Challenges for numerical methods in biochemistry

Can we solve 20-dimensional PDE?

Tomáš Vejchodský

vejchod@math.cas.cz

Wolfson Centre
for Mathematical Biology
Mathematical Institute

Institute of Mathematics
Academy of Sciences
Czech Republic

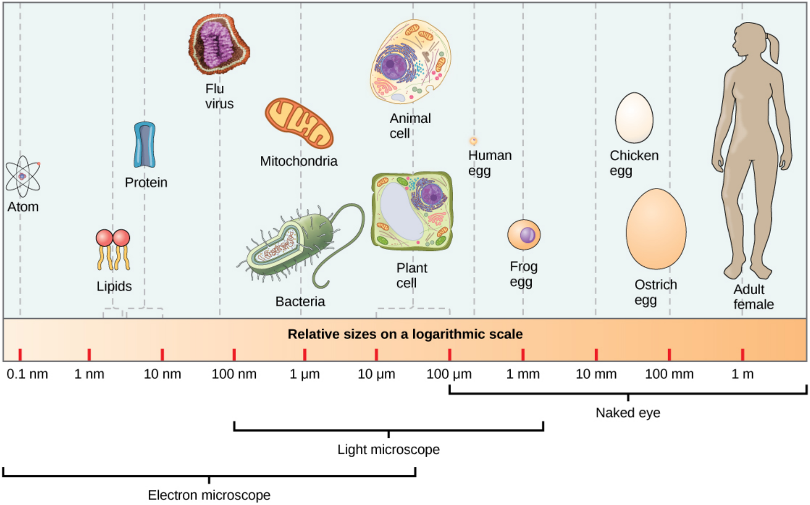
PANM 17, Dolní Maxov, 8–13 June 2014

How does the life function?

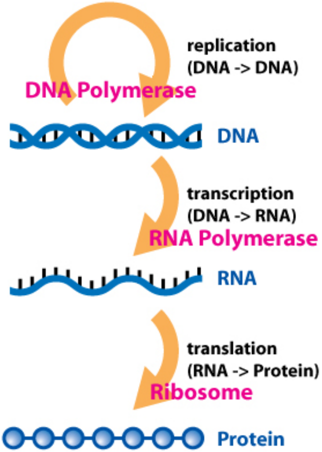
How does the life function?

How does the life function?

Levels and scales



Central dogma of molecular biology

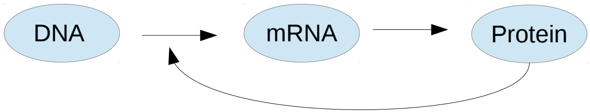


Gene regulatory networks

Protein production

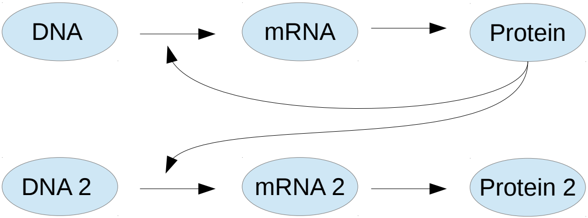
Gene regulatory networks

Feedback loops (transcription factors)



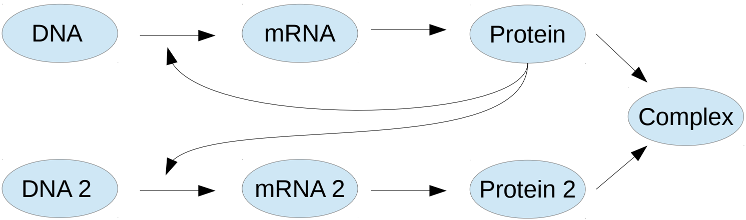
Gene regulatory networks

Feedback loops (transcription factors)

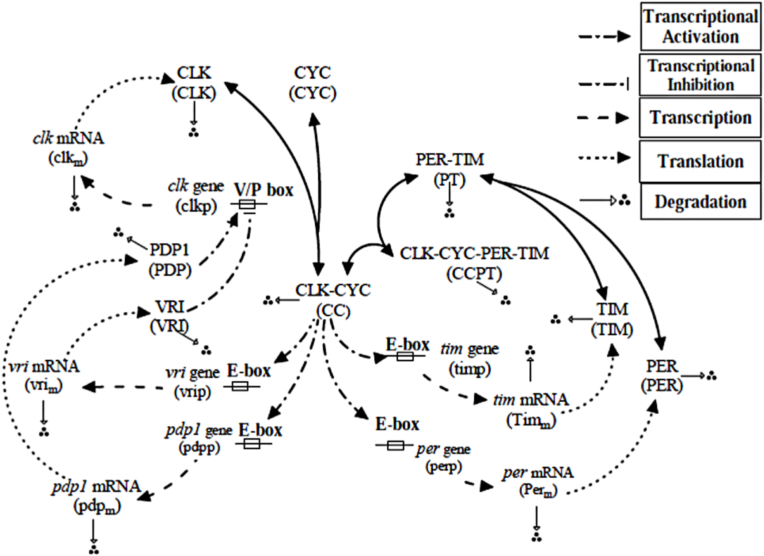


Gene regulatory networks

Feedback loops (transcription factors)

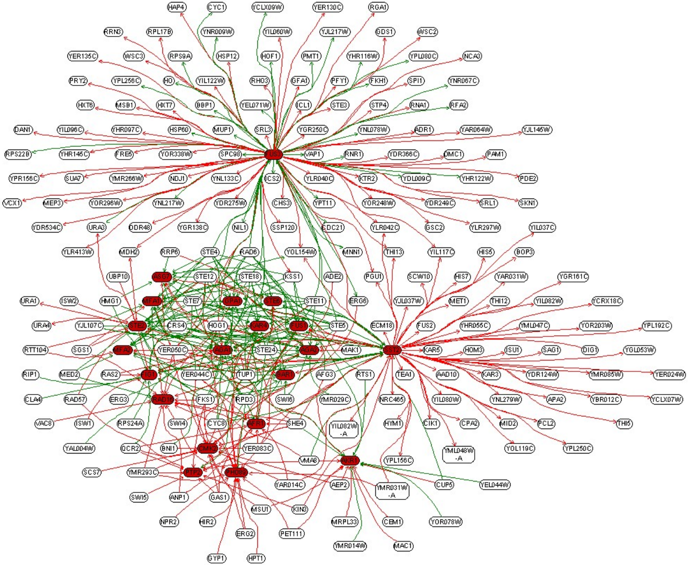


Gene regulatory networks



Circadian rhythms in Drosophila [Xie, Kulasiri, 2007]

