

Reduction and Identification of Models: Basic Ideas and Applications

**Leonid Kalachev
University of Montana**

Together with Heikki Haario (Lappeenranta University of Technology)

Dynamic Models: in general terms, any models formulated in terms of time dependent differential or difference equations.

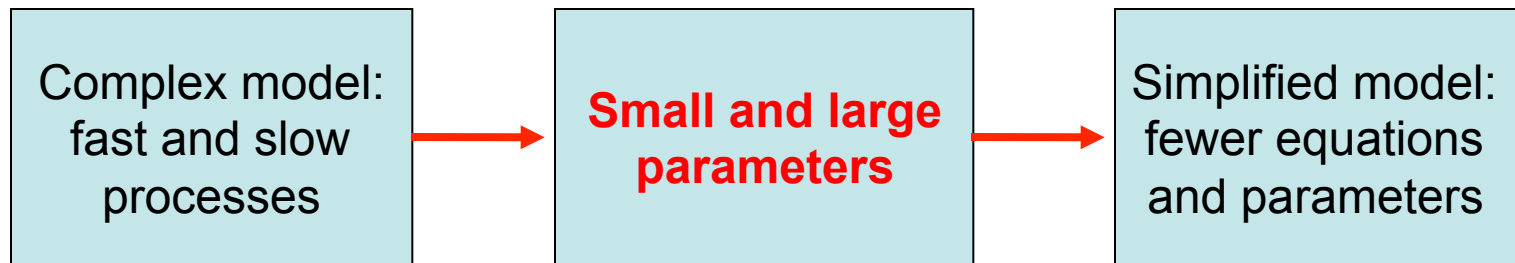
Modeling any process that may be schematically represented as a reaction scheme: chemical and bio-chemical kinetics, population dynamics, aggregation and growth/decay processes, etc.

Homogeneous models: species are well mixed, variables are only time dependent; no spatial dependence.

Heterogeneous models: variables are time dependent and also spatially dependent.

WHY Reduction and Identification?

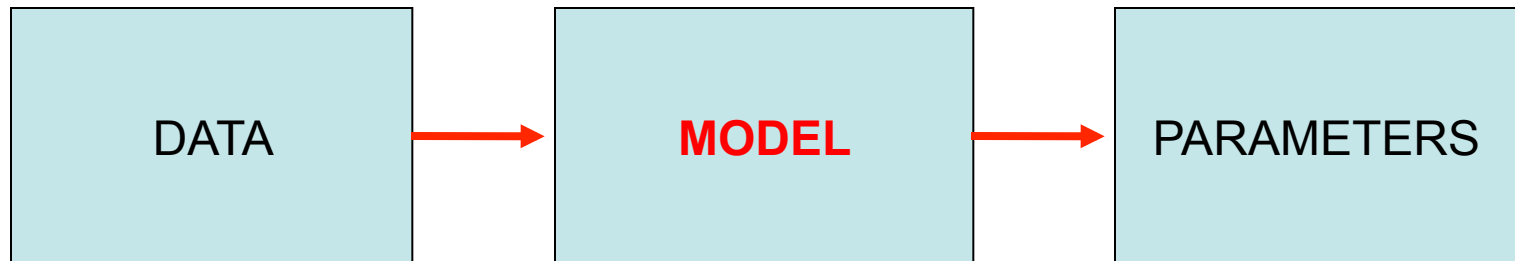
1. Reduction: parameters in a complex model are known; some parameters are small or large; reduction may be performed to produce a simpler model that is easier to understand and interpret.



**Reduction algorithms and their justification
is the topic of Applied Perturbation Analysis!**

WHY Reduction and Identification?

2. Identification: given the data and a hypothetical model; model parameters must be estimated by fitting model solution to the data.



Where do the models come from?

WHY Reduction and Identification?

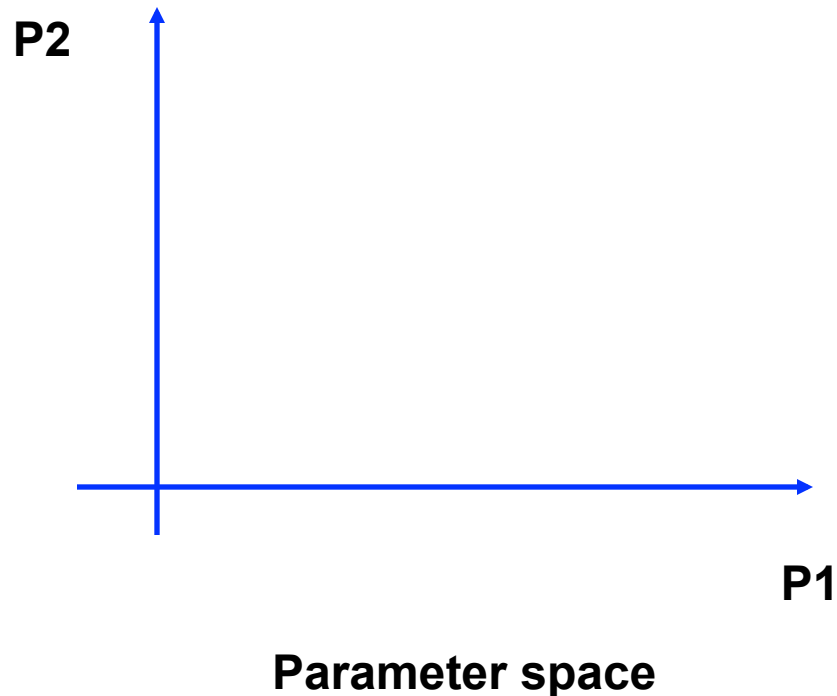
3. Reduction and Identification: How many meaningful models fit the data? What is the minimal number of parameters that needs to be included in a model?

An “optimal model” is derived from the basic principles (physics, chemistry, etc.) and contains the smallest possible number of parameters that still allows one to describe the experimental data.

How to get the “optimal model”: via repeated application of model reduction techniques together with analysis of reliability regions for model parameter values.

