Chemometric Methods for the Kinetic Hard-modelling of Spectroscopic Data

ETH Zurich, March 23rd 2009

Julien Billeter

julien.billeter@chem.ethz.ch

ETH Zürich, Institute for Chemical and Bioengineering,
Safety and Environmental Technology Group, Zürich, Switzerland.

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Safety and Environmental Technology Group

Content

Part 1

Chemometrics and kinetic hard-modelling

Part 2

Uncertainties and error propagation

Part 3

Rank deficiency and spectral validation

Part 1

Chemometrics and kinetic hard-modelling

Part 2

Uncertainties and error propagation

Part 3

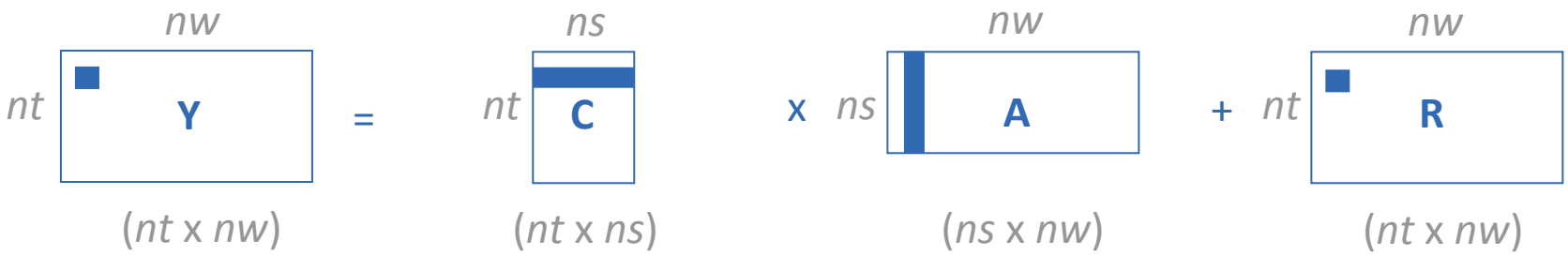
Rank deficiency and spectral validation

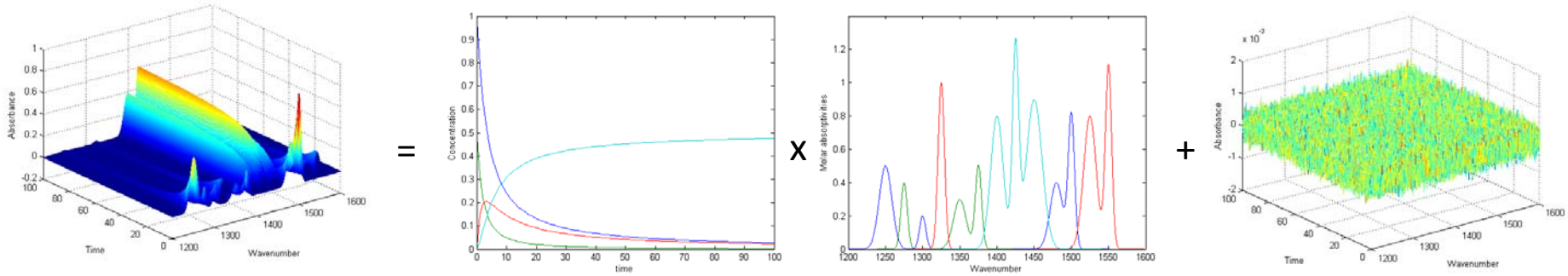
What is Chemometrics?

- Chemometrics is the chemical discipline that uses mathematical and statistical methods to:
  - (a) design or select optimal measurement procedures and experiments, and
  - (b) provide maximum chemical information by analyzing chemical data

Matthias Otto (2007)
From: Chemometrics – Statistics and Computer Application in Analytical Chemistry

Beer's law





$$A + B \xrightarrow{k_1} C$$

$$A + C \xrightarrow{k_2} D$$

Direct fitting

Beer's law

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A}$$

Modelled part

$\mathbf{C}_{\text{modelled}} = \text{function}(\mathbf{k})$

$\mathbf{A}_{\text{modelled}} = \text{function}(\boldsymbol{\theta})$

Linear counter part

- Implicit calibration
- Explicit calibration

$\mathbf{A} = \mathbf{C}^{+}_{\text{modelled}}\mathbf{Y}$ [1]

$\mathbf{A} = (\mathbf{C}^{+}\mathbf{Y})_{\text{calibration}}$

$\mathbf{C} = \mathbf{Y}\mathbf{A}^{+}_{\text{modelled}}$ [2]

$\mathbf{C} = (\mathbf{Y}\mathbf{A}^{+})_{\text{calibration}}$

Least squares fitting

$$\min_{\mathbf{k}} \left\| \mathbf{Y} - \mathbf{C}_{\text{modelled}}\mathbf{A} \right\|^2$$

$$\min_{\boldsymbol{\theta}} \left\| \mathbf{Y} - \mathbf{C}\mathbf{A}_{\text{modelled}} \right\|^2$$

$$\min_{\mathbf{k}} \left\| \mathbf{C} - \mathbf{C}_{\text{modelled}} \right\|^2$$

[1] $\mathbf{C}^{+} = \left(\mathbf{C}^{\mathsf{T}}\mathbf{C}\right)^{-1}\mathbf{C}^{\mathsf{T}}$ [2] $\mathbf{A}^{+} = \mathbf{A}^{\mathsf{T}}\left(\mathbf{A}\mathbf{A}^{\mathsf{T}}\right)^{-1}$

Indirect (inverse) fitting

Inverse model

$$\mathbf{C} = \mathbf{Y}\mathbf{B}_C \qquad \mathbf{A} = \mathbf{B}_A\mathbf{Y}$$

Modelled part

$$\mathbf{C}_{modelled} = \text{function}(\mathbf{k}) \qquad \mathbf{A}_{modelled} = \text{function}(\boldsymbol{\theta})$$

Linear counter part

- Implicit calibration
- Explicit calibration

$$\mathbf{B}_C = \mathbf{Y}^{+}\mathbf{C}_{modelled} \qquad \mathbf{B}_A = \mathbf{A}_{modelled}\mathbf{Y}^{+}$$

$$\mathbf{B}_C = (\mathbf{Y}^{+}\mathbf{C})_{calibration} \qquad \mathbf{B}_A = (\mathbf{A}\mathbf{Y}^{+})_{calibration}$$

Least squares fitting

$$\min_{\mathbf{k}} \left\| \mathbf{C}_{modelled} - \mathbf{Y}\mathbf{B}_C \right\|^2$$

$$\min_{\boldsymbol{\theta}} \left\| \mathbf{A}_{modelled} - \mathbf{B}_A\mathbf{Y} \right\|^2$$

$$\min_{\mathbf{k}} \left\| \mathbf{C} - \mathbf{C}_{modelled} \right\|^2$$

Modelling concentrations (C) or spectra (A) ?

Modelling **C**

Based on a first-principles model (the kinetic rate law)

Depends on a limited number of kinetic parameters, e.g. rate constants **k** (1 x *nr*)

Modelling **A**

Based on the modelling of **A** using Gaussian functions

Depends on a large number of parameters θ, which are difficult to determine

Requires a subsequent modelling of **C** to determine kinetic parameters, e.g. rate constants **k** (1 x *nr*)

Fitting method used in Parts 2 and 3

- Direct fitting by modelling concentrations (**C**)

$$\min_{\mathbf{k}} \left\| \mathbf{Y} - \mathbf{C}_{\text{modelled}}(\mathbf{k})\mathbf{A} \right\|^2$$

- using implicit and explicit calibration of the pure component spectra (**A**)
  - Implicit calibration: (i.e. calibration-free)

$$\mathbf{A} = \mathbf{C}^{+}_{\text{modelled}}(\mathbf{k})\mathbf{Y}$$

  - Explicit calibration: (also called Strategy 2)

$$\mathbf{A} = \left(\mathbf{C}^{+}\mathbf{Y}\right)_{\text{calibration}}$$

Modelling concentration profiles using the rate law

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

$$\mathbf{k} = \begin{bmatrix} k_1 & k_2 \end{bmatrix}$$

Modelling concentration profiles using the rate law

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

$$\mathbf{k} = [k_1 \; k_2]$$

A	B	C	D
0	0	1	0
0	0	0	1

P (*nr* x *ns*)

-

A	B	C	D
1	1	0	0
1	0	1	0

E (*nr* x *ns*)

=

A	B	C	D
-1	-1	1	0
-1	0	-1	1

N (*nr* x *ns*)

Modelling concentration profiles using the rate law

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

$$\mathbf{k} = [k_1 \; k_2]$$

A	B	C	D
0	0	1	0
0	0	0	1

P (*nr* x *ns*)

-

A	B	C	D
1	1	0	0
1	0	1	0

E (*nr* x *ns*)

=

A	B	C	D
-1	-1	1	0
-1	0	-1	1

N (*nr* x *ns*)

***nr* = 2 kinetic rate laws**

$$\begin{cases} \dfrac{dx_{t,1}}{dt} = k_1 c_{t,A}^{e_{1,1}} c_{t,B}^{e_{1,2}} c_{t,C}^{e_{1,3}} c_{t,D}^{e_{1,4}} = k_1 c_{t,A}^{1} c_{t,B}^{1} \\ \dfrac{dx_{t,2}}{dt} = k_2 c_{t,A}^{e_{2,1}} c_{t,B}^{e_{2,2}} c_{t,C}^{e_{2,3}} c_{t,D}^{e_{2,4}} = k_2 c_{t,A}^{1} c_{t,C}^{1} \end{cases}$$

Modelling concentration profiles using the rate law

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

$$\mathbf{k} = [k_1 \; k_2]$$

	A	B	C	D
	0	0	1	0
	0	0	0	1

P (*nr* x *ns*)

-

	A	B	C	D
	1	1	0	0
	1	0	1	0

E (*nr* x *ns*)

=

	A	B	C	D
	-1	-1	1	0
	-1	0	-1	1

N (*nr* x *ns*)

***ns* = 4 concentration profiles (System of ODE)**

$$\begin{cases} \frac{dc_{t,A}}{dt} = n_{1,1}\frac{dx_{t,1}}{dt} + n_{2,1}\frac{dx_{t,2}}{dt} = -1\cdot\frac{dx_{t,1}}{dt} - 1\cdot\frac{dx_{t,2}}{dt} \\ \frac{dc_{t,B}}{dt} = n_{1,2}\frac{dx_{t,1}}{dt} + n_{2,2}\frac{dx_{t,2}}{dt} = -1\cdot\frac{dx_{t,1}}{dt} + 0\cdot\frac{dx_{t,2}}{dt} \\ \frac{dc_{t,C}}{dt} = n_{1,3}\frac{dx_{t,1}}{dt} + n_{2,3}\frac{dx_{t,2}}{dt} = 1\cdot\frac{dx_{t,1}}{dt} - 1\cdot\frac{dx_{t,2}}{dt} \\ \frac{dc_{t,D}}{dt} = n_{1,4}\frac{dx_{t,1}}{dt} + n_{2,4}\frac{dx_{t,2}}{dt} = 0\cdot\frac{dx_{t,1}}{dt} + 1\cdot\frac{dx_{t,2}}{dt} \end{cases}$$

