Estimating Parameters and Hidden Variables in a Nonlinear State-space Model of Biological Networks

Minh Quach, Nicolas Brunel and Florence d'Alché-Buc

IBISC FRE CNRS 2873
University of Evry and Genopole, France

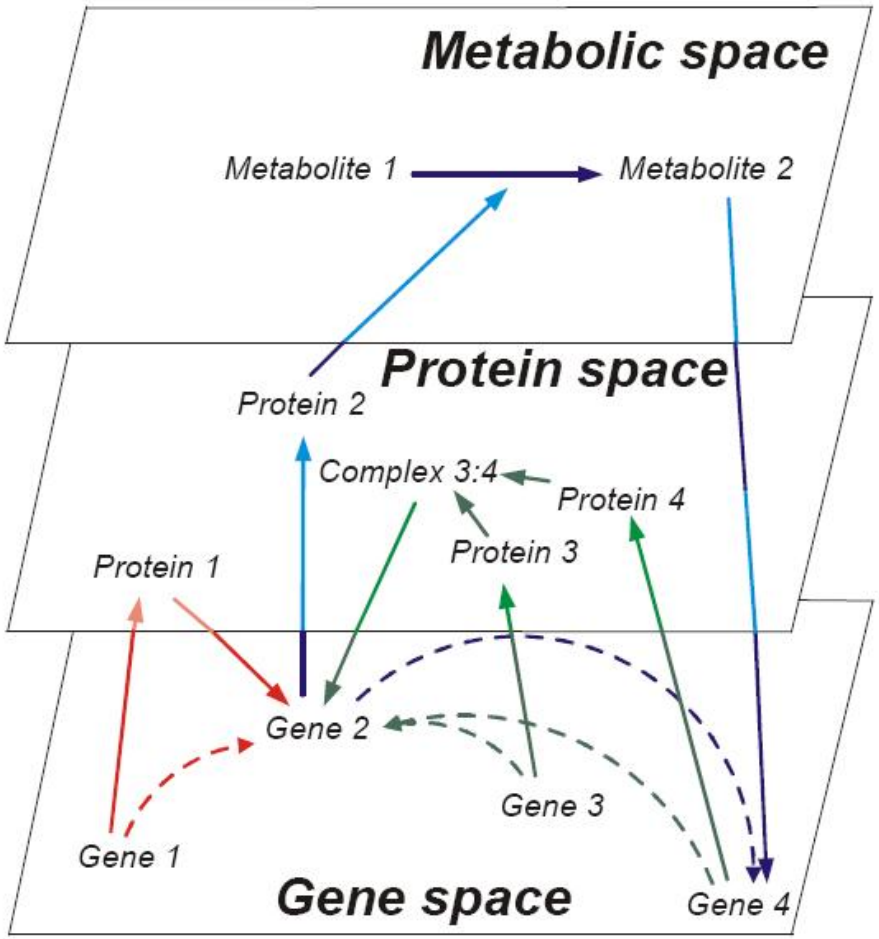
28th March 07 / PESB workshop

Outline

Outline

Biological networks

- Multi-scale
  - Regulatory networks
  - Metabolic networks
  - Signaling pathways
- Mathematical model
  - stochastic nature
  - dynamical systems
- Challenge
  - nonlinear
  - partially observed



Quantitative models of Biological Networks

System of ODE's

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t); \theta)$$

- $\mathbf{x}(t)$: state variables at time t
  - protein concentrations
  - mRNA concentrations
  - metabolite concentrations
- $\mathbf{f}$: encodes the structure of the system
  - nonlinear function
  - Michaelis-Menten kinetics
  - Mass action kinetics
  - ...
- θ: parameter set (kinetic parameters, rate constants,...)
- $\mathbf{u}(t)$: input variables at time t

Reverse Engineering of Biological Networks

Given

- An ODE model :

$$\frac{d\mathbf{x}(\mathbf{t})}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t); \theta)$$

- A partially and noisy observation model:

$$\mathbf{y}(t) = \mathbf{H}(\mathbf{x}(t), \mathbf{u}(t); \theta) + \epsilon(t)$$

  where $\mathbf{H}$ is a nonlinear observation function, $\epsilon(t)$ is a i.i.d noise
- A sequence of observed data : $\mathbf{y}_{1:K} = \{\mathbf{y}_1, ..., \mathbf{y}_K\}$ at time $t_1, t_2, ..., t_k$

Goal

- Estimation of parameters θ
- Estimation of states $\mathbf{x}(t)$

Reverse Engineering of Biological Networks

Given

- An ODE model :

$$\frac{d\mathbf{x}(\mathbf{t})}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t); \theta)$$

- A partially and noisy observation model:

$$\mathbf{y}(t) = \mathbf{H}(\mathbf{x}(t), \mathbf{u}(t); \theta) + \epsilon(t)$$

  where $\mathbf{H}$ is a nonlinear observation function, $\epsilon(t)$ is a i.i.d noise
- A sequence of observed data : $\mathbf{y}_{1:K} = \{\mathbf{y}_1, ..., \mathbf{y}_K\}$ at time $t_1, t_2, ..., t_k$

