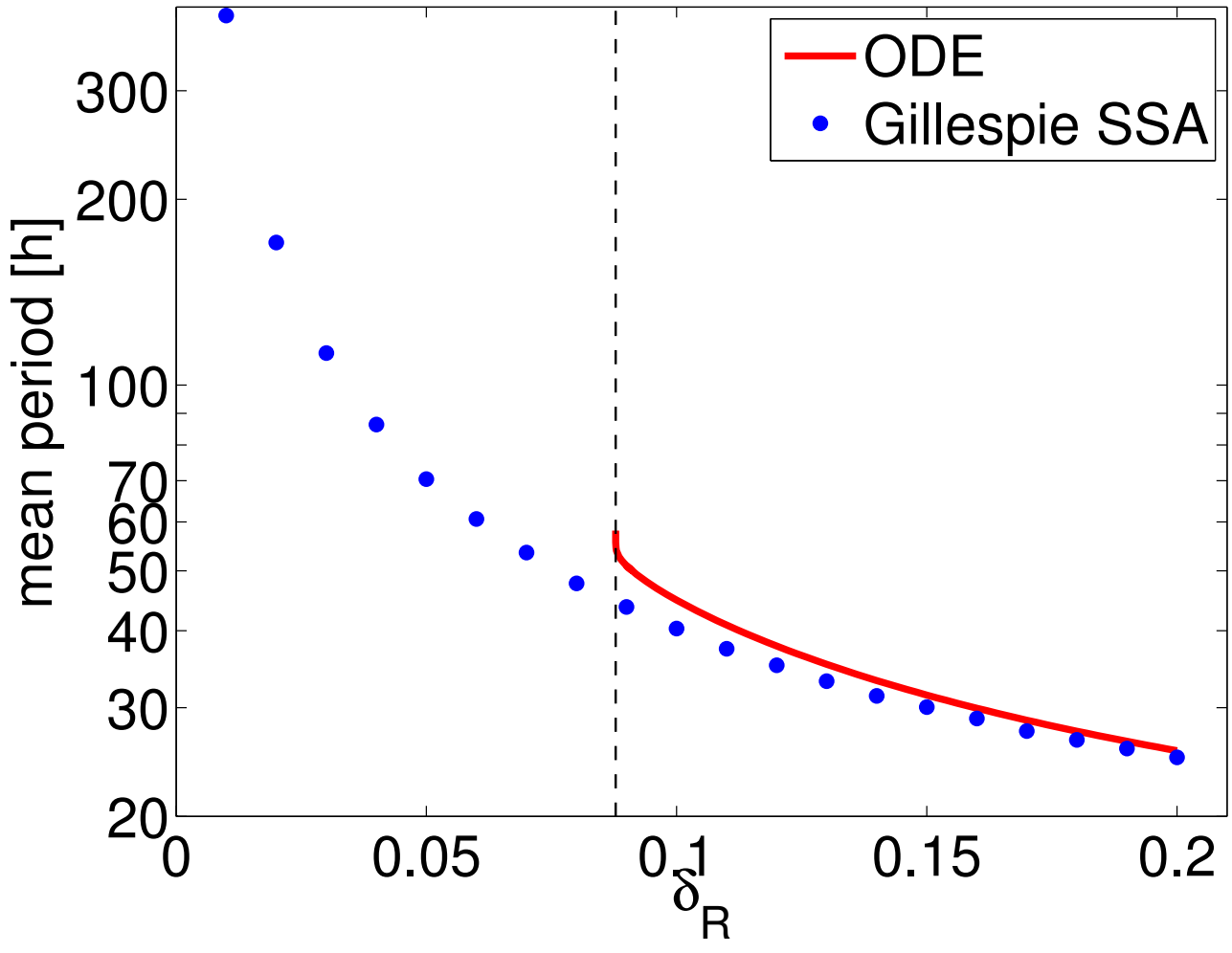


$\delta_R = 0.05$



Stochastic bifurcations

VKBL model: Mean period vs. δ_R

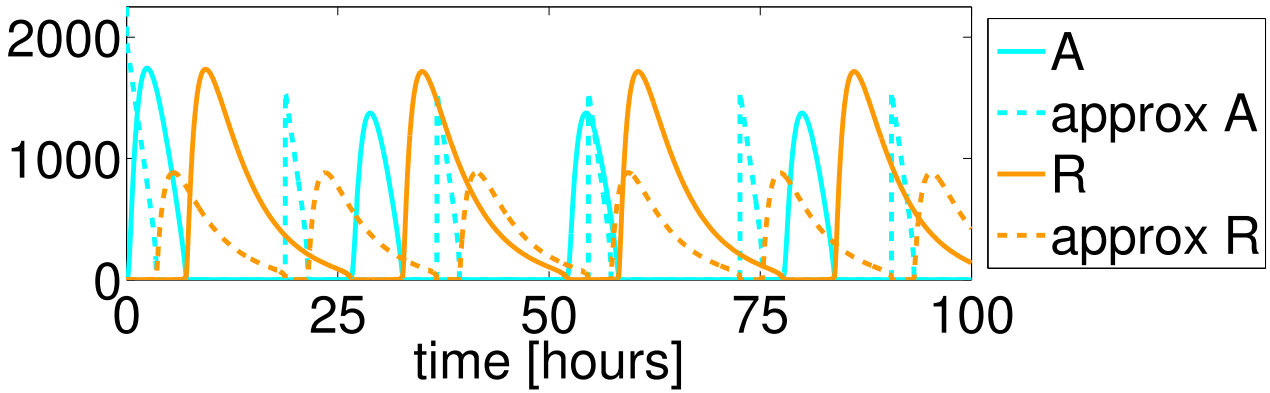


3. Modelling and numerics

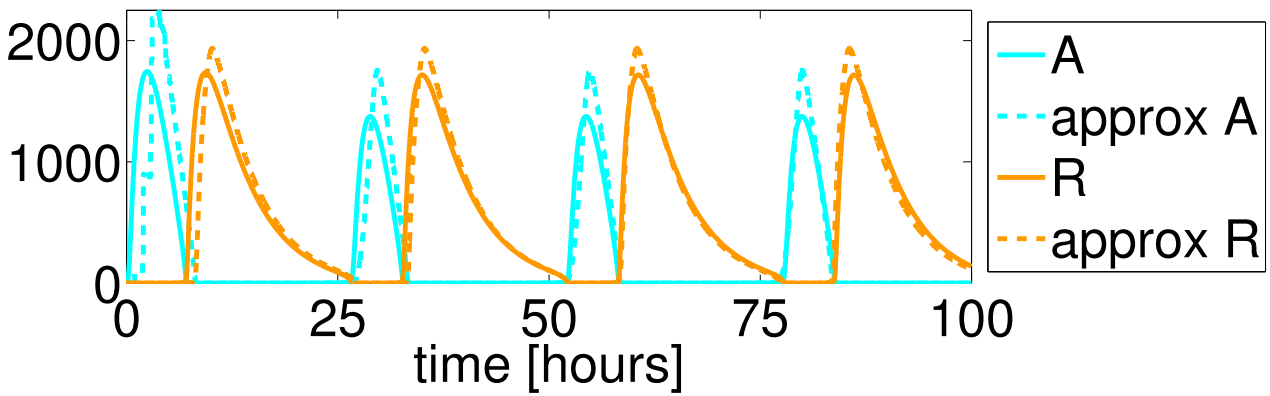
- Model reduction
- Stationary distribution by CME and FPE
- High-dimensional problems

Model reduction – VKBL model of circadian rhythms

- Original model (9 species, 16 reactions) $\Rightarrow$ 9 ODE
- Quasi-steady state assumption ($7\times$) $\Rightarrow$ 2 ODE



- Delayed quasi-steady state assumptions ($7\times$) $\Rightarrow$ 2 DDE



[T.V., Radek Erban, Philip Maini, 2014]

Stationary distribution

$$p_s(\mathbf{x}) = \lim_{t \to \infty} p(\mathbf{x}, t)$$

(i) Histogram:

- many realizations of stochastic simulation algorithm

Stationary distribution

$$p_s(\mathbf{x}) = \lim_{t\to\infty} p(\mathbf{x}, t)$$

(i) Histogram:

- many realizations of stochastic simulation algorithm

(ii) Stationary chemical master equation:

Example (Protein production):

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}p_{n,m} = {} & \alpha_R D p_{n-1,m} - \alpha_R D p_{n,m} + \delta_R(n+1)p_{n+1,m} - \delta_R n p_{n,m} \\ & + \alpha_P n p_{n,m-1} - \alpha_P n p_{n,m} + \delta_P(m+1)p_{n,m+1} - \delta_P m p_{n,m} \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad n, m = 0, 1, 2, \ldots\end{aligned}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{p} = A\mathbf{p}$$

$$\mathbf{0} = A\mathbf{p}$$

Stationary distribution

$$p_s(\mathbf{x}) = \lim_{t \to \infty} p(\mathbf{x}, t)$$

(i) Histogram:

- many realizations of stochastic simulation algorithm

(ii) Stationary chemical master equation: $A\mathbf{p} = \mathbf{0}$

- A is large, sparse
- $-A$ is M-matrix
- $\mathbf{1}^T A = \mathbf{0}$

Stationary distribution

$$p_s(\mathbf{x}) = \lim_{t \to \infty} p(\mathbf{x}, t)$$

(i) Histogram:

- many realizations of stochastic simulation algorithm

(ii) Stationary chemical master equation: $A\mathbf{p} = \mathbf{0}$

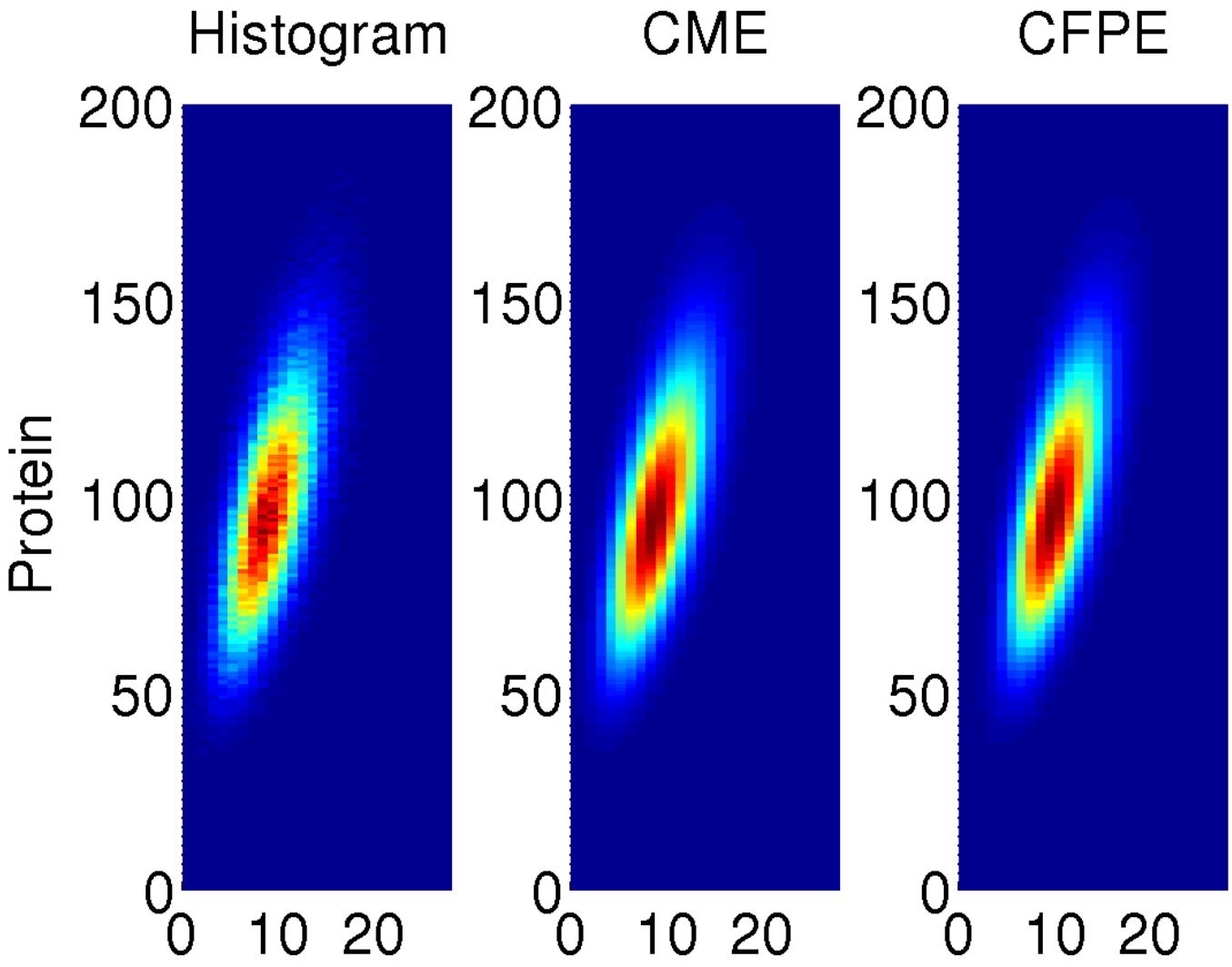
- A is large, sparse
- $-A$ is M-matrix
- $\mathbf{1}^T A = \mathbf{0}$

