Biochemical systems with stochastic parameter variation

Wolfram Liebermeister
MPI für molekulare Genetik

Oberseminar 18. November 2004

Humboldt-Universität, Theoretische Biophysik

Causes of uncertainty

- Measurement uncertainty
- Biological variation

  External conditions, genotype, differential expression ...
- Unknown parameters

  Rough guess:
  K_M values between 10^{-4} and 100 mM (Brenda database)

Modelling objectives

- Prediction with missing/uncertain knowledge:
  Probabilistic statements
- Mean values in cell populations
- Bayesian parameter estimation
- Robustness studies
  Which properties are (almost) determined by the model structure?

Outline

- Choice of parameter distributions
- Direct computation of variable distribution
- Stochastic parameter fluctuations

Outline

- **Choice of parameter distributions**
- **Direct computation of variable distribution**
- **Stochastic parameter fluctuations**

The log-normal distribution

- Definition:
  x is log-normally distributed $\Leftrightarrow$ $\ln x$ is normally distributed

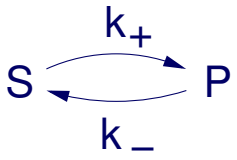
- x, y log-normal $\rightarrow$ $z = x^a \cdot y^b$ also log-normal

- Joint distributions of several variables:
  $\ln \mathbf{x}$ is multivariate normal with $\langle \ln \mathbf{x} \rangle, \mathrm{cov}(\ln \mathbf{x})$

- Parameter distribution defined by
  standard parameters $\mathbf{p}^0$ and coefficients of variation σ_{p_i}/p_i^0

  $\rightarrow$ mean values $\langle \ln \mathbf{p} \rangle$ and variances $\mathrm{var}(\ln \mathbf{p})$

Consistent joint distributions. Example 1

Mass-action kinetics 

$$v = k_+ [S] - k_- [P]$$

$$\begin{aligned} k_+/k_- &= q = \mathrm{e}^{-\beta\Delta g} && (= \text{equilibrium constant}) \\ \sqrt{k_+ \cdot k_-} &= ru && (= \text{ratio} \cdot \text{enzyme activity}) \end{aligned}$$

This can be fulfilled by choosing g, r, u, and setting

$$\begin{aligned} \ln q &= -\beta\Delta g \\ \ln k_+ &= \ln u + 1/2 \ln q + \ln r \\ \ln k_- &= \ln u - 1/2 \ln q + \ln r \end{aligned}$$

Linear combination of logarithms

$$\begin{aligned} (\ln k_+, \ln k_-)^{\mathrm{T}} &= \mathbf{a} \ln u + \mathbf{b} \ln q + \mathbf{c} \ln r \\ \text{with} \quad \mathbf{a} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{b} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

Compute mean and cov for $(\ln k_+, \ln k_-)^{\mathrm{T}}$ from mean and variance for $g, \ln r, \ln u$

Consistent joint distributions. Example 2

Reversible Michaelis-Menten kinetics $v = \frac{V^{\rightarrow}_{\max}/K^{\rightarrow}_{\mathrm{M}}[S] - V^{\leftarrow}_{\max}/K^{\leftarrow}_{\mathrm{M}}[P]}{1+[S]/K^{\rightarrow}_{\mathrm{M}}+[P]/K^{\leftarrow}_{\mathrm{M}}}$

Assumptions about parameters (with $\mathbf{z}$ multivariate standard normal):

$$\begin{aligned} \frac{V^{\rightarrow}_{\max}}{K^{\rightarrow}_{\mathrm{M}}}\frac{K^{\leftarrow}_{\mathrm{M}}}{V^{\leftarrow}_{\max}} &= q \\ \sqrt{V^{\rightarrow}_{\max}\cdot V^{\leftarrow}_{\max}} &= uV^{*}_{\max} \end{aligned} \quad \text{with} \quad \begin{aligned} \ln K^{\rightarrow}_{\mathrm{M}} &= \mu^{\mathrm{S}}_{K} + \sigma_K z_1 \\ \ln K^{\leftarrow}_{\mathrm{M}} &= \mu^{\mathrm{P}}_{K} + \sigma_K z_2 \\ \ln V^{*}_{\max} &= \mu_V + \sigma_V z_3, \end{aligned}$$