Goal

- Estimation of parameters θ
- Estimation of states $\mathbf{x}(t)$

Outline

Nonlinear State-Space Model

Continuous time ODE model

$$\begin{aligned} \frac{d\mathbf{x}(\mathbf{t})}{dt} &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t); \theta) \\ \mathbf{y}(t) &= \mathbf{H}(\mathbf{x}(t), \mathbf{u}(t); \theta) + \epsilon(t) \end{aligned}$$

The corresponding discrete-time state-space model

The system at discrete-time points $t_1, ..., t_K$

$$\begin{aligned} \mathbf{x}(t_{k+1}) &= \mathbf{F}(\mathbf{x}(t_k), \mathbf{u}; \theta) \\ \mathbf{y}(t_k) &= \mathbf{H}(\mathbf{x}(t_k), \mathbf{u}(t_k); \theta) + \epsilon(t_k) \end{aligned}$$

with

$$\mathbf{F}(\mathbf{x}(t_k), \mathbf{u}; \theta) = \mathbf{x}(t_k) + \int_{t_k}^{t_{k+1}} \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau); \theta) d\tau$$

Nonlinear State-Space Model

Continuous time ODE model

$$\begin{aligned} \frac{d\mathbf{x}(\mathbf{t})}{dt} &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t); \theta) \\ \mathbf{y}(t) &= \mathbf{H}(\mathbf{x}(t), \mathbf{u}(t); \theta) + \epsilon(t) \end{aligned}$$

The corresponding discrete-time state-space model

The system at discrete-time points $t_1, ..., t_K$

$$\begin{aligned} \mathbf{x}(t_{k+1}) &= \mathbf{F}(\mathbf{x}(t_k), \mathbf{u}; \theta) \\ \mathbf{y}(t_k) &= \mathbf{H}(\mathbf{x}(t_k), \mathbf{u}(t_k); \theta) + \epsilon(t_k) \end{aligned}$$

with

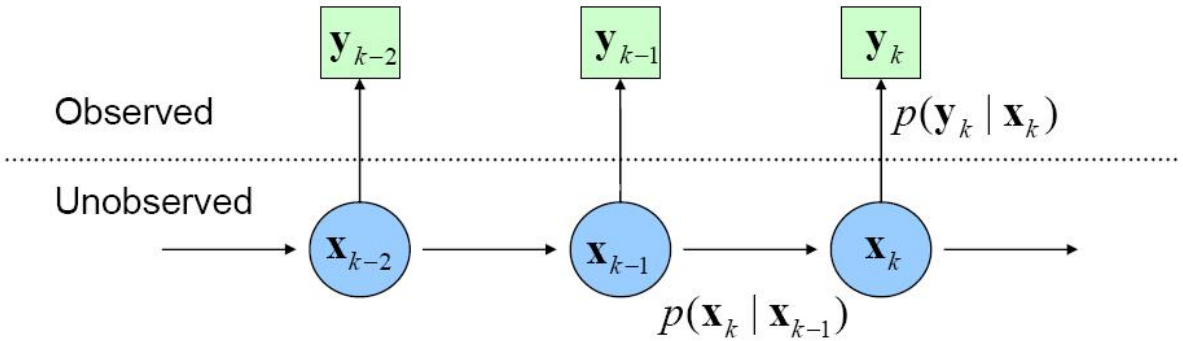
$$\mathbf{F}(\mathbf{x}(t_k), \mathbf{u}; \theta) = \mathbf{x}(t_k) + \int_{t_k}^{t_{k+1}} \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau); \theta) d\tau$$

Outline

Bayesian inference

Given:

- Prior distribution over the initial state and parameters: $p(\mathbf{x}_1, \theta)$
- A transition model: $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \theta)$
- An observation model: $p(\mathbf{y}_k|\mathbf{x}_k, \theta)$
- A sequence of observations: $\mathbf{y}_{1:K} = \{\mathbf{y}_1, ..., \mathbf{y}_K\}$



