# Accounting for Extrinsic Variability
in the Estimation of Stochastic Rate Constants

Heinz Koeppl
BISON Group - Biomolecular Signaling & cONrol
Automatic Control Lab (ALC), ETH Zurich
Switzerland

*…**joint work** with*

C. Zechner (ALC/ETHZ)
Arnab Ganguly (ALC/ETHZ)
Serge Pelet (IBC/ETHZ)
Matthias Peter (IBC/ETHZ)

I

■ Modeling chemical kinetics by a CTMC

■ Cell-to-cell variability - extrinsic noise

■ Statistical inference - recursive estimation scheme

■ Modeling case study: osmo-stress response in yeast

■ Conclusions

Compartment Ω $t = t_0$

Population encoding

$2\times$ $3\times$ $2\times$ $1\times$ $1\times$

$\mathbf{X}(t_0) = (2 \quad 3 \quad 2 \quad 1 \quad 1)$
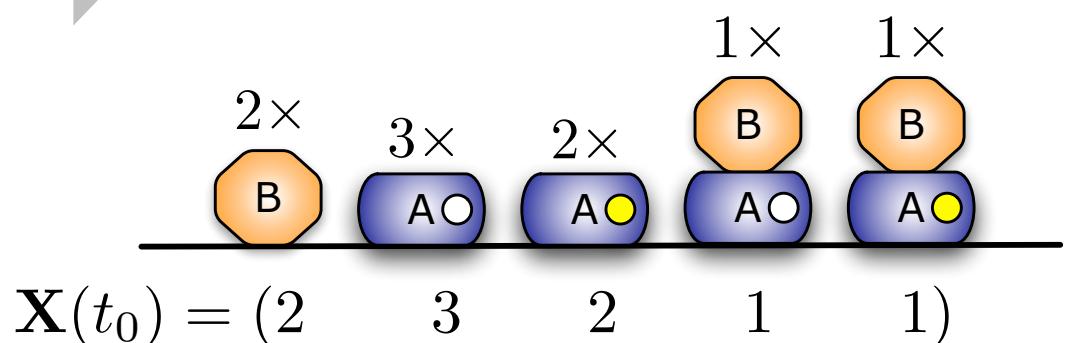
state vector of species
multiplicity

$$A_u + B \xrightarrow{c} A_u : B$$
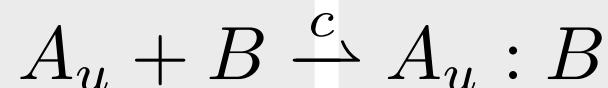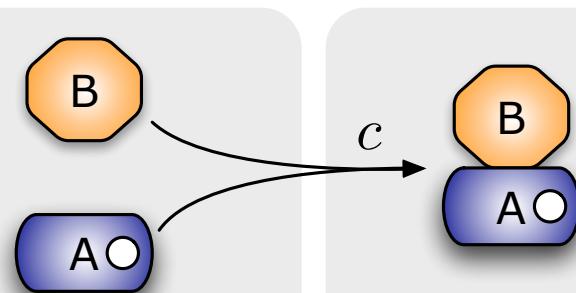
Population state at t
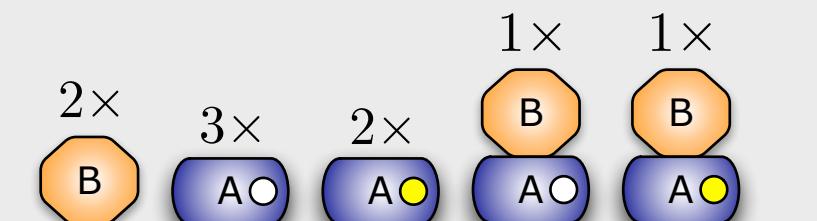
2× 3× 2× 1× 1×

$$\mathbf{X}(t) = (2 \quad 3 \quad 2 \quad 1 \quad 1)$$

Population state at $t + \Delta$
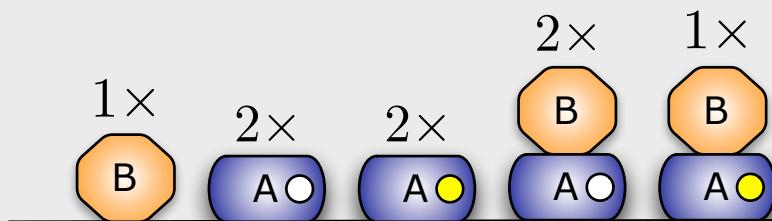
1× 2× 2× 2× 1×

$$\mathbf{X}(t + \Delta) = (1 \quad 2 \quad 2 \quad 2 \quad 1)$$

$$\boldsymbol{\delta} = (-1 \quad -1 \quad 0 \quad 1 \quad 0)$$

state-change vector of reaction

4

Population state at t

Population state at $t + \Delta$

$2\times$ $3\times$ $2\times$ $1\times$ $1\times$
B A A B B
A A A

$\mathbf{X}(t) = (2 \quad 3 \quad 2 \quad 1 \quad 1)$

$1\times$ $2\times$ $2\times$ $2\times$ $1\times$
B A A B B
A A A

$\mathbf{X}(t + \Delta) = (1 \quad 2 \quad 2 \quad 2 \quad 1)$

Propensity or intensity of reaction/jump
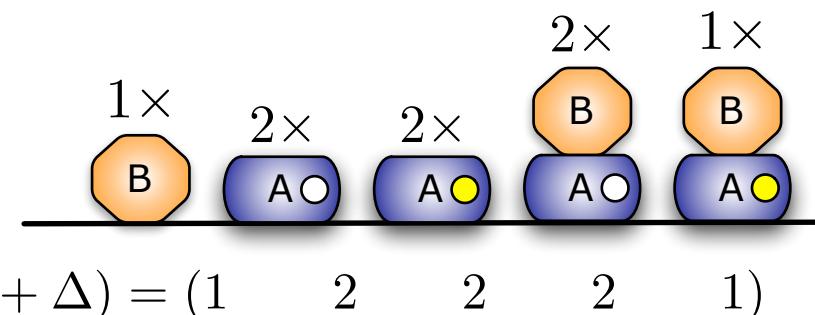
$$P[\mathbf{X}(t + \Delta) = \mathbf{x} + \boldsymbol{\delta} \,|\, \mathbf{X}(t) = \mathbf{x}] = a(\mathbf{x})\Delta + o(\Delta)$$

$$= c \cdot x_1 x_2 \Delta + o(\Delta)$$

$$= c \cdot 2 \cdot 3\Delta + o(\Delta)$$

Transition probability of a continuous-time Markov chain

In how many ways B and Au can combine?

$$f(\Delta) \in o(\Delta) \quad \text{if} \quad \lim_{\Delta \to 0} \frac{f(\Delta)}{\Delta} = 0$$

5

$$\mathbf{X}(t) = \mathbf{x}$$

$$(2 \quad 3 \quad 2 \quad 1 \quad 1)$$

$$R_i$$

$$\mathbf{X}(t + \Delta) = \mathbf{x} + \boldsymbol{\delta}_i$$

$$(1 \quad 2 \quad 2 \quad 2 \quad 1)$$


c_i
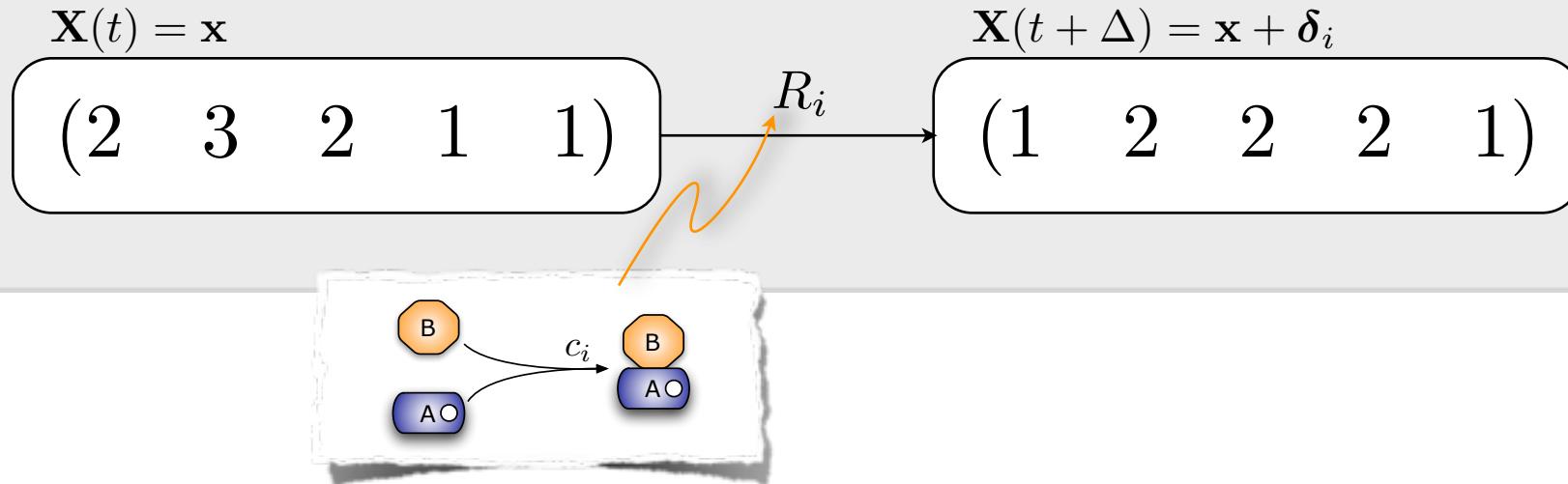
Jump times

$$\tau_k = \inf\{s > \tau_{k-1} | \mathbf{X}(s) \neq \mathbf{X}(\tau_{k-1})\}$$

Waiting times

$$(\tau_k - \tau_{k-1} | \mathbf{X}(\tau_{k-1}) = \mathbf{x}) \sim Exp\{a_0(\mathbf{x})\}$$