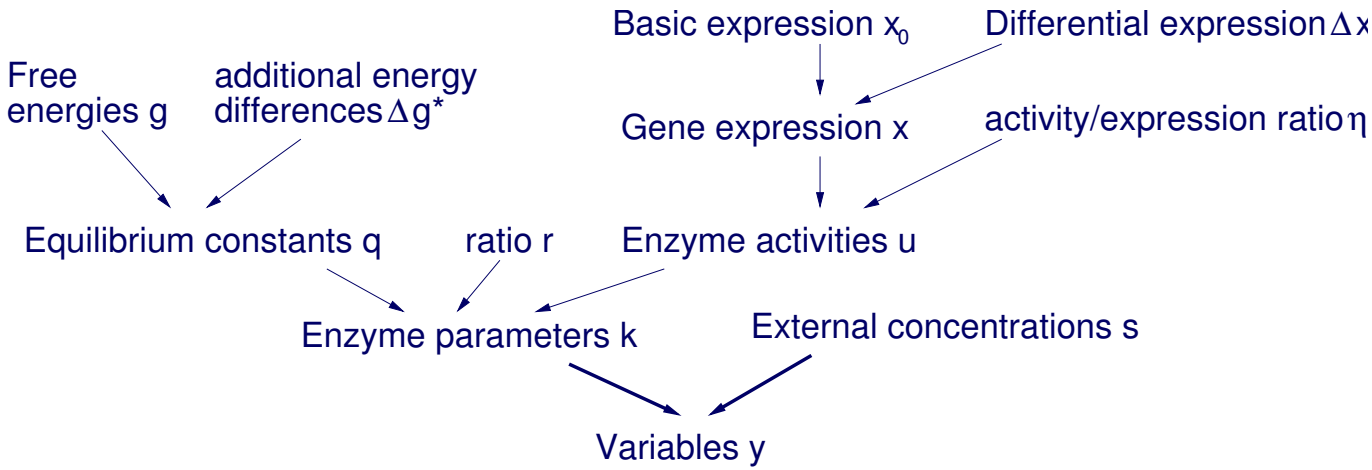
Linear combination of logarithms

$$(\ln K^{\rightarrow}_{\mathrm{M}}, \ln K^{\leftarrow}_{\mathrm{M}}, \ln V^{\rightarrow}_{\max}, \ln V^{\leftarrow}_{\max})^{\mathrm{T}} = \mathbf{a}\ln u + \mathbf{b}\ln q + C\,\mathbf{z} + \mathbf{d}$$

with $\mathbf{a} = (0,0,1,1)^{\mathrm{T}}$, $\mathbf{b} = (0,0,1/2,-1/2)^{\mathrm{T}}$, $\mathbf{d} = (\mu^{\mathrm{S}}_{K}, \mu^{\mathrm{P}}_{K}, \mu_V, \mu_V)^{\mathrm{T}}$

$$C = \begin{pmatrix} \sigma_K & 0 & 0 \\ 0 & \sigma_K & 0 \\ 1/2\sigma_K & -1/2\sigma_K & \sigma_V \\ -1/2\sigma_K & 1/2\sigma_K & \sigma_V \end{pmatrix}$$

Scheme for dependent parameters



Distribution of all parameters: multiplicative relations yield

$$\ln \mathbf{k} \quad = \quad A\ \ln \mathbf{u} + B\ \ln \mathbf{q} + C\mathbf{z} + \mathbf{d}, \qquad \ln \mathbf{q} = -\beta N^{\mathrm{T}} \mathbf{g}$$

Compute means and covariances

$$\begin{aligned} \langle \ln \mathbf{k} \rangle &= A\ \langle \ln \mathbf{u} \rangle + B\ \langle \ln \mathbf{q} \rangle + \mathbf{d} \\ \mathrm{cov}(\ln \mathbf{k}) &= A\ \mathrm{cov}(\ln \mathbf{u})\ A^{\mathrm{T}} + B\ \mathrm{cov}(\ln \mathbf{q})\ B^{\mathrm{T}} + CC^{\mathrm{T}} \end{aligned}$$

Consistent description by log-normal distributions!