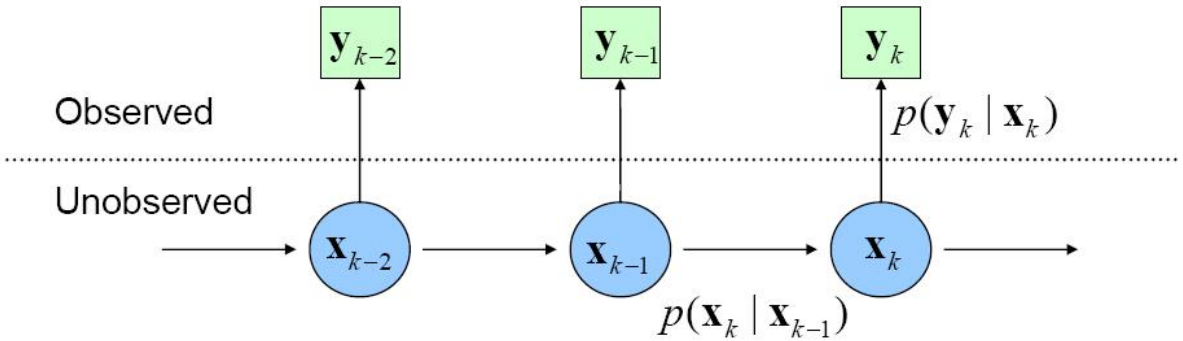
Estimating the posterior distributions

- The filtering distribution: $p(\mathbf{x}_k, \theta|\mathbf{y}_{1:k})$
- The smoothing distribution: $p(\mathbf{x}_k, \theta|\mathbf{y}_{1:K})$

Bayesian inference

Given:

- Prior distribution over the initial state and parameters: $p(\mathbf{x}_1, \theta)$
- A transition model: $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \theta)$
- An observation model: $p(\mathbf{y}_k|\mathbf{x}_k, \theta)$
- A sequence of observations: $\mathbf{y}_{1:K} = \{\mathbf{y}_1, ..., \mathbf{y}_K\}$



Estimating the posterior distributions

- The filtering distribution: $p(\mathbf{x}_k, \theta|\mathbf{y}_{1:k})$
- The smoothing distribution: $p(\mathbf{x}_k, \theta|\mathbf{y}_{1:K})$

Recursive Bayesian Filtering

Suppose that θ is known, recursively calculate the filtering distribution of the states $p(\mathbf{x}_k|\mathbf{y}_{1:k})$

Two steps

1. Prediction: $p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k}) = \int p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k})d\mathbf{x}_k$
2. Update:

$$p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k+1}) = \frac{p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})}{p(\mathbf{y}_{k+1}|\mathbf{y}_{1:k})}$$

where:

$$p(\mathbf{y}_{k+1}|\mathbf{y}_{1:k}) = \int p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})d\mathbf{x}_{k+1}$$

- Analytical solution obtained only when **F**, **H** are linear and $p(\mathbf{x}_1)$ and ϵ are Gaussian $\rightarrow$ Kalman Filter
- When **F**, **H** are nonlinear, the integrals are usually intractable. Approximate solutions are needed!

Recursive Bayesian Filtering

Suppose that θ is known, recursively calculate the filtering distribution of the states $p(\mathbf{x}_k|\mathbf{y}_{1:k})$

Two steps

1. Prediction: $p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k}) = \int p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k})d\mathbf{x}_k$
2. Update:

$$p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k+1}) = \frac{p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})}{p(\mathbf{y}_{k+1}|\mathbf{y}_{1:k})}$$

where:

$$p(\mathbf{y}_{k+1}|\mathbf{y}_{1:k}) = \int p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})d\mathbf{x}_{k+1}$$

- Analytical solution obtained only when **F**, **H** are linear and $p(\mathbf{x}_1)$ and ϵ are Gaussian $\rightarrow$ Kalman Filter
- When **F**, **H** are nonlinear, the integrals are usually intractable. Approximate solutions are needed!

Recursive Bayesian Filtering

Suppose that θ is known, recursively calculate the filtering distribution of the states $p(\mathbf{x}_k|\mathbf{y}_{1:k})$

Two steps

1. Prediction: $p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k}) = \int p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k})d\mathbf{x}_k$
2. Update:

$$p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k+1}) = \frac{p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})}{p(\mathbf{y}_{k+1}|\mathbf{y}_{1:k})}$$

where:

$$p(\mathbf{y}_{k+1}|\mathbf{y}_{1:k}) = \int p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k})d\mathbf{x}_{k+1}$$

- Analytical solution obtained only when **F**, **H** are linear and $p(\mathbf{x}_1)$ and ϵ are Gaussian $\rightarrow$ Kalman Filter
- When **F**, **H** are nonlinear, the integrals are usually intractable. Approximate solutions are needed!

Nonlinear SSM → Approximate Solutions

- Gaussian Approximations:
  - Extended Kalman Filter [Jazwinski 1970]
  - Unscented Kalman Filter [Julier and Uhlmann 1995-2000]
- Sequential Monte Carlo Methods [Gordon 1996, Doucet 1998]
  - Particle filters
- Variational Methods [Ghahramani 1999, Valpola 2002]

Gaussian Approximations

Basic problem: Nonlinear transformation of a random variable:

$$\mathbf{y} = \mathbf{F}(\mathbf{x})$$

- Given:

$$\overline{\mathbf{x}} = E(\mathbf{x}) \qquad \mathbf{P}_{\mathbf{x}} = E\left[(\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^{\top}\right]$$

- Find:

$$\overline{\mathbf{y}} = E(\mathbf{y}) \qquad \mathbf{P}_{\mathbf{y}} = E\left[(\mathbf{y} - \overline{\mathbf{y}})(\mathbf{y} - \overline{\mathbf{y}})^{\top}\right]$$