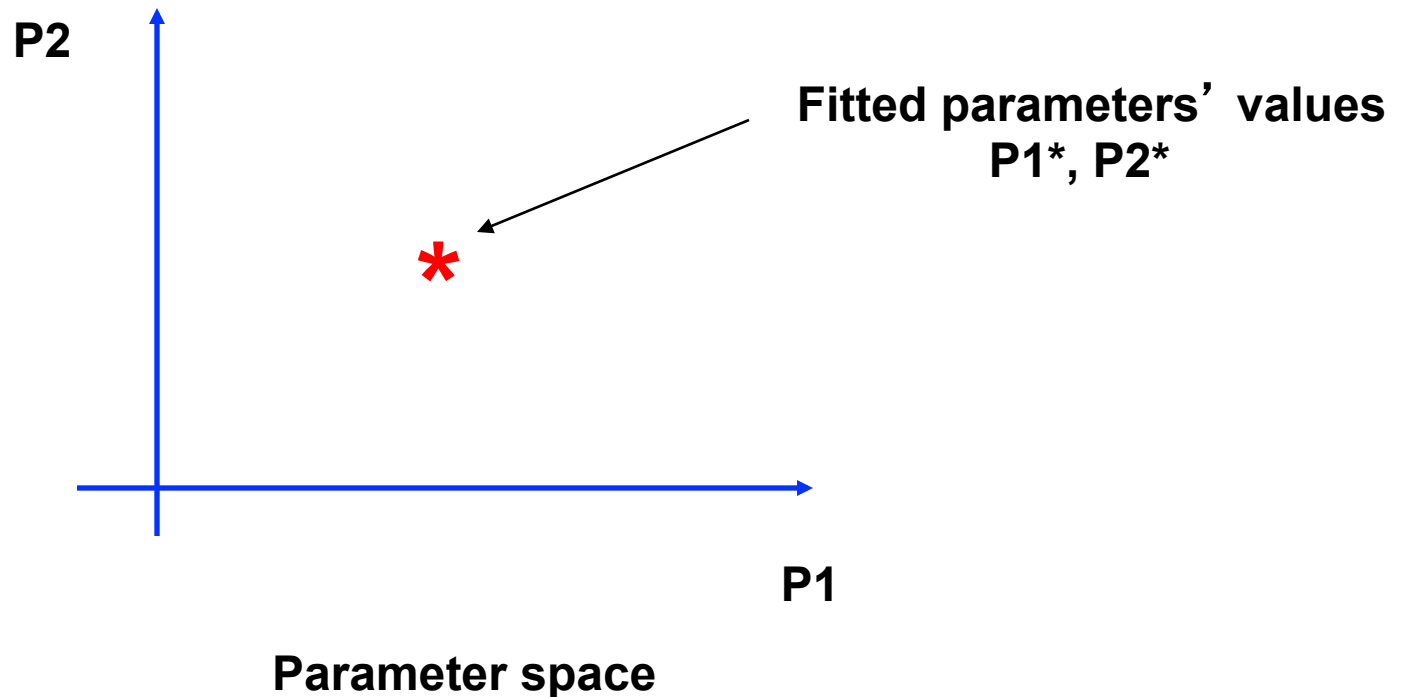
**Assume that we have some model that fits the data.
What is an indication of the need to use model reduction?**

**Combination of STATISTICS / APPLIED MATH APPROACHES:
BOOTSTRAPING & MCMC (Markov Chain Monte-Carlo)**



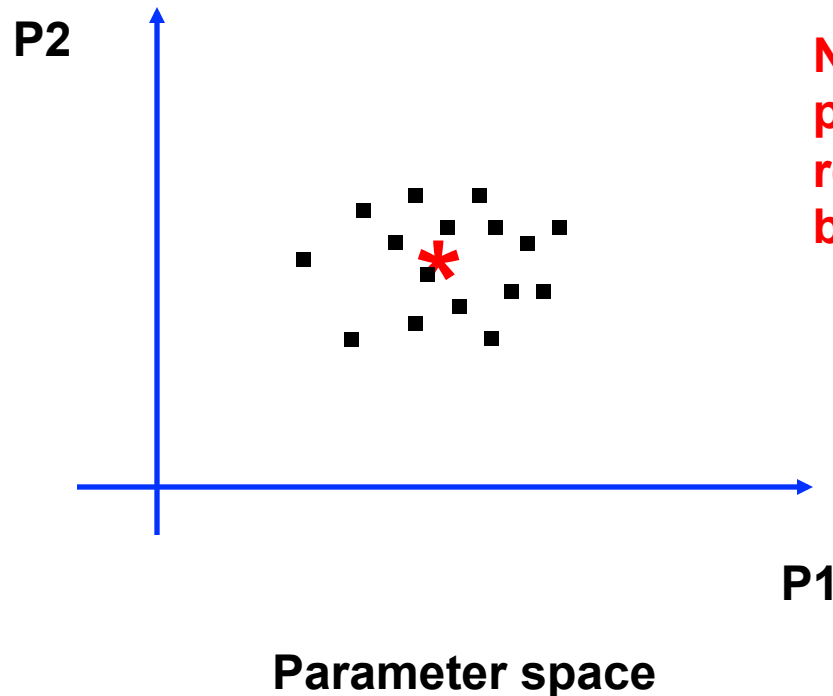
Assume that we have some model that fits the data.
What is an indication of the need to use model reduction?

**Combination of STATISTICS / APPLIED MATH APPROACHES:
BOOTSTRAPING & MCMC (Markov Chain Monte-Carlo)**



Assume that we have some model that fits the data.
What is an indication of the need to use model reduction?

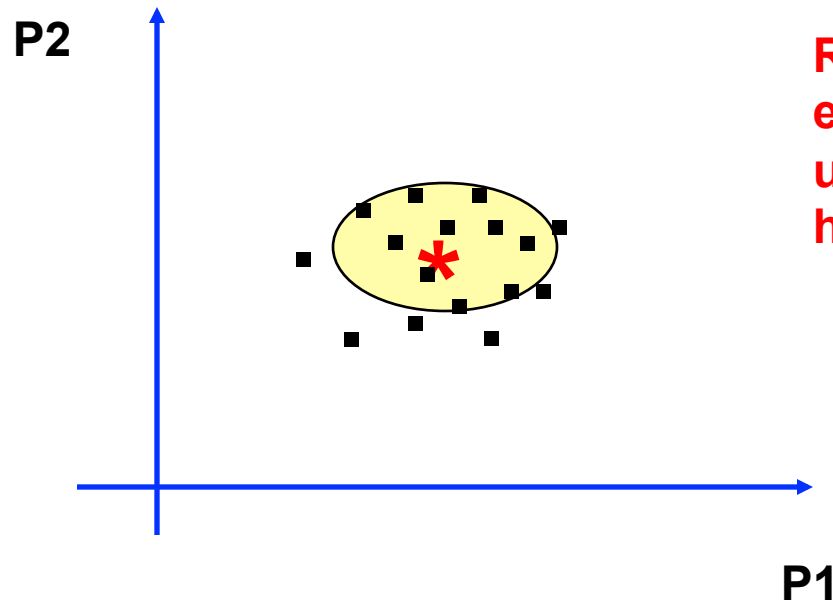
**Combination of STATISTICS / APPLIED MATH APPROACHES:
BOOTSTRAPING & MCMC (Markov Chain Monte-Carlo)**



**New parameter values
produces by fitting
re-sampled data or
by Markov Chain !**

Assume that we have some model that fits the data.
What is an indication of the need to use model reduction?