Outline

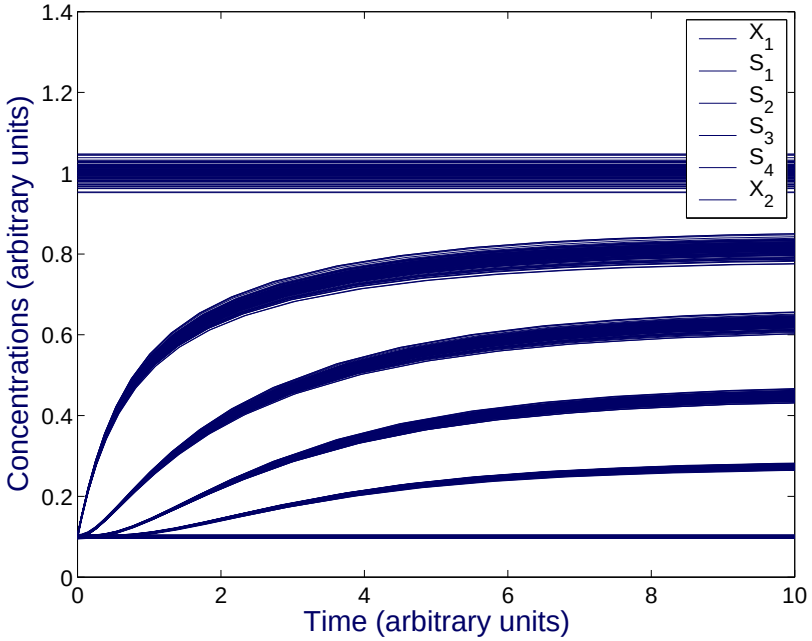
- Choice of parameter distributions
- **Direct computation of variable distribution**
- Stochastic parameter fluctuations

Distribution of variables

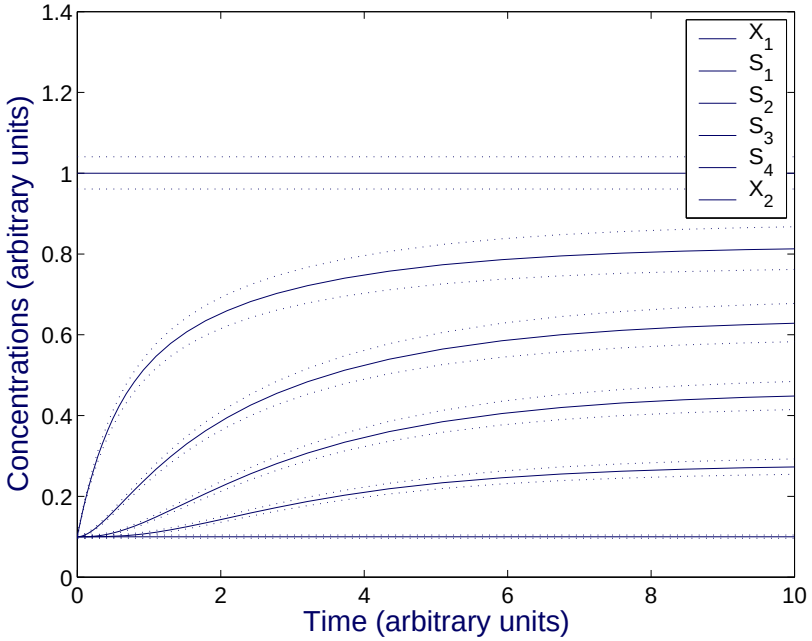
Example: Linear chain

$$X_1 \rightarrow S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_4 \rightarrow X_2$$