Total intensity

$$a_0(\mathbf{x}) = a_1(\mathbf{x}) + \cdots + a_M(\mathbf{x})$$

$\mathbf{X}(t)$

τ_{k-1} τ_k t

6

$$\mathbf{X}(t) = \mathbf{x}$$

$$(2 \quad 3 \quad 2 \quad 1 \quad 1)$$

$$\mathbf{X}(t+\Delta) = \mathbf{x} + \boldsymbol{\delta}_1$$

$$(1 \quad 3 \quad 1 \quad 1 \quad 2)$$

R_1

R_i

$$\mathbf{X}(t+\Delta) = \mathbf{x} + \boldsymbol{\delta}_i$$

$$(1 \quad 2 \quad 2 \quad 2 \quad 1)$$

R_M

$$\mathbf{X}(t+\Delta) = \mathbf{x} + \boldsymbol{\delta}_M$$

$$(2 \quad 3 \quad 2 \quad 0 \quad 2)$$

Jump type at k-th jump

$$P[r_k = R_i | \mathbf{X}(\tau_{k-1}) = \mathbf{x}] = \frac{a_i(\mathbf{x})}{a_0(\mathbf{x})}$$
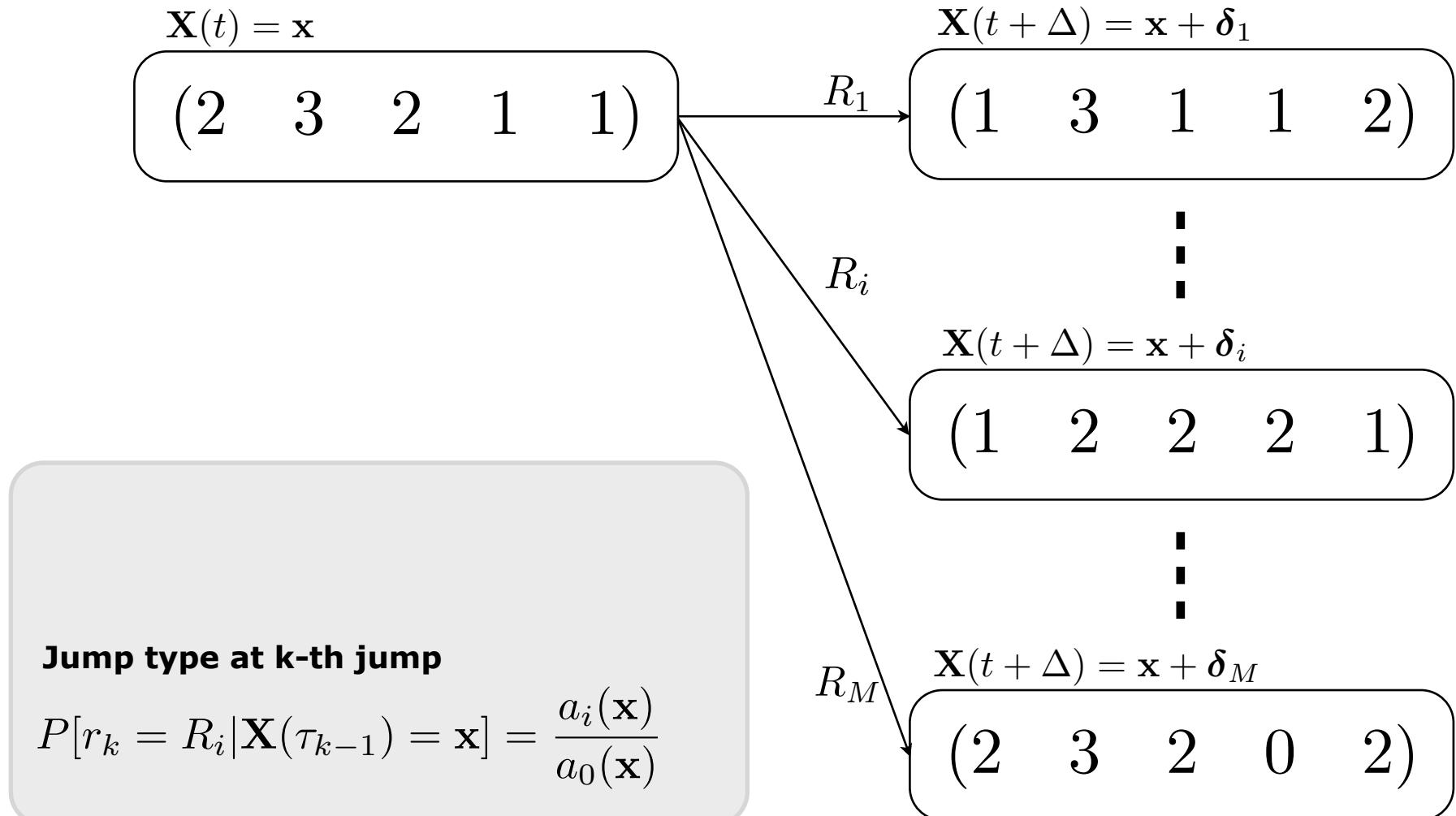
■ Observe a sample path (through the transition system)

$(\mathbf{x}_1, \tau_1)$

$(2 \quad 3 \quad 2 \quad 1 \quad 1)$

$\xrightarrow{R_1}$

$(\mathbf{x}_2, \tau_2)$

$(1 \quad 3 \quad 1 \quad 1 \quad 2)$

$(1 \quad 2 \quad 2 \quad 2 \quad 1)$

$(\mathbf{x}_3, \tau_3)$

$(0 \quad 2 \quad 1 \quad 2 \quad 2)$

$(1 \quad 2 \quad 2 \quad 1 \quad 2)$

$(\mathbf{x}_4, \tau_4)$

$(0 \quad 2 \quad 1 \quad 1 \quad 3)$
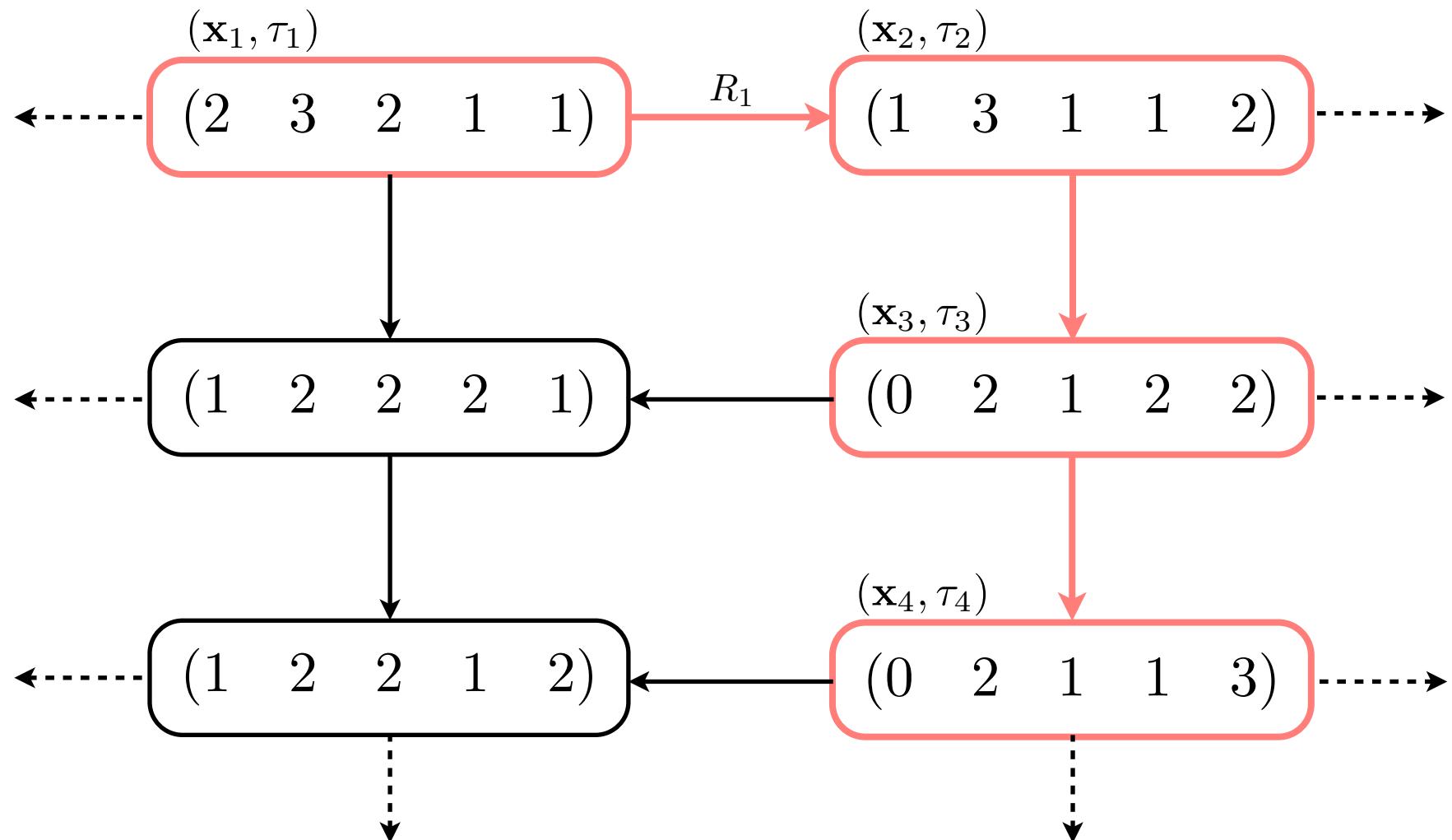
■ **Density of path**

$$\pi = (\mathbf{X}_{[\tau_1, \tau_4)} = \mathbf{x}_{[\tau_1, \tau_4)}) = \{(\mathbf{x}_1, \tau_1), (\mathbf{x}_2, \tau_2), (\mathbf{x}_3, \tau_3), (\mathbf{x}_4, \tau_4)\}$$

$(\mathbf{x}_1, \tau_1)$

$$(2 \quad 3 \quad 2 \quad 1 \quad 1)$$

R_1

$(\mathbf{x}_2, \tau_2)$

$$(1 \quad 3 \quad 1 \quad 1 \quad 2)$$

$(\mathbf{x}_3, \tau_3)$

$$(0 \quad 2 \quad 1 \quad 2 \quad 2)$$

$(\mathbf{x}_4, \tau_4)$

$$(0 \quad 2 \quad 1 \quad 1 \quad 3)$$

$$
\begin{aligned}
p(\pi) = & \, p(\mathbf{x}_1) \\
& \times a_0(\mathbf{x}_1) \exp[-a_0(\mathbf{x}_1)(\tau_2 - \tau_1)] \frac{a_1(\mathbf{x}_1)}{a_0(\mathbf{x}_1)} \\
& \times a_0(\mathbf{x}_2) \exp[-a_0(\mathbf{x}_2)(\tau_3 - \tau_2)] \frac{a_2(\mathbf{x}_2)}{a_0(\mathbf{x}_2)} \\
& \times a_0(\mathbf{x}_3) \exp[-a_0(\mathbf{x}_3)(\tau_4 - \tau_3)] \frac{a_3(\mathbf{x}_3)}{a_0(\mathbf{x}_3)}
\end{aligned}
$$
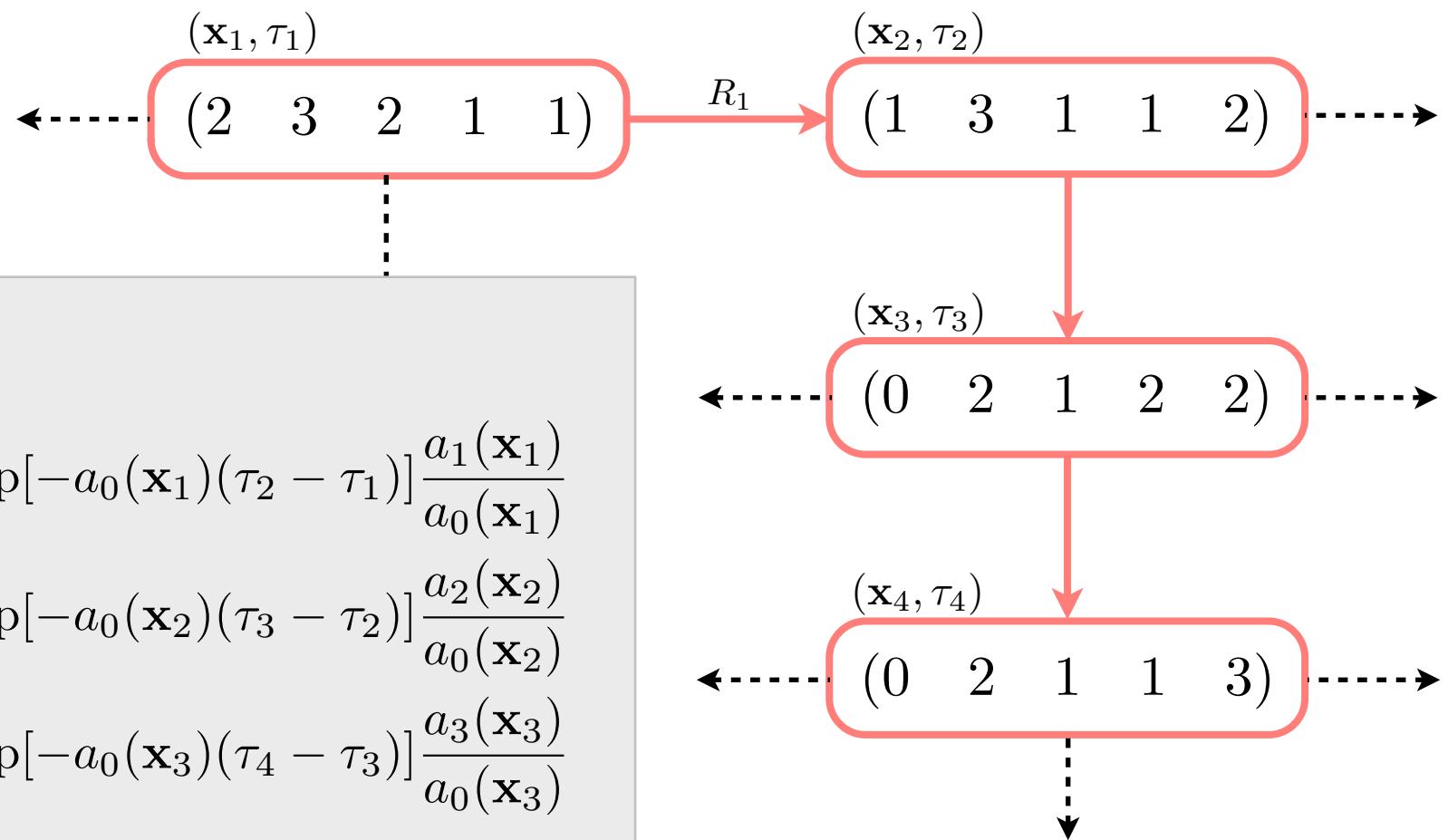
■ Mass-action propensities $\quad a_i(\mathbf{x}) = c_i f_i(\mathbf{x})$