***nr* = 2 kinetic rate laws**

$$\begin{cases} \frac{dx_{t,1}}{dt} = k_1 c_{t,A}^{e_{1,1}} c_{t,B}^{e_{1,2}} c_{t,C}^{e_{1,3}} c_{t,D}^{e_{1,4}} = k_1 c_{t,A}^{1} c_{t,B}^{1} \\ \frac{dx_{t,2}}{dt} = k_2 c_{t,A}^{e_{2,1}} c_{t,B}^{e_{2,2}} c_{t,C}^{e_{2,3}} c_{t,D}^{e_{2,4}} = k_2 c_{t,A}^{1} c_{t,C}^{1} \end{cases}$$

Modelling concentration profiles using the rate law

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

$$\mathbf{k} = [k_1 \; k_2]$$

	A	B	C	D
	0	0	1	0
	0	0	0	1

P (*nr* x *ns*)

-

	A	B	C	D
	1	1	0	0
	1	0	1	0

E (*nr* x *ns*)

=

	A	B	C	D
	-1	-1	1	0
	-1	0	-1	1

N (*nr* x *ns*)

***ns* = 4 concentration profiles (System of ODE)**

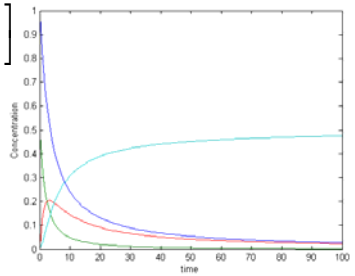
$$\begin{cases} \frac{dc_{t,A}}{dt} = n_{1,1}\frac{dx_{t,1}}{dt} + n_{2,1}\frac{dx_{t,2}}{dt} = -1\cdot\frac{dx_{t,1}}{dt} - 1\cdot\frac{dx_{t,2}}{dt} \\ \frac{dc_{t,B}}{dt} = n_{1,2}\frac{dx_{t,1}}{dt} + n_{2,2}\frac{dx_{t,2}}{dt} = -1\cdot\frac{dx_{t,1}}{dt} + 0\cdot\frac{dx_{t,2}}{dt} \\ \frac{dc_{t,C}}{dt} = n_{1,3}\frac{dx_{t,1}}{dt} + n_{2,3}\frac{dx_{t,2}}{dt} = 1\cdot\frac{dx_{t,1}}{dt} - 1\cdot\frac{dx_{t,2}}{dt} \\ \frac{dc_{t,D}}{dt} = n_{1,4}\frac{dx_{t,1}}{dt} + n_{2,4}\frac{dx_{t,2}}{dt} = 0\cdot\frac{dx_{t,1}}{dt} + 1\cdot\frac{dx_{t,2}}{dt} \end{cases}$$

***nr* = 2 kinetic rate laws**

$$\begin{cases} \frac{dx_{t,1}}{dt} = k_1 c_{t,A}^{e_{1,1}} c_{t,B}^{e_{1,2}} c_{t,C}^{e_{1,3}} c_{t,D}^{e_{1,4}} = k_1 c_{t,A}^{1} c_{t,B}^{1} \\ \frac{dx_{t,2}}{dt} = k_2 c_{t,A}^{e_{2,1}} c_{t,B}^{e_{2,2}} c_{t,C}^{e_{2,3}} c_{t,D}^{e_{2,4}} = k_2 c_{t,A}^{1} c_{t,C}^{1} \end{cases}$$

Numerical integration

$$\mathbf{C} = [\mathbf{c}_{t,A} \; \mathbf{c}_{t,B} \; \mathbf{c}_{t,C} \; \mathbf{c}_{t,D}]$$



The System of ODE is integrated with initial concentrations $\mathbf{c_0} = [c_{0,A} \; c_{0,B} \; c_{0,C} \; c_{0,D}]$

Part 1

Chemometrics and kinetic hard-modelling

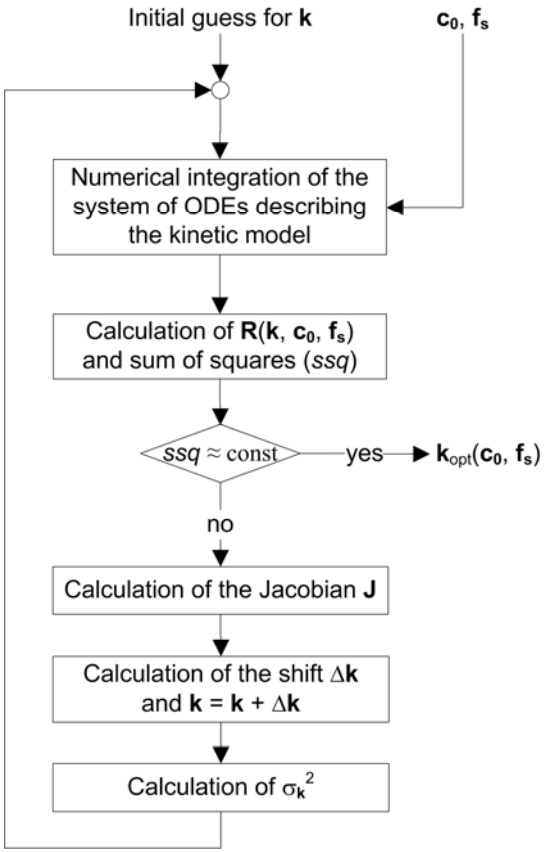
Part 2

Uncertainties and error propagation

Part 3

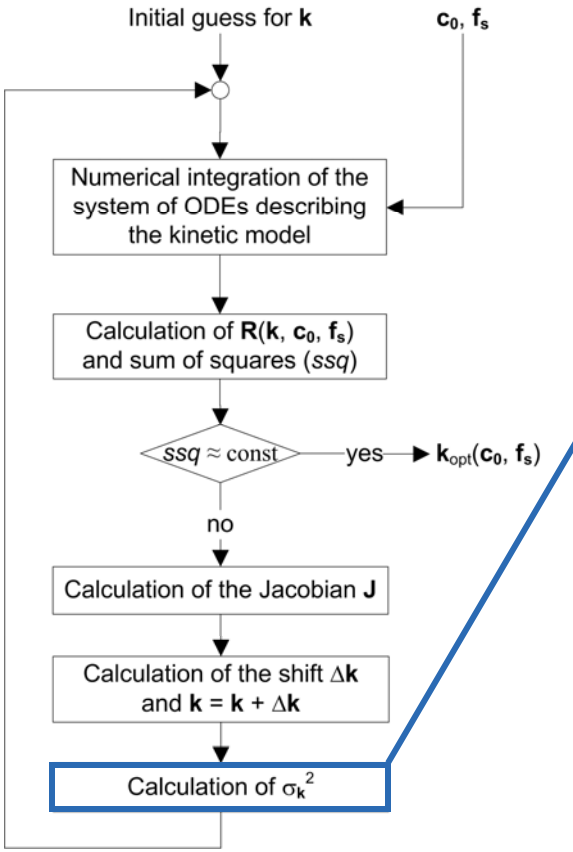
Rank deficiency and spectral validation

Newton-Gauss algorithm $\min_{\mathbf{k}} \left\| \mathbf{Y} - \mathbf{C}_{\text{modelled}} \mathbf{A} \right\|^2$



Newton-Gauss algorithm

$$\min_{\mathbf{k}} \left\| \mathbf{Y} - \mathbf{C}_{\text{modelled}} \mathbf{A} \right\|^2$$



Newton-Gauss algorithm delivers an estimate of the uncertainties in the optimised parameters (k)

$$\boldsymbol{\sigma}_{\mathbf{k}}^2 = \operatorname{diag}\left(\mathbf{H}^{-1}\right)\sigma_{\mathbf{r}}^2$$

$\boldsymbol{\sigma}_{\mathbf{k}}^2$ — Variance of the optimised parameters **k** (1 x *nr*)

$\operatorname{diag}(\cdot)$ — Operator extracting a vector of diagonal elements from a matrix

$\mathbf{H} = \left(\frac{\partial \mathbf{R}}{\partial \mathbf{k}}\right)^{\mathsf{T}} \left(\frac{\partial \mathbf{R}}{\partial \mathbf{k}}\right)$ — Sensitivity of the residuals **R** with respect to the optimised parameters **k**

$\sigma_{\mathbf{r}}^2$ — Variance of the residuals

Error propagation in Newton-Gauss algorithm

Classical estimation of uncertainties: $\boldsymbol{\sigma}_{\mathbf{k}}^2 = \mathrm{diag}\left(\mathbf{H}^{-1}\right)\sigma_{\mathbf{r}}^2$

Problem: this estimation ($\boldsymbol{\sigma}_{\mathbf{k}}^2$) is lower than the variance calculated by repetition of the experiments.

Error propagation in Newton-Gauss algorithm

Classical estimation of uncertainties: $\boldsymbol{\sigma}_{\mathbf{k}}^2 = \mathrm{diag}\left(\mathbf{H}^{-1}\right)\sigma_{\mathbf{r}}^2$

Problem: this estimation ($\boldsymbol{\sigma}_{\mathbf{k}}^2$) is lower than the variance calculated by repetition of the experiments.

Uncertainties and error propagation (easily extendable):

$\boldsymbol{\sigma}_{\mathbf{k}}^2 =$

$\boldsymbol{\sigma}_{\mathbf{k}}^2 =$

Error propagation in Newton-Gauss algorithm

Classical estimation of uncertainties: $\boldsymbol{\sigma}_{\mathbf{k}}^2 = \mathrm{diag}\left(\mathbf{H}^{-1}\right)\sigma_{\mathbf{r}}^2$

Problem: this estimation ($\boldsymbol{\sigma}_{\mathbf{k}}^2$) is lower than the variance calculated by repetition of the experiments.

Uncertainties and error propagation (easily extendable):

$$\boldsymbol{\sigma}_{\mathbf{k}}^2 = \boxed{\mathrm{diag}\left(\mathbf{H}^{-1}\right)\sigma_{\mathbf{r}}^2}$$

$$\boldsymbol{\sigma}_{\mathbf{k}}^2 = \qquad \boldsymbol{\sigma}_{\mathbf{k,r}}^2$$

Error propagation in Newton-Gauss algorithm

Classical estimation of uncertainties: $\boldsymbol{\sigma}_{\mathbf{k}}^2 = \mathrm{diag}\left(\mathbf{H}^{-1}\right)\sigma_{\mathbf{r}}^2$

Problem: this estimation ($\boldsymbol{\sigma}_{\mathbf{k}}^2$) is lower than the variance calculated by repetition of the experiments.