**Combination of STATISTICS / APPLIED MATH APPROACHES:
BOOTSTRAPING & MCMC (Markov Chain Monte-Carlo)**

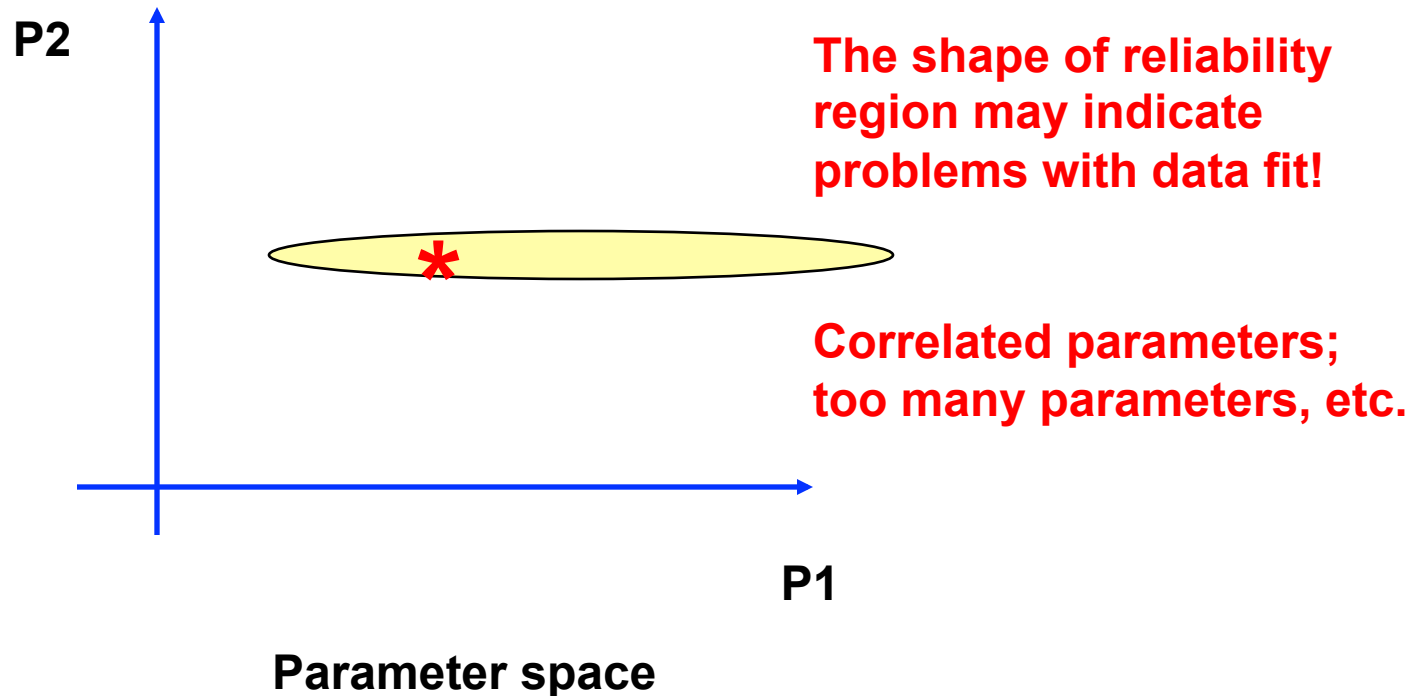


**Reliability region
estimated, e.g.,
using smoothed
histogram!**

Parameter space

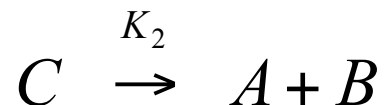
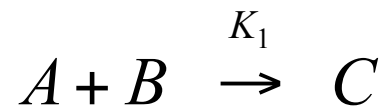
Assume that we have some model that fits the data.
What is an indication of the need to use model reduction?

**Combination of STATISTICS / APPLIED MATH APPROACHES:
BOOTSTRAPING & MCMC (Markov Chain Monte-Carlo)**



Simple illustrative example: Math models in chemical kinetics are usually formulated in terms of differential equations. What are the expected results of model reduction?

We start with the simplest reaction mechanism (e.g., generic receptor):



Also assume that reactions take place in a well mixed environment (well stirred tank reactor, small size tissue sample, etc.).

Using the Law of Mass Action, we may write a system of ordinary differential equations (ODEs) describing behavior of concentrations $[A]$, $[B]$, and $[C]$ of species A , B , and C , respectively:

$$\frac{d[A]}{dt} = -K_1[A] \cdot [B] + K_2[C],$$

$$\frac{d[B]}{dt} = -K_1[A] \cdot [B] + K_2[C],$$

$$\frac{d[C]}{dt} = +K_1[A] \cdot [B] - K_2[C],$$

$$0 \leq t \leq T.$$

Initial conditions, e.g., at time $t = 0$:

$$[A] = A^*, [B] = B^* < A^*, [C] = 0.$$

Non-dimensionalization:

$$u = [A] / A^*, \quad v = [B] / A^*, \quad w = [C] / A^*$$

$$k_1 = K_1 \cdot T \cdot A^*, \quad k_2 = K_2 \cdot T, \quad v^* = B^* / A^*, \quad \tau = t / T$$

The scaled model has the form (now all the parameters are non-dimensional):

$$\begin{aligned} \frac{du}{d\tau} &= -k_1 \cdot u \cdot v + k_2 w, & u(0) &= 1, \\ \frac{dv}{d\tau} &= -k_1 \cdot u \cdot v + k_2 w, & v(0) &= v^*, \\ \frac{dw}{d\tau} &= +k_1 \cdot u \cdot v - k_2 w, & w(0) &= 0. \end{aligned}$$

$$0 \leq \tau \leq 1.$$

Now non-dimensional characteristic reaction times $1/k_1$ and $1/k_2$ may be compared with the time interval of interest $[0,1]$. Short characteristic times (compared to $[0,1]$) correspond to fast processes and long characteristic times correspond to slow processes.

We may have several possibilities:

(a) Forward and reverse reactions are moderate (of order $O(1)$).

$$k_1 = O(1), \quad k_2 = O(1)$$

(b) Forward and reverse reactions are slow.

$$k_1 \ll 1, \quad k_2 \ll 1$$

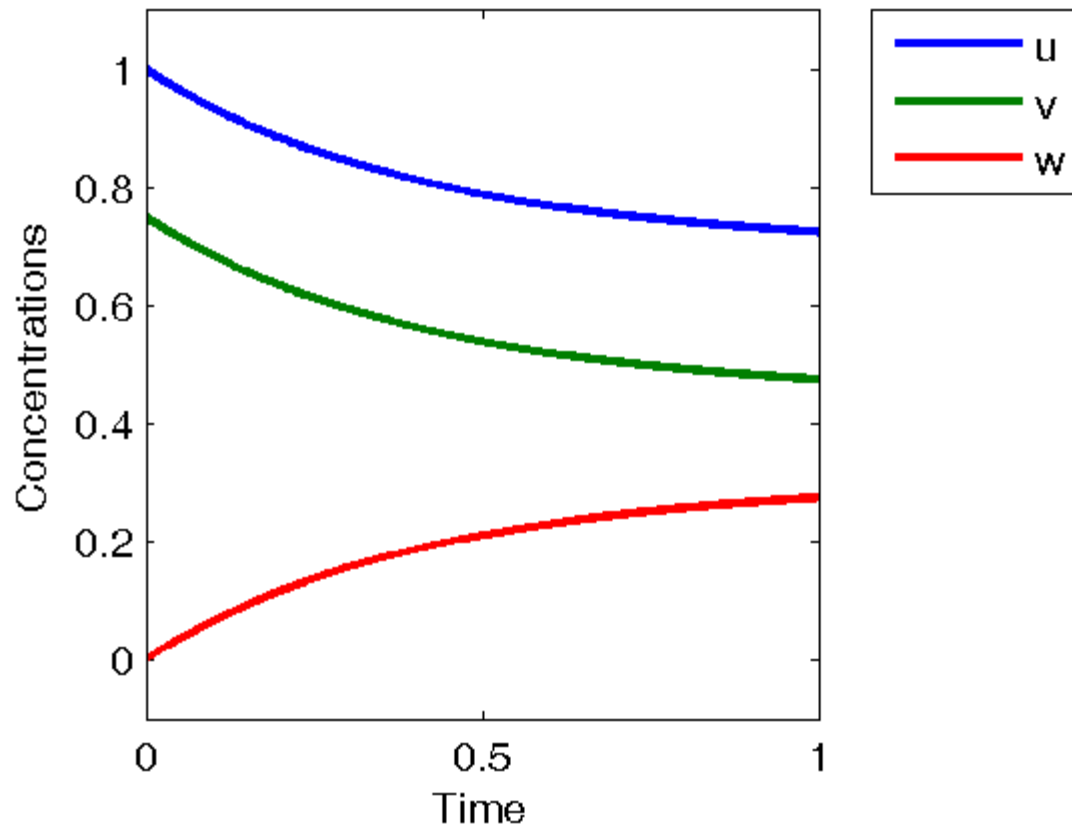
(c) Forward and reverse reactions are fast.