Parameters independent log-normal



Monte Carlo simulations

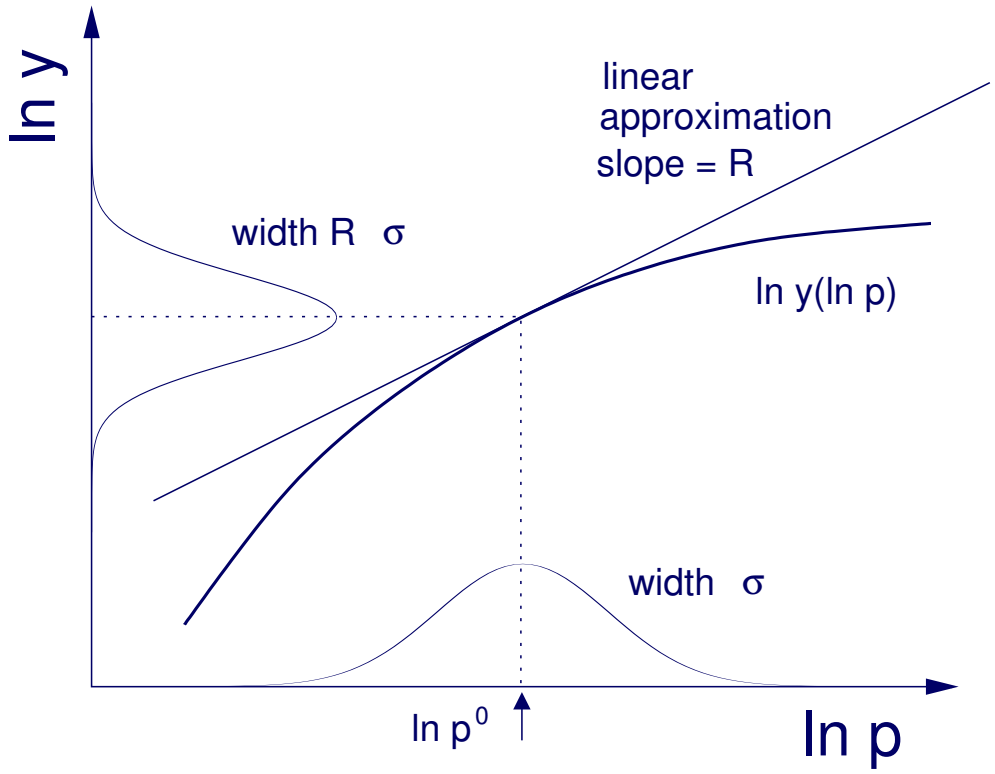


Quantiles

Expansion method: Idea

Distribution of system variable $\mathbf{y}$

(steady state flux, concentration, control coefficient, time courses, oscillation characteristics, signal characteristics, ...)



Expansion method: Results

Mean and covariance of $\ln \mathbf{y}$ can be computed by

First order expansion

$$\begin{aligned} \langle \ln y_i \rangle &= \ln(y_i(\mathbf{p}^0)) \\ \mathrm{cov}(\ln \mathbf{y}) &= \hat{R}^{\mathrm{Y}} \mathrm{cov}(\ln \mathbf{p}) (\hat{R}^{\mathrm{Y}})^{\mathrm{T}} \end{aligned}$$

$\hat{R}^{\mathrm{Y}}$: (normalised) response coefficients

Log-normal parameters yield log-normal variables.

