A note on inference for reaction kinetics

with monomolecular reactions

Manfred Opper and Andreas Ruttor

TU Berlin, Computer Science

Stochastic Reaction Systems

- Discrete states $\mathbf{n} = (n_1, \ldots, n_d)$, n_i number of molecules of type i

- Markov (jump) dynamics

$$P\left(\mathbf{n}', t + \Delta t | \mathbf{n}, t\right) \simeq \delta_{\mathbf{n}',\mathbf{n}} + \Delta t \; f_\theta\left(\mathbf{n}' | \mathbf{n}\right)$$

  for $\Delta t \to 0$ with rate function $f_\theta(\mathbf{n}'|\mathbf{n})$.

- Master equation

$$\frac{dP\left(\mathbf{n}, t\right)}{dt} = \sum_{\mathbf{n}' \neq \mathbf{n}} \left[P\left(\mathbf{n}', t\right) f_\theta\left(\mathbf{n}|\mathbf{n}'\right) - P\left(\mathbf{n}, t\right) f_\theta\left(\mathbf{n}'|\mathbf{n}\right)\right].$$

- Rates for mass action kinetics

$$f_\theta(\mathbf{n}|\mathbf{n}') = \sum_r \theta_r \prod_j \frac{n_j!}{(n_j - p_{rj})!} \, \delta_{n'_j, n_j + q_{rj} - p_{rj}}$$

Inference Problem

- Given **noisy observations** $D \equiv \mathbf{y}_1, \ldots, \mathbf{y}_K$ of **hidden process** $\mathbf{n}(t_i)$ at times t_i for $i = 1, \ldots, K$

  Estimate **system parameters** θ contained in rates $f_\theta(\mathbf{n}'|\mathbf{n})$

- Methods based on exact computations of the Likelihood

$$p(D|\theta) = E\left\{ \prod_{k=1}^{K} p(\mathbf{y}_k|\mathbf{n}(t_k)) \right\}$$

  are usually intractable !

Some popular inference methods

- MC sampling of posterior process
- MC sampling within diffusion approximation
- Weak (linear) noise approximations
- Variational Mean field approximations

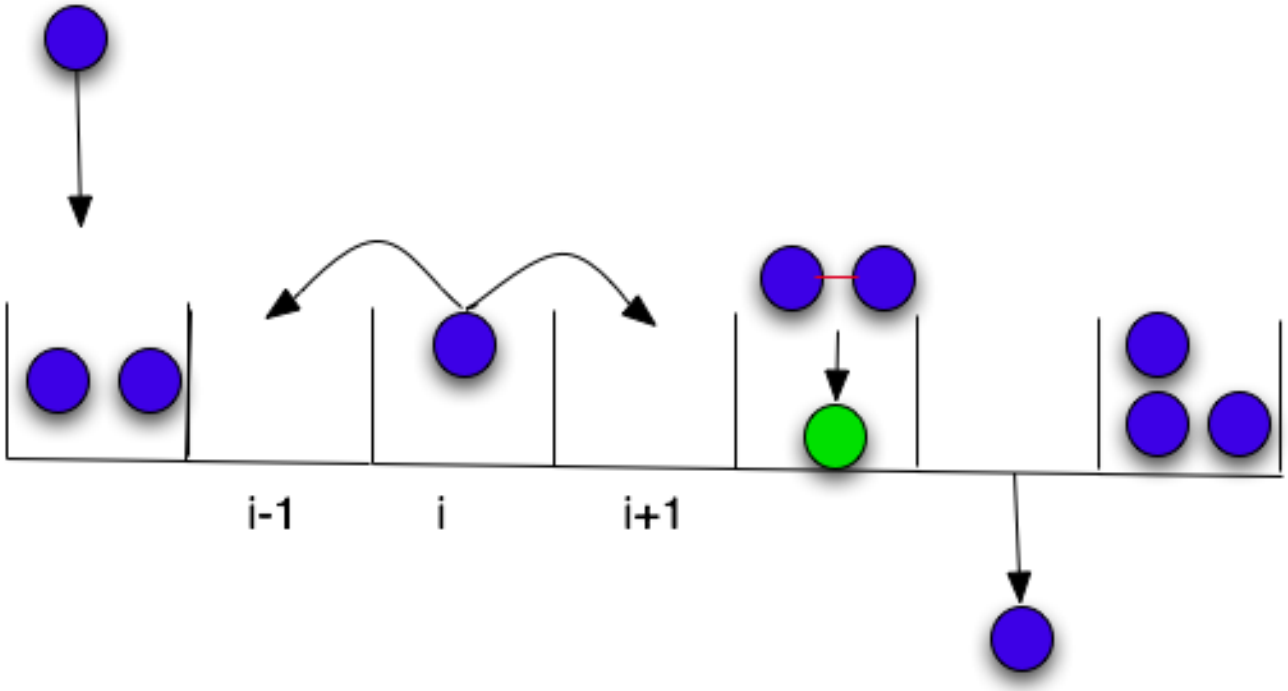
Monomolecular reaction systems

- Allowed reactions and their **rates**

$$
\begin{array}{rlll}
C_{jk} n_j & \text{with} & n'_k = n_k + 1 & \\
 & \text{and} & n'_j = n_j - 1 & \text{Exchange} \\
C_{0j} & \text{with} & n'_j = n_j + 1 & \text{Creation} \\
C_{j0} n_j & \text{with} & n'_j = n_j - 1 & \text{Annihilation}
\end{array}
$$

No real chemical reactions !