$$k_1 \gg 1, \quad k_2 \gg 1$$

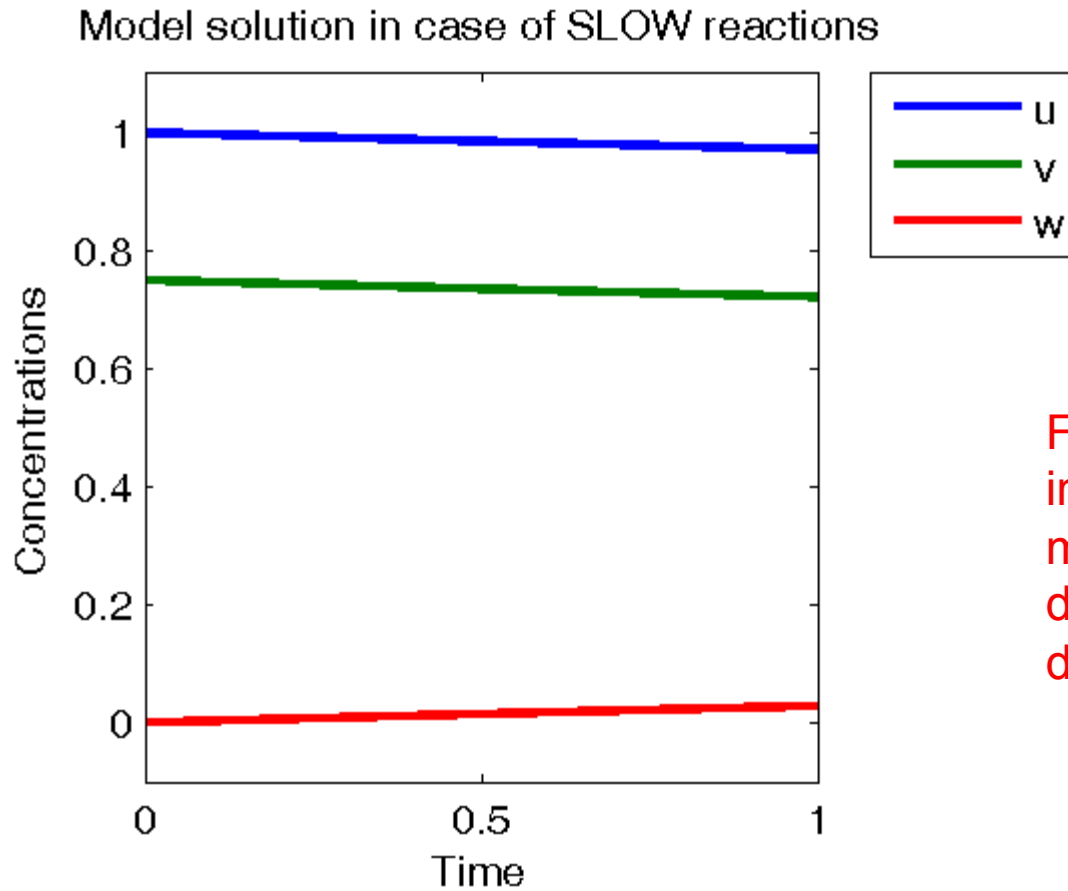
(d) Other: $k_1 \gg 1, k_2 \ll 1$; $k_1 \ll 1, k_2 \gg 1$, etc.

(a) Moderate forward and reverse reactions: $k_1 = O(1)$, $k_2 = O(1)$

Model solution in case of MODERATE reactions



(b) Slow forward and reverse reactions: $k_1 \ll 1, k_2 \ll 1$



For the time interval of interest $[0,1]$ a simpler model may be used to describe concentrations' dependence on time!

In this case we may re-scale the coefficients to obtain:

$$\begin{aligned}\frac{du}{d\tau} &= \varepsilon (-\tilde{k}_1 \cdot u \cdot v + \tilde{k}_2 w), & u(0) &= 1, \\ \frac{dv}{d\tau} &= \varepsilon (-\tilde{k}_1 \cdot u \cdot v + \tilde{k}_2 w), & v(0) &= v^*, \\ \frac{dw}{d\tau} &= \varepsilon (+\tilde{k}_1 \cdot u \cdot v - \tilde{k}_2 w), & w(0) &= 0. \\ & & 0 \leq \tau &\leq 1.\end{aligned}$$

Here $0 < \varepsilon \ll 1$ is a small parameter!

$$\tilde{k}_1 = k_1 / \varepsilon = O(1), \quad \tilde{k}_2 = k_2 / \varepsilon = O(1)$$

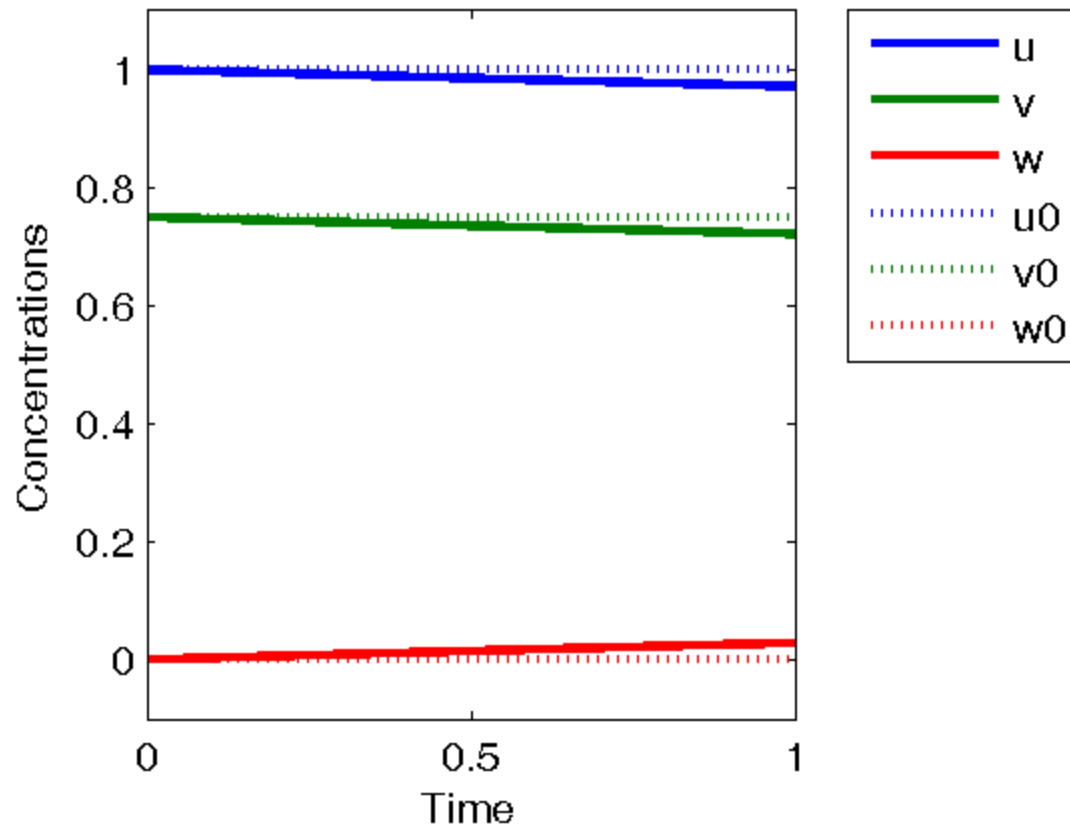
In the limit as small parameter tends to zero, we obtain:

$$\begin{aligned} \frac{du}{d\tau} = 0, & \quad u(0) = 1, & \quad u(\tau) = 1, \\ \frac{dv}{d\tau} = 0, & \quad v(0) = v^*, & \quad \text{So, } v(\tau) = v^*, \\ \frac{dw}{d\tau} = 0. & \quad w(0) = 0. & \quad w(\tau) = 0. \end{aligned}$$