monoms in x

$(\mathbf{x}_1, \tau_1)$

$$(2 \quad 3 \quad 2 \quad 1 \quad 1)$$

R_1

$(\mathbf{x}_2, \tau_2)$

$$(1 \quad 3 \quad 1 \quad 1 \quad 2)$$

$(\mathbf{x}_3, \tau_3)$

$$(0 \quad 2 \quad 1 \quad 2 \quad 2)$$

$(\mathbf{x}_4, \tau_4)$

$$(0 \quad 2 \quad 1 \quad 1 \quad 3)$$
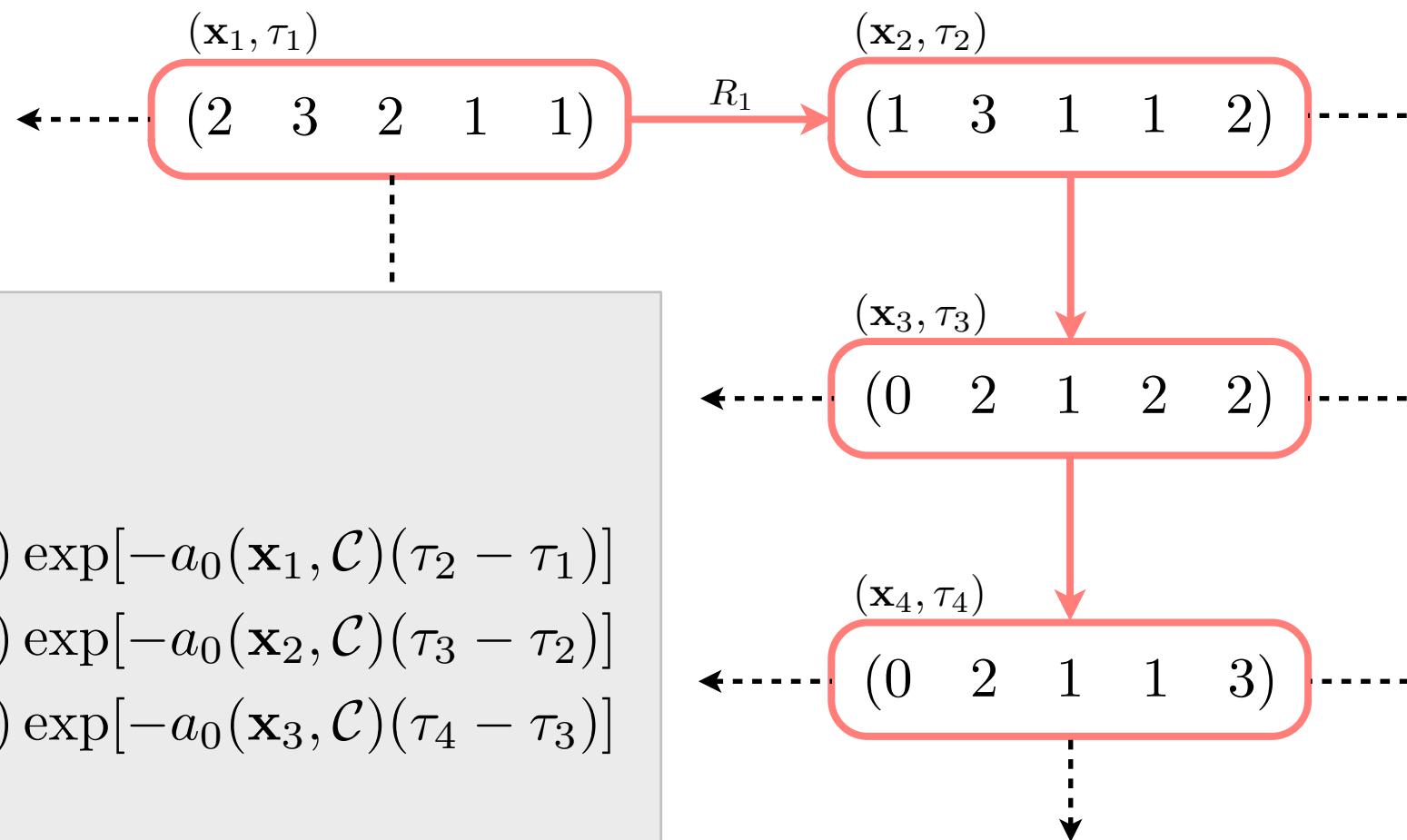
$$
\begin{aligned}
p(\pi|\mathcal{C}) = {} & p(\mathbf{x}_1) \\
& \times c_1 f_1(\mathbf{x}_1) \exp[-a_0(\mathbf{x}_1, \mathcal{C})(\tau_2 - \tau_1)] \\
& \times c_2 f_2(\mathbf{x}_2) \exp[-a_0(\mathbf{x}_2, \mathcal{C})(\tau_3 - \tau_2)] \\
& \times c_3 f_3(\mathbf{x}_3) \exp[-a_0(\mathbf{x}_3, \mathcal{C})(\tau_4 - \tau_3)]
\end{aligned}
$$

■ Take a general path

$$\pi = (\mathbf{X}_{[\tau_1, \tau_J)} = \mathbf{x}_{[\tau_1, \tau_J)}) = \{(\mathbf{x}_1, \tau_1), \ldots, (\mathbf{x}_J, \tau_J)\}$$

number of occurrences of reaction m in the path

$$p(\pi|\mathcal{C}) = p(\mathbf{x}_1) \left[\prod_{m=1}^{M} c_m^{r_m} \right] \left[\prod_{j=1}^{J} f_{m_j}(\mathbf{x}(\tau_j)) \right] \exp\left\{ -\int_{\tau_1}^{\tau_J} a_0(\mathbf{x}(s), \mathcal{C}) ds \right\}$$

$$\propto \left[\prod_{m=1}^{M} c_m^{r_m} \right] \exp\left\{ -\int_{\tau_1}^{\tau_J} a_0(\mathbf{x}(s), \mathcal{C}) ds \right\}$$

Factorization of density

$$p(\pi|\mathcal{C}) \propto g(\pi|\mathcal{C}) = \prod_{m=1}^{M} g_m(\pi|c_m)$$

Gamma distribution

$$\Gamma(a, b) \propto x^{a-1} e^{-bx}$$

$$g_m(\pi|c_m) = c_m^{r_m} \exp\left\{ -c_m \int_{\tau_1}^{\tau_J} f_m(\mathbf{x}(s)) ds \right\}$$

11

■ Assume prior conjugate to path density (likelihood)

$$g_m(\pi|c_m) = c_m^{r_m} \exp\left\{-c_m \int_{\tau_1}^{\tau_J} f_m(\mathbf{x}(s))\mathrm{d}s\right\}$$

■ Independent Gamma priors $\quad c_m \sim \Gamma(a_m, b_m)$

Posterior

$$c_m|\pi \sim \Gamma\left(a_m + r_m, b_m + \int_{\tau_1}^{\tau_J} f_m(\mathbf{x}(s))\mathrm{d}s\right)$$

■ Given a sample path we can sample parameters and also compute MAP estimates.

■ Modeling chemical kinetics by a CTMC

■ Cell-to-cell variability - extrinsic noise

■ Statistical inference - recursive estimation scheme

■ Modeling case study: osmo-stress response in yeast

■ Conclusions

Single-cell time-lapsed microscopy

Flow cytometry measurement
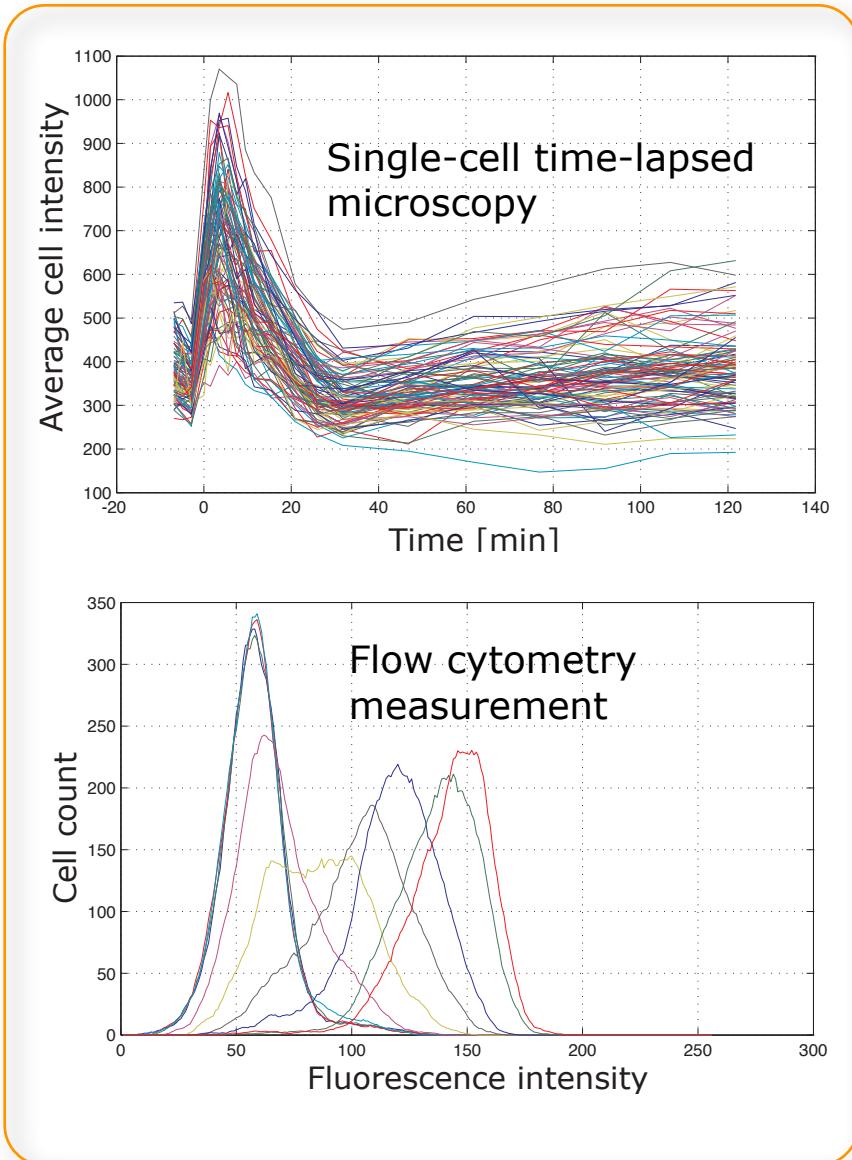
■ New measurement technologies resolve protein dynamics at single cell level.

■ Sources of cell-to-cell variability in isogenic cell population?

■ Tempting to attribute variability solely to stochastic chemical kinetics.

■ Large extrinsic components, volume, cell-cycle stage, concentration.

Engineered yeast strain

Stimuli

Signal

DNA

| P | CFP | | P | YFP |

CFP   YFP

Proteins
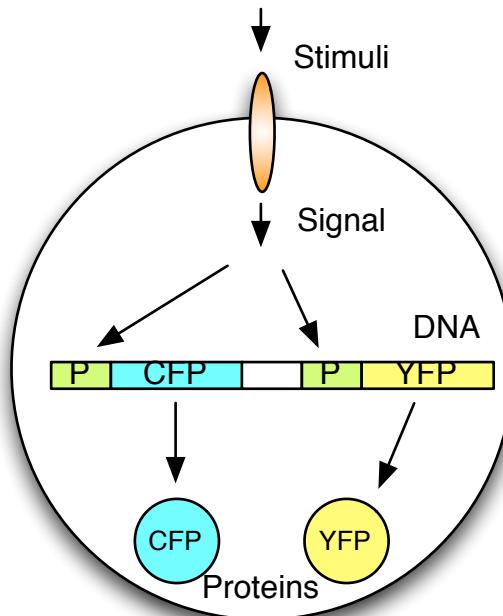
cell-population
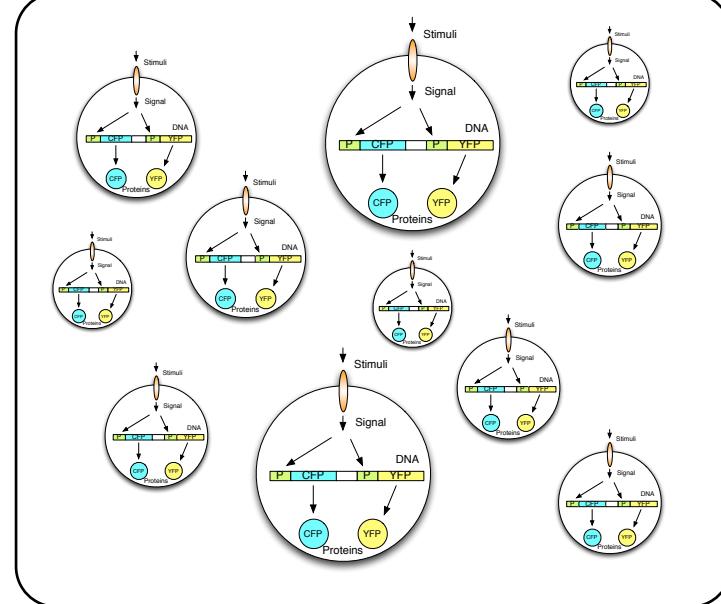
Strong correlation - weaker stochastic component

YFP – F.U.