Nonlinear transformation

Table: EKF vs UKF

EKF	UKF
Linearize $\mathbf{F}$: $\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{x}}$	Compute sigma-points: $\{\mathcal{X}_i\} = \{\bar{\mathbf{x}} + \gamma\sqrt{\mathbf{P_x}} \quad \bar{\mathbf{x}} - \gamma\sqrt{\mathbf{P_x}}\}$
$\bar{\mathbf{y}} = \mathbf{F}(\bar{\mathbf{x}}) \qquad \mathbf{P_y} = \mathbf{A}^\top \mathbf{P_x} \mathbf{A}$	Transform sigma-points: $\mathcal{Y}_i = \mathbf{F}(\mathcal{X}_i)$
	Reconstruct posterior statistics: $\bar{\mathbf{y}} = \sum_i \alpha_i \mathcal{Y}_i \quad \mathbf{P_y} = \sum_i \alpha_i (\mathcal{Y}_i - \bar{\mathbf{y}})(\mathcal{Y}_i - \bar{\mathbf{y}})^\top$

Unscented Kalman Filter

- Deterministic sampling method, number of sigma points is small $\rightarrow$ fast
- No need to calculate derivatives (Jacobians, Hessians, etc.)
- Exact to 2nd order of Taylor series expansion for both mean and covariance.
- Can be extended to capture higher-order statistics (skew, kurtosis, etc.)

Parameter Estimation

- Augmented state approach

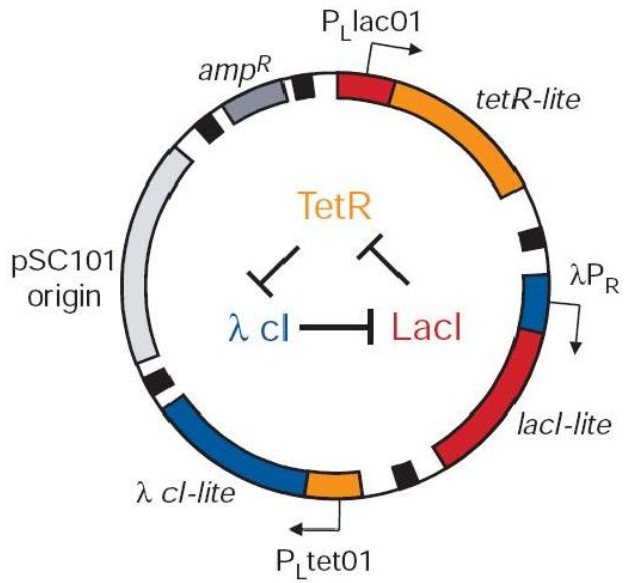
$$\begin{aligned} \theta_{k+1} &= \theta_k \\ \mathbf{x}(t_{k+1}) &= \mathbf{F}(\mathbf{x}(t_k), \mathbf{u}; \theta_k) \\ \mathbf{y}(t_k) &= \mathbf{H}(\mathbf{x}(t_k), \mathbf{u}(t_k); \theta_k) + \epsilon(t_k) \end{aligned}$$

- Joint state and parameter estimation

Outline

Repressilator

[Elowitz, Nature 2000]



$$\frac{dr_1}{dt} = v_1^{max}\frac{k_{12}^n}{k_{12}^n + p_2^n} - k_1^{mRNA} r_1$$

$$\frac{dr_2}{dt} = v_2^{max}\frac{k_{23}^n}{k_{23}^n + p_3^n} - k_2^{mRNA} r_2$$

$$\frac{dr_3}{dt} = v_3^{max}\frac{k_{31}^n}{k_{31}^n + p_1^n} - k_3^{mRNA} r_3$$