Approximation satisfies the initial conditions; close to the original solution in the interval of interest:

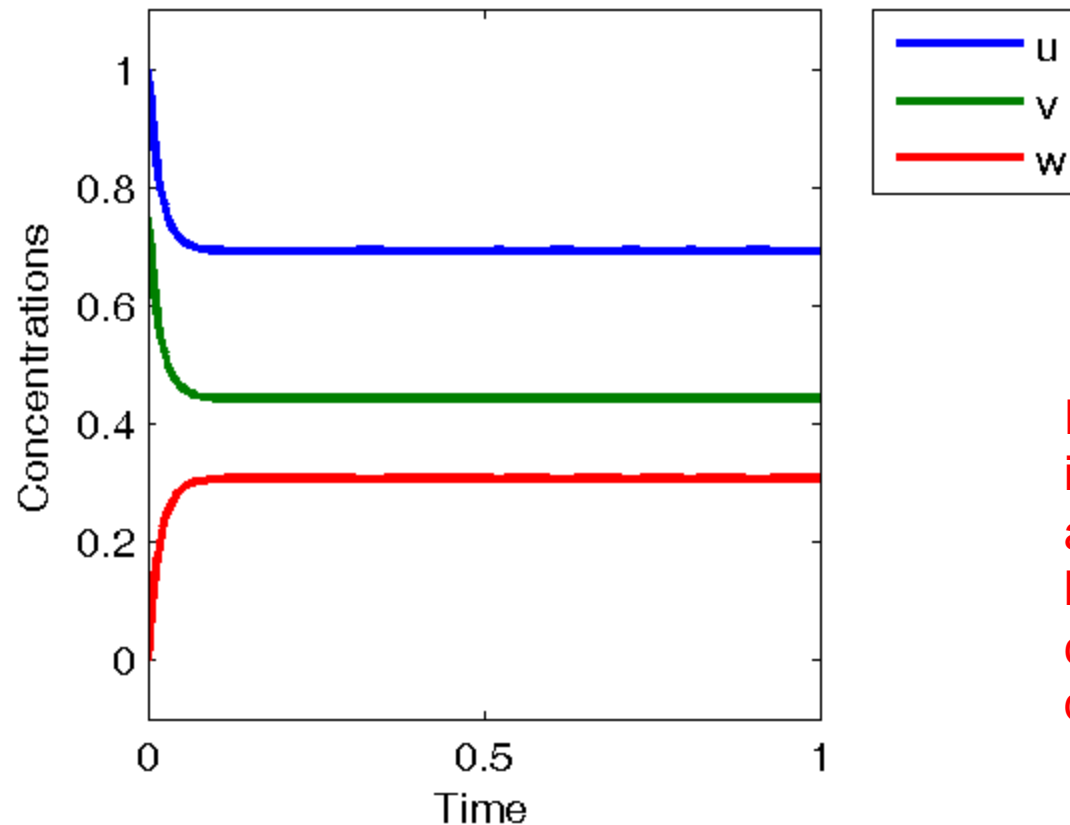
REGULARLY PERTURBED PROBLEM!

Comparison: original and reduced models



(c) Fast forward and reverse reactions: $k_1 \gg 1, k_2 \gg 1$

Model solution in case of FAST reactions



For the interior of time interval of interest $[0, 1]$ a simpler model may be used to describe concentrations' dependence on time!

In this case we may re-scale the coefficients to obtain:

$$\begin{aligned}\varepsilon \frac{du}{d\tau} &= -\tilde{k}_1 \cdot u \cdot v + \tilde{k}_2 w, & u(0) &= 1, \\ \varepsilon \frac{dv}{d\tau} &= -\tilde{k}_1 \cdot u \cdot v + \tilde{k}_2 w, & v(0) &= v^*, \\ \varepsilon \frac{dw}{d\tau} &= +\tilde{k}_1 \cdot u \cdot v - \tilde{k}_2 w, & w(0) &= 0.\end{aligned}$$

$$0 \leq \tau \leq 1.$$

Here $0 < \varepsilon \ll 1$ is a small parameter!

$$\tilde{k}_1 = k_1 \cdot \varepsilon = O(1), \quad \tilde{k}_2 = k_2 \cdot \varepsilon = O(1)$$

In the limit as small parameter tends to zero, we obtain:

$$0 = -\tilde{k}_1 \cdot u \cdot v + \tilde{k}_2 w,$$

Note: initial conditions are NOT satisfied!

$$0 < \tau \leq 1.$$

!!!

The reduced model consists of the above equation and two more equations (conservation of mass):

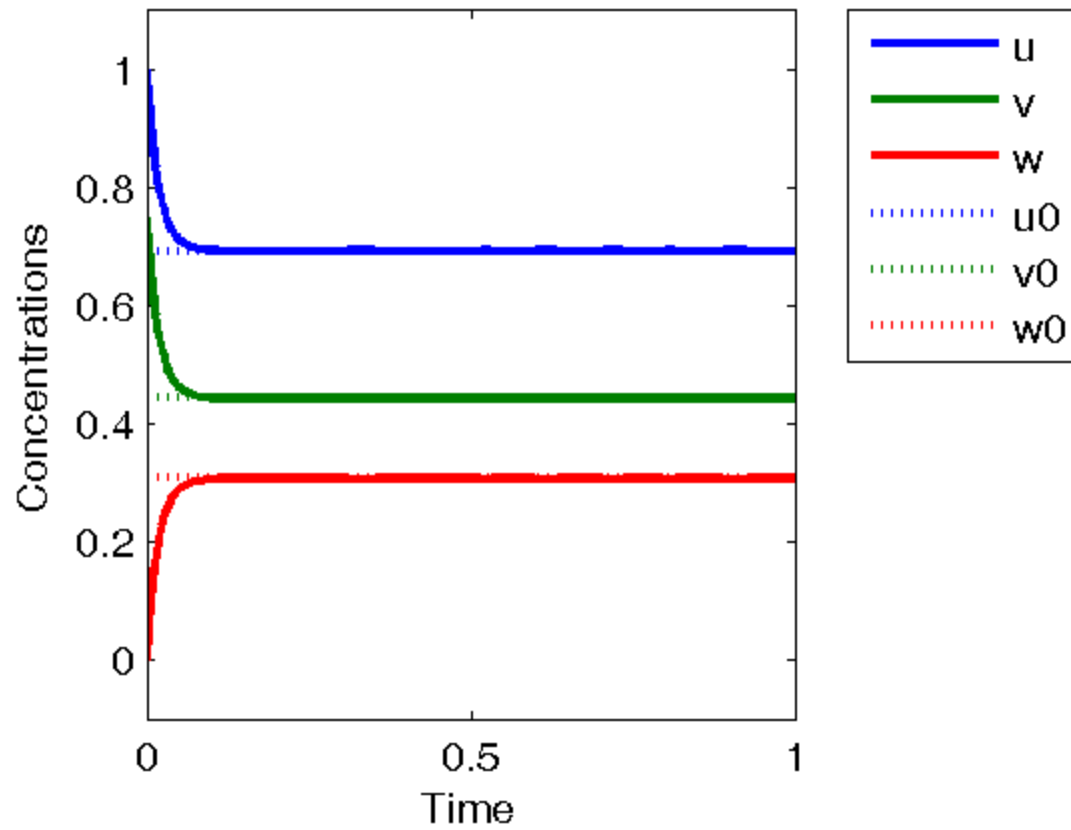
$$0 = -u \cdot v + (\tilde{k}_2 / \tilde{k}_1) w, \quad u - v = 1 - v^*, \quad u + w = 1.$$

!!!