CFP – F.U.

Electronics and computer system

Fluidics system

Optics

Lasers

Detectors

Reaction list

$$HOG1^P + GENE^{STL1} \underset{c_2}{\overset{c_1}{\rightleftharpoons}} HOG1^P : GENE^{STL1}$$

$$HOG1^P : GENE^{STL1} \overset{c_3}{\rightarrow} HOG1^P : GENE^{STL1} + STL1$$

$$HOG1^P + RSC \underset{c_5}{\overset{c_4}{\rightleftharpoons}} HOG1^P : RSC$$

$$HOG1^P : RSC + GENE^{STL1} \underset{c_7}{\overset{c_6}{\rightleftharpoons}} HOG1^P : RSC : GENE^{STL1}$$

$$HOG1^P : RSC : GENE^{STL1} \overset{c_8}{\rightarrow} HOG1^P : RSC : GENE^{STL1} + STL1$$

$$STL1 \overset{c_9}{\rightarrow} 0$$

$$HOG1^P + PTP2 \underset{c_{11}}{\overset{c_{10}}{\rightleftharpoons}} HOG1^P : PTP2$$

$$HOG1^P : PTP2 \overset{c_{12}}{\rightarrow} HOG1 + PTP2,$$
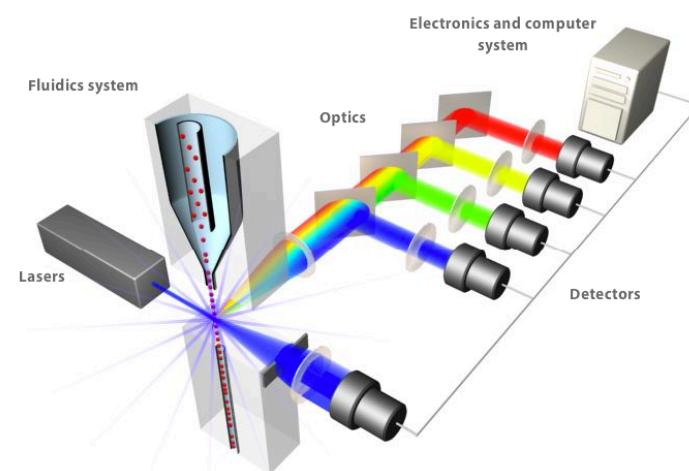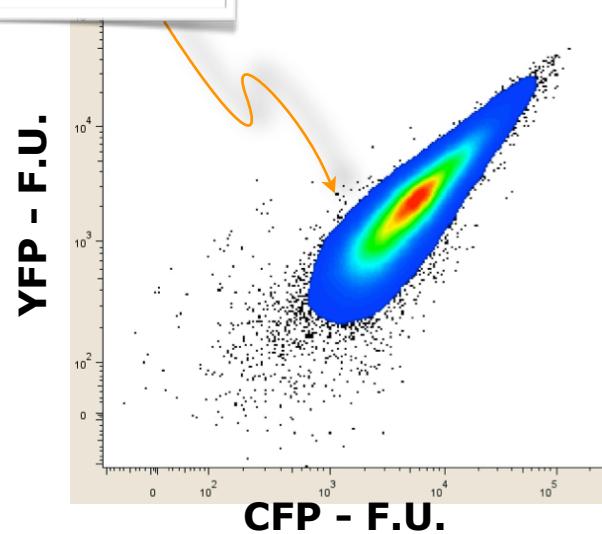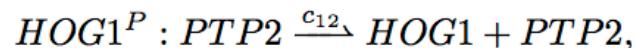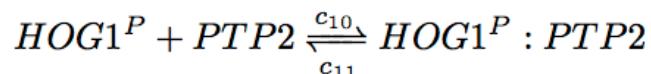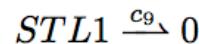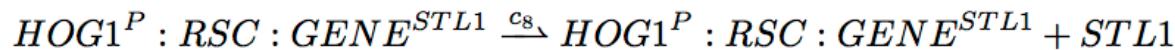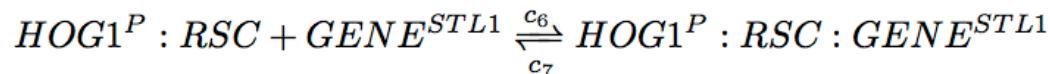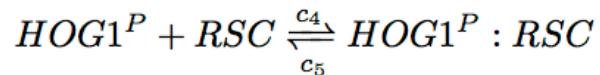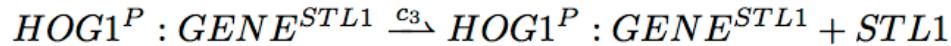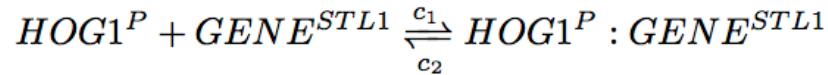
Every cell corresponds to one
realization of that process

Extrinsic components

$$\mathrm{d}c_i(t) = \gamma(c_i^0 - c_i(t))\mathrm{d}t + \mathrm{d}W(t)$$

Ornstein-Uhlenbeck
process for rate constants

■ Rate constants **fluctuate independently** around same mean.

■ Ignoring common **lower-dimensional causes** of cell-to-cell variability.

■ Cannot account for **strong temporal correlation** to the cell-state (e.g. cell-cycle position).

■ Mechanistic rate constants **determined by biophysics** of interacting molecules - should be invariant.

■ Treat mechanistic **rate-constants as invariant** across of an isogenic cell population.

■ Incorporate **variability in total protein count** and concentration, i.e. in mass-conservation constraints.

■ For fast processes such as signaling, **temporal variation can be ignored**.

■ Variability in protein count or concentration result in variability of **rate parameters in aggregated rate-laws**.

Example

Uni-molecular conversion $\quad A \underset{c_2}{\overset{c_1}{\rightleftharpoons}} B$

Mass-conservation $\qquad b = A + B$

Markov chains

$b^1 = 4$ (0 4) — (1 3) — (2 2) — (3 1) — (4 0)   Cell 1

with transitions $4c_2, 3c_2, 2c_2, 1c_2$ (top) and $1c_1, 2c_1, 3c_1, 4c_1$ (bottom)

$b^m = 2$ (0 2) — (1 1) — (2 0)   Cell m

with transitions $2c_2, 1c_2$ (top) and $1c_1, 2c_1$ (bottom)

Same rate-constants!
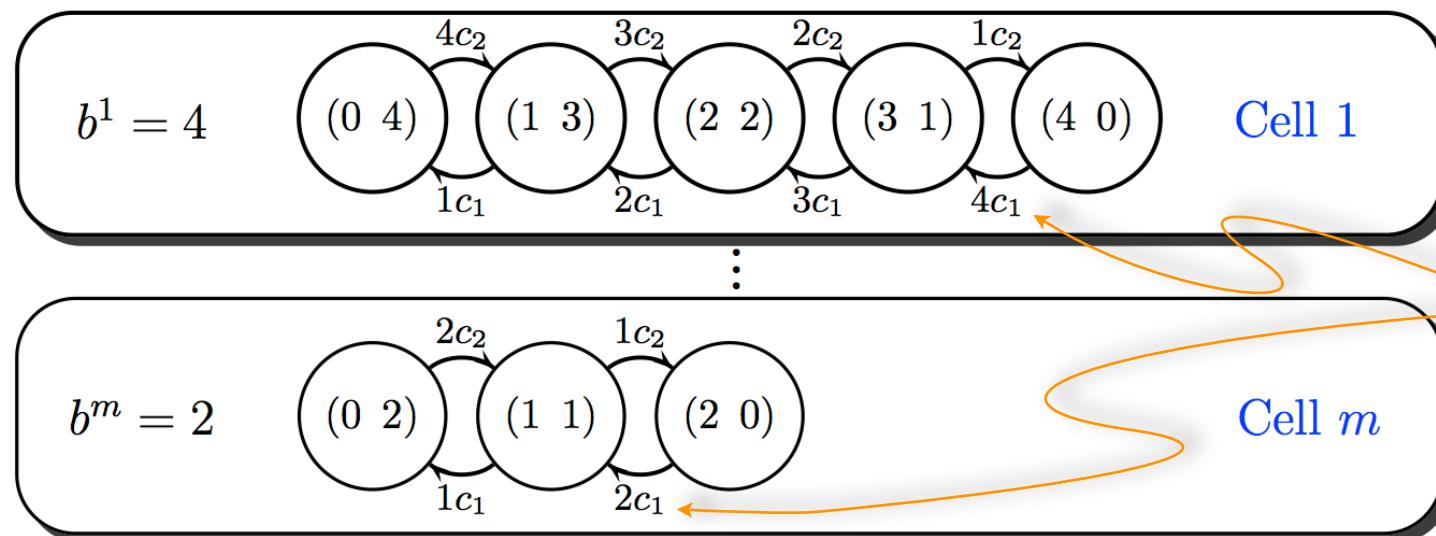
■ **Goal:** Estimate common parameters by observing sample paths from a heterogeneous population of Markov-chains

Stochastic rate constants

Hyper-parameters capturing cell-to-cell variability

Chemical species

$$p(\text{Parameters, Extrinsic Variability, States}|\text{Measurement})$$
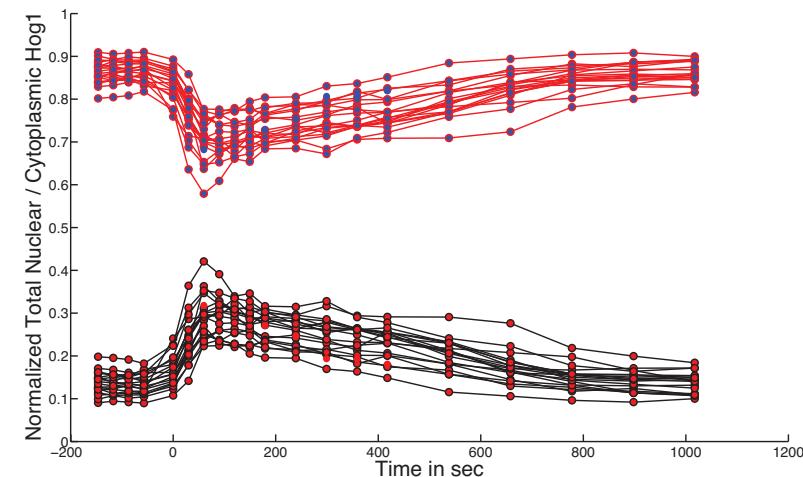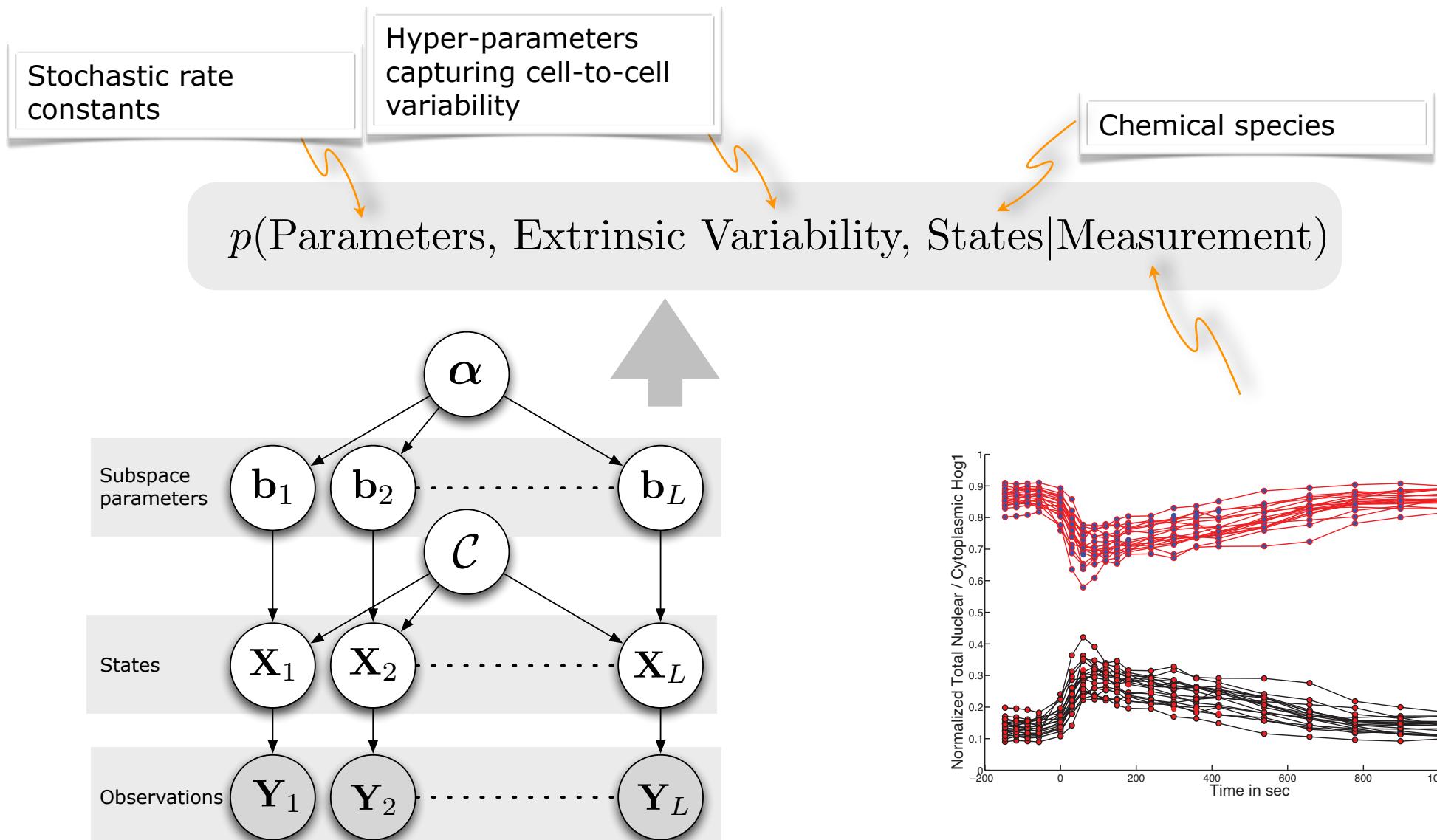
$$p(\mathcal{C}, \alpha, \mathbf{X}_{[t_1, t_l]}|\mathbf{Y}(t_1), \dots, \mathbf{Y}(t_l)) = \int p(\mathcal{C}, \mathcal{B}, \alpha, \mathbf{X}_{[t_1, t_l]}|\mathbf{Y}(t_1), \dots, \mathbf{Y}(t_l))d\mathcal{B}$$