Gene regulatory networks



Neighbourhood of mating response genes in yeast

[Rung, Schlitt, et al, 2002]

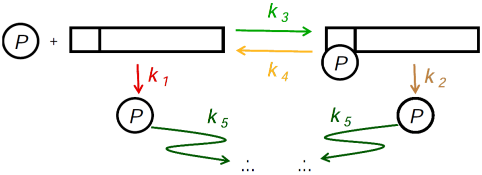
Outline

1. Chemical systems
    - Deterministic models
    - Stochastic models
2. Solution methods
    - Stochastic simulation algorithms
    - Chemical master equation
    - Chemical Fokker-Planck equation
3. Higher-dimensional problems
    - Tensor methods
    - Example
4. Conclusions

1. Chemical systems

Chemical systems

Example: Protein production



Chemical equations:

$$G \xrightarrow{k_1} G + P \qquad G + P \xrightarrow{k_3} G^* \qquad P \xrightarrow{k_5} \emptyset$$

$$G^* \xrightarrow{k_2} G^* + P \qquad G^* \xrightarrow{k_4} G + P$$

Deterministic model – law of mass action

Protein production system

$$G \xrightarrow{k_1} G + P \qquad G + P \xrightarrow{k_3} G^* \qquad P \xrightarrow{k_5} \emptyset$$

$$G^* \xrightarrow{k_2} G^* + P \qquad G^* \xrightarrow{k_4} G + P$$

Mass action ODE:

Conservation law: $G + G^* = H$

$$\frac{\mathrm{d}G}{\mathrm{d}t} = -\frac{k_3}{\Omega} GP + k_4(H - G)$$

$$\frac{\mathrm{d}P}{\mathrm{d}t} = k_1 G + k_2(H - G) - \frac{k_3}{\Omega} GP + k_4(H - G) - k_5 P$$

Initial condition:

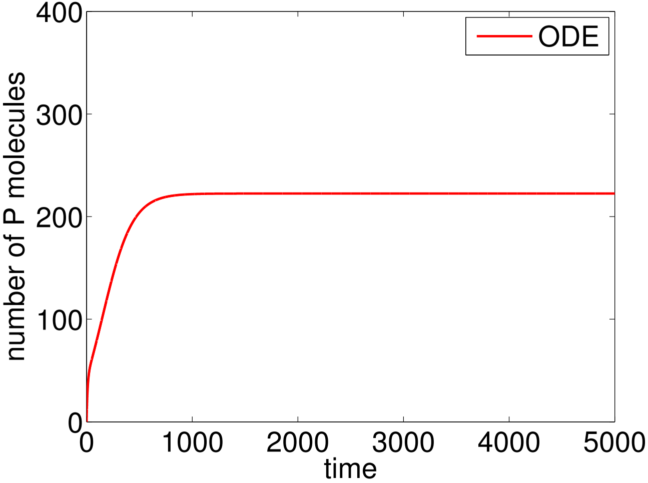
$G(0) = H,\ P(0) = 0$

Notation:

$G = G(t)$ ... number of DNA molecules

$P = P(t)$ ... number of protein molecules

Deterministic model – law of mass action



$$H = 1,\ k_1 = 5,\ k_2 = 30,\ k_3 = 3 \cdot 10^{-5}/H,\ k_4 = 0.003,\ k_5 = 0.1$$

Stochastic model

Protein production system:

$$G \xrightarrow{k_1} G + P \qquad G + P \xrightarrow{k_3} G^* \qquad P \xrightarrow{k_5} \emptyset$$

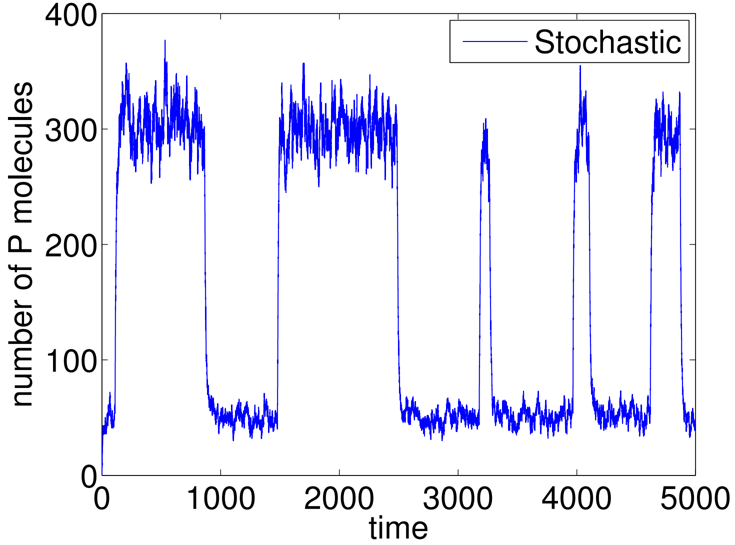
$$G^* \xrightarrow{k_2} G^* + P \qquad G^* \xrightarrow{k_4} G + P$$