Uncertainties and error propagation (easily extendable):

$$\boldsymbol{\sigma}_{\mathbf{k}}^2 = \mathrm{diag}\left(\mathbf{H}^{-1}\right)\sigma_{\mathbf{r}}^2 + \mathrm{diag}\left[\left(\frac{\partial \mathbf{k}}{\partial \mathbf{c_0}}\right)^{\mathrm{T}} \mathrm{DIAG}\left(\boldsymbol{\sigma}_{\mathbf{c_0}}^2\right)\left(\frac{\partial \mathbf{k}}{\partial \mathbf{c_0}}\right)\right]$$

$$\boldsymbol{\sigma}_{\mathbf{k}}^2 = \boldsymbol{\sigma}_{\mathbf{k,r}}^2 + \boldsymbol{\sigma}_{\mathbf{k},\mathbf{c_0}}^2$$

DIAG = operator generating a diagonal matrix from the corresponding vector argument

Error propagation in Newton-Gauss algorithm

Classical estimation of uncertainties:

$$\boldsymbol{\sigma}_{\mathbf{k}}^2 = \mathrm{diag}\left(\mathbf{H}^{-1}\right)\sigma_{\mathbf{r}}^2$$

Problem: this estimation ($\boldsymbol{\sigma}_{\mathbf{k}}^2$) is lower than the variance calculated by repetition of the experiments.

Uncertainties and error propagation (easily extendable):

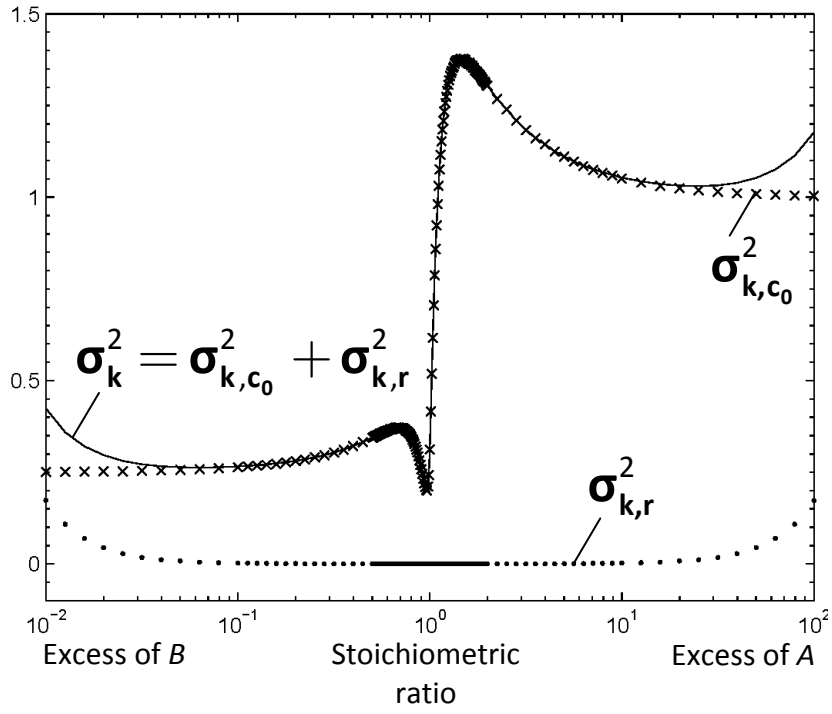
$$\boldsymbol{\sigma}_{\mathbf{k}}^2 = \mathrm{diag}\left(\mathbf{H}^{-1}\right)\sigma_{\mathbf{r}}^2 + \mathrm{diag}\left[\left(\frac{\partial \mathbf{k}}{\partial \mathbf{c_0}}\right)^{\mathrm{T}} \mathrm{DIAG}\left(\boldsymbol{\sigma}_{\mathbf{c_0}}^2\right)\left(\frac{\partial \mathbf{k}}{\partial \mathbf{c_0}}\right)\right] + \mathrm{diag}\left[\left(\frac{\partial \mathbf{k}}{\partial \mathbf{f_s}}\right)^{\mathrm{T}} \mathrm{DIAG}\left(\boldsymbol{\sigma}_{\mathbf{f_s}}^2\right)\left(\frac{\partial \mathbf{k}}{\partial \mathbf{f_s}}\right)\right]$$

$$\boldsymbol{\sigma}_{\mathbf{k}}^2 = \boldsymbol{\sigma}_{\mathbf{k,r}}^2 + \boldsymbol{\sigma}_{\mathbf{k},\mathbf{c_0}}^2 + \boldsymbol{\sigma}_{\mathbf{k,f}}^2$$

DIAG = operator generating a diagonal matrix from the corresponding vector argument

Simulation: $A + B \xrightarrow{k_1} C$

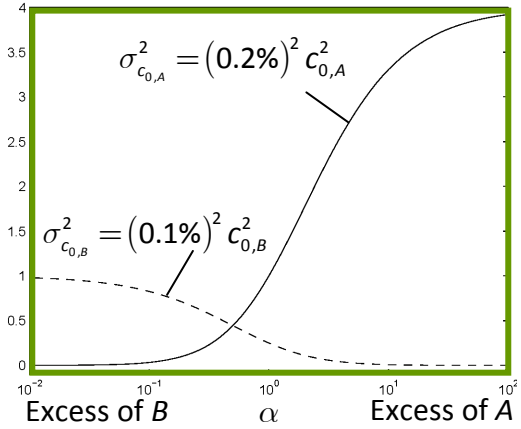
$$\sigma_k^2 = \sigma_{k,r}^2 + \sigma_{k,c_0}^2 \quad \text{with} \quad \sigma_{k,c_0}^2 = \text{diag}\left[\left(\frac{\partial k}{\partial c_0}\right)^T \text{DIAG}\left(\sigma_{c_0}^2\right)\left(\frac{\partial k}{\partial c_0}\right)\right]$$



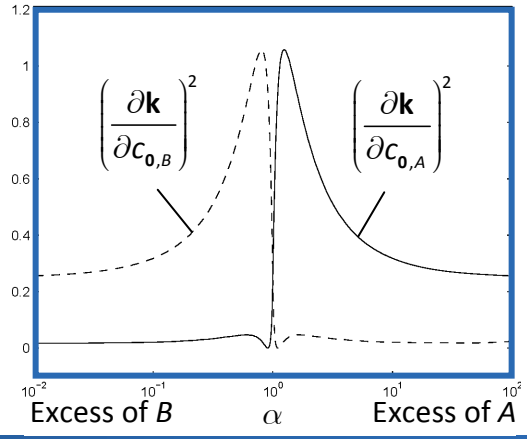
$$\alpha = c_{0,A} / c_{0,B} \quad \text{with } c_{0,A} + c_{0,B} = 1 \text{ mol} \cdot \text{L}^{-1}$$

Different uncertainties in the initial concentrations due to sampling

Asymmetry in σ_{k,c_0}^2



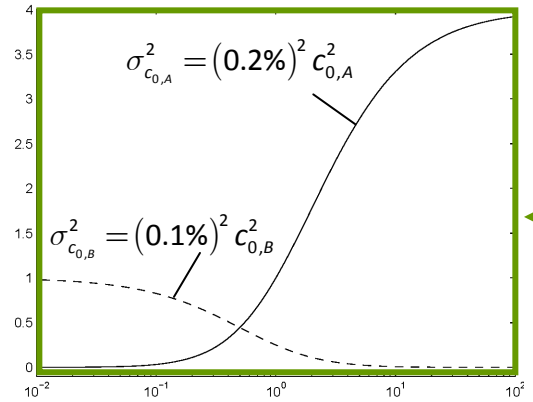
Maxima / Minima in σ_{k,c_0}^2



Simulation: optimal experimental conditions

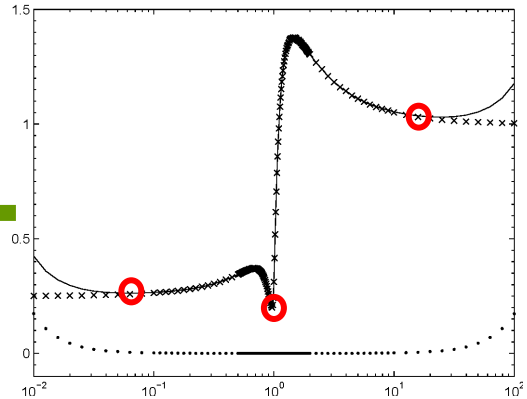
$$A + B \xrightarrow{k_1} C$$

$$\boldsymbol{\sigma}_{\mathbf{k}}^2 = \boldsymbol{\sigma}_{\mathbf{k,r}}^2 + \boldsymbol{\sigma}_{\mathbf{k},\mathbf{c_0}}^2 \approx \boldsymbol{\sigma}_{\mathbf{k},\mathbf{c_0}}^2 = \mathrm{diag}\left[\left(\frac{\partial \mathbf{k}}{\partial \mathbf{c_0}}\right)^{\mathrm{T}} \mathrm{DIAG}\left(\boldsymbol{\sigma}_{\mathbf{c_0}}^2\right)\left(\frac{\partial \mathbf{k}}{\partial \mathbf{c_0}}\right)\right]$$



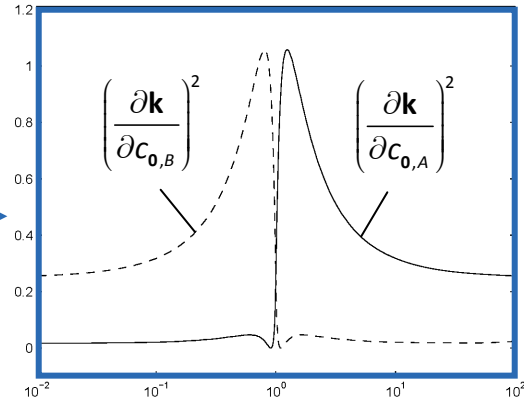
Excess of *B* $\alpha = c_{0,A} / c_{0,B}$ Excess of *A*

Excess in the species with the lowest uncertainty in its initial concentration (here *B*)



Excess of *B* $\alpha = c_{0,A} / c_{0,B}$ Excess of *A*

This curve was validated by Monte-Carlo sampling (10 000 points)

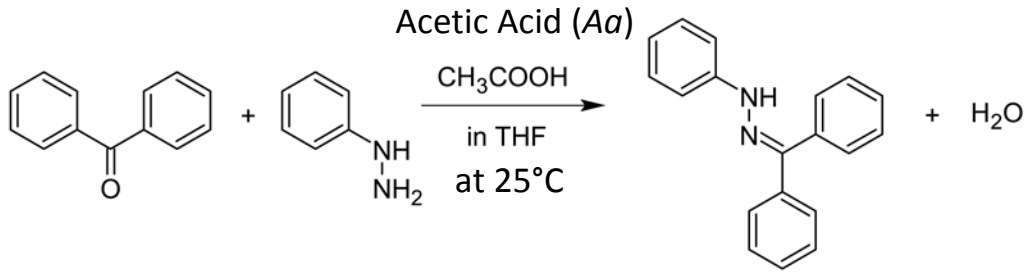


Excess of *B* $\alpha = c_{0,A} / c_{0,B}$ Excess of *A*

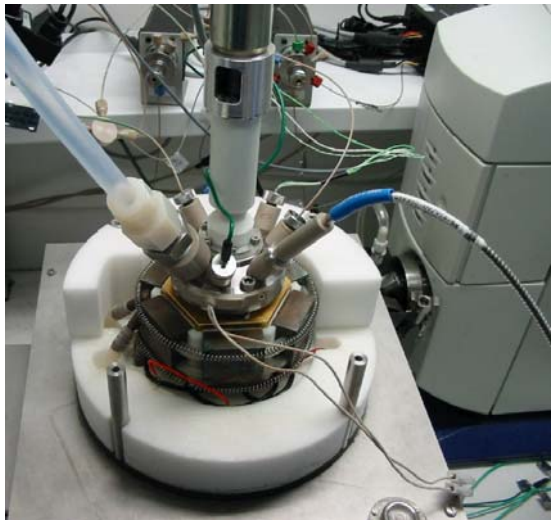
Under exact stoichiometric conditions ($\alpha = 1$)
Under pseudo-first order conditions ($\alpha \in]\infty, 10] \cup [0.1, 0[$)