Analytically intractable!

20

Data is very uncertain

Only a small number of species is captured
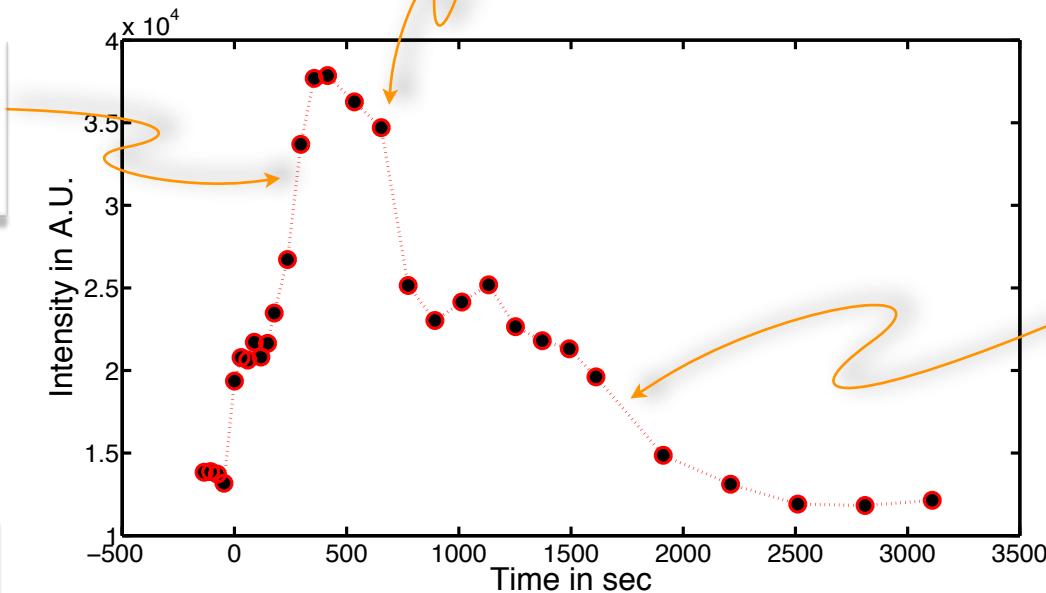
Discrete-time measurements

General State Space model

State transition kernel:

$$p(\mathbf{X}(t_l)|\mathbf{X}(t_{l-1}), \mathcal{C})$$

CME

Measurement likelihood:

$$p(\mathbf{Y}(t_l)|\mathbf{X}(t_l)) = \mathcal{N}\left(\mathbf{W}\mathbf{X}(t_l), \sigma^2\right)$$

■ One could use MCMC to sample from the state posterior

■ Infeasible if parameters are unknown

Prediction step:

$$p\left(\mathbf{X}(t_l),\mathcal{C}|\mathbf{Y}(t_1),\ldots,\mathbf{Y}(t_{l-1})\right) =$$

$$\int p\left(\mathbf{X}(t_l)|\mathbf{X}(t_{l-1}),\mathcal{C}\right) \times p\left(\mathbf{X}(t_{l-1}),\mathcal{C}|\mathbf{Y}(t_1),\ldots,\mathbf{Y}(t_{l-1})\right) d\mathbf{X}(t_{l-1})$$

Correction step:

$$p\left(\mathbf{X}(t_l),\mathcal{C}|\mathbf{Y}(t_1),\ldots,\mathbf{Y}(t_l)\right) \propto p\left(\mathbf{Y}(t_l)|\mathbf{X}(t_l)\right) \times p\left(\mathbf{X}(t_l),\mathcal{C}|\mathbf{Y}(t_1),\ldots,\mathbf{Y}(t_{l-1})\right)$$

Bayesian Recursion
for sequential state
estimation

■ Analytically tractable only for linear + fully Gaussian models: **Kalman filter**

■ Integral can be approximated using sampling methods: **Sequential Monte Carlo**

$$p\left(\mathbf{X}(t_{l-1}), \mathcal{C} | \mathbf{Y}(t_1), \ldots, \mathbf{Y}(t_{l-1})\right)$$

$$p\left(\mathbf{X}(t_l), \mathcal{C} | \mathbf{Y}(t_1), \ldots, \mathbf{Y}(t_{l-1})\right)$$

$$p\left(\mathbf{Y}(t_l) | \mathbf{X}(t_l)\right)$$

$$p\left(\mathbf{X}(t_l), \mathcal{C} | \mathbf{Y}(t_1), \ldots, \mathbf{Y}(t_l)\right)$$

t_{l-1}

t_l
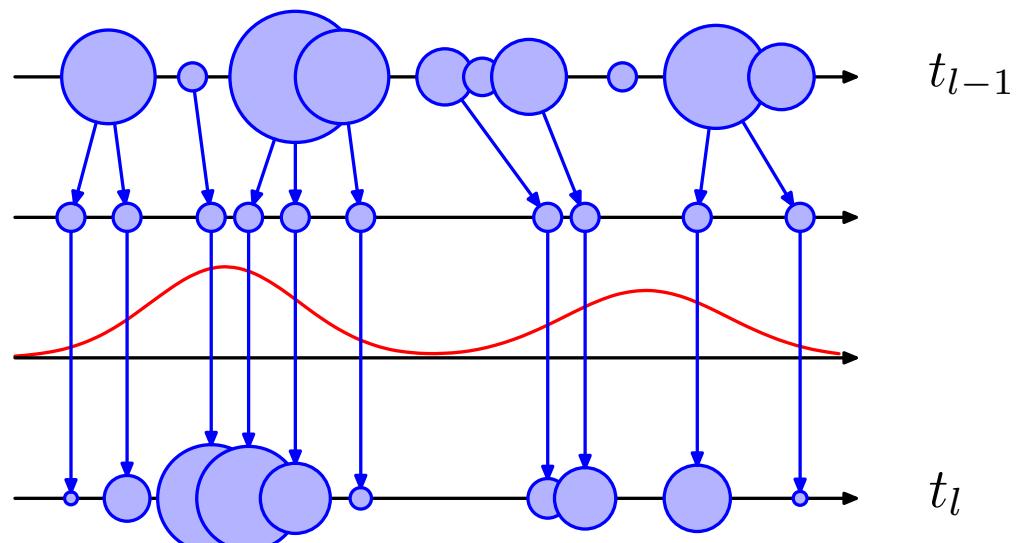
Integral approximated using **importance sampling**

Particle approximation:

$$p\left(\mathbf{X}(t_l), \mathcal{C} | \mathbf{Y}(t_1), \ldots, \mathbf{Y}(t_{l-1})\right) \approx \sum_{i=1}^{P} w^{(i)} p\left(\mathbf{X}(t_l) | \mathbf{X}^{(i)}(t_{l-1}), \mathcal{C}\right)$$

Importance weights:

$$w^{(i)} \propto p\left(\mathbf{Y}(t_{l-1}) | \mathbf{X}^{(i)}(t_{l-1})\right)$$

Transition kernel as proposal distribution: **bootstrap filter** (Gillespie simulation)

23

http://www.bison.ethz.ch/seqMC.mp4

Markov chain Monte Carlo:

Chain with target posterior as it's stationary distribution
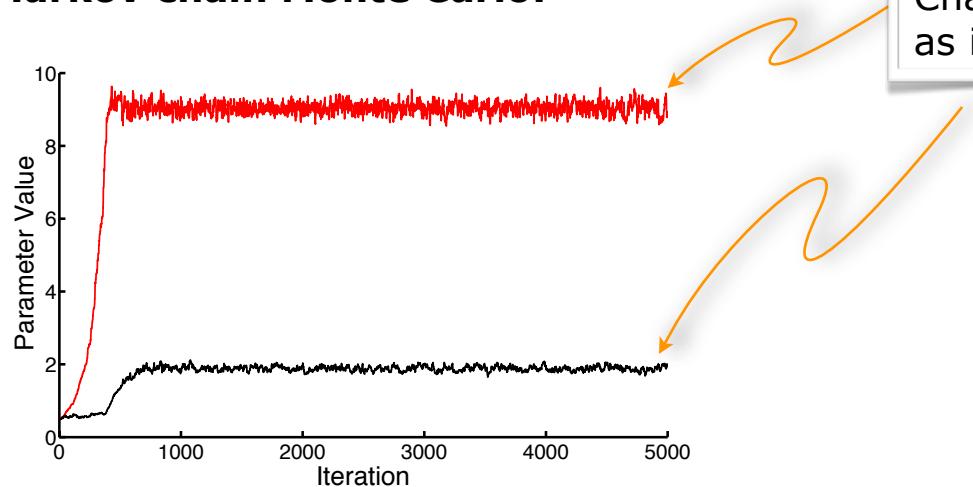
Detailed Balance Condition:

$$K(\mathcal{P}^{j+1}|\mathcal{P}^j)p(\mathcal{P}^j|\mathcal{Y}) = K(\mathcal{P}^j|\mathcal{P}^{j+1})p(\mathcal{P}^{j+1}|\mathcal{Y})$$

Transition kernel of the Markov chain

**Metropolis-Hastings Sampler**

Proposal density

Known from model

$$a = \min\left\{1, \frac{q(\mathcal{P}^j|\mathcal{P}^{j+1})p(\mathcal{P}^{j+1}|\mathcal{Y})}{q(\mathcal{P}^{j+1}|\mathcal{P}^j)p(\mathcal{P}^j|\mathcal{Y})}\right\}$$

Acceptance probability

M-H fulfills detailed balance:

$$K(\mathcal{P}^{j+1}|\mathcal{P}^j)p(\mathcal{P}^j|\mathcal{Y}) = a \times q(\mathcal{P}^{j+1}|\mathcal{P}^j)p(\mathcal{P}^j|\mathcal{Y})$$
$$= \min\left\{q(\mathcal{P}^{j+1}|\mathcal{P}^j)p(\mathcal{P}^j|\mathcal{Y}), q(\mathcal{P}^j|\mathcal{P}^{j+1})p(\mathcal{P}^{j+1}|\mathcal{Y})\right\}$$
$$= K(\mathcal{P}^j|\mathcal{P}^{j+1})p(\mathcal{P}^{j+1}|\mathcal{Y})$$
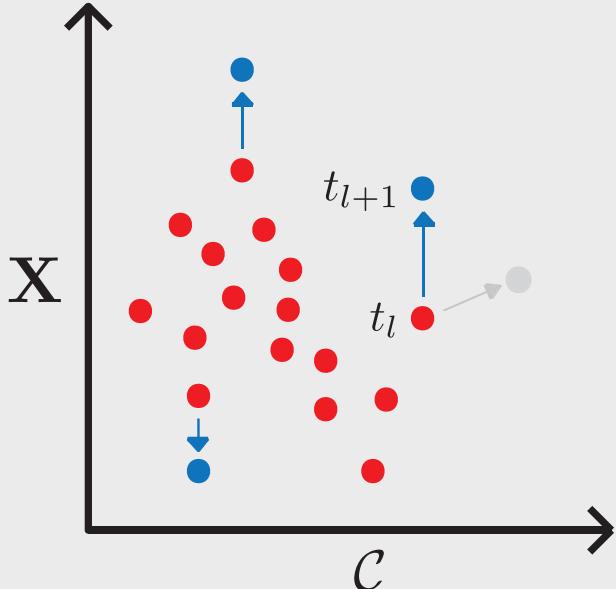
- Parameters cannot take values different from their initialization

- Particles degenerate!