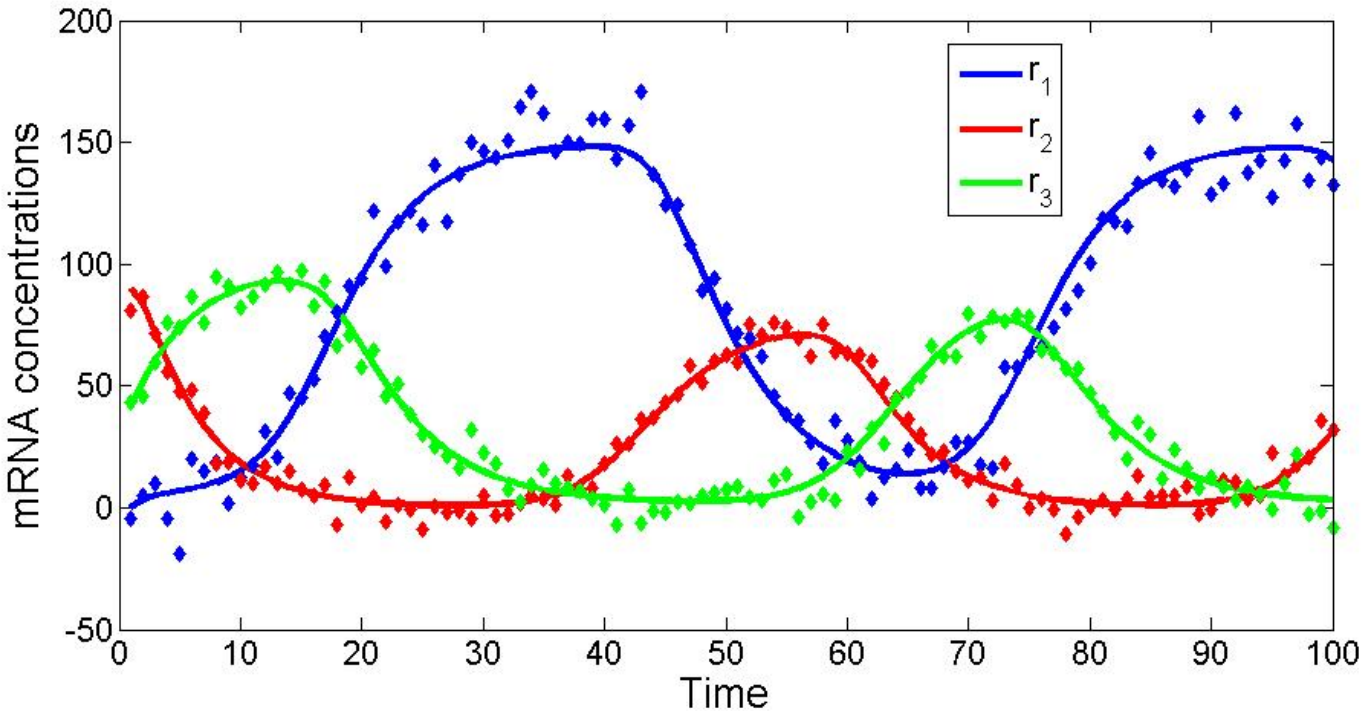
$$\frac{dp_1}{dt} = k_1 r_1 - k_1^{protein} p_1$$

$$\frac{dp_2}{dt} = k_2 r_2 - k_2^{protein} p_2$$

$$\frac{dp_3}{dt} = k_3 r_3 - k_3^{protein} p_3$$

- mRNAs are observed, proteins are hidden
- mRNA and protein degradation rate constants are supposed to be known
- Estimate 9 parameters