Experimental reaction

Overall reaction:



Benzophenone (*B*) + Phenylhydrazine (*P*) → Benzophenone-Phenylhydrazone (*BP*) + Water



Reactor: CRC.v4 with FT-IR and UV-vis

Kinetic mechanism:

$$B + P + Aa \xrightarrow{k} BP + Aa$$

Experimental conditions:

25°C, dosing *Aa*, followed in mid-IR (1200–1650 cm^{-1}) and UV-vis (240–400 nm).
$c_{0,B}$ = 0.40033 (±0.292%) mol·L^{-1}, $c_{0,P}$ = 1.19737 (±0.292%) mol·L^{-1}, $c_{0,Aa}$ = 0 mol·L^{-1}
$c_{dos,Aa}$ = 17.48376 (±0%) mol·L^{-1}, dosing rate = 8.17 (±0.14%) mL·min^{-1} in 0.6 min.

17 repetitions

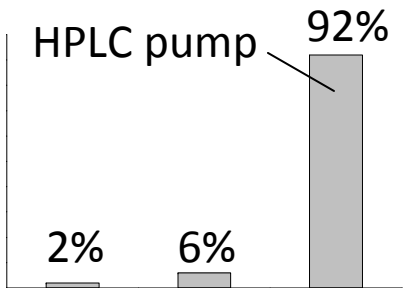
Experimental results

	UV-vis		mid-IR	
	k [a]	σ_k [a]	**k** [a]	σ_k [a]
Experimental (mean and standard deviation over 17 repetitions)	1.76(8)	**0.02(8)**	1.73(9)	**0.05(4)**
Predicted by error propagation	-	**0.02(3)**	-	**0.02(2)**
Literature	1.40 [b]	-	1.51 [b]	-
	1.65 [c]	-	1.65 [c]	-

[a] in $L^2 \cdot mol^{-2} \cdot s^{-1} \times 10^{-4}$

[b] Carvalho et al., Talanta, 68 (2006), 1190-1200

[c] Billeter et al., Chemom. Intell. Lab. Syst., (2009), submitted



$$\sigma_k^2 = \sigma_{k,r}^2 + \sigma_{k,c_0}^2 + \sigma_{k,f}^2$$

Part 1

Chemometrics and kinetic hard-modelling

Part 2

Uncertainties and error propagation

Part 3

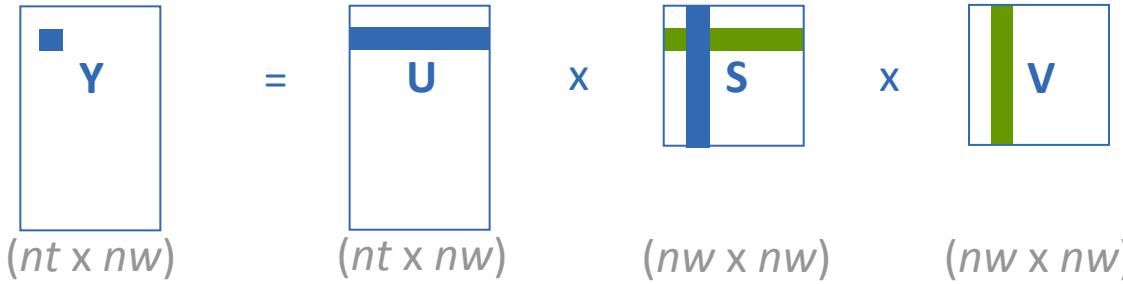
Rank deficiency and spectral validation

Singular Value Decomposition (SVD)

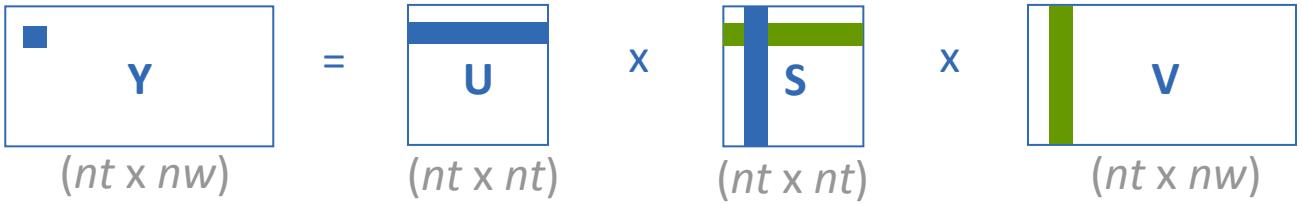
$$\mathbf{Y} = \mathbf{U}\,\mathbf{S}\,\mathbf{V}$$

U : Matrix of orthonormal column eigenvectors
S : Matrix of singular values
V : Matrix of orthonormal row eigenvectors

nw < nt



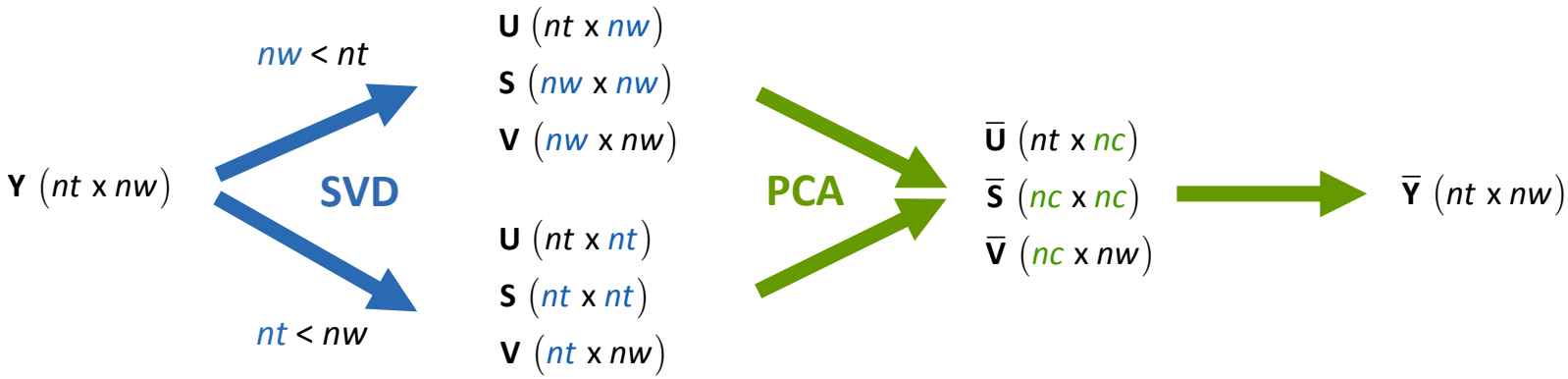
nt < nw



Principal Component Analysis (PCA)

$$\mathbf{Y} = \mathbf{U}\,\mathbf{S}\,\mathbf{V} \qquad \mathbf{R} = \mathbf{Y} - \bar{\mathbf{Y}} = \text{noise}$$

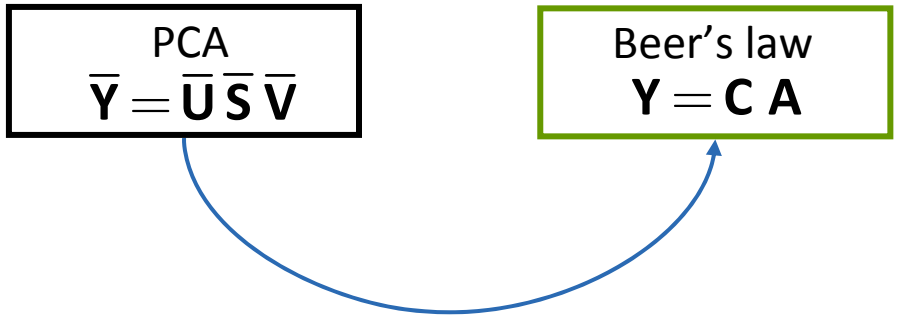
Reduction of the dimensionality to *nc*,
i.e. the number of significant singular values (or eigenvectors)



nt or *nw* factors

nc factors

Target Factor Analysis (TFA)



nc = number of significant factors

ns = number of reactive species

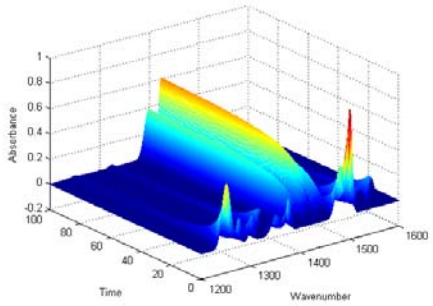
TFA: Relationship between PCA and Beer's law

$$\bar{\mathbf{C}}\,(nt \times nc) = \bar{\mathbf{U}}\,(nt \times nc)\,\mathbf{T} \neq \mathbf{C}\,(nt \times ns)$$
$$\bar{\mathbf{A}}\,(nc \times nw) = \mathbf{T}^{-1}\,\bar{\mathbf{S}}\,\bar{\mathbf{V}}\,(nc \times nw) \neq \mathbf{A}\,(ns \times nw)$$

Where **T** is a transformation matrix of dimensions (*nc* x *nc*)