- Additional MCMC moves can be used to diversify particles

Sample trajectory from the particle distribution and ignore corresponding parameter (**marginalization**)

Sample new parameter given the trajectory from Gamma

$$\{\mathbf{X}_{[t_1, t_l]}, \mathcal{C}\} \sim p(\mathbf{X}_{[t_1, t_l])}|\mathbf{Y}(t_1), \ldots, \mathbf{Y}(t_l))p(\mathcal{C}|\mathbf{X}_{[t_1, t_l]}))$$

Still a valid sample from the target posterior!

Bootstrap filter for state & parameter estimation:

1: **for** $t_l \in \{t_2, \ldots, t_L\}$ **do**
2:      Resample particles according to their weights
3:
4:      **for** each particle $i \in \{1, \ldots, P\}$ **do**
5:          **for** each rate constant $c_k^i \in \mathcal{C}^i$ **do**
6:              Update parameter posterior (Γ-distribution)
7:              Sample new c_k^i from posterior
8:          **end for**
9:
10:          Forward simulate path to next measurement time point using proposed
             parameter values
11:
12:          Recompute particle weight
13:      **end for**
14:      Normalize particle weights
15: **end for**

Include MCMC step for parameters after particle resampling

Gillespie!

Measure "goodness" of simulated subpath

Must be a valid probability distribution

27

Bootstrap filter for state & parameters - multiple cells

1: **for** $t_l \in \{t_2, \dots, t_L\}$ **do**
2:     Resample particles according to their weights
3:
4:     **for** each particle $i \in \{1, \dots, P\}$ **do**
5:
6:         **for** each rate constant $c_k^i \in \mathcal{C}^i$ **do**
7:             Update parameter posterior (Γ-distribution)
8:             Sample new c_k^i from posterior
9:         **end for**
10:
11:         **for** each cell $m \in \{1, \dots, L\}$ **do**
12:             Update posterior over conservation constants (discrete density)
13:             Sample new conservation constant
14:         **end for**
15:
16:     Sample hyperparameters using M-H
17:
18:     Forward simulate path to next measurement time point using proposed parameter values
19:
20:     Recompute particle weight
21:     **end for**
22:   Normalize particle weights
23: **end for**

All cells are used to update the posterior!

Additionally we sample the "extrinsic" variables - sufficient statistics

...and the statistics (**extrinsic noise**)

■ Modeling chemical kinetics by a CTMC

■ Cell-to-cell variability - extrinsic noise

■ Statistical inference - recursive estimation scheme

■ Modeling case study: osmo-stress response in yeast

■ Conclusions

[Movie description]
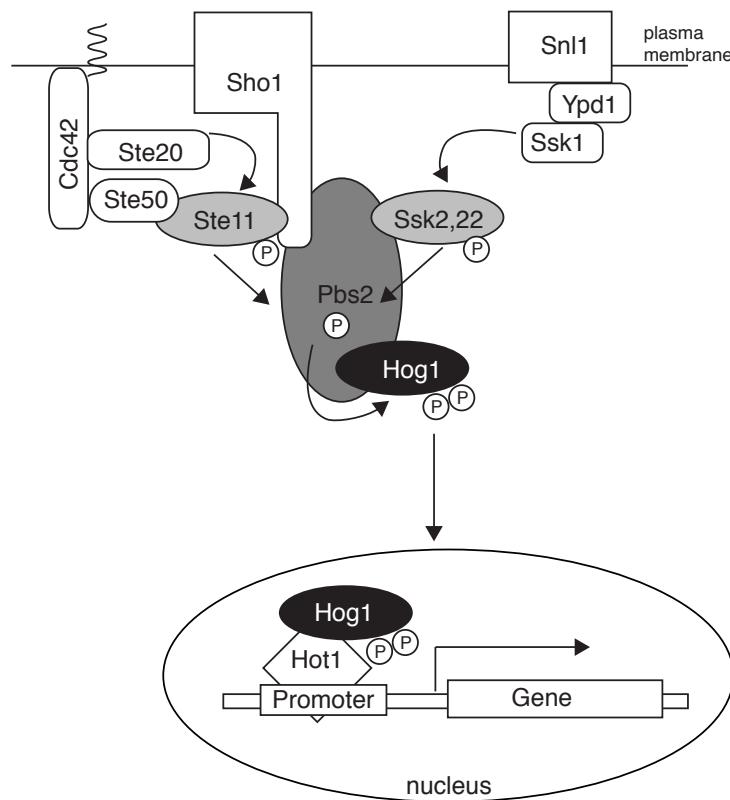
| Position of nucleus	Expression of stress proteins
Abundance and location of phosphorylated Hog1 protein	Overlay of all channels
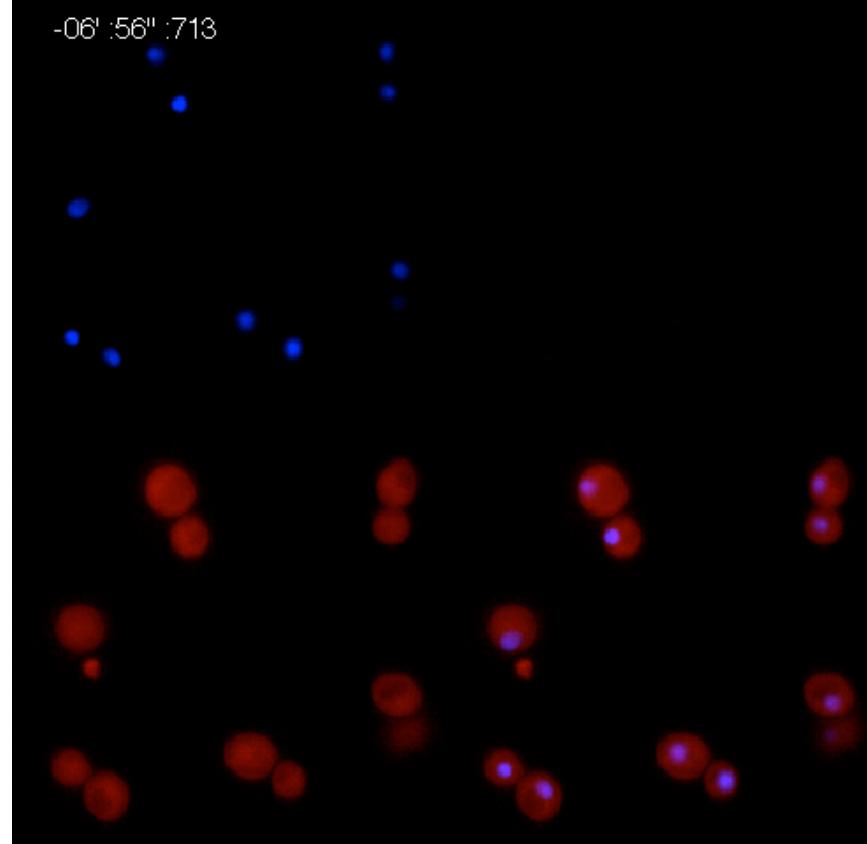
Observables: total HOG proteins and nuclear HOG proteins
(and expression of reporter gene)