(iii) Stationary chemical Fokker-Planck equation: $0 = \mathrm{div}(\mathcal{A}\nabla p - \mathbf{b}p)$

- convection-diffusion equation
- no-flux boundary conditions
- might not be elliptic

Stationary distribution

Example 1: Protein production



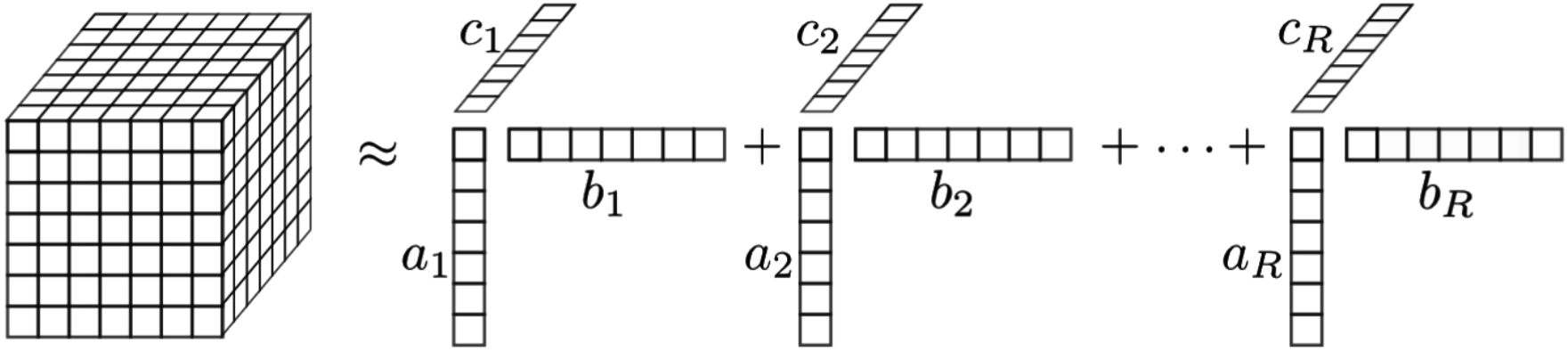
Stationary distribution

Example 2: VKBL model of circadian rhythms

- 9 chemical species $\Rightarrow$ 9 dimensions !?