Discrete state continuous time Markov process

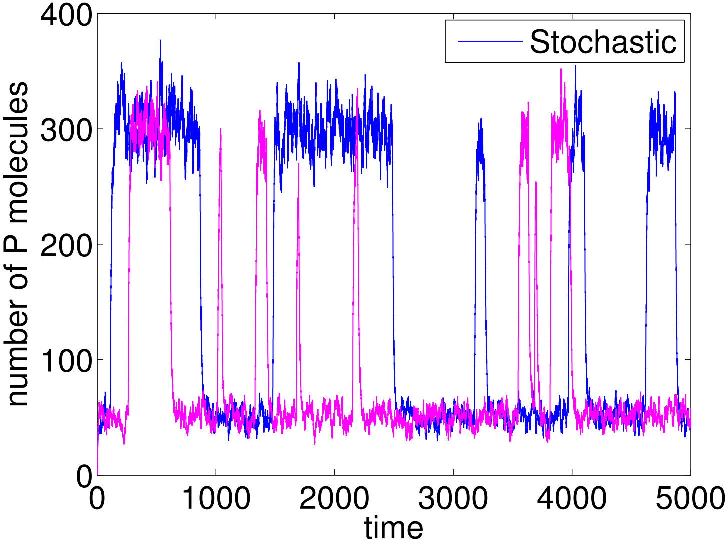
- State at time t: $[G, P]$
- Change of state:

reaction	state at $t + \mathrm{d}t$	with probability
1.	$[G, P+1]$	$k_1\, G\, \mathrm{d}t$
2.	$[G, P+1]$	$k_2 (H - G)\, \mathrm{d}t$
3.	$[G-1, P-1]$	$k_3\, GP\, \mathrm{d}t$
4.	$[G+1, P+1]$	$k_4 (H - G)\, \mathrm{d}t$
5.	$[G, P-1]$	$k_5\, P\, \mathrm{d}t$
–	$[G, P]$	otherwise