enzymatic reactions

Ptp3

NaCl

p p

Hog1

Hog1

p

Ptp3

Nuclear envelope

Hot1

STL1

Ptp2

p p

Hog1

Hog1

p p

Hot1 | Hog1

GPD1

c_1

p p

Hot1 | Hog1

GPD1

$+$

Gpd1

RSC

RSC | p p

Hot1 | Hog1

GPD1

c_2

RSC | p p

Hot1 | Hog1

GPD1

$+$

Gpd1
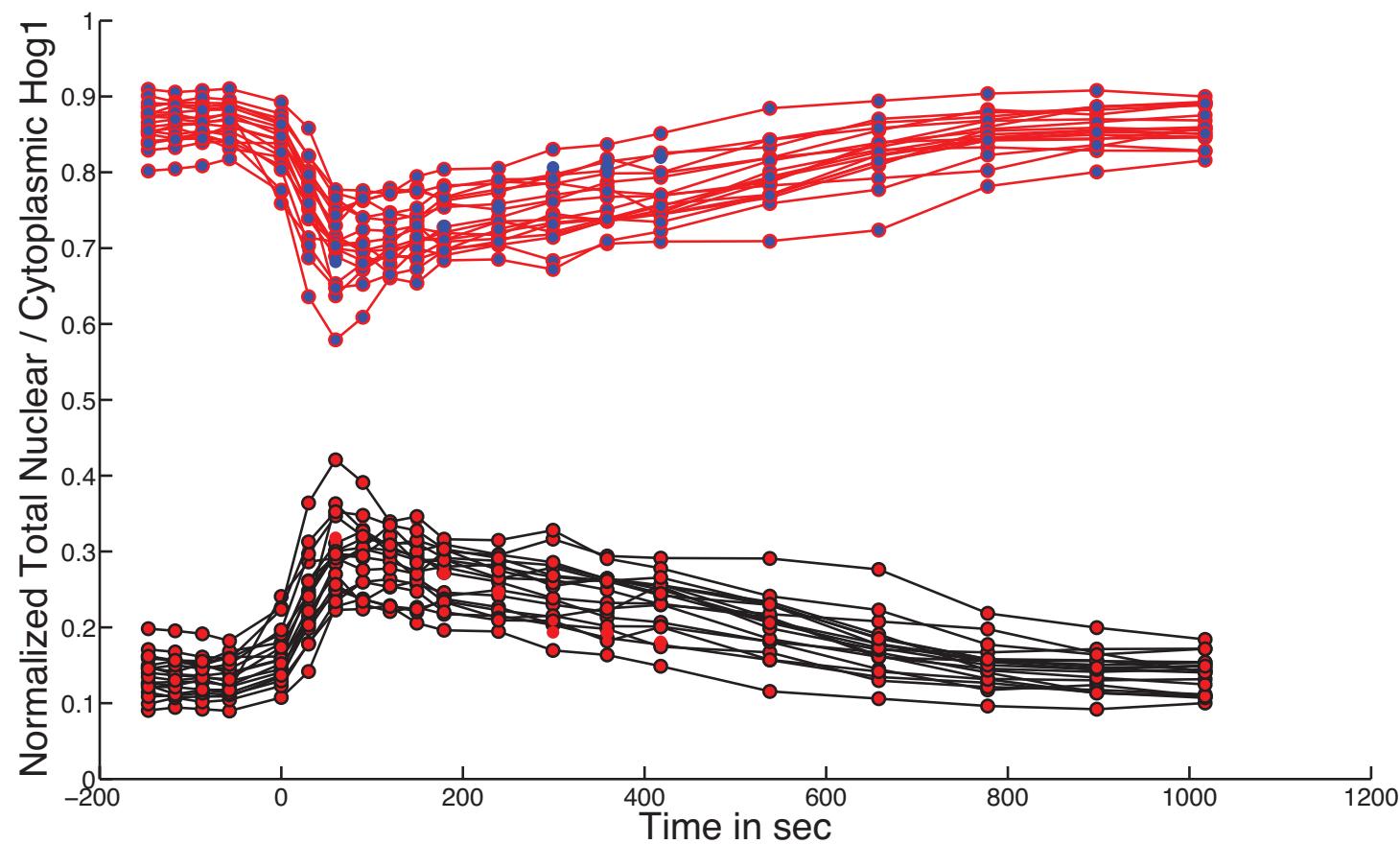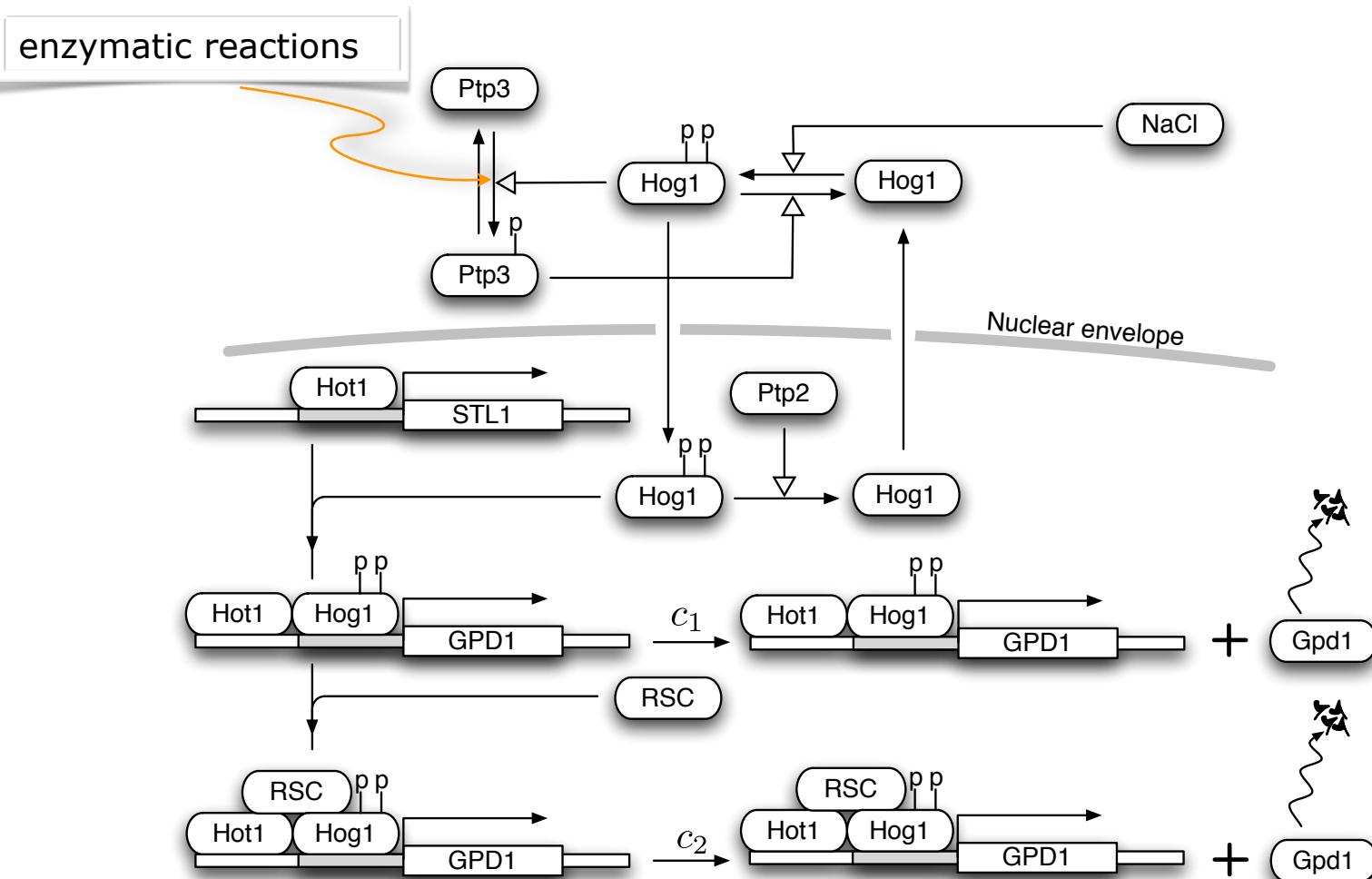
...11 states, 16 kinetic rate constants (with log-normal priors)

Estimation of all rate constants
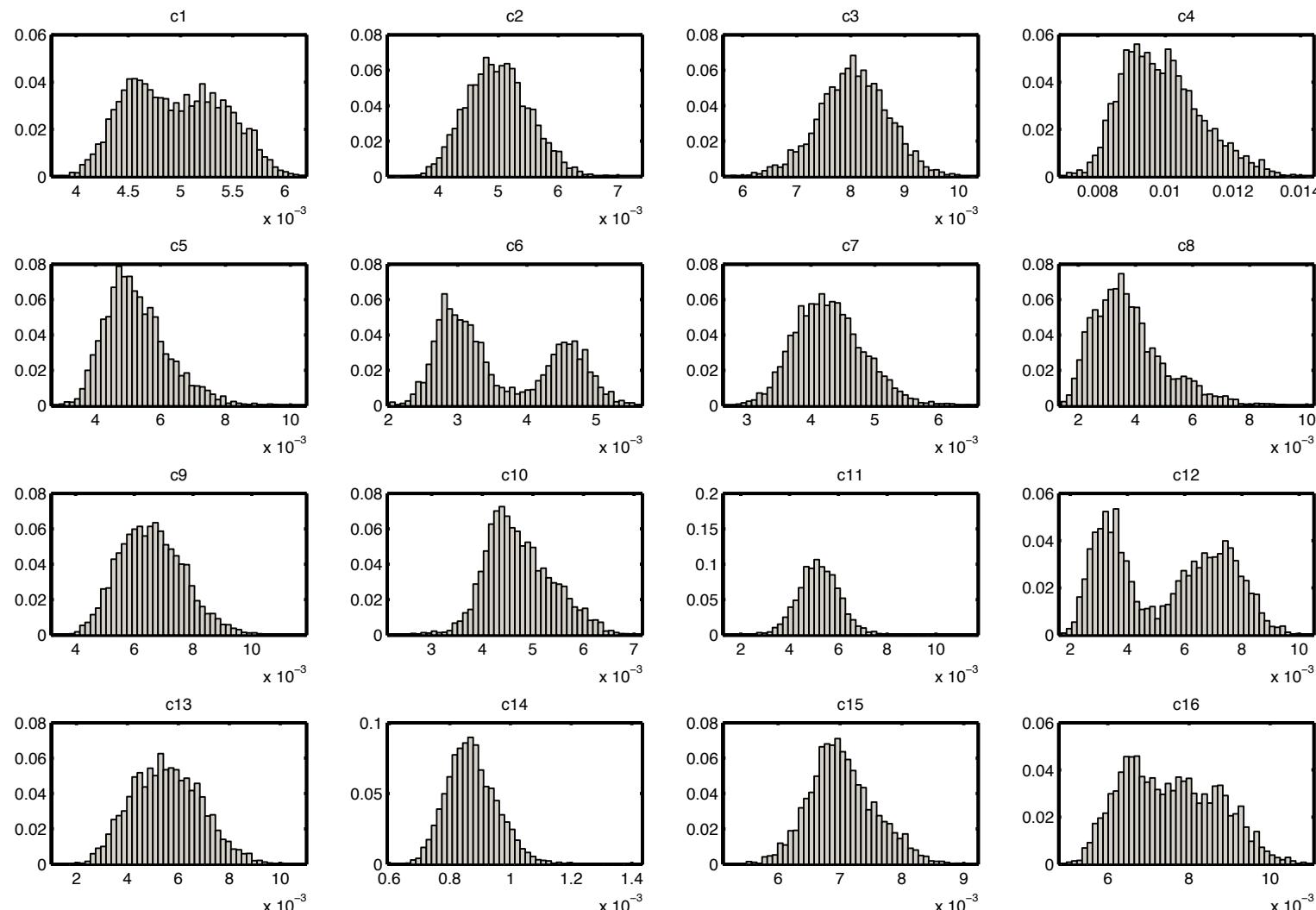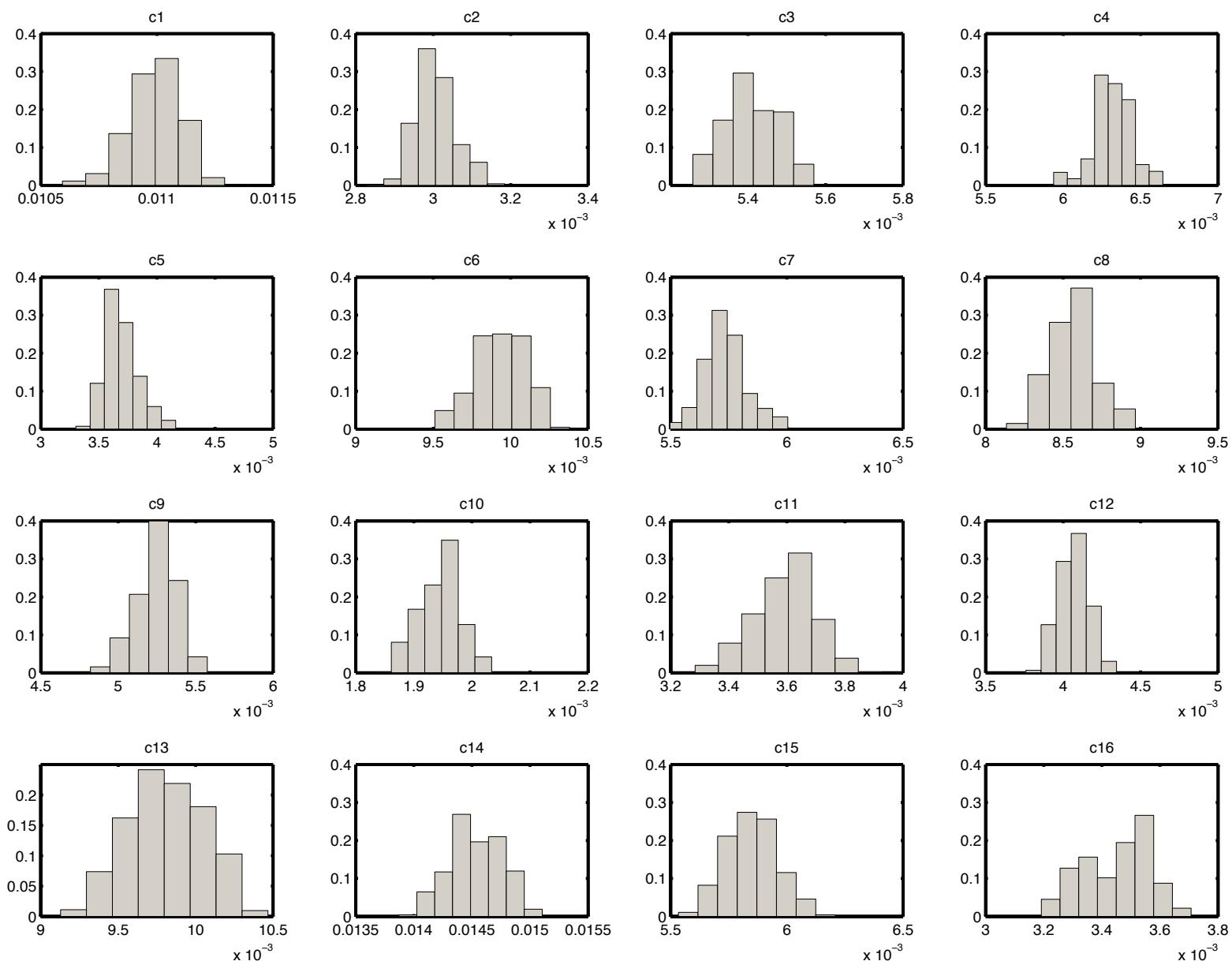
Marginal posterior: Phosed Hog1 translocation rate