Rank deficiency in spectroscopy

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

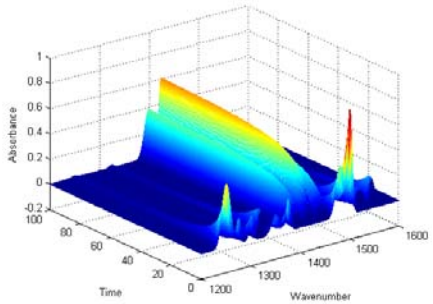


Y

Rank deficiency in spectroscopy

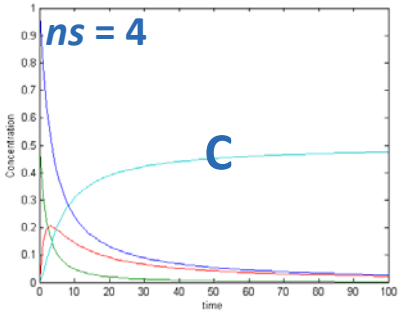
Beer's law (*ns* = 4 species) Y = CA

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

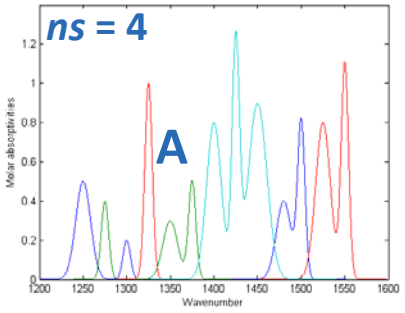


Y

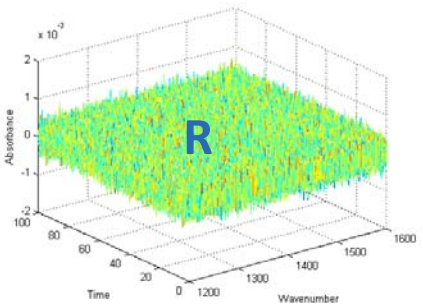
=



x



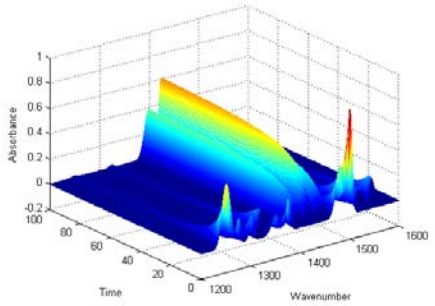
+

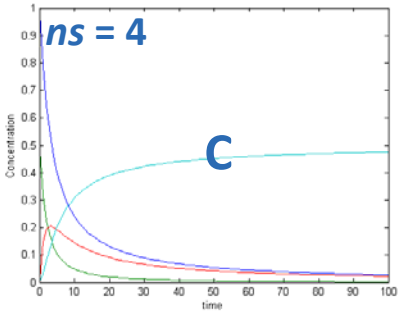


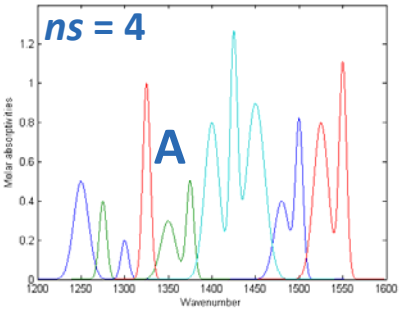
Rank deficiency in spectroscopy

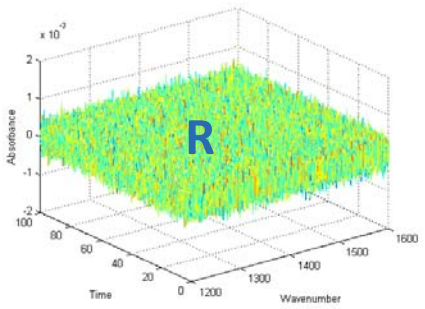
Beer's law (*ns* = 4 species) Y = CA

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

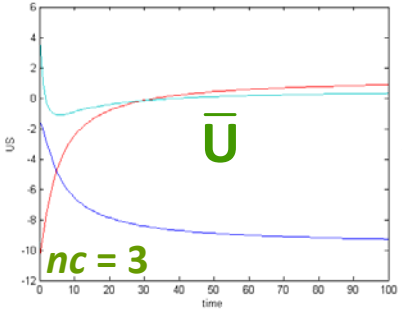


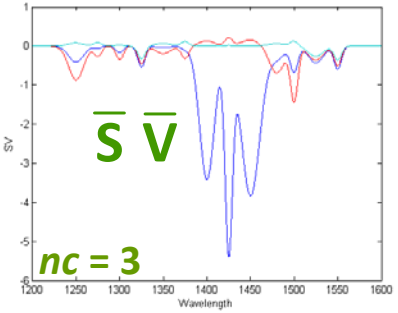
Y = 

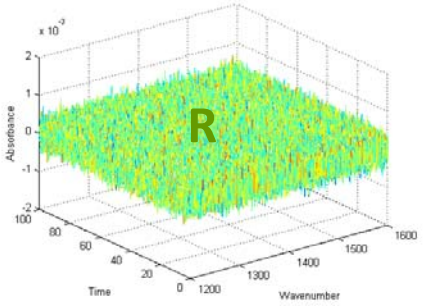
x 

+ 

PCA (*nc* = 3 factors) $\bar{Y} = \bar{U}\,\bar{S}\,\bar{V}$

= 

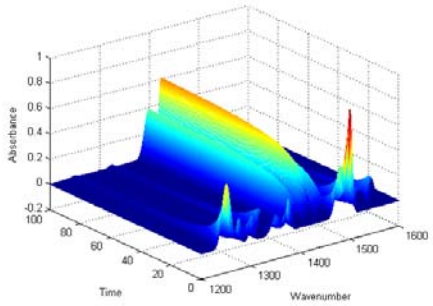
x 

+ 

Rank deficiency in spectroscopy

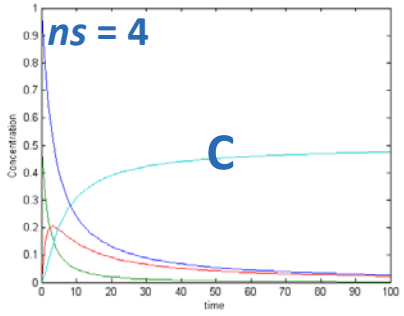
Beer's law (*ns* = 4 species) Y = CA

$$\begin{cases} A + B \xrightarrow{k_1} C \\ A + C \xrightarrow{k_2} D \end{cases}$$

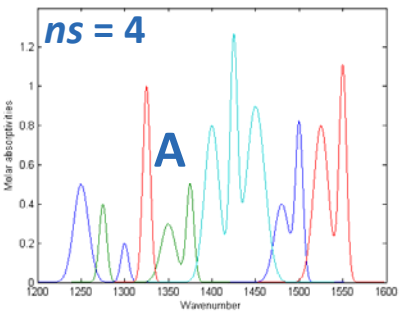


Y

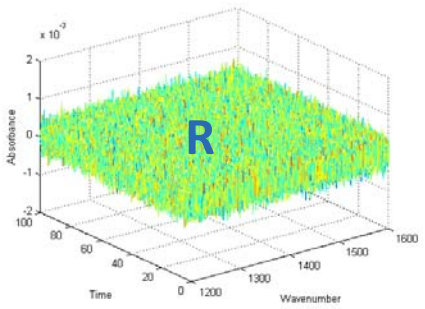
=



x

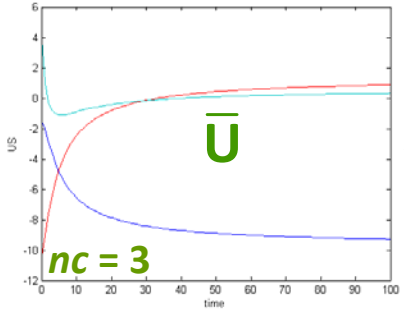


+

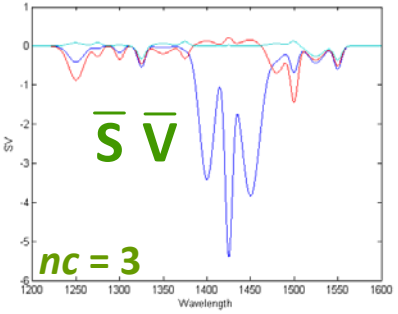


PCA (*nc* = 3 factors) $\bar{Y} = \bar{U}\,\bar{S}\,\bar{V}$

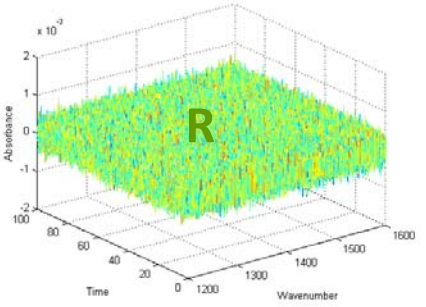
=



x



+



Spectroscopic data **Y** are rank deficient when:

significant factors (*nc*) in PCA < number of reactive species (*ns*) ⇔ rank(**Y**) < *ns*

Sources and problems of rank deficiency

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A}$$

Rank deficiencies in **Y** is due to:

Linear dependencies in **C** and/or Linear dependencies in **A**

Mathematical dilemma in case of implicit calibration

A cannot be computed by $\mathbf{C}^{+}\mathbf{Y}$ as **A** is not unique

Not discussed here

All spectra in **A** are assumed to be linearly independent

Example: two species that are consumed or generated at the same rate

rank(**Y**) = min [rank(**C**), rank(**A**)] = rank(**C**)

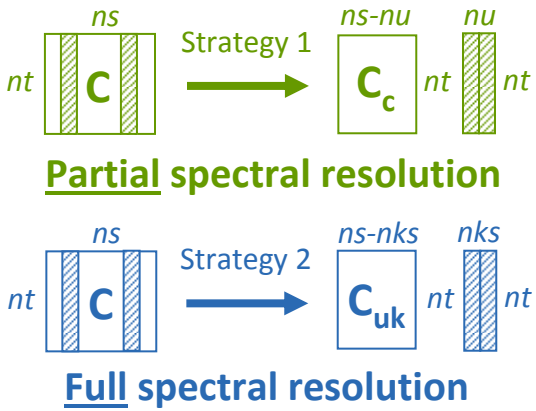
Strategies to treat rank deficiency in kinetic hard-modelling by implicit calibration

Strategies to treat rank deficiency in kinetic hard-modelling by implicit calibration

- **Model (Beer's law) reduction**

  Strategy 1: define *nu* uncoloured species

  Strategy 2: include *nks* known spectra in the analysis (explicit calibration)



Strategies to treat rank deficiency in kinetic hard-modelling by implicit calibration

- **Model (Beer's law) reduction**

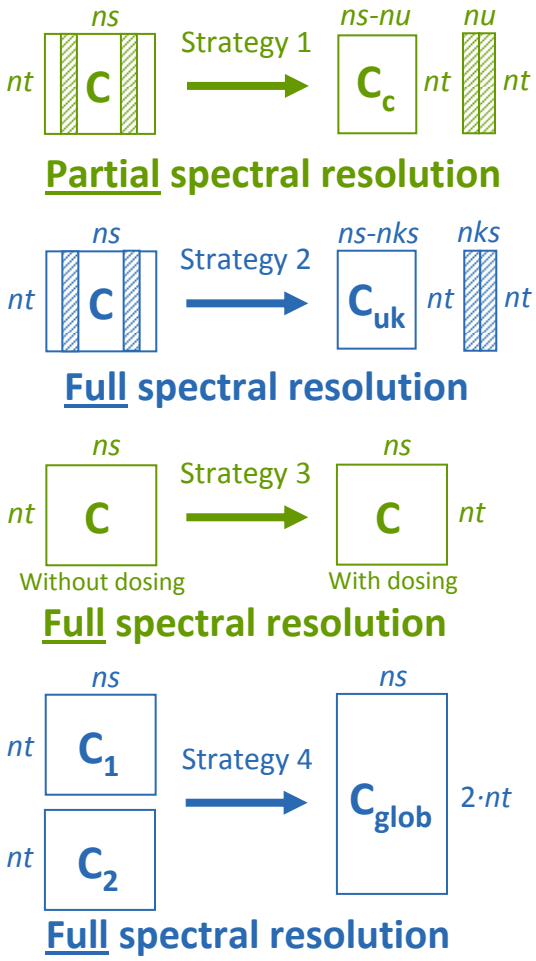
  Strategy 1: define *nu* uncoloured species