Synthetic data



Parameter Estimation

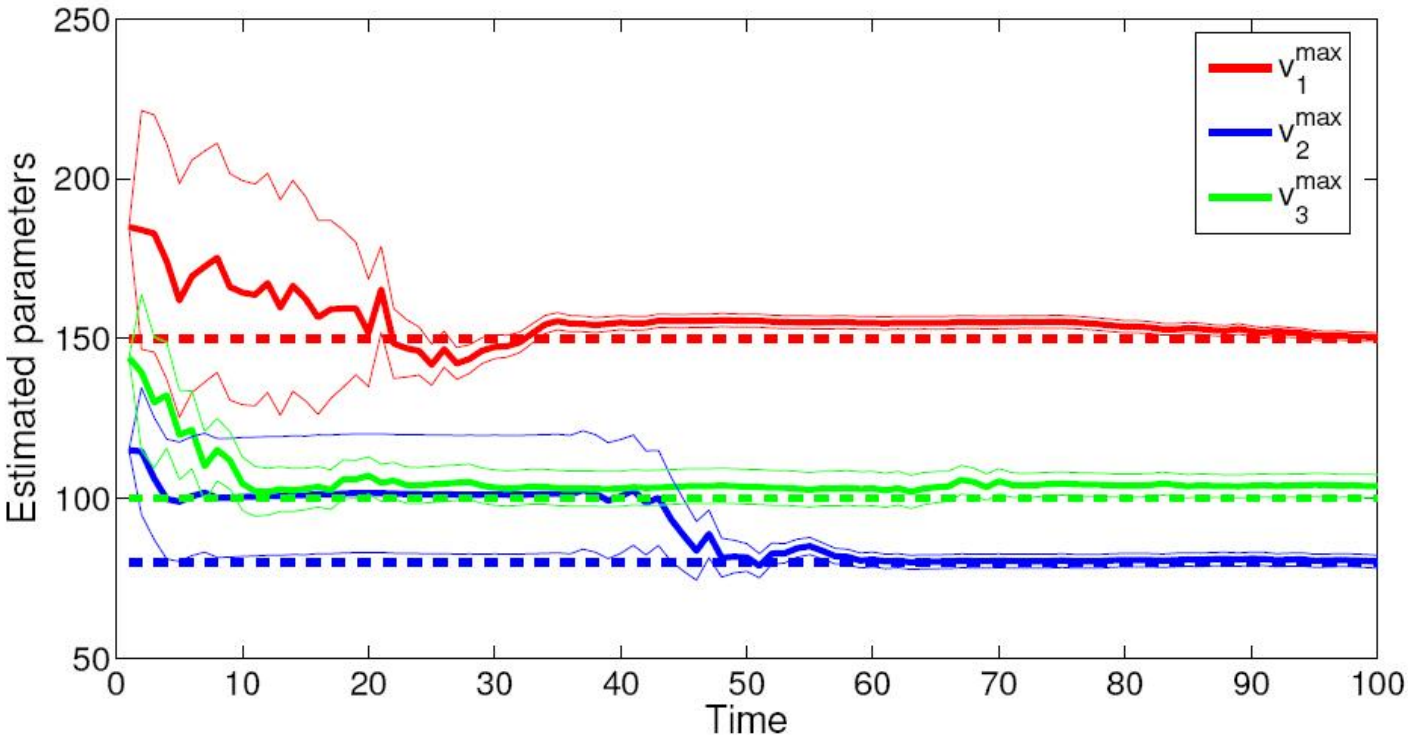


Fig. 4. Recursive estimation of the maximal rate of Michaelis-Menten kinetics through time for the case $S = 1$ and sampling time $\Delta_t = 0.2$ (corresponds to 100 data points). Dash lines: true parameters. Solid lines: Estimated parameters along with their confidence intervals

Parameter Estimation

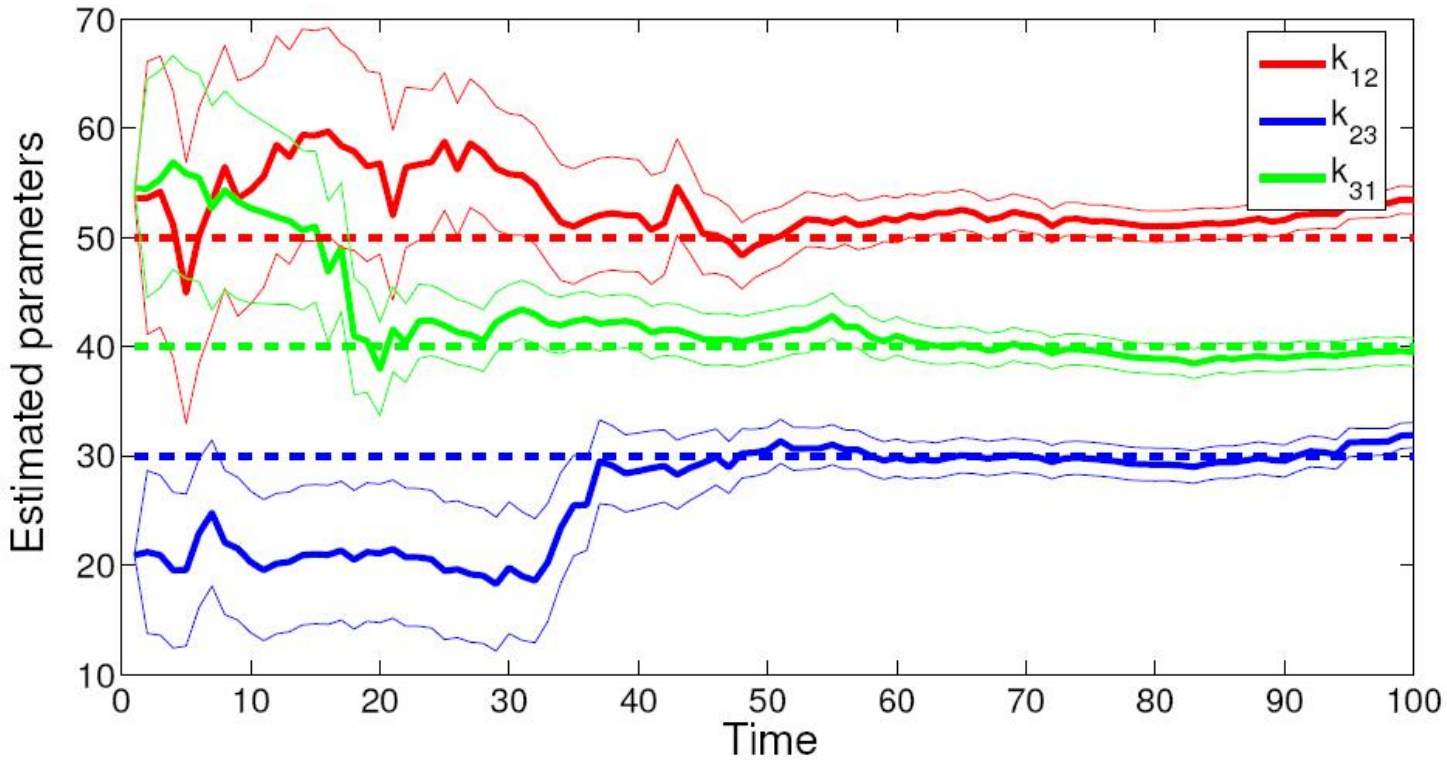


Fig. 5. Recursive estimation of Michaelis constants k_{12}, k_{23}, k_{31} through time.