Multiple-cell estimate

Single-cell estimate

$$x\,10^{-3}sec^{-1}$$

cell-population

$p(b)$

σ

μ

$\log(b)$

total copy

σ

μ
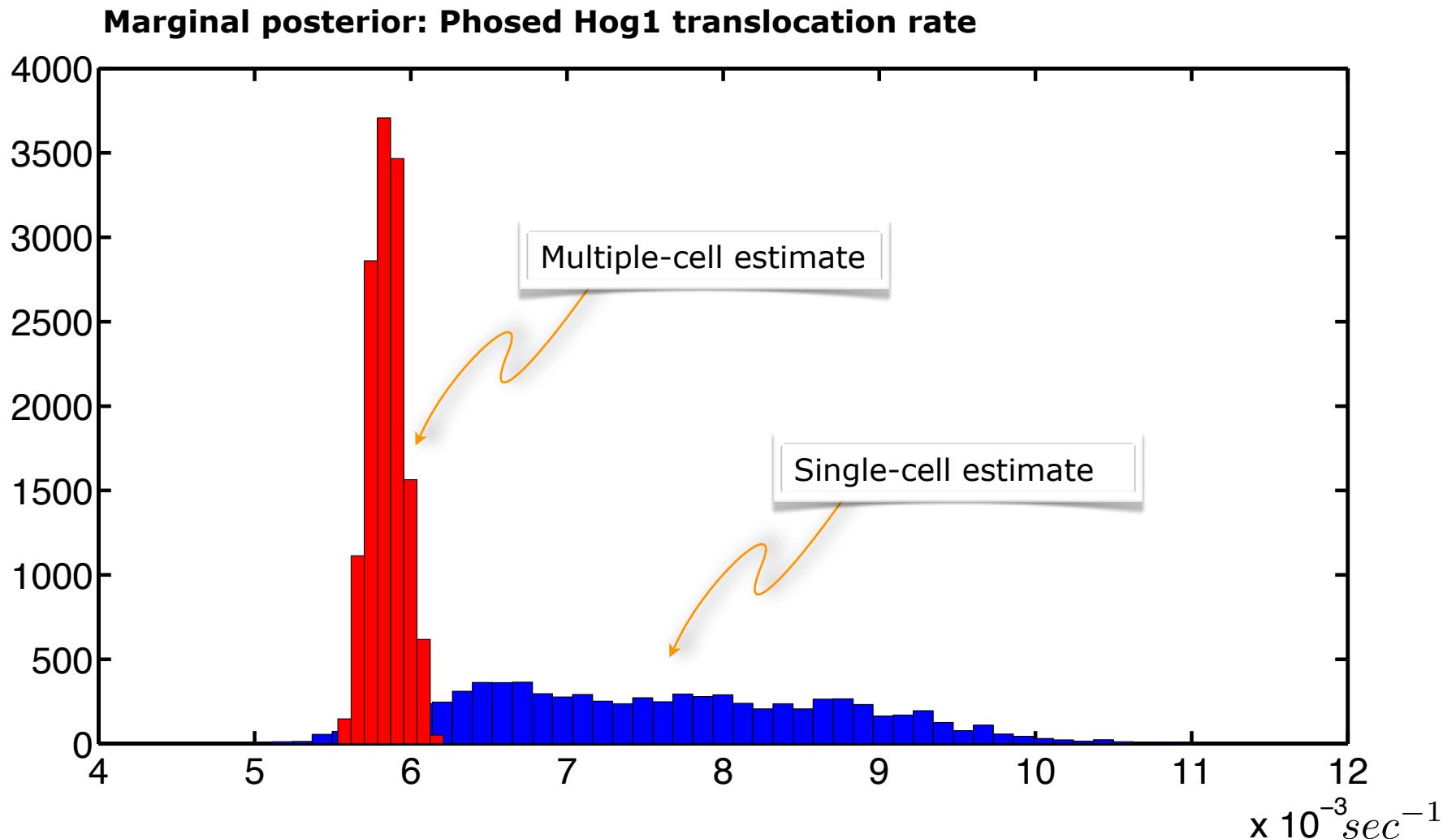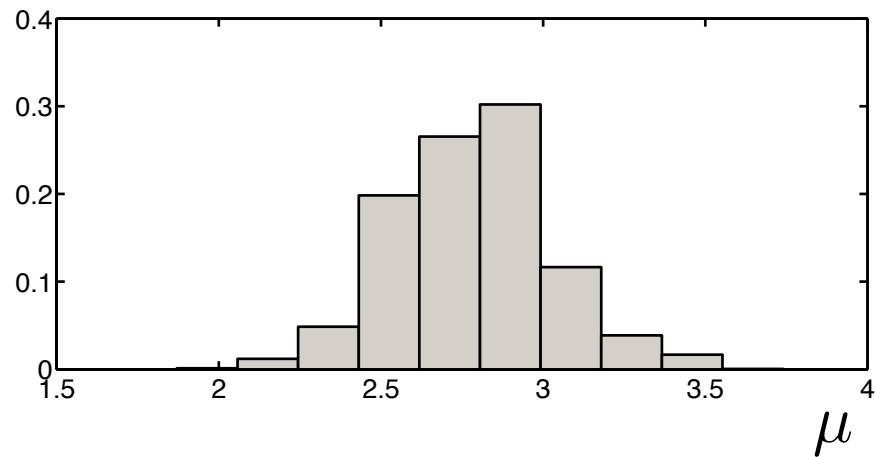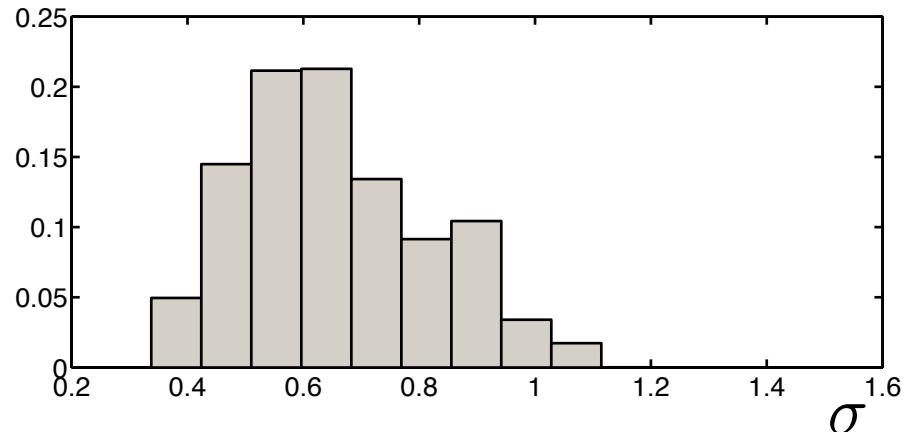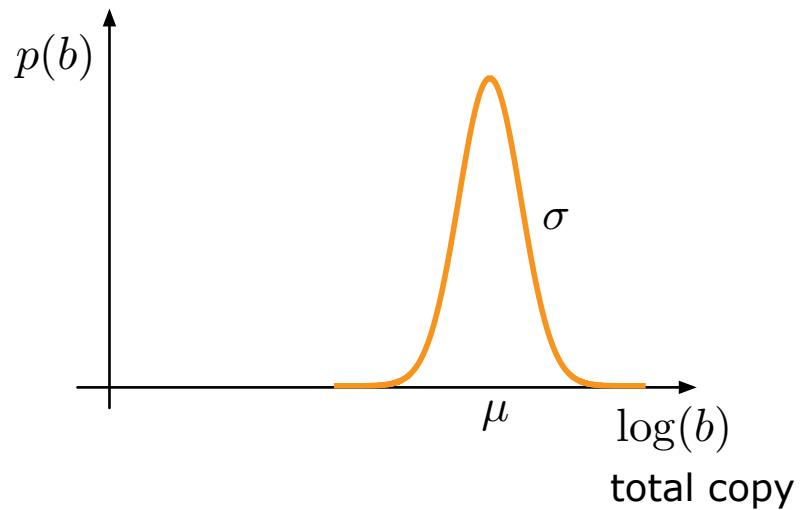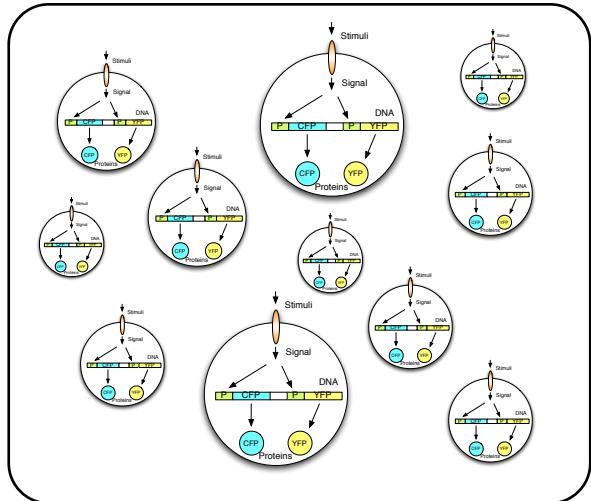
■ Artificial homogenization of cell-population - remaining variability intrinsic noise/stochastic kinetics.
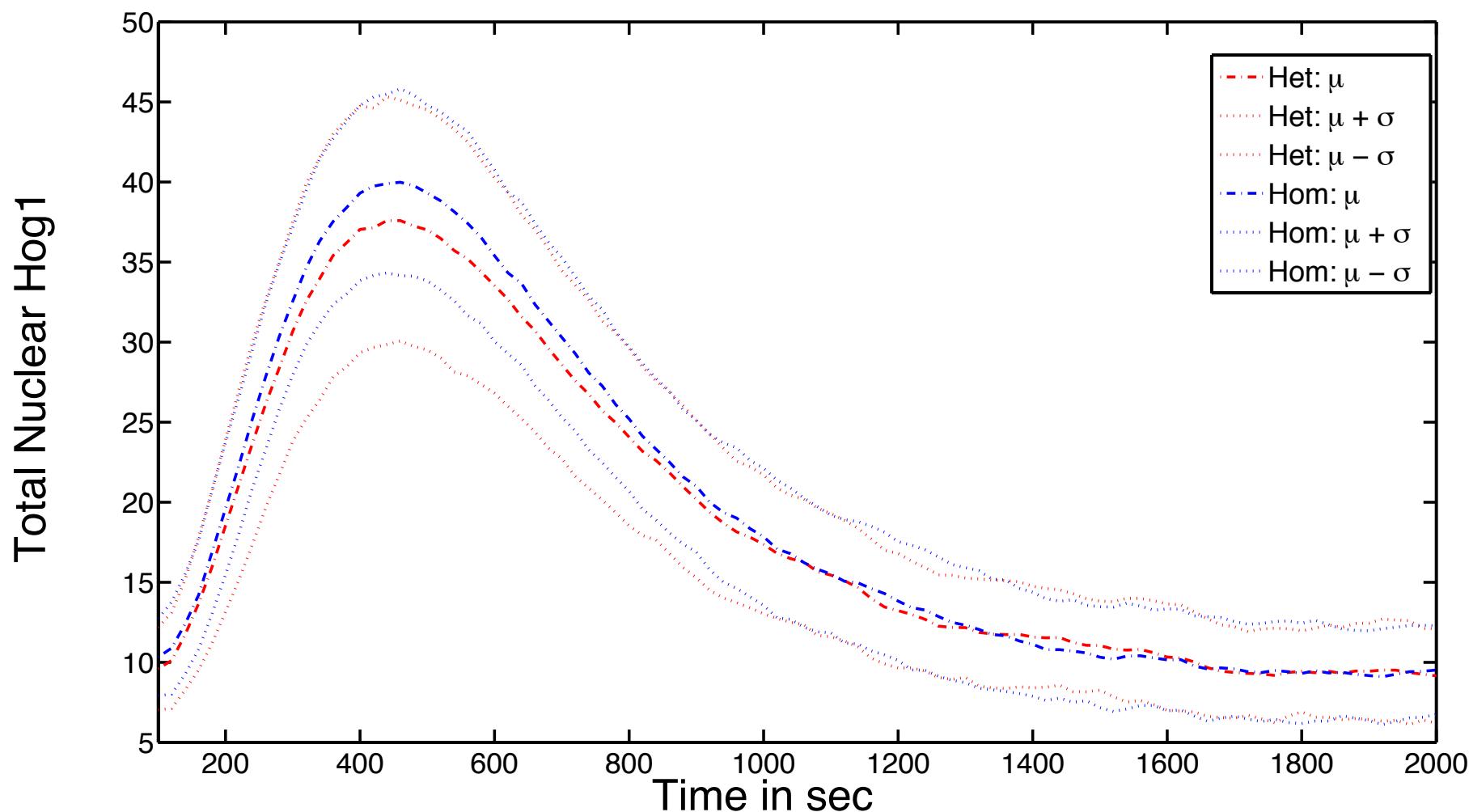
■ Single-cell technologies reveal **large cell-to-cell variability** - what are the important contributions to this variability?

■ Complementary approach to capture cell-to-cell variability - variability in total **protein count or concentration**.

■ Approach to estimate this contribution to the variability - allows **quantifying intrinsic component** (molecular noise).

■ Applied to real-world data and situation - **low dimensional readout** and high measurement noise.

■ ToDo: Better proposal density for SMC - **conditional Gillespie**.

■ Arbitrary stress profiles - microfluidics

DA-MF / 001a / 2011-04-12
unger@control.ee.ethz.ch