Curse of dimensionality: $\mathcal{O}(N^d)$

Tensor methods

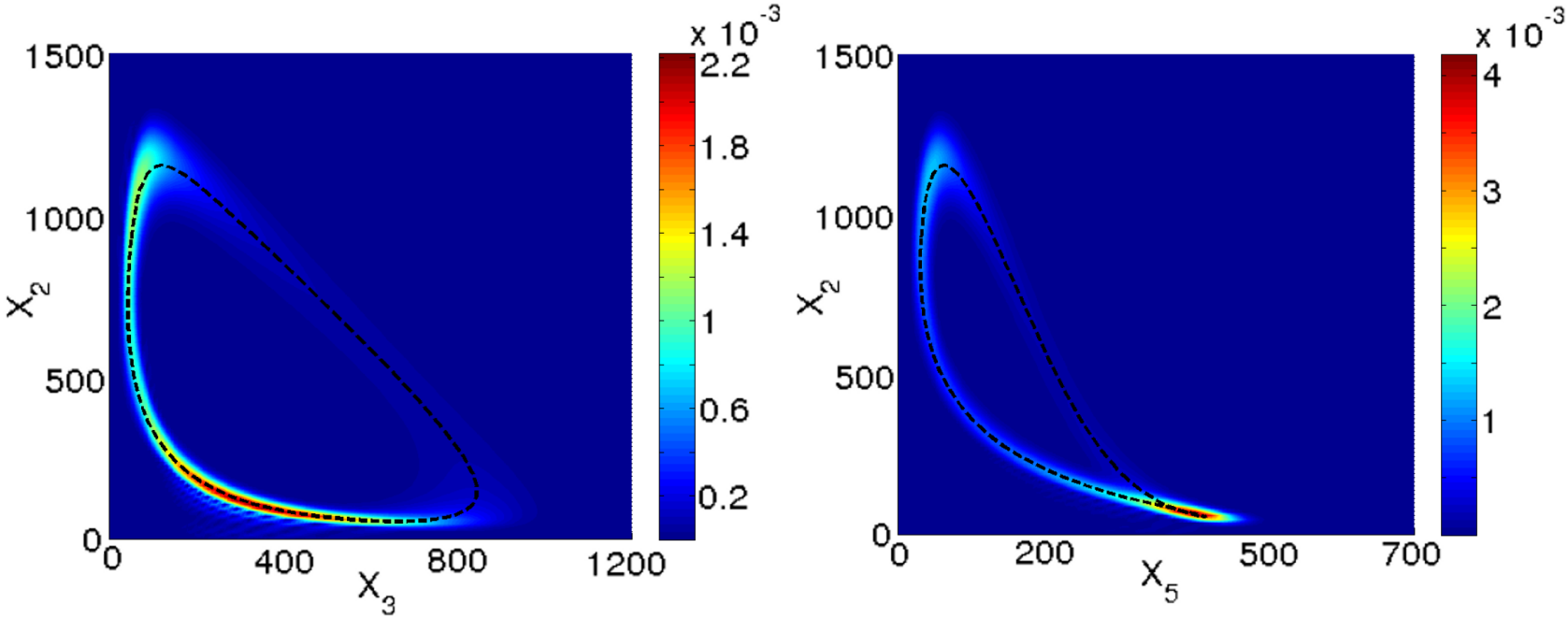


$$\mathcal{O}(N^d) \approx \mathcal{O}(RNd)$$

[Shuohao Liao, T.V., Radek Erban]

Example

- cdc2 and cyclin interactions [J. Tyson, 1991]
- 6-dimensional chemical Fokker-Planck equation



[Shuohao Liao]

Conclusions

- Mathematical models in biochemistry
- Deterministic – mass-action
- Stochastic – Markov process
- CME and CFPE
- Stochastic bifurcations
- Model reduction
- Tensor methods for higher-dimensional problems

Acknowledgement

My collaborators: Radek Erban, Philip K. Maini, and Shuohao Liao

Marie Curie Fellowship, StochDetBioModel

EUROPEAN
COMMISSION

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. 328008.

Thank you for your attention

Tomáš Vejchodský

vejchod@math.cas.cz

Wolfson Centre
for Mathematical Biology
Mathematical Institute

Institute of Mathematics
Academy of Sciences
Czech Republic

AIME@CZ, Prague, 11–12 March 2014