Approximation **does not satisfy the initial conditions**; it is close to the original solution in the **interior of interval of interest**:

SINGULARLY PERTURBED PROBLEM!

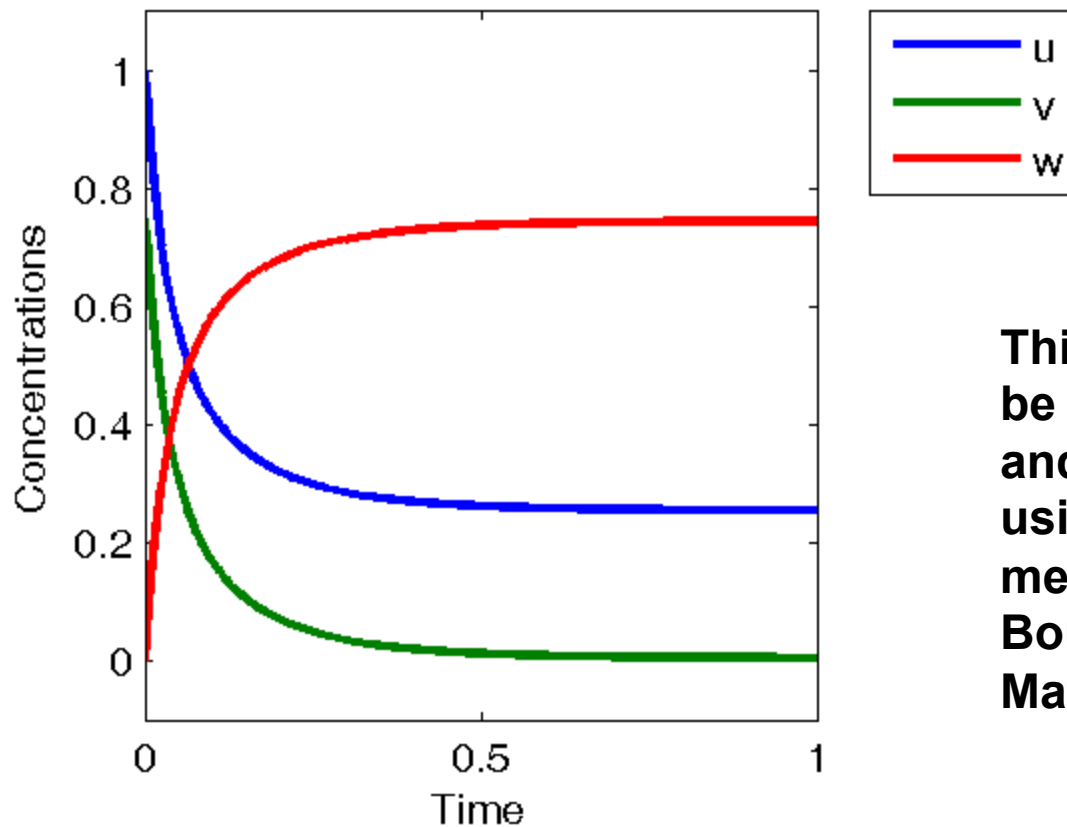
Comparison: original and reduced models



(c) Some other cases: e.g., fast forward and slow reverse reactions

$$k_1 \gg 1, \quad k_2 \ll 1$$

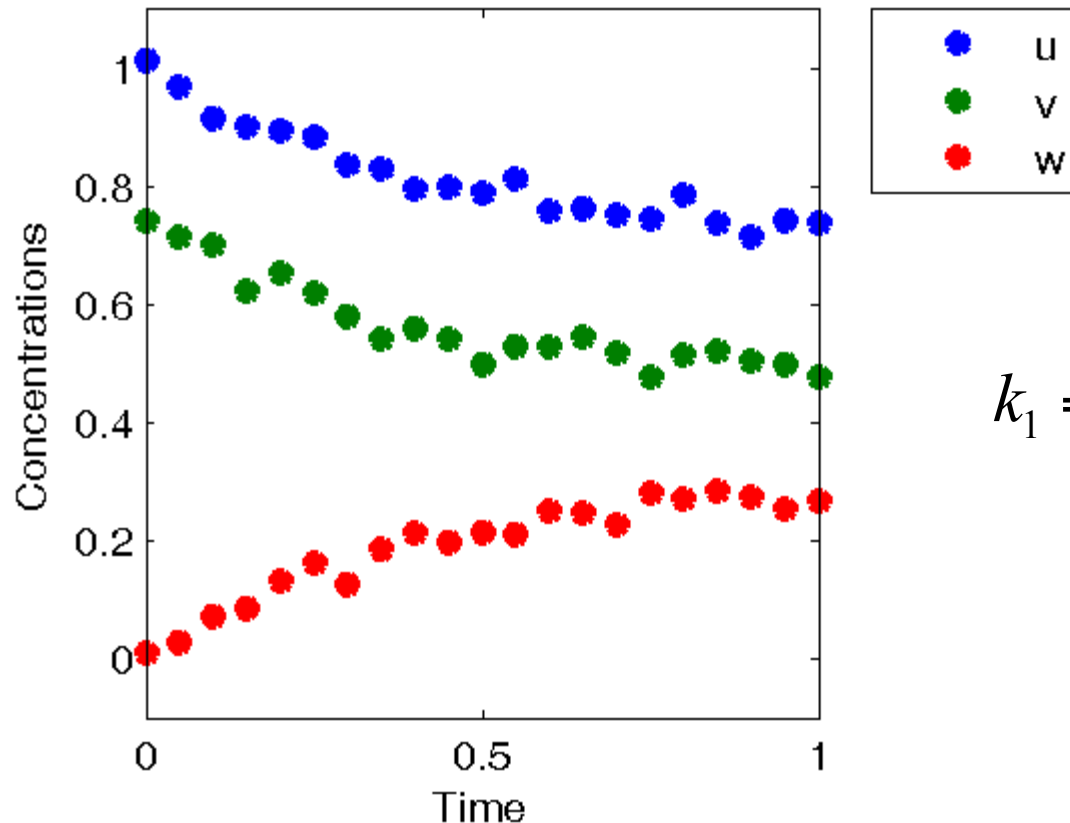
FAST FORWARD and SLOW REVERSE reactions



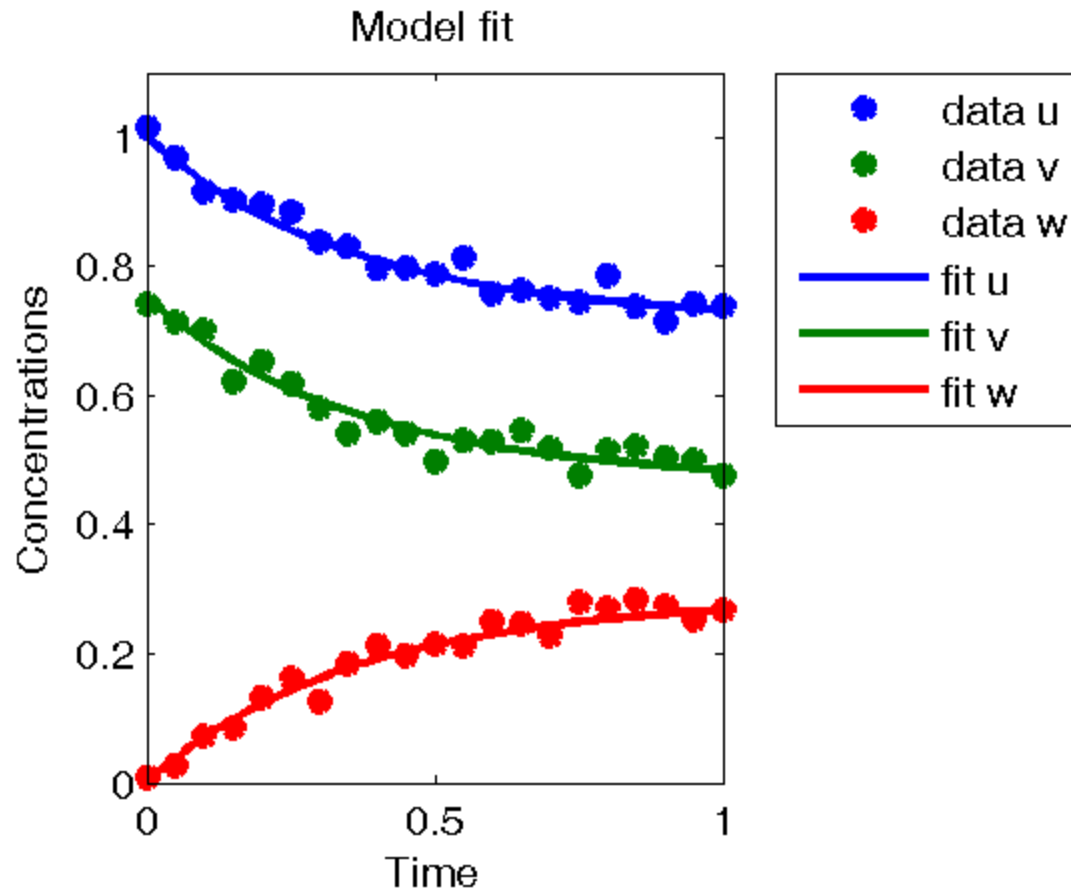