  Strategy 2: include *nks* known spectra in the analysis (explicit calibration)

- **Rank augmentation**

  Strategy 3: dose one or more species in *nf* dosing steps

  Strategy 4: perform *ne* additional experiments by varying the initial concentrations (second order global analysis in global mode)



Strategies to treat rank deficiency in kinetic hard-modelling by implicit calibration

- **Model (Beer's law) reduction**

  Strategy 1: define *nu* uncoloured species

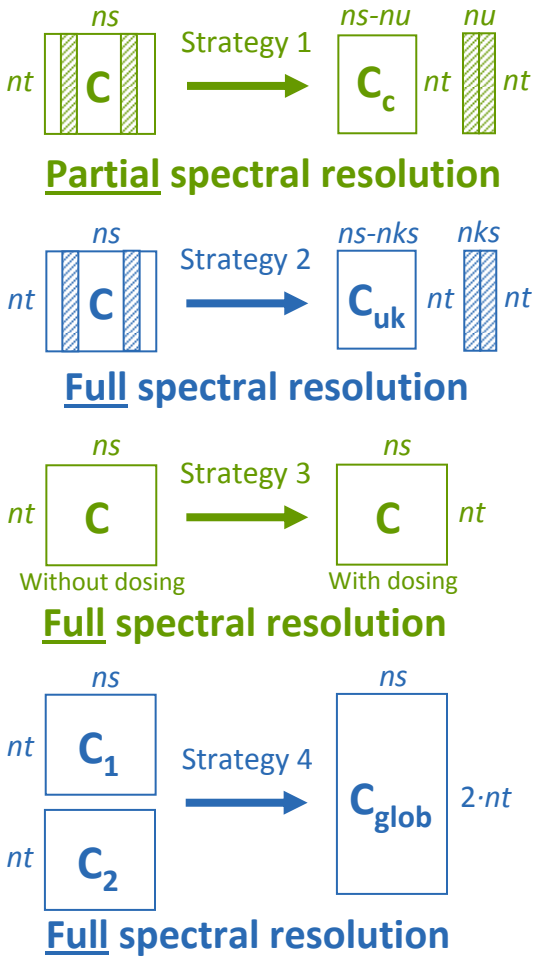
  Strategy 2: include *nks* known spectra in the analysis (explicit calibration)

- **Rank augmentation**

  Strategy 3: dose one or more species in *nf* dosing steps

  Strategy 4: perform *ne* additional experiments by varying the initial concentrations (second order global analysis in global mode)

How to define the species to include in these four Strategies ?



Rank and kernel of the concentration matrix (C)

- ***nc* = rank (Y) = rank (C)**

  number of linearly independent columns or rows of **Y** or **C**

  The rank of **Y** or **C** defines the maximum number of columns (species) to keep in Strategy 1 (*ns* – *nu*) and Strategy 2 (*ns* – *nks*)

- **ker C**

  vector space spanned by the vectors forming the null space **0** when multiplied by **C**

  e.g. ker $\mathbf{C} = \begin{bmatrix} 0.8 & 0.1 \\ -0.3 & -0.7 \\ 0 & 0 \\ 0.5 & -0.6 \end{bmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$

  (1) ker **C** defines a mass balance equation: **C** (ker **C**) = **0**

  (2) ker **C** defines which columns of **C** are:

  (a) linearly dependent (rows comprised by non-zeros entries in ker **C**) or

  (b) linearly independent (rows comprised by zeros in ker **C**)

Interpretation of the kernel of C

	Species without zero rows in ker C	Species with zero rows in ker C
kernel of C	**linearly dependent from the other species**	**linearly independent from the other species**
Strategy 1:	define this species as uncoloured	define this species as coloured
Strategy 2:	providing its pure spectrum avoids rank deficiency	providing its pure spectrum does not avoid rank deficiency
Strategy 3:	dosing this species breaks rank deficiency	dosing this species does not break rank deficiency
Strategy 4:	varying its initial concentration breaks rank deficiency	varying its initial concentration does not break rank deficiency

A time invariant matrix Ω equivalent to C

$$\boldsymbol{\Omega} = \begin{bmatrix} \left(\mu\mathbf{1}\right)^{\bullet\mathbf{E}^T}\,\mathrm{DIAG}\left(\mathbf{k}\right)\mathbf{N} \\ \hline \mathbf{c_0} \\ \hline \mathbf{C_{in}} \\ \hline \mathbf{C_0^{ne}} \end{bmatrix} \quad \left(ns + 1 + nf + ne \times ns - nu - nks\right)$$

μ	an arbitrary positive scalar different from 0 and 1
1 (*ns* x *nr*)	matrix comprised of ones
E (*nr* x *ns*)	matrix of reactant coefficients
$\bullet\mathbf{E}^T$	element-wise raise to the power of $\mathbf{E}^T$
DIAG	operator generating a diagonal matrix from a vector argument

k (1 x *nr*)	vector of rate constants
N (*nr* x *ns*)	matrix of stoichiometric coefficients
$\mathbf{c_0}$ (1 x *ns*)	vector of initial concentrations
$\mathbf{C_{in}}$ (*nf* x *ns*)	matrix of the dosing concentrations corresponding to the *nf* dosing steps
$\mathbf{C_0^{ne}}$ (*ne* x *ns*)	matrix of the varied initial concentrations corresponding to the *ne* additional experiments

Validation of this time invariant matrix:

☑ Equivalence can be mathematically proven

☑ Extensively tested on various mechanisms

Advantages of the time invariant approach:

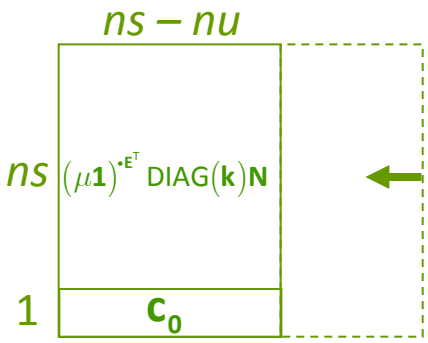
- No numerical integration required
- Analytical (symbolic) relationship between the experimental conditions ($\mathbf{c_0}$, $\mathbf{C_{in}}$,…)

Strategies to treat rank deficiency applied to Ω

Model reduction

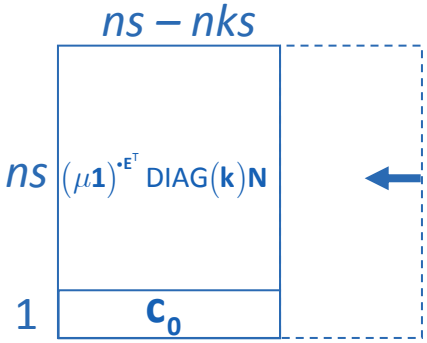
Strategy 1

nu uncoloured species



Strategy 2

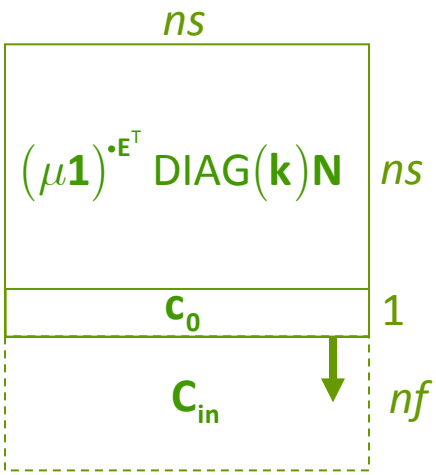
nks known pure spectra



Rank augmentation

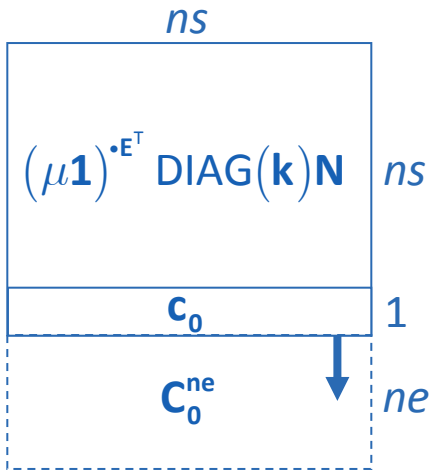
Strategy 3

nf dosing steps

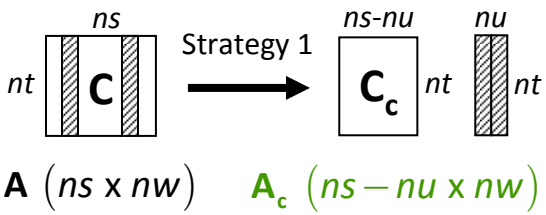


Strategy 4

ne varied initial concentrations



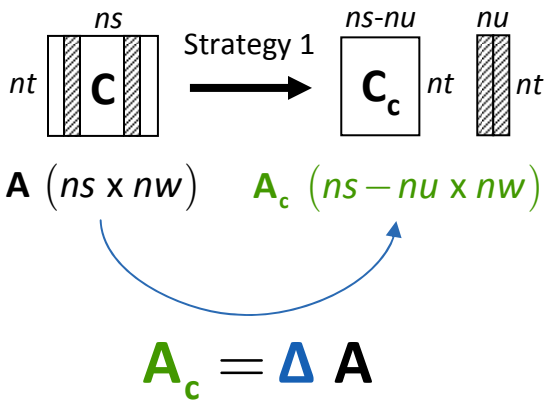
Spectral consequence of Strategy 1 (defining uncoloured species)



Spectral contribution of the *nu* uncoloured species is linearly transferred into the fitted pure component spectra of the coloured species.

The **fitted component spectra A_c** of the (*ns*–*nu*) coloured species are comprised of **linear combinations** of the *ns* **true pure component spectra A**.

Spectral consequence of Strategy 1 (defining uncoloured species)



$$\mathbf{A_c} = \mathbf{\Delta}\ \mathbf{A}$$

(Spectral balance)

Spectral contribution of the *nu* uncoloured species is linearly transferred into the fitted pure component spectra of the coloured species.