State Estimation

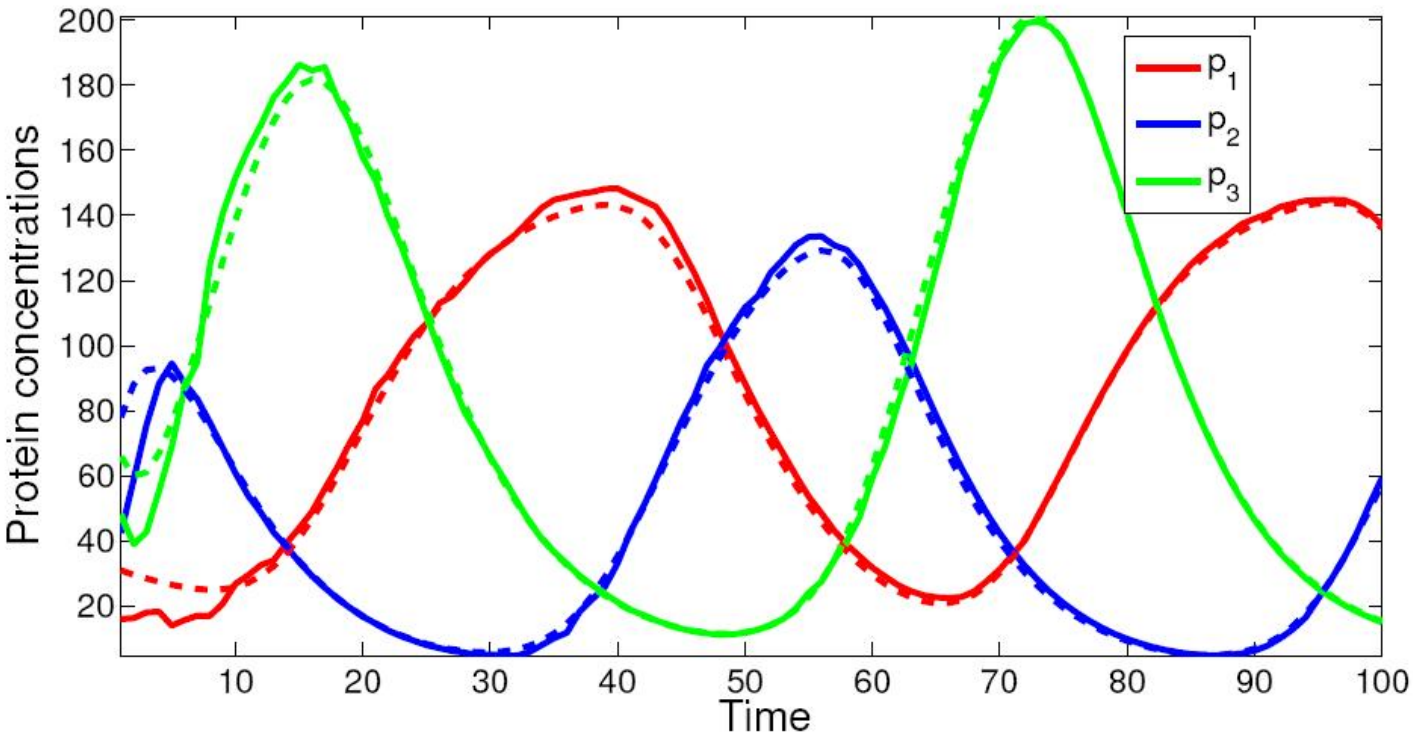
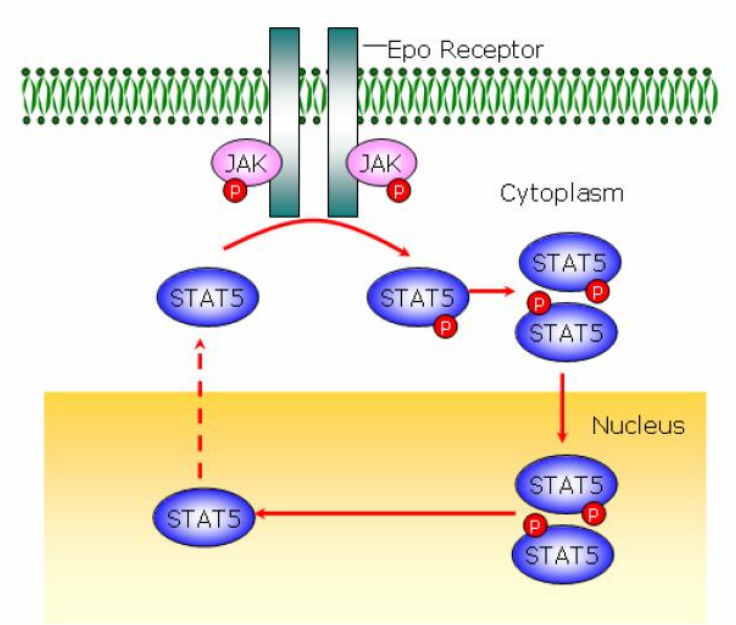


Fig. 3. The evolution of the true (dashed) and estimated (solid) protein concentrations.