This model turns out to be **singularly perturbed** and may also be reduced using one of the asymptotic methods: **Boundary Function Method, Matching Technique, etc.**

Simple illustrative example (continuation): Identification of chemical kinetics models formulated in terms of differential equations. Reliability regions for model parameters ($k_1, k_2, v(0)$): comparison of moderate and fast reactions cases.

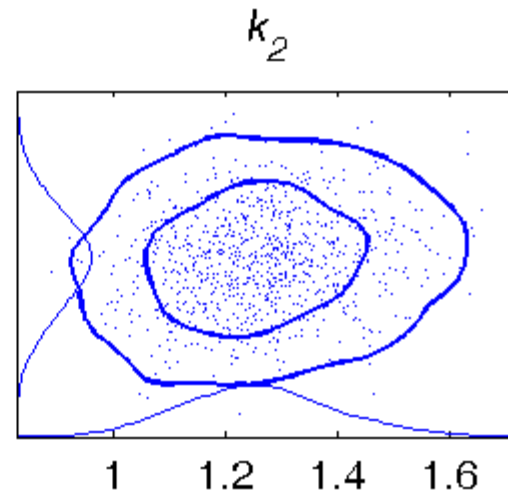
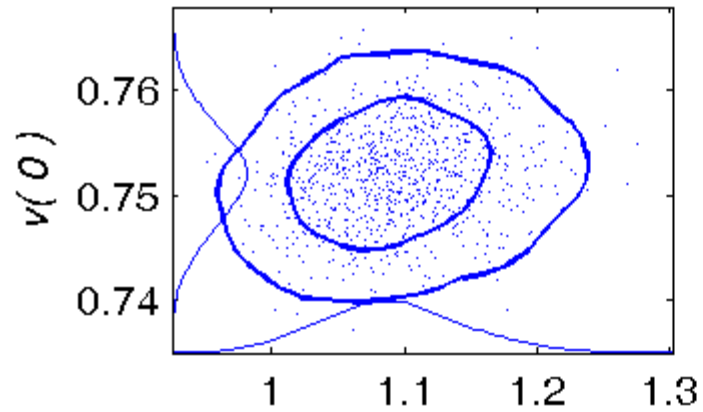
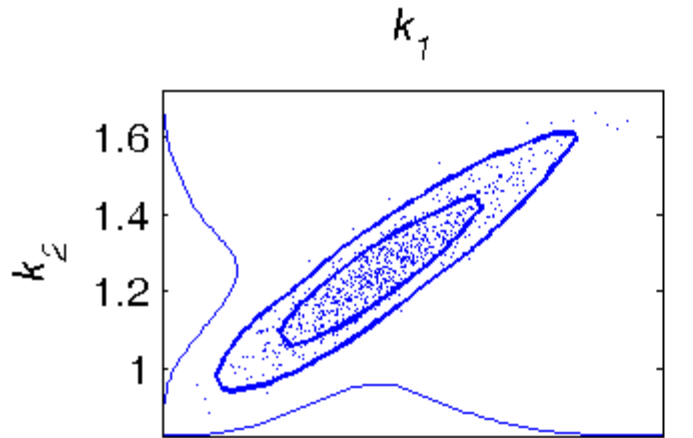
(a) Model solution: MODERATE reactions + noise