- Discussed as a model for the dynamics of the density of *Bicoid* (morphogen) protein in *Drosophila* in: YF Wu, E Myasnikova, J Reinitz *BMC Systems Biology, 2010, 1 (52)*. Parameter inference using a variational mean field approach by Dewar, Kadirkamanathan, Opper & Sanguinetti, *BMC Systems biology, 2010, 4:21*.



Some known results

- For monomolecular reactions $\mathbf{m}(t) = E[\mathbf{n}(t)]$ obeys the 'classical' linear (!) rate equations

$$\frac{dm_i}{dt} = C_{0i} + \sum_{j=1}^{M} C_{ji} m_j - \sum_{j=0}^{M} C_{ij} m_i$$

- Transition probability $P(\mathbf{n}'; t' | \mathbf{n}; t)$ is a d - fold convolution of multivariate Poisson and multinomial probabilities. (T. Jahnke, W. Huisinga, *J. Math. Biol. 54: 1-26, 2007*)

The likelihood

- We would like to compute (Gaussian noise on the data)

$$p(D|\theta) = \frac{1}{(\sqrt{2\pi\sigma^2})^{Kd}} E\left\{\exp\left[-\frac{1}{2\sigma^2}\sum_k ||\mathbf{y}_k - \mathbf{n}(t_k)||^2\right]\right\}$$

- But we can only compute something simpler

$$E\left[\exp\left\{\sum_{k=1}^{K} \boldsymbol{\phi}_k^\top \mathbf{n}(t_k)\right\}\right]$$

Solution

- Define

$$\Psi_t(\mathbf{n}) \doteq E\left[\exp\left\{\sum_{k:t_k>t} \boldsymbol{\phi}_k^\top \mathbf{n}(t_k)\right\} | \mathbf{n}_t = \mathbf{n}\right]$$

- Between 'observations' $\boldsymbol{\phi}_k$, the function $\Psi_t(\mathbf{n})$ fulfils the backward equation

$$\frac{d}{dt}\Psi_t(\mathbf{n}) = \sum_{\mathbf{n}' \neq \mathbf{n}} f(\mathbf{n}'|\mathbf{n}) \left[\Psi_t(\mathbf{n}) - \Psi_t(\mathbf{n}')\right]$$

- End condition: $\psi_T(\mathbf{n}) = 1$ & jump conditions

$$\Psi_{t_k^-}(\mathbf{n}) = \Psi_{t_k^+}(\mathbf{n})\ e^{\boldsymbol{\phi}_k^\top \mathbf{n}(t_k)}$$

Solution cont'd

- Solution is of the form

$$\Psi_t(\mathbf{n}) = \exp\left[a(t) + \mathbf{b}(t)^\top \mathbf{n}\right]$$

- The functions $a(t)$ and $\mathbf{r}(t) = \left(e^{b_1(t)}, \ldots, e^{b_d(t)}\right)$ obey a set of linear ODEs which can be solved analytically to give

$$\begin{aligned} \mathbf{r}(t) &= \ln\left\{1 + e^{(t_k - t)\mathbf{C}}(\mathbf{r}(t_k) - 1)\right\} \qquad t_{k-1} < t < t_k \\ \mathbf{r}(t_k^-) &= \mathbf{r}(t_k^+) e^{\boldsymbol{\phi}(t_k)} \qquad\qquad k = 1, \ldots, K \end{aligned}$$

- After solving for $\mathbf{b}(t)$, one can obtain

$$a(t) = \sum_j C_{0j} \int_t^T (r_j(t) - 1) dt$$

A machine learning 'trick'

Use the convex duality transformation

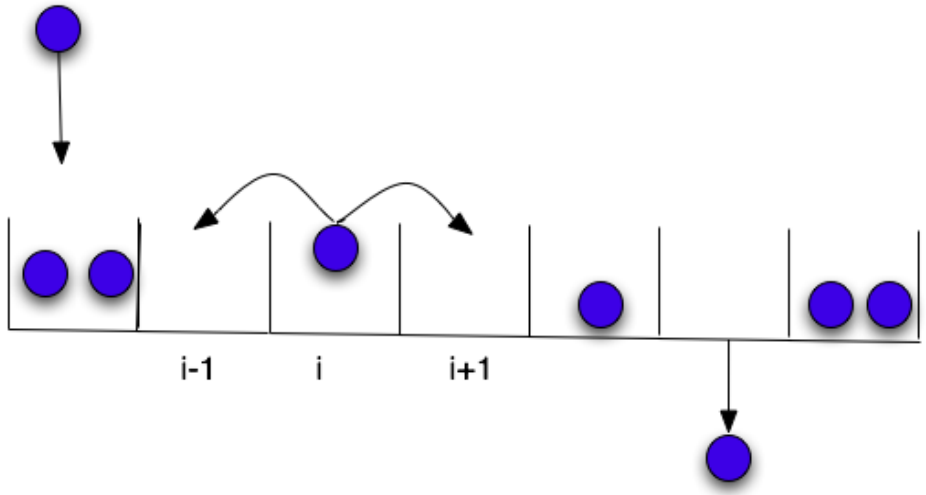
$$\frac{1}{2\sigma^2}||\mathbf{y}-\mathbf{n}||^2 = \max_{\boldsymbol{\phi}}\left\{-\frac{\sigma^2}{2}||\boldsymbol{\phi}||^2 + \boldsymbol{\phi}^\top(\mathbf{y}-\mathbf{n})\right\} \ .$$