The **fitted component spectra $\mathbf{A_c}$** of the (*ns*–*nu*) coloured species are comprised of **linear combinations** of the *ns* **true pure component spectra A**.

$$\mathbf{\Delta}\ (ns - nu \times ns) = \mathbf{C_c^+ C} = \left(\mathbf{\Omega}\big|_{\substack{\text{comprised of}\\ \text{coloured species}}}\right)^+ \mathbf{\Omega}\big|_{\substack{\text{comprised of}\\ \text{all species}}}$$

Example: calculation of the kernel

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\boldsymbol{\Omega} = \left[\frac{(\mu\mathbf{1})^{\bullet \mathbf{E}^T}\ \mathrm{DIAG}(\mathbf{k})\mathbf{N}}{\mathbf{c_0}}\right] = \begin{array}{c} \begin{matrix} A & B & C & D \end{matrix} \\ \left[\begin{array}{cccc} -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{array}\right] \end{array}$$

$$\ker\ \boldsymbol{\Omega} = \begin{bmatrix} -\alpha \\ 1 \\ 1-\alpha \\ 1-2\alpha \end{bmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

Matlab code (8 lines)

```
>> syms c0A alpha mu k1 k2
>> N       = [-1 -1 1 0; -1 0 -1 1];
>> E       = [1 1 0 0; 1 0 1 0];
>> k       = [k1, k2];
>> c0      = [c0A, alpha*c0A, 0, 0];
>> one     = ones(size(E'));
>> omega   = [(mu*one).^(E')*diag(k)*N; c0];
>> null(omega)
ans =
      -alpha
           1
       1-alpha
     1-2*alpha
```

$$\alpha = \frac{c_{0,B}}{c_{0,A}}$$

Example: calculation of the kernel

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\boldsymbol{\Omega} = \left[\frac{(\mu\mathbf{1})^{\bullet \mathbf{E}^T}\ \mathrm{DIAG}(\mathbf{k})\mathbf{N}}{\mathbf{c_0}}\right] = \begin{array}{c} \begin{matrix} A & B & C & D \end{matrix} \\ \left[\begin{array}{cccc} -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{array}\right] \end{array}$$

$$\ker\ \boldsymbol{\Omega} = \begin{array}{c} \left[\begin{array}{c} -\alpha \\ 1 \\ 1-\alpha \\ 1-2\alpha \end{array}\right] \end{array} \begin{array}{l} A \\ B \\ C \\ D \end{array}$$

If $\alpha = 1$, species C is linearly independent from the others

Matlab code (8 lines)

```
>> syms c0A alpha mu k1 k2
>> N          = [-1 -1 1 0; -1 0 -1 1];
>> E          = [1 1 0 0; 1 0 1 0];
>> k          = [k1, k2];
>> c0         = [c0A, alpha*c0A, 0, 0];
>> one        = ones(size(E'));
>> omega      = [(mu*one).^(E')*diag(k)*N; c0];
>> null(omega)
ans =
      -alpha
           1
     1-alpha
   1-2*alpha
```

$$\alpha = \frac{c_{0,B}}{c_{0,A}}$$

Example: calculation of the kernel

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\boldsymbol{\Omega} = \left[\frac{(\mu\mathbf{1})^{\bullet \mathbf{E}^T}\ \mathrm{DIAG}(\mathbf{k})\mathbf{N}}{\mathbf{c_0}}\right] = \left[\begin{array}{cccc} A & B & C & D \\ -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{array}\right]$$

$$\ker \boldsymbol{\Omega} = \begin{bmatrix} -\alpha \\ 1 \\ 1-\alpha \\ 1-2\alpha \end{bmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

If $\alpha = 1$, species *C* is linearly independent from the others

If $\alpha = 0.5$, species *D* is linearly independent from the others

Matlab code (8 lines)

```
>> syms c0A alpha mu k1 k2
>> N       = [-1 -1 1 0; -1 0 -1 1];
>> E       = [1 1 0 0; 1 0 1 0];
>> k       = [k1, k2];
>> c0      = [c0A, alpha*c0A, 0, 0];
>> one     = ones(size(E'));
>> omega   = [(mu*one).^(E')*diag(k)*N; c0];
>> null(omega)
ans =
      -alpha
           1
     1-alpha
   1-2*alpha
```

$$\alpha = \frac{c_{0,B}}{c_{0,A}}$$

Example: calculation of the kernel

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\boldsymbol{\Omega} = \left[\frac{(\mu\mathbf{1})^{\bullet \mathbf{E}^T}\,\mathrm{DIAG}(\mathbf{k})\mathbf{N}}{\mathbf{c_0}}\right] = \left[\begin{array}{cccc} A & B & C & D \\ -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{array}\right]$$

$$\ker \boldsymbol{\Omega} = \begin{bmatrix} -\alpha \\ 1 \\ 1-\alpha \\ 1-2\alpha \end{bmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

If $\alpha = 1$, species *C* is linearly independent from the others

If $\alpha = 0.5$, species *D* is linearly independent from the others

If $\alpha \neq 1$ or 0.5, all species are linearly dependent from the others

Matlab code (8 lines)

```
>> syms c0A alpha mu k1 k2
>> N       = [-1 -1 1 0; -1 0 -1 1];
>> E       = [1 1 0 0; 1 0 1 0];
>> k       = [k1, k2];
>> c0      = [c0A, alpha*c0A, 0, 0];
>> one     = ones(size(E'));
>> omega   = [(mu*one).^(E')*diag(k)*N; c0];
>> null(omega)
ans =
      -alpha
           1
       1-alpha
     1-2*alpha
```

$$\alpha = \frac{c_{0,B}}{c_{0,A}}$$

Example: design of experiments

$$A + B \xrightarrow{k_1} C$$
$$A + C \xrightarrow{k_2} D$$

$$\ker \mathbf{\Omega}\ (4 \times 1) = \begin{bmatrix} -\alpha \\ 1 \\ 1-\alpha \\ 1-2\alpha \end{bmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

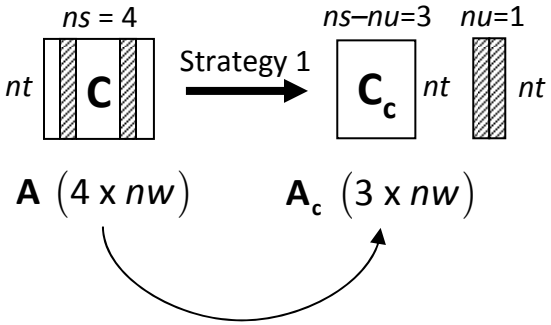
ns = 4 species

Dimension of the kernel is one, i.e. only one species has to be considered in Strategies 1 – 4

Strategy	$\alpha = 1$				$\alpha = 0.5$				$\alpha \neq 1$ or 0.5			
	A	***B***	***C***	***D***	***A***	***B***	***C***	***D***	***A***	***B***	***C***	***D***
(1): define one uncoloured species	0/1	0/1	0	0/1	0/1	0/1	0/1	0	0/1	0/1	0/1	0/1
(2): provide one known pure spectrum	0/1	0/1	0	0/1	0/1	0/1	0/1	0	0/1	0/1	0/1	0/1
(3): dose one species	0/1	0/1	0	0/1	0/1	0/1	0/1	0	0/1	0/1	0/1	0/1
(4): vary one initial concentration	0/1	0/1	0	0/1	0/1	0/1	0/1	0	0/1	0/1	0/1	0/1

1: uncoloured, pure spectrum provided, dosed or initial concentration varied
0: coloured, pure spectrum not provided, not dosed or initial concentration not varied

Example: spectral consequence of Strategy 1 (defining uncoloured species)



$$\mathbf{A_c} = \mathbf{\Delta}\,\mathbf{A}$$

(Spectral balance)

Let's define species *A* uncoloured

All species

$$\mathbf{\Omega}\big|_{\text{comprised of all species}} = \begin{array}{c} \begin{matrix} A & B & C & D \end{matrix} \\ \left[\begin{array}{cccc} -\mu k_1 - \mu k_2 & -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 - k_2 & -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 - \mu k_2 & -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 - k_2 & -k_1 & k_1 - k_2 & k_2 \\ \hline c_{0,A} & \alpha c_{0,A} & 0 & 0 \end{array}\right] \end{array}$$

Coloured species

$$\mathbf{\Omega}\big|_{\text{comprised of coloured species}} = \begin{array}{c} \begin{matrix} 'B' & 'C' & 'D' \end{matrix} \\ \left[\begin{array}{ccc} -\mu k_1 & \mu k_1 - \mu k_2 & \mu k_2 \\ -\mu k_1 & \mu k_1 - k_2 & k_2 \\ -k_1 & k_1 - \mu k_2 & \mu k_2 \\ -k_1 & k_1 - k_2 & k_2 \\ \hline \alpha c_{0,A} & 0 & 0 \end{array}\right] \end{array}$$

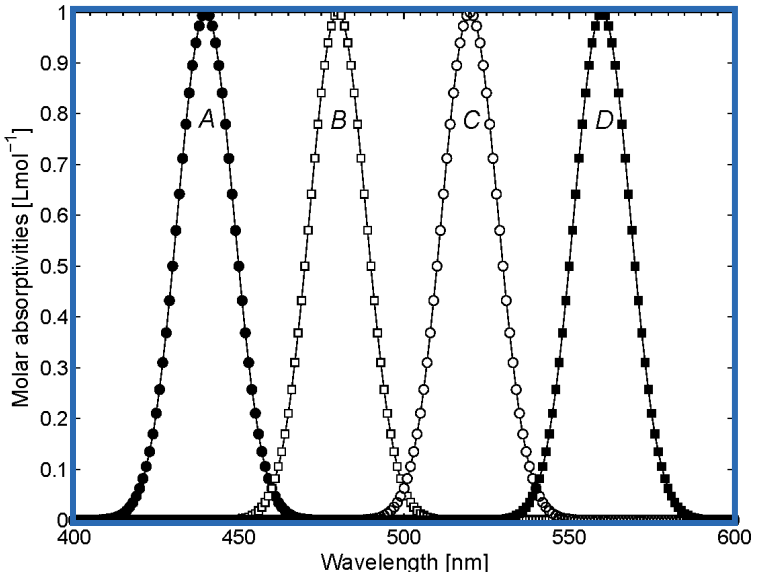
All species

$$\mathbf{\Delta}\ (3 \text{ x } 4) = \left(\mathbf{\Omega}\big|_{\text{comprised of coloured species}}\right)^{+} \mathbf{\Omega}\big|_{\text{comprised of all species}} = \begin{array}{c} \begin{matrix} A & B & C & D \end{matrix} \\ \left[\begin{array}{cccc} \alpha^{-1} & 1 & 0 & 0 \\ \alpha^{-1} - 1 & 0 & 1 & 0 \\ \alpha^{-1} - 2 & 0 & 0 & 1 \end{array}\right] \end{array} \begin{matrix} 'B' \\ 'C' \\ 'D' \end{matrix}$$

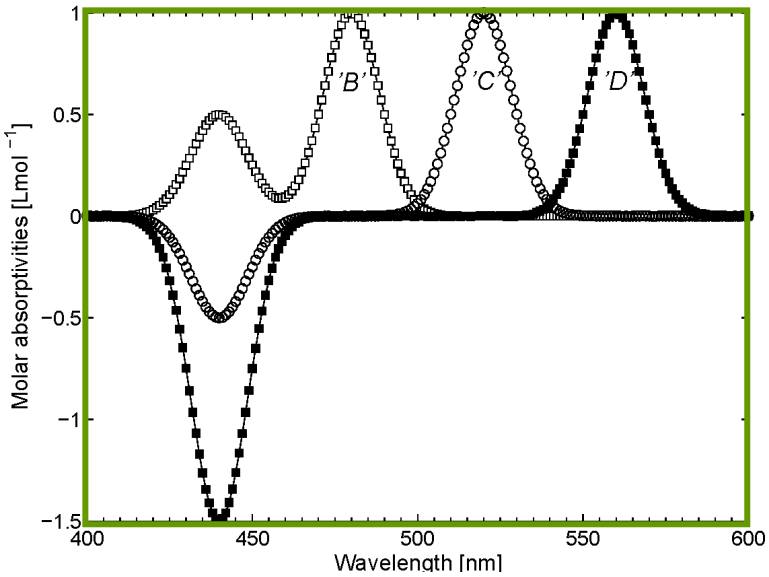
Coloured species

Example: spectral consequence of Strategy 1 (defining uncoloured species)