Expansion method: Results

Mean and covariance of $\ln \mathbf{y}$ can be computed by

First order expansion

$$\begin{aligned} \langle \ln y_i \rangle &= \ln(y_i(\mathbf{p}^0)) \\ \mathrm{cov}(\ln \mathbf{y}) &= \hat{R}^{\mathrm{Y}} \mathrm{cov}(\ln \mathbf{p}) (\hat{R}^{\mathrm{Y}})^{\mathrm{T}} \end{aligned}$$

$\hat{R}^{\mathrm{Y}}$: (normalised) response coefficients

Log-normal parameters yield log-normal variables.

Second order expansion (with the shortcut $Q := \mathrm{cov}(\ln \mathbf{p})$)

$$\begin{aligned} \langle \ln y_i \rangle &= \ln(y_i(\mathbf{p}^0)) + \frac{1}{2} \sum_{mn} \hat{R}^{(2)\mathrm{Y}}_{lmn} Q_{jk} \\ \mathrm{cov}(\Delta \ln y_l, \Delta \ln y_k) &= \sum_{mn} \hat{R}^{\mathrm{Y}}_{lm} Q_{mn} \hat{R}^{\mathrm{Y}}_{kn} \\ &\quad + \frac{1}{4} \sum_{mnrs} \hat{R}^{(2)\mathrm{Y}}_{lmn} \hat{R}^{(2)\mathrm{Y}}_{krs} (Q_{ms} Q_{nr} + Q_{mr} Q_{ns}) \end{aligned}$$

Ratios and order relations among variables

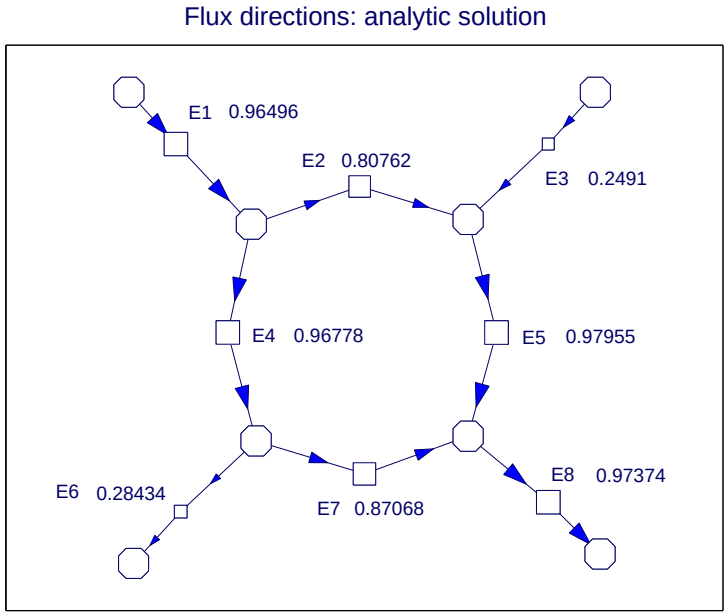
Probabilities of **binary variables** derived from log-normal system variables y, y_1, y_2:

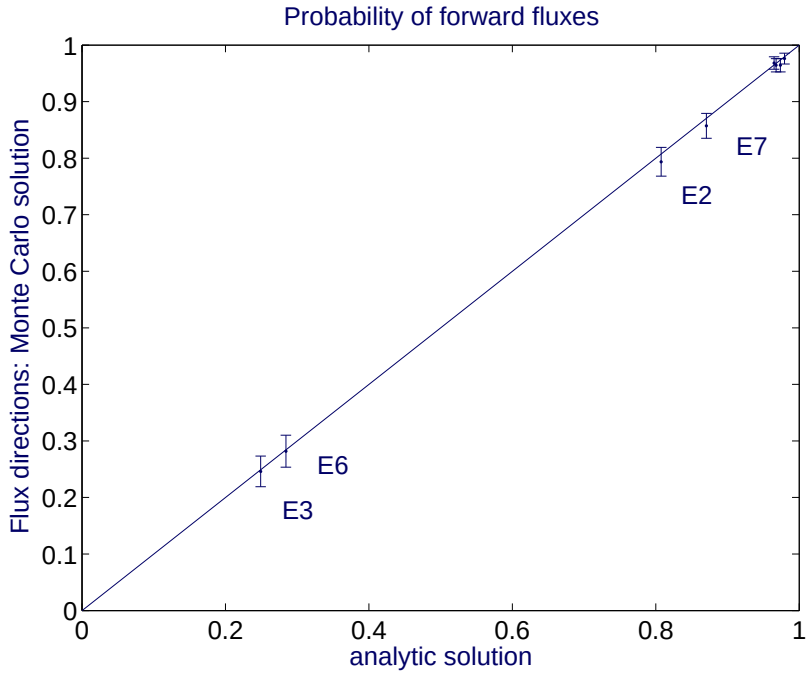
$$\begin{aligned}
\mathrm{Prob}(y > a) &= \Phi\left(\frac{\langle \ln y \rangle - \log a}{\mathrm{var}(\ln y)^{1/2}}\right) \\
\mathrm{Prob}(y_1 > y_2) &= \Phi\left(\frac{\langle \ln y_1 \rangle - \langle \ln y_2 \rangle}{(\mathrm{var}(\ln y_1) + \mathrm{var}(\ln y_2) - 2\ \mathrm{cov}(\ln y_1, \ln y_2))^{1/2}}\right) \\
\mathrm{Prob}(y_1/y_2 > a) &= \Phi\left(\frac{\langle \ln y_1 \rangle - \langle \ln y_2 \rangle - \ln a}{(\mathrm{var}(\ln y_1) + \mathrm{var}(\ln y_2) - 2\ \mathrm{cov}(\ln y_1, \ln y_2))^{1/2}}\right)
\end{aligned}$$

where $\Phi(\cdot)$ is the cumulative density of the standard normal distribution

Example: Probability of flux directions

1. determine joint distribution of forward and backward fluxes
2. compute Prob(forward flux $>$ backward flux)





Other example:

Fold changes of metabolite concentrations by differential expression

Variability and sensitivity

The linear expansion yields measures of ...

- **Variability**

$$\mathrm{cov}(\ln \mathbf{y}) \approx \hat{R}^{\mathrm{Y}} \mathrm{cov}(\ln \mathbf{p}) (\hat{R}^{\mathrm{Y}})^{\mathrm{T}}$$

  Log-normal distributions: coefficient of variation reads $\frac{\sigma_y}{\langle y \rangle} = \sqrt{\mathrm{e}^{\mathrm{var}(\ln y)} - 1}$.

- **Stochastical sensitivity** (Krewski et al., 1995)

$$\begin{aligned} U(\ln y_l | \ln p_i) &:= \mathrm{var}_a(\mathcal{E}(\ln y_l | \ln p_i = a)) \\ &\approx (\sum_j \hat{R}^{\mathrm{Y}}_{lj} \ \mathrm{cov}(\ln p_j, \ln p_i))^2 / \mathrm{var}(\ln p_i) \end{aligned}$$

- **Covariance between parameter and variable**

$$\mathrm{cov}(\ln y_l, \ln p_i) \approx \sum_j \hat{R}^{\mathrm{Y}}_{lj} \ \mathrm{cov}(\ln p_j, \ln p_i)$$

Expansion method: Comments

- Many independent parameters:

  First-order expansion:
  For any parameter distribution, variable distributions
  tend to become log-normal (central limit theorem).

- Second-order expansion

  ... to check the quality of first-order expansion

  ... shows that population mean $\neq$ results from mean parameters!