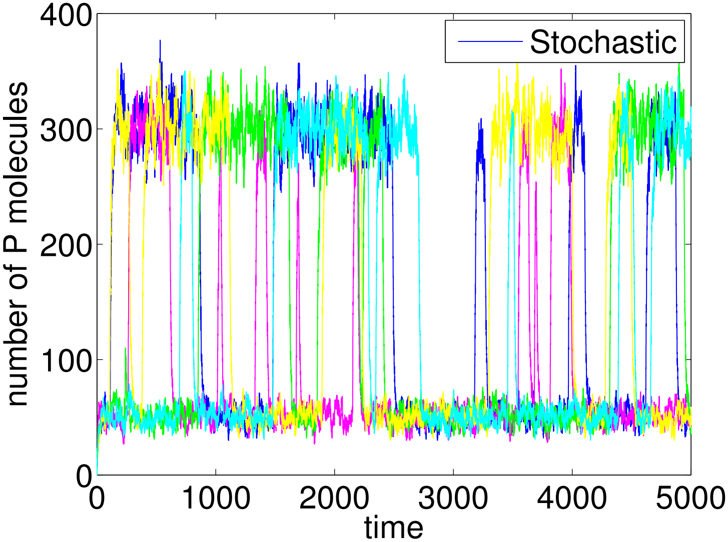
Stochastic model – Gillespie algorithm



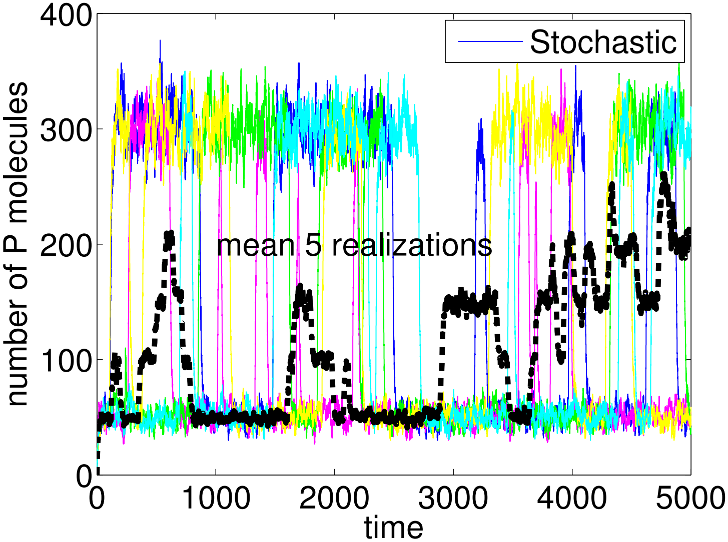
Stochastic model – Gillespie algorithm



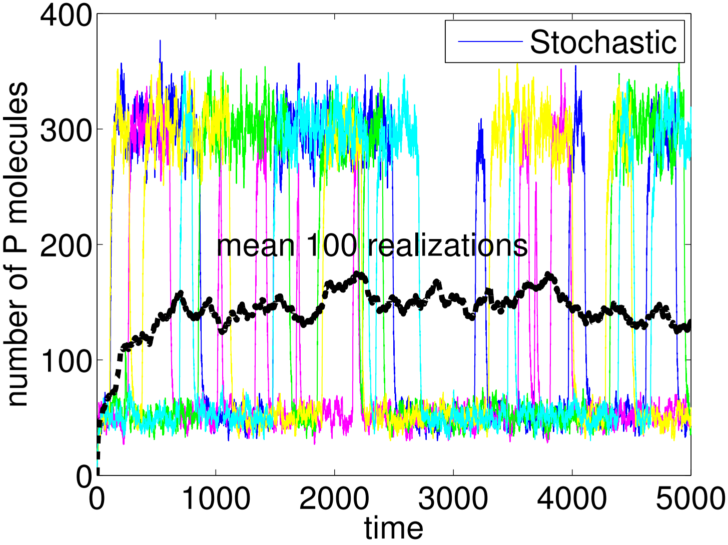
Stochastic model – Gillespie algorithm



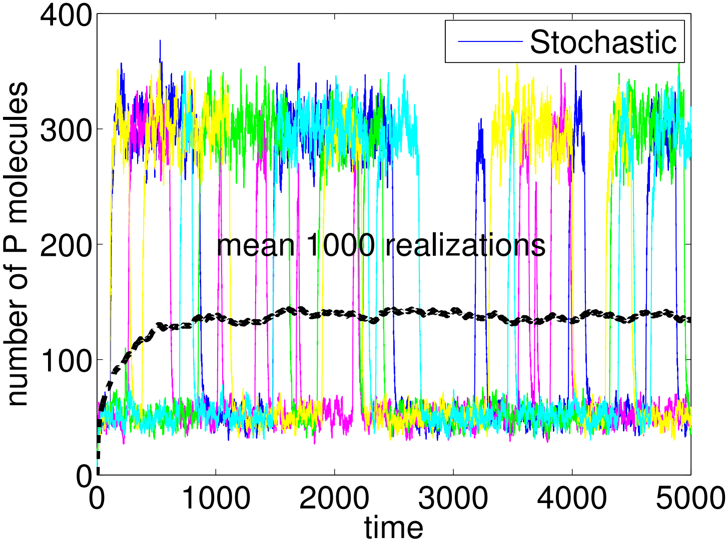
Stochastic model – Gillespie algorithm



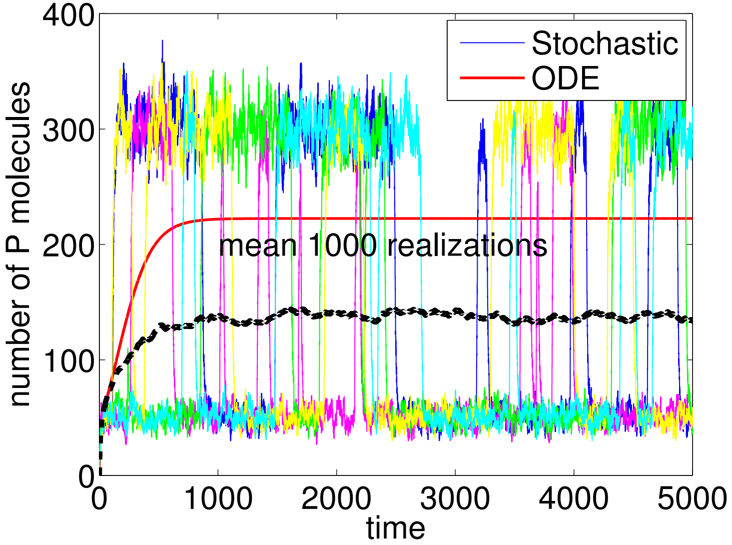
Stochastic model – Gillespie algorithm



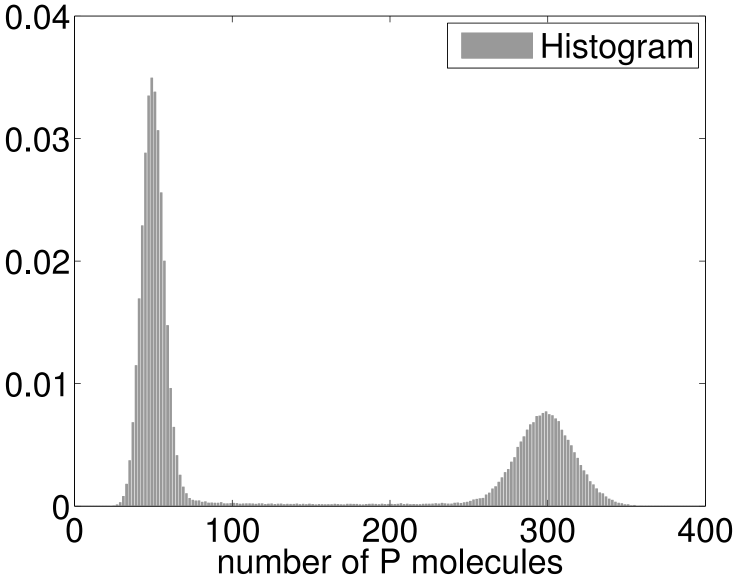
Stochastic model – Gillespie algorithm



Stochastic model – Gillespie algorithm



Stochastic model – histogram (200 000 realizations)



Stochastic model – analysis

Protein production system:

$$G \xrightarrow{k_1} G + P \qquad G + P \xrightarrow{k_3} G^* \qquad P \xrightarrow{k_5} \emptyset$$

$$G^* \xrightarrow{k_2} G^* + P \qquad G^* \xrightarrow{k_4} G + P$$

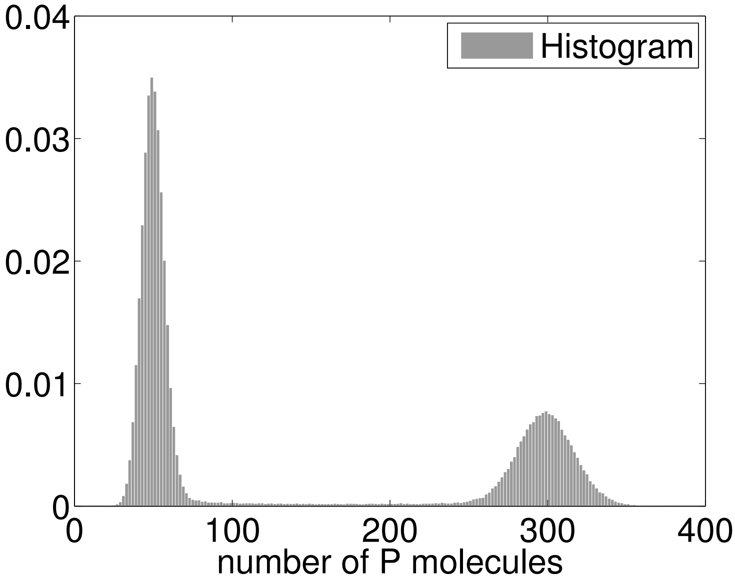
Definition: $p_{n,m}(t) = \Pr[G(t) = n, P(t) = m]$

Chemical master equation (CME):