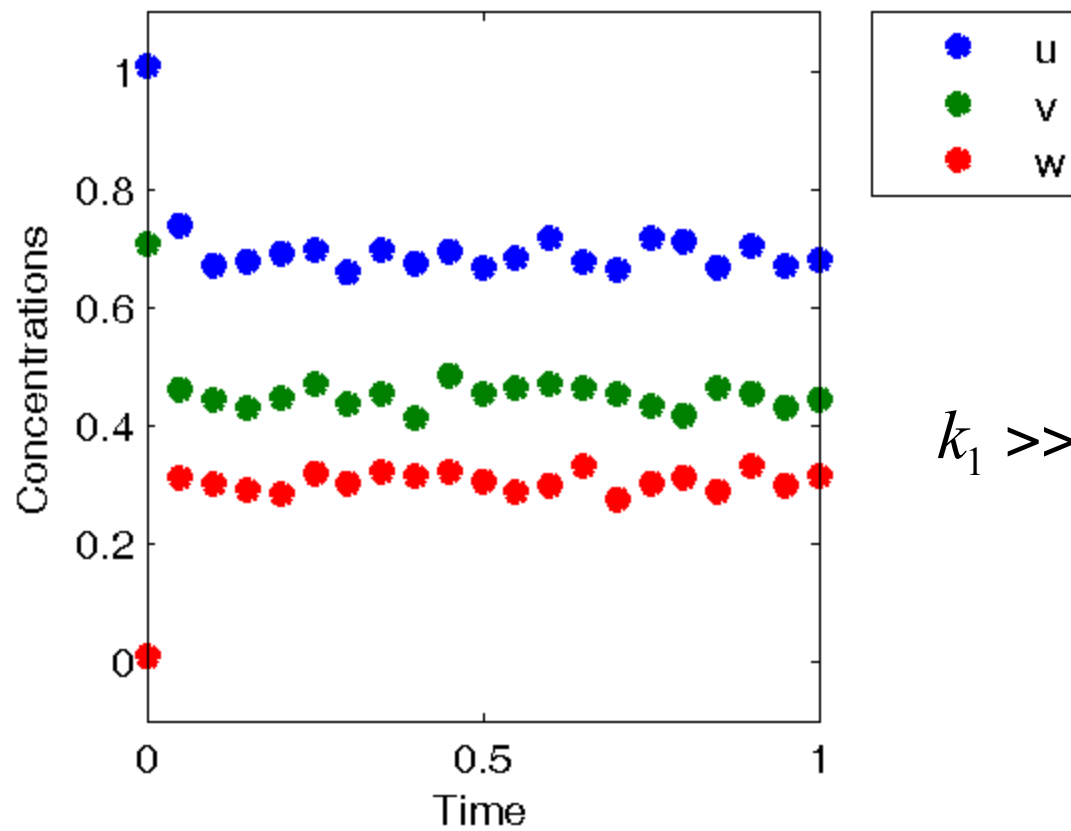
$$k_1 = O(1), \quad k_2 = O(1)$$



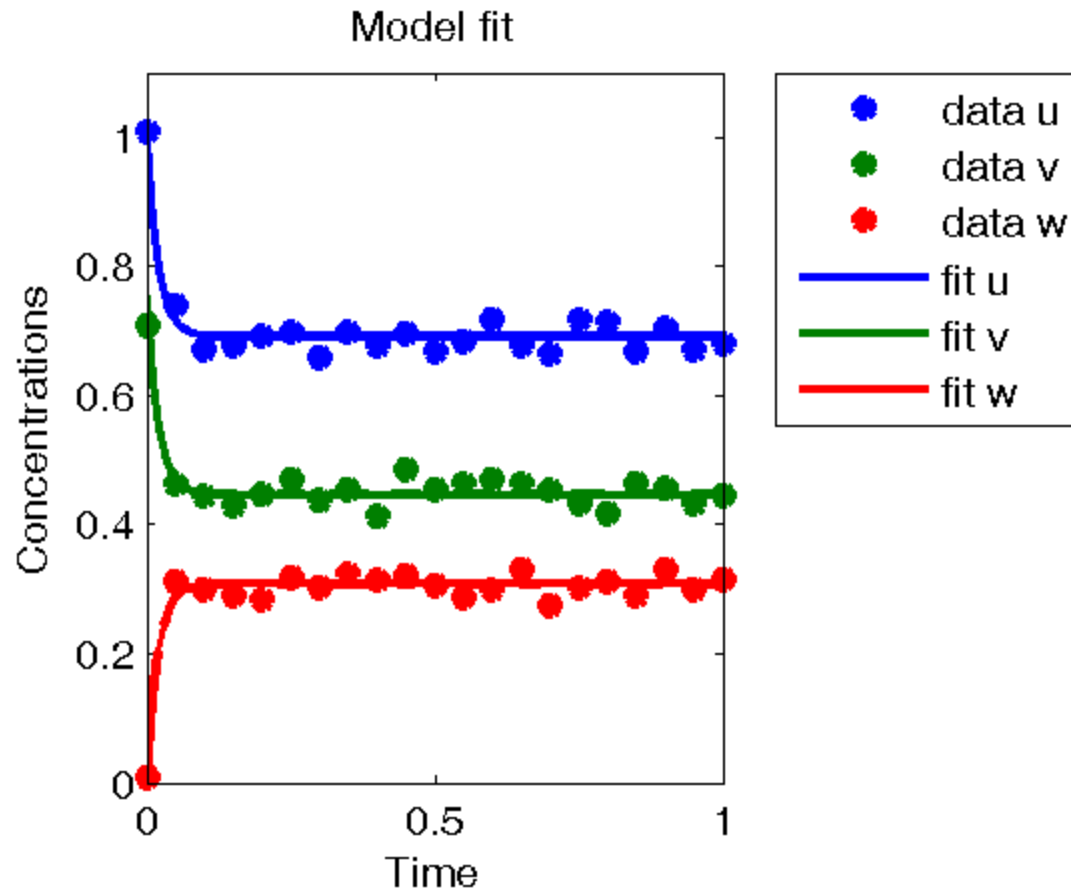
Shapes of reliability regions projections:



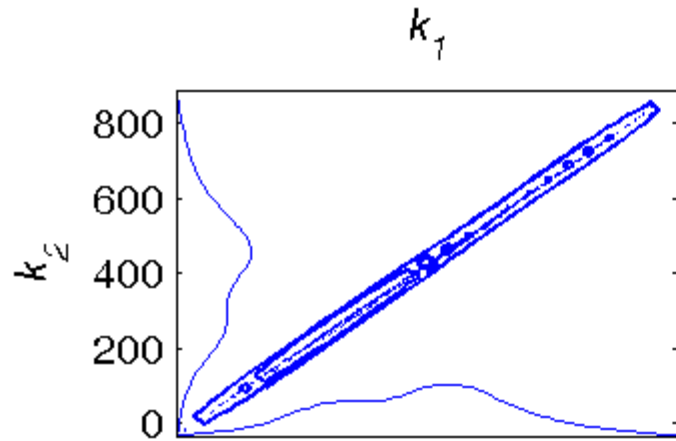
(b) Model solution: FAST reactions + noise



$$k_1 \gg 1, \quad k_2 \gg 1$$

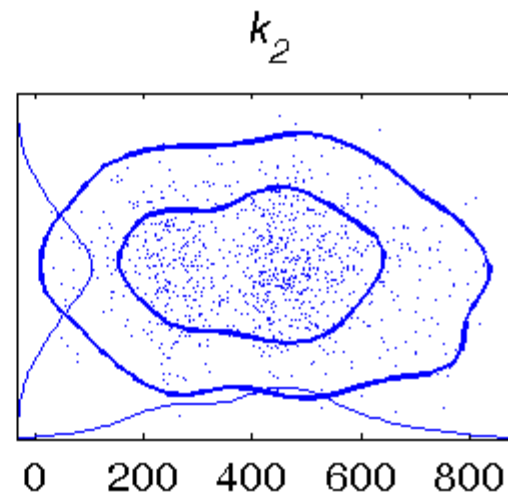
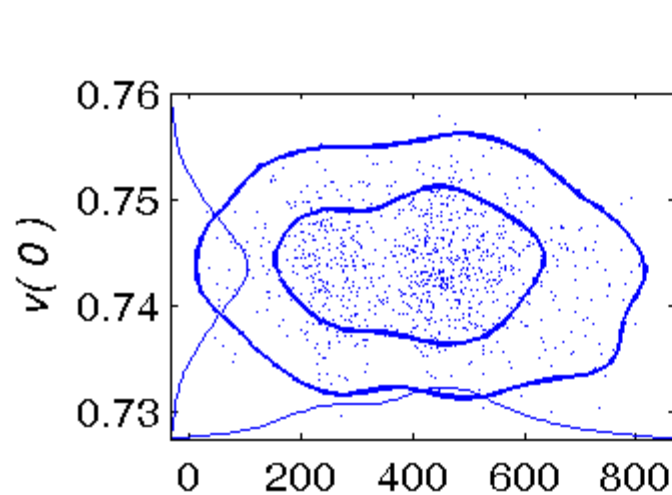


Shapes of reliability regions projections:



Two parameters are correlated;
only the ratio may be estimated:

$$k = k_2 / k_1$$



Thank you!