- Bayesian parameter estimation

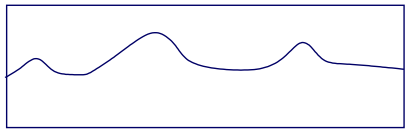
  Improve parameter distributions by using experimental data

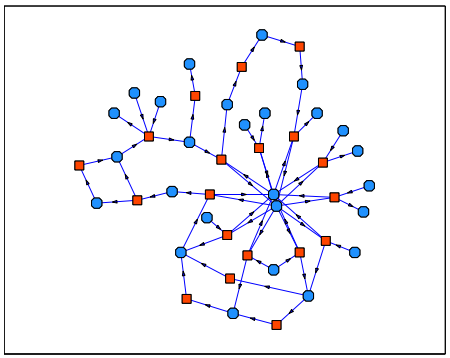
  Explicit formulae for the posterior follows from the first-order expansion

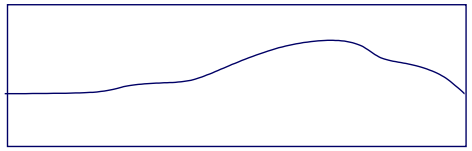
Outline

- Choice of parameter distributions
- Direct computation of variable distribution
- **Stochastic parameter fluctuations**

Time-dependent parameter perturbations







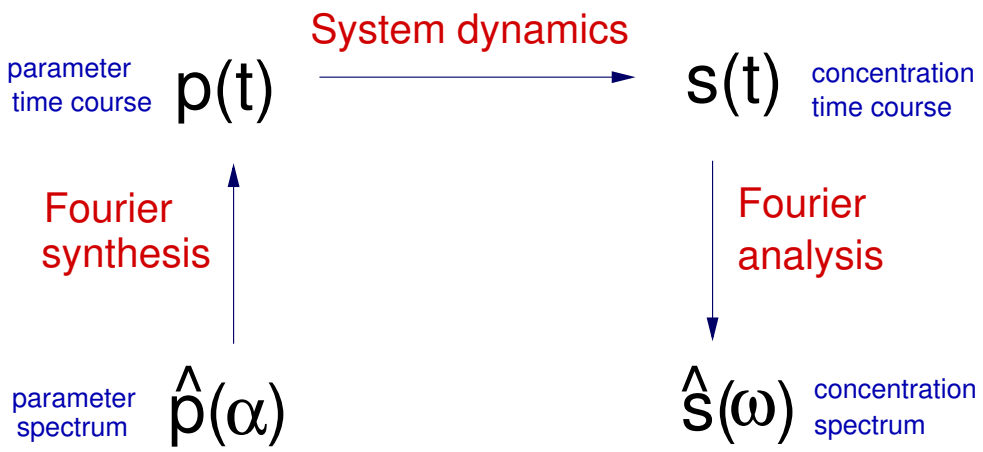
p(t') → s(t), j(t)

parameter time courses

variable time courses

Parameter perturbation of frequency $\omega \longrightarrow$ Forced oscillations of $\mathbf{s}(t), \mathbf{j}(t)$

Spectral response coefficients



Spectral response coefficients

Definition

$$R^{\mathrm{S}}_{lm}(\omega) := \frac{\delta \hat{s}_l(\hat{\mathbf{p}}(\cdot),\omega)}{\delta \hat{p}_m(\omega)}$$

Formula

$$R^{\mathrm{S}}(\omega) = -(N\epsilon^{\mathrm{S}} - \mathrm{i}\omega\mathbf{I})^{-1} N\epsilon^{\mathrm{P}}$$

Spectral density of concentration fluctuations

$$\mathcal{S}^{\mathrm{S}}(\omega) \quad = \quad R^{\mathrm{S}}(\omega)\ \mathcal{S}^{\mathrm{P}}(\omega)\ (R^{\mathrm{S}}(\omega))^{\dagger}$$

Stochastic parameter fluctuations

Stochastic fluctuations due to small particle numbers:

Chemical Langevin equation (see Gillespie, 2000)

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{x}_i(t) = N_{ik}\ a_k(\bar{\mathbf{x}}(t)) + N_{ik}\ \sqrt{a_k(\bar{\mathbf{x}}(t))}\ \eta_k(t)$$

$\bar{x}_i$: molecule numbers, a_k: propensity functions, η_k: white standard noise

Fluctuations around steady state:

1. Linearise system at mean concentrations $\mathbf{x}^0$
2. Represent the fluctuations by virtual noise parameters with elasticities

$$\epsilon^{\mathrm{P}} := (n \text{ liters/mol})^{-1/2}\ \mathrm{diag}(\mathbf{v}(\mathbf{x}^0))^{1/2}$$

n liters/mol $= N_{\mathrm{A}}\Omega$ (n: average molecule number at 1 mol/liter)

Ω: system volume, N_{A}: Avogadro's constant

White Gaussian noise $\rightarrow$ spectral density reads $\mathcal{S}^{\mathrm{S}}(\omega) = R^{\mathrm{S}}(\omega)\ (R^{\mathrm{S}}(\omega))^{\dagger}$

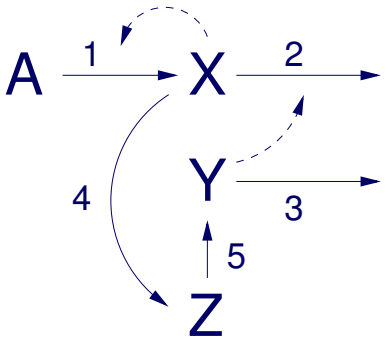
Resonance

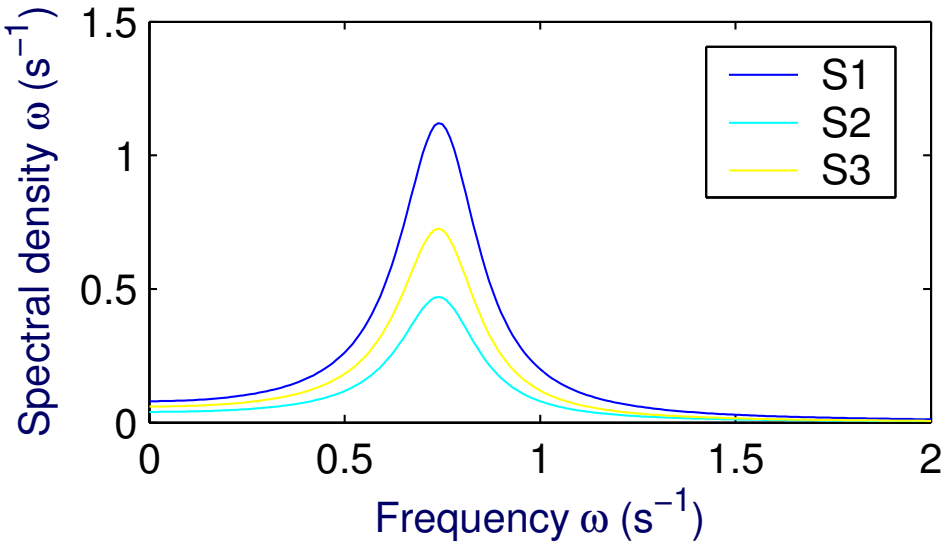
Example:

Minimal reaction system with Hopf bifurcation (Wilhelm + Heinrich, 1995)

Stable steady state:

Small volume $(3.9\ \text{nm})^3$ (≈ 100 molecules of each substance)
$\Rightarrow$ spontaneous fluctuations of concentrations





Resonance near system's own oscillation frequency $\omega_0 \approx 0.75\ \text{s}^{-1}$

Conclusions

- **Choice of parameter distributions**
  - Thermodynamic constraints $\longrightarrow$ correlated parameter distributions
  - Graphical scheme $\longrightarrow$ easy computation of distributions
  - Log-normal distribution is convenient, but not necessary

- **Monte Carlo or expansion method?**
  - Monte Carlo is simple and general, but error is $\mathcal{O}(\sqrt{1/n})$
  - Expansion relies on small parameter variance

  But: Analytical formulae, showing relations to model structure

- **Thanks to ...**

  Dr. Edda Klipp, MPI MolGen
  Christoph Wierling, MPI MolGen