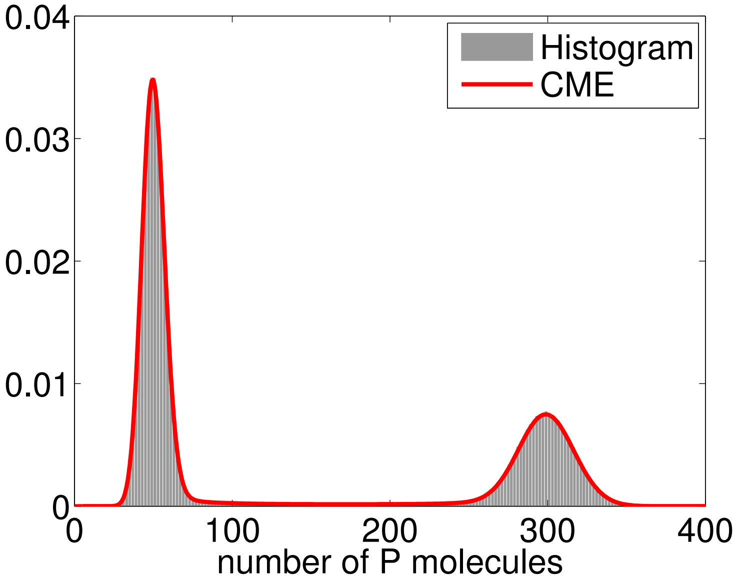
$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} p_{n,m} = {} & k_1 n p_{n,m-1} - k_1 n p_{n,m} \\ & + k_2 (H-n) p_{n,m-1} - k_2 (H-n) p_{n,m} \\ & + k_3 (n+1)(m+1) p_{n+1,m+1} - k_3 n m p_{n,m} \\ & + k_4 (H-n+1) p_{n-1,m-1} - k_4 (H-n) m p_{n,m} \\ & + k_5 (m+1) p_{n,m+1} - k_5 m p_{n,m}\end{aligned}$$

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

Stochastic model – histogram



Stochastic model – histogram



Chemical Fokker–Planck equation (CFPE)

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

Chemical Fokker-Planck equation (CFPE):

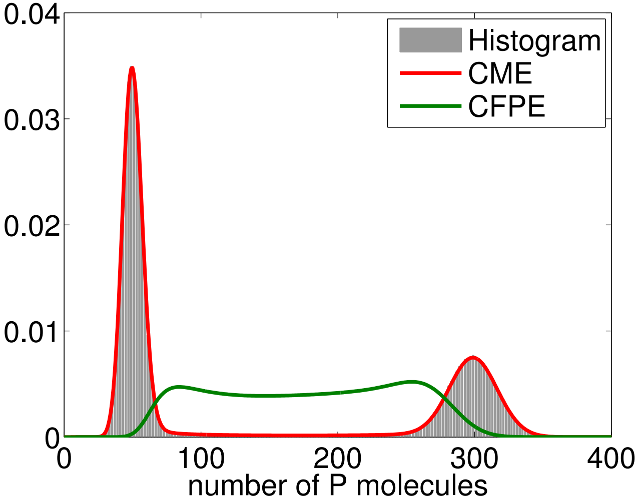
$$\frac{\partial p}{\partial t} = \mathrm{div}(\mathcal{A}\nabla p - \mathbf{b}p), \qquad (x, y) \in (0, H) \times (0, \infty)$$

where

$$\mathcal{A} = \frac{1}{2}\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{12} & \mathcal{A}_{22} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 - \partial\mathcal{A}_{11}/\partial x - \partial\mathcal{A}_{12}/\partial y \\ \tilde{b}_2 - \partial\mathcal{A}_{12}/\partial x - \partial\mathcal{A}_{22}/\partial y \end{bmatrix},$$

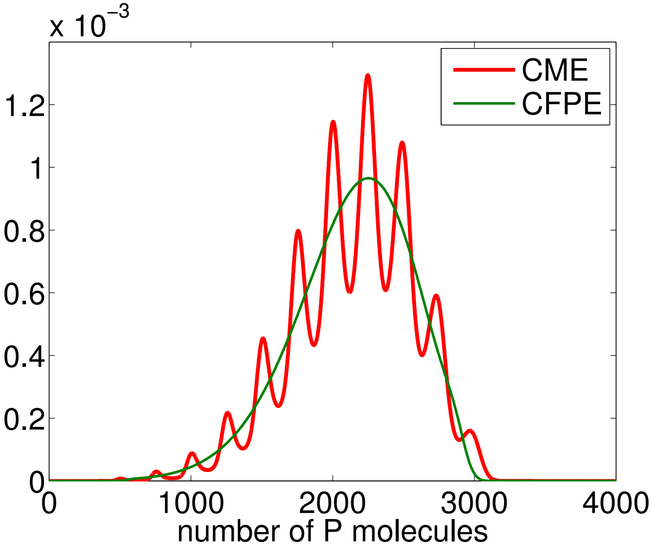
$$\tilde{b}_1 = -k_3xy + k_4(H - x)$$
$$\tilde{b}_2 = k_1x + k_2(H - x) - k_3xy + k_4(H - x) - k_5y$$
$$A_{11} = A_{12} = (k_3xy + k_4(H - x))/2$$
$$A_{22} = (k_1x + k_2(H - x) + k_3xy + k_4(H - x) + k_5y)/2$$

Chemical Fokker–Planck equation (CFPE)



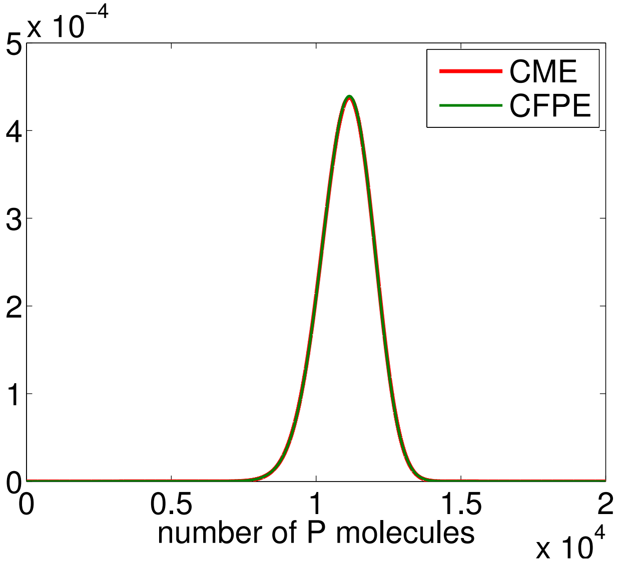
$$H = 1, \quad k_1 = 5,\ k_2 = 30,\ k_3 = 3 \cdot 10^{-5}/H,\ k_4 = 0.003,\ k_5 = 0.1$$

Chemical Fokker–Planck equation (CFPE)



$H = 10,\ \ k_1 = 5,\ k_2 = 30,\ k_3 = 3 \cdot 10^{-5}/H,\ k_4 = 0.003,\ k_5 = 0.1$