Outline

JAK-STAT signaling pathway

[Swameye, PNAS 2003]



- ODE:

$$\begin{cases} \dot{x}_1(t) &= -a_1 x_1(t)u(t) + 2a_4 x_4(t)1_{\{t\geq\tau\}} \\ \dot{x}_2(t) &= a_1 x_1(t)u(t) - 2a_4 x_2^2(t) \\ \dot{x}_3(t) &= -a_3 x_3(t) + x_2^2(t) \\ \dot{x}_4(t) &= a_3 x_3(t) - a_4 x_4(t)1_{\{t\geq\tau\}} \end{cases}$$

- Observed variables

$$\begin{aligned} y_1 &= x_2 + 2x_3 \\ y_2 &= x_1 + x_2 + 2x_3 \end{aligned}$$

- Experimental data: 16 time points
- $\theta = (a_1, a_3, a_4)^\top$ is the parameters to be estimated

Prediction vs Experimental data

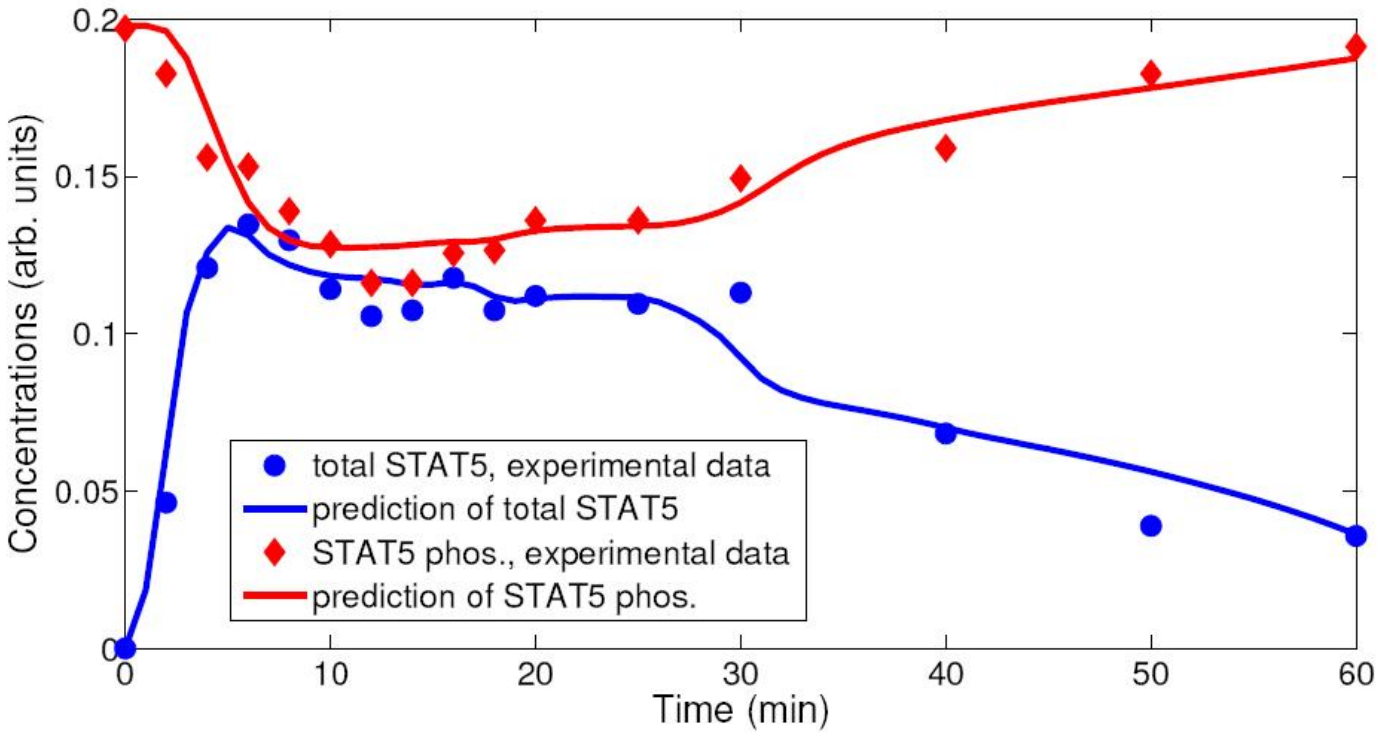


Fig. 12. Prediction of STAT5 phosphorylation and total amount of STAT5.

Conclusion

- Conclusion
  - A general framework based on nonlinear state-space models for describing biological networks
  - Bayesian inference based on UKF for estimating parameters and hidden states from noisy and partially observed data
- Ongoing work
  - Unscented Kalman smoothing
  - Particle smoothing