to bound the true likelihood

$$-\ln E\left\{\exp\left[-\frac{1}{2\sigma^2}\sum_k ||\mathbf{y}_k - \mathbf{n}(t_k)||^2\right]\right\} \geq$$

$$\max_{\{\boldsymbol{\phi}\}_{k=1}^K}\left\{-\frac{\sigma^2}{2}\sum_k ||\boldsymbol{\phi}_k||^2 + \sum_k \boldsymbol{\phi}_k^\top \mathbf{y} - \ln E\left[\exp\left(\sum_k \boldsymbol{\phi}_k^\top \mathbf{n}(t_k)\right)\right]\right\}$$

Preliminary Results on simulated data



8 compartments (states), 11 observations

θ	estimate	true
decay	0.02787	0.027
diffusion	16.508	17.200
creation	27.504	30.000

Exact representation

$$p(D|\theta) = \frac{1}{(\sqrt{2\pi\sigma^2})^{Kd}} E\left\{\exp\left[-\frac{1}{2\sigma^2}\sum_k ||\mathbf{y}_k - \mathbf{n}(t_k)||^2\right]\right\}$$

$$= \left(\frac{1}{2\pi}\right)^{Kd} \int \prod_k d\boldsymbol{\psi}_k \; e^{-\frac{1}{2}\sigma^2 \sum_k ||\boldsymbol{\psi}_k||^2 - i\sum_k \boldsymbol{\psi}_k^\top \mathbf{y} + \mathcal{K}(i\boldsymbol{\psi})}$$

where

$$\mathcal{K}(\boldsymbol{\phi}) = \ln E\left\{\exp\left[\sum_k \boldsymbol{\phi}^\top \mathbf{n}(t_k)\right]\right\}$$

is the *cumulant generating function* of the set of variables $\mathbf{n}(t_k)_{k=1,\ldots,K}$.

Saddle point approximation to integral

$$p(D|\theta) \approx \left(\frac{1}{\sqrt{2\pi}}\right)^{Kd} \exp\left\{\frac{\sigma^2}{2}\sum_k ||\hat{\boldsymbol{\phi}}_k||^2 - \sum_k \hat{\boldsymbol{\phi}}_k^\top \mathbf{y} + \mathcal{K}(\hat{\boldsymbol{\phi}})\right\} \times$$

$$\times \frac{1}{\sqrt{|\mathcal{K}''(\hat{\boldsymbol{\phi}}) + \sigma^2 \mathbf{I}|}}$$

where $\mathcal{K}''$ is the matrix of second derivatives of $\mathcal{K}$.

Example

$\mathbf{n}(0) = 0$, single noise free ($\sigma = 0$) observation $\mathbf{y}$ at time T:

- Exact result:

$$p(\mathbf{y}|\theta) = \prod_i \left(e^{-\lambda_i} \frac{\lambda_i^{y_i}}{y_i!} \right)$$

  with

$$\lambda_j = \sum_k C_{0k} \left[(e^{T\mathbf{C}} - \mathbf{I})\mathbf{C}^{-1} \right]_{kj}$$

- The approximation based on the 'auxiliary' likelihood

$$E\left\{ e^{\boldsymbol{\phi}^\top \mathbf{n}(T)} \right\} = \exp\left[\sum_i \lambda_i e^{(\phi_i - 1)} \right]$$

  leads to

$$-\ln p(\mathbf{y}|\theta) \approx \sum_i \left(\lambda_i + y_i \ln \frac{y_i}{\lambda_i e} \right)$$

The correction from the fluctuations around the saddlepoint give an extra

$$\frac{d}{2}\ln(2\pi) + \frac{1}{2}\sum_i \ln y_i$$

Which is equivalent to using **Stirling's Approximation**

$$\ln x! \approx \left(x + \frac{1}{2}\right)\ln x - x + \frac{1}{2}\ln(2\pi)$$

in the exact result !

- Note that

$$\frac{1}{12x+1} \leq \ln x! - \text{Stirling} \leq \frac{1}{12x}$$

Plans for the future

- Bayesian approach: $p(\theta|D)$
- Extension to $d \to \infty$ (continuous space)
- Perturbation theory for cumulant generating function to approximately treat chemical reactions.