Chemical Fokker–Planck equation (CFPE)



$H = 50, \ k_1 = 5, \ k_2 = 30, \ k_3 = 3 \cdot 10^{-5}/H, \ k_4 = 0.003, \ k_5 = 0.1$

Chemical Langevin equation (CLE)

Stochastic differential equation:

$$\mathrm{d}\mathbf{X}(t) = \tilde{\mathbf{b}}(\mathbf{X}(t))\mathrm{d}t + C(\mathbf{X}(t))\mathrm{d}\mathbf{W}, \qquad \text{where } \mathcal{A} = C^T C$$

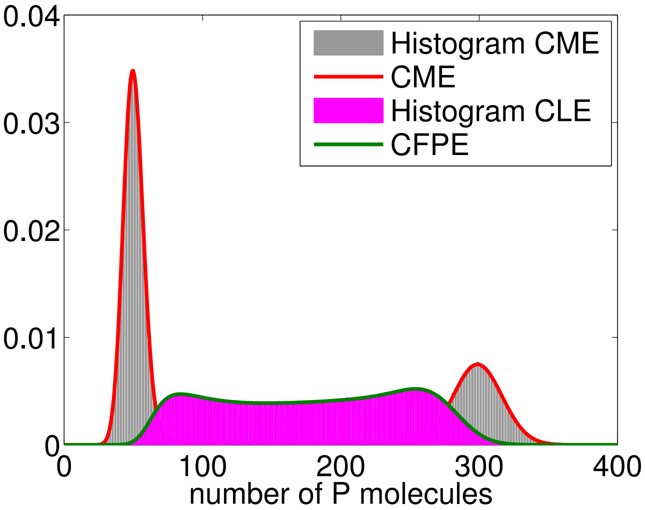
Euler–Maruyama method:

$$\mathbf{X}(t+\Delta t) = \mathbf{X}(t) + \tilde{\mathbf{b}}(\mathbf{X}(t))\Delta t + C(\mathbf{X}(t))\boldsymbol{\xi}\sqrt{\Delta t}$$

Notation:

$$\mathbf{X} = \begin{bmatrix} X \\ Y \end{bmatrix}, \quad \tilde{\mathbf{b}} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{bmatrix}, \quad \boldsymbol{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}, \quad \xi_i \sim \mathcal{N}(0,1), \; i = 1, 2$$

Chemical Langevin equation (CLE)



$$H = 1, \quad k_1 = 5,\ k_2 = 30,\ k_3 = 3 \cdot 10^{-5}/H,\ k_4 = 0.003,\ k_5 = 0.1$$

Models of chemical dynamics – summary

Deterministic:

- Reaction ODEs

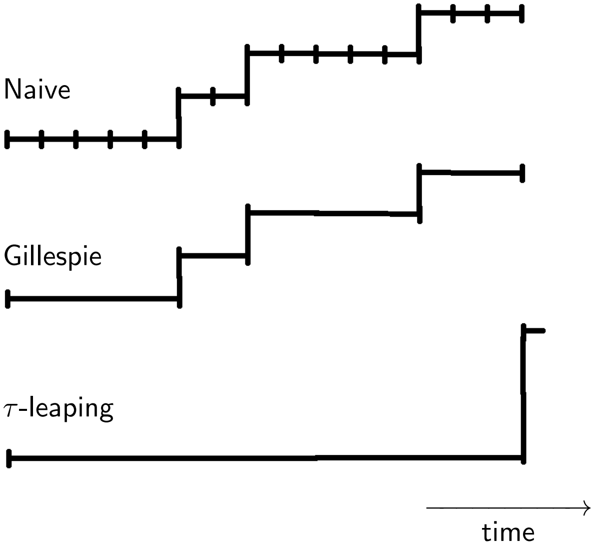
Stochastic:

- Exact description:
  - Discrete state continuous time Markov process
  - Chemical master equation
- Approximation:
  - Continuous state continuous time Markov process (Chemical Langevin equation)
  - Chemical Fokker-Planck equation

2. Solution methods
   - Stochastic simulation algorithms
   - Chemical master equation
   - Chemical Fokker–Planck equation

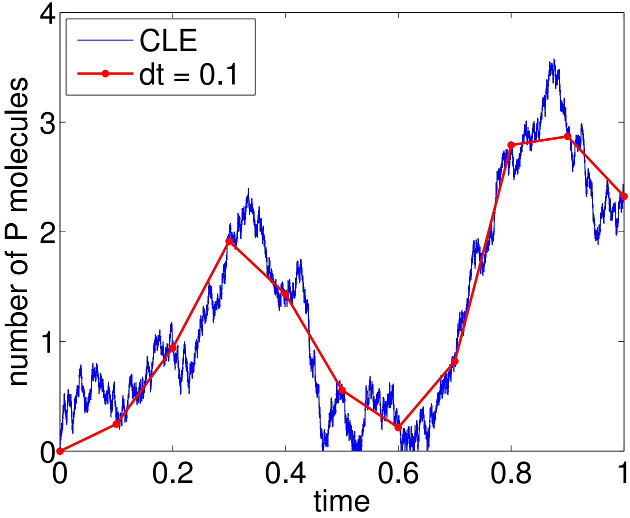
Stochastic simulation algorithms

Chemical master equation



Stochastic simulation algorithms

Chemical Langevin equation: Euler–Maruyama method



Chemical master equation

Example (Protein production):

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}p_{n,m} = {} & k_1 n p_{n,m-1} - k_1 n p_{n,m} \\ & + k_2(H-n)p_{n,m-1} - k_2(H-n)p_{n,m} \\ & + k_3(n+1)(m+1)p_{n+1,m+1} - k_3 nmp_{n,m} \\ & + k_4(H-n+1)p_{n-1,m-1} - k_4(H-n)mp_{n,m} \\ & + k_5(m+1)p_{n,m+1} - k_5 mp_{n,m}\end{aligned}$$

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

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

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

Properties:

- A is sparse, infinite $\Rightarrow$ truncation (FSP method)
- $-A$ is M-matrix
- $\mathbf{1}^T A = \mathbf{0}$

Chemical Fokker–Planck equation

Evolutionary:

$$\begin{aligned} \frac{\partial p}{\partial t} &= \operatorname{div}(\mathcal{A}\nabla p - \mathbf{b}p), && \text{in } \Omega \subset (0,H)\times(0,\infty) \\ p(x,y,0) &= p_0(x,y) && \text{at } t=0 \\ (\mathcal{A}\nabla p - \mathbf{b}p)\cdot n &= 0 && \text{on } \partial\Omega \end{aligned}$$

Conservative:

$$\frac{\partial}{\partial t}\int_\Omega p\,\mathrm{d}x = \int_\Omega \frac{\partial p}{\partial t}\,\mathrm{d}x = \int_\Omega \operatorname{div}(\mathcal{A}\nabla p - \mathbf{b}p)\,\mathrm{d}x = \int_{\partial\Omega}(\mathcal{A}\nabla p - \mathbf{b}p)\cdot n\,\mathrm{d}x = 0$$

$$\Rightarrow \quad \int_\Omega p\,\mathrm{d}x = \int_\Omega p_0\,\mathrm{d}x = 1$$

Chemical Fokker–Planck equation

Stationary:

$$\begin{aligned} \operatorname{div}(\mathcal{A}\nabla p - \mathbf{b}p) &= 0 \quad \text{in } \Omega \subset (0,H)\times(0,\infty) \\ (\mathcal{A}\nabla p - \mathbf{b}p)\cdot n &= 0 \quad \text{on } \partial\Omega, \\ \int_\Omega p \,\mathrm{d}x &= 1 \end{aligned}$$

Adjoint:

$$\begin{aligned} \operatorname{div}\mathcal{A}\nabla z + \mathbf{b}\cdot\nabla z &= 0 \quad \text{in } \Omega \\ (\mathcal{A}\nabla z)\cdot n &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

Solution: $z = 1$

Not elliptic:

- $\mathcal{A}(x,y)$ not positive definite for some (x,y)

3. Higher-dimensional problems
    - Tensor methods
    - Example

General chemical system

$$\sum_{i=1}^{N} \nu_{ji}^{\mathrm{r}} X_i \xrightarrow{k_j} \sum_{i=1}^{N} \nu_{ji}^{\mathrm{p}} X_i, \qquad j = 1, 2, \ldots, M$$

Notation:

- Well mixed reactor: N chemical species, M reactions
- $\mathbf{X} = [X_1, \ldots, X_N]$, $X_i(t) =$ number of molecules, $i = 1, \ldots, N$
- $\alpha_j(\mathbf{x}) = k_j \prod_{i=1}^{N} \binom{x_i}{\nu_{ji}^{\mathrm{r}}}$, $j = 1, \ldots, M$, are propensities
- $\nu_{ji} = \nu_{ji}^{\mathrm{p}} - \nu_{ji}^{\mathrm{r}}$, change of X_i during reaction R_j,
- $\boldsymbol{\nu}_j = [\nu_{j1}, \ldots, \nu_{jN}]$

Higher-dimensional problems

Chemical master equation

$$\frac{\partial}{\partial t} p(\mathbf{x}, t) = \sum_{j=1}^{M} \left[\alpha_j(\mathbf{x} - \boldsymbol{\nu}_j) p(\mathbf{x} - \boldsymbol{\nu}_j, t) - \alpha_j(\mathbf{x}) p(\mathbf{x}, t) \right], \quad \mathbf{x} \in \mathbb{N}_0^N$$

Chemical Fokker–Planck equation

$$\frac{\partial p}{\partial t}(\mathbf{x}, t) = \operatorname{div}\left[\mathcal{A}(\mathbf{x}) \nabla p(\mathbf{x}, t) - \mathbf{b}(\mathbf{x}) p(\mathbf{x}, t) \right], \quad \mathbf{x} \in \Omega \subset [0, \infty)^N$$

where

$$\mathcal{A} = \frac{1}{2} \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{12} & \mathcal{A}_{22} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 - \partial \mathcal{A}_{11}/\partial x - \partial \mathcal{A}_{12}/\partial y \\ \tilde{b}_2 - \partial \mathcal{A}_{12}/\partial x - \partial \mathcal{A}_{22}/\partial y \end{bmatrix},$$

$$\tilde{b}_i(\mathbf{x}) = \sum_{j=1}^{M} \nu_{ji} \alpha_j(\mathbf{x}), \quad \mathcal{A}_{ik}(\mathbf{x}) = \sum_{j=1}^{M} \nu_{ji} \nu_{jk} \alpha_j(\mathbf{x})$$

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

Tensor methods

Let $\Omega = [0, L_1] \times \cdots \times [0, L_N]$

N-dimensional grid: $\mathbf{x}_{i_1,\ldots,i_N} = (x_{i_1}, \ldots, x_{i_N})$

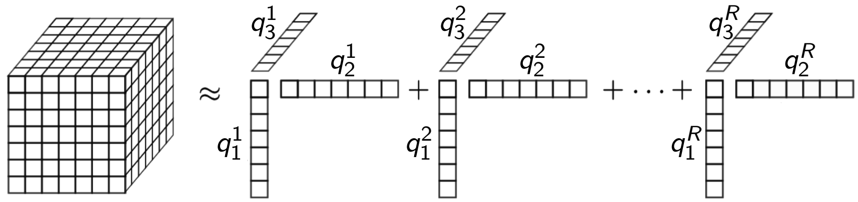
- $x_{i_d} = (i_d - 1)h_d \quad \Rightarrow n_d$ nodes in each dimension
- $i_d = 1, 2, \ldots, n_d, \; d = 1, 2, \ldots, N$
- $h_d = L_d/(n_d - 1), \; d = 1, 2, \ldots, N$

Tensor: $p(\mathbf{x}_{i_1,\ldots,i_N}) = \mathbf{p}_{i_1,\ldots,i_N} \in \mathbb{R}^{n_1 \times \cdots \times n_N}$

Canonical decomposition:

$$\underbrace{\mathbf{p}_{i_1,\ldots,i_N}}_{n^N \text{ entries}} \approx \underbrace{\sum_{r=1}^{R} q^r_{1,i_1} q^r_{2,i_2} \ldots q^r_{N,i_N}}_{nNR \text{ entries}}, \quad \text{where } q^r_d \in \mathbb{R}^{n_d}, d = 1, 2, \ldots, N$$

Tensor methods



Canonical decomposition:

$$\underbrace{\mathbf{p}_{i_1,\ldots,i_N}}_{n^N \text{ entries}} \approx \underbrace{\sum_{r=1}^{R} q^r_{1,i_1} q^r_{2,i_2} \ldots q^r_{N,i_N}}_{nNR \text{ entries}}, \quad \text{where } q^r_d \in \mathbb{R}^{n_d}, d = 1, 2, \ldots, N$$

Arithmetic operations

Let

- $\mathbf{p}_{i_1,\dots,i_N} = \sum_{r=1}^{R} q^r_{1,i_1} q^r_{2,i_2} \dots q^r_{N,i_N}$
- $\mathbf{u}_{i_1,\dots,i_N} = \sum_{s=1}^{S} v^s_{1,i_1} v^s_{2,i_2} \dots v^s_{N,i_N}$

Addition: $\mathcal{O}(1)$

$$\mathbf{p}_{i_1,\dots,i_N} + \mathbf{u}_{i_1,\dots,i_N} = \sum_{r=1}^{R} q^r_{1,i_1} q^r_{2,i_2} \dots q^r_{N,i_N} + \sum_{s=1}^{S} v^s_{1,i_1} v^s_{2,i_2} \dots v^s_{N,i_N}$$

$\Rightarrow$ Rank: $R + S$

Arithmetic operations

Let

- $\mathbf{p}_{i_1,\dots,i_N} = \sum_{r=1}^{R} q^r_{1,i_1} q^r_{2,i_2} \dots q^r_{N,i_N}$
- $\mathbf{u}_{i_1,\dots,i_N} = \sum_{s=1}^{S} v^s_{1,i_1} v^s_{2,i_2} \dots v^s_{N,i_N}$

Multiplication by scalar: $\mathcal{O}(n)$

$$\alpha \mathbf{p}_{i_1,\dots,i_N} = \sum_{r=1}^{R} (\alpha q^r_{1,i_1}) q^r_{2,i_2} \dots q^r_{N,i_N}$$

Arithmetic operations

Let

- $\mathbf{p}_{i_1,\ldots,i_N} = \sum_{r=1}^{R} q^r_{1,i_1} q^r_{2,i_2} \ldots q^r_{N,i_N}$
- $\mathbf{u}_{i_1,\ldots,i_N} = \sum_{s=1}^{S} v^s_{1,i_1} v^s_{2,i_2} \ldots v^s_{N,i_N}$

Scalar product: $\mathcal{O}(nNRS)$

$$\mathbf{p} \cdot \mathbf{u} = \sum_{r=1}^{R} \sum_{s=1}^{S} (q^r_1 \cdot v^s_1) \cdots (q^r_N \cdot v^s_N)$$

Arithmetic operations

Let

- $\mathbf{p}_{i_1,\dots,i_N} = \sum_{r=1}^{R} q^r_{1,i_1} q^r_{2,i_2} \dots q^r_{N,i_N}$
- $\mathbf{u}_{i_1,\dots,i_N} = \sum_{s=1}^{S} v^s_{1,i_1} v^s_{2,i_2} \dots v^s_{N,i_N}$

Derivative: $\mathcal{O}(n)$

$$\frac{\partial}{\partial x_1} p(\mathbf{x}_{i_1,\dots,i_N}) \approx \sum_{r=1}^{R} \frac{q^r_{1,i_1+1} - q^r_{1,i_1-1}}{2h_1} q^r_{2,i_2} \dots q^r_{N,i_N}$$

Arithmetic operations

Let

- $\mathbf{p}_{i_1,\ldots,i_N} = \sum_{r=1}^{R} q^r_{1,i_1} q^r_{2,i_2} \ldots q^r_{N,i_N}$
- $\mathbf{u}_{i_1,\ldots,i_N} = \sum_{s=1}^{S} v^s_{1,i_1} v^s_{2,i_2} \ldots v^s_{N,i_N}$

Tensor truncation:

$$\sum_{r=1}^{R} q^r_{1,i_1} q^r_{2,i_2} \ldots q^r_{N,i_N} \approx \sum_{s=1}^{S} v^s_{1,i_1} v^s_{2,i_2} \ldots v^s_{N,i_N} \quad \text{with } S < R$$

$\Rightarrow$ Numerical stability $\Rightarrow$ Tensor train format

Example: 20-dimensional Laplacian

$$\begin{aligned} -\Delta u(x_1, \ldots, x_{20}) &= 1 \quad x_i \in (0,1) \\ u(x_1, \ldots, x_{20}) &= 0 \quad \text{if } \exists i :: x_i = 0 \\ \frac{\partial}{\partial x_i} u(x_1, \ldots, x_{20}) &= 0 \quad \text{if } x_i = 1 \end{aligned}$$

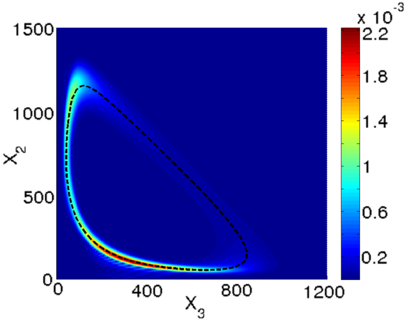
Plot:

$$\tilde{u}(x_{19}, x_{20}) = \int_0^1 \cdots \int_0^1 u(x_1, \ldots, x_{20}) \, \mathrm{d}x_1 \ldots \mathrm{d}x_{18}$$

[Shuohao Liao]

Example

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



[Shuohao Liao]

Conclusions

- Mathematical models of (bio)chemical systems
- Deterministic – mass-action
- Stochastic – Markov process, stochastic differential equation
- CME and CFPE
- Aspects of numerical solution
- Tensor methods for higher-dimensional problems

Acknowledgement

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

PANM 17, Dolní Maxov, 8–13 June 2014