True: **A** (4 x *nw*)

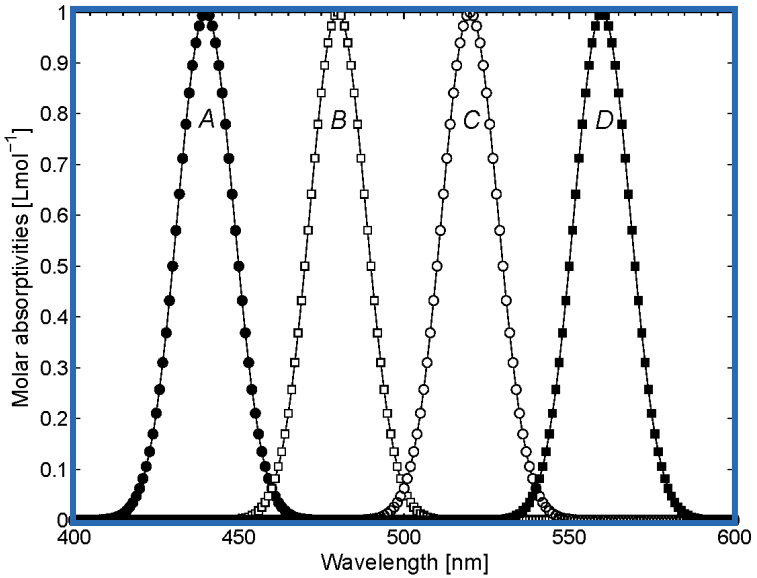


Fitted: $\mathbf{A}_c$ (3 x *nw*)

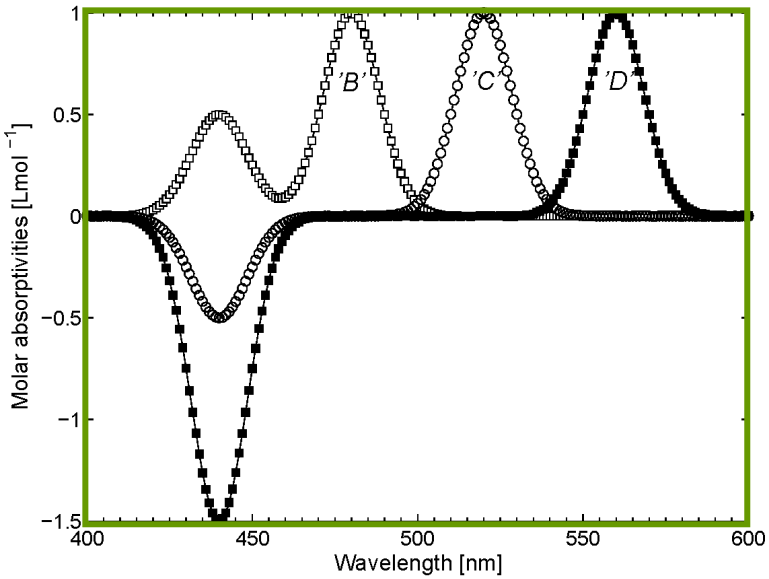


Example: spectral consequence of Strategy 1 (defining uncoloured species)

True: **A** (4 x *nw*)



Fitted: $\mathbf{A}_c$ (3 x *nw*)



$$\begin{bmatrix} \mathbf{a}_{c\ 'B',:} \\ \mathbf{a}_{c\ 'C',:} \\ \mathbf{a}_{c\ 'D',:} \end{bmatrix} = \mathbf{\Delta}\ \mathbf{A} = \begin{bmatrix} \alpha^{-1} & 1 & 0 & 0 \\ \alpha^{-1}-1 & 0 & 1 & 0 \\ \alpha^{-1}-2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{a}_{A,:} \\ \mathbf{a}_{B,:} \\ \mathbf{a}_{C,:} \\ \mathbf{a}_{D,:} \end{bmatrix} = \begin{bmatrix} 0.5 \\ -0.5 \\ -1.5 \end{bmatrix} \mathbf{a}_{A,:} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \mathbf{a}_{B,:} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \mathbf{a}_{C,:} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mathbf{a}_{D,:} \quad \text{when } \alpha = 2$$

Kinetic and spectral validation

- In kinetic hard-modelling, the validation of a kinetic mechanism is based on:

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A}$$

  - **(C) The kinetic consistency** and the reproducibility of fitted kinetic parameters under different experimental conditions.
  - **(A) The spectral consistency** of fitted pure component spectra compared to independently measured ones.

Kinetic and spectral validation

- In kinetic hard-modelling, the validation of a kinetic mechanism is based on:

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A}$$

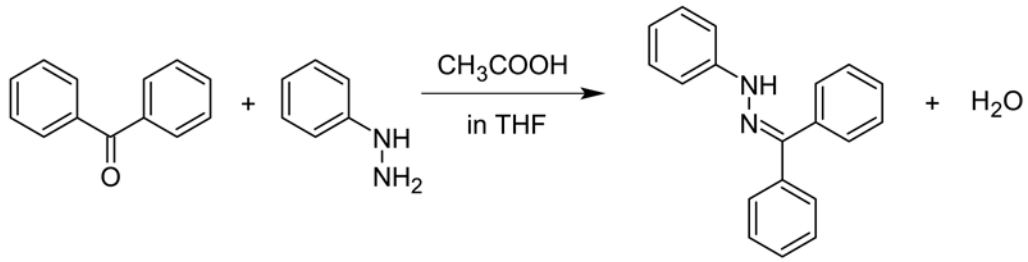
  - **(C) The kinetic consistency** and the reproducibility of fitted kinetic parameters under different experimental conditions.
  - **(A) The spectral consistency** of fitted pure component spectra compared to independently measured ones.

The spectral validation is facilitated, when Strategy 1 is used, as linear combinations of true pure component spectra can now be explained (**Δ**) !

$$\mathbf{\Delta}\ (ns - nu \times ns) = \mathbf{C}_c^{+}\mathbf{C} = \left(\mathbf{\Omega}\big|_{\substack{\text{comprised of}\\ \text{coloured species}}}\right)^{+} \mathbf{\Omega}\big|_{\substack{\text{comprised of}\\ \text{all species}}}$$

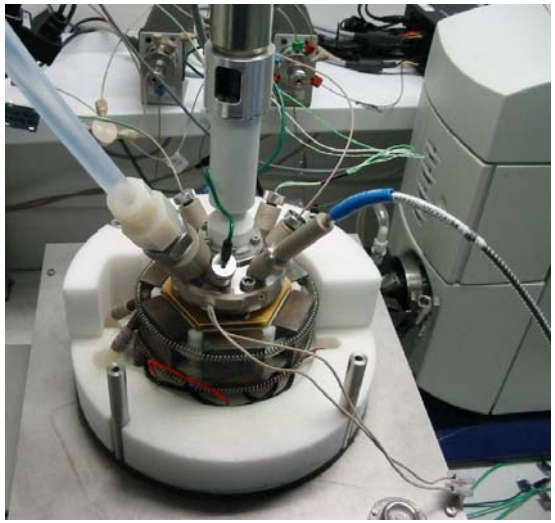
Experimental reaction

Overall reaction:



Kinetic mechanism:

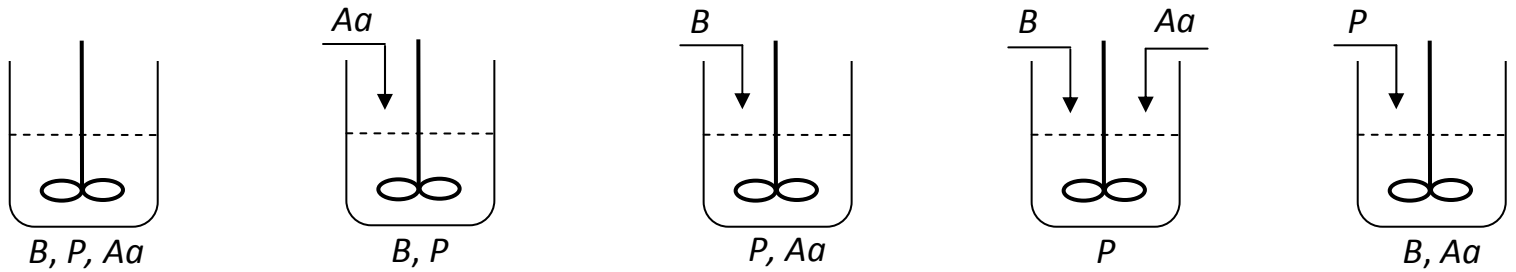
$$B + P + Aa \xrightarrow{k} BP + Aa$$



Reactor: CRC.v4 with FT-IR and UV-vis

Experimental conditions:

25°C, followed in mid-IR (1200–1650 cm^{-1}) and UV-vis (240–400 nm).



Kinetic validation of the model

Experimental conditions	Strategy [a]	Mid-IR		UV-vis	
		k [b]	Published k [b]	k [b]	Published k [b]
Batch conditions	only (1)	**1.58** (± 0.02)	**1.74** (± 0.05) [c]	**1.63** (± 0.02)	**1.77** (± 0.03) [c]
	(1) + (2)	**1.58** (± 0.02)	**1.40** [d]	**1.63** (± 0.02)	**1.51** [d]
	(1) + (4)	**1.57** (± 0.02)		**1.62** (± 0.02)	
Dosing *Aa*	(1) + (3)	**1.60** (± 0.04)		**1.65** (± 0.04)	
Dosing *B*	(1) + (3)	**1.60** (± 0.03)		**1.64** (± 0.03)	
	(1) + (3)	**1.55** (± 0.02)		**1.57** (± 0.02)	
Dosing *B* + *Aa*	only (3)	**1.59** (± 0.06)		**1.64** (± 0.06)	
Dosing *P*	(1) + (3)	**1.62** (± 0.06)		**1.67** (± 0.06)	

a) (1): defining uncoloured species, (2): providing known pure spectra, (3): dosing, (4): second order global analysis
b) in $L^2 \cdot mol^{-2} \cdot s^{-1} \times 10^{-4}$
c) Billeter et al., Chemom. Intell. Lab. Syst., 93 (2008), 120-131
c) Carvalho et al., Talanta, 68 (2006), 1190-1200

$$B + P + Aa \xrightarrow{k_1} BP + Aa$$

Kinetic validation of the model

Experimental conditions	Strategy [a]	Mid-IR		UV-vis	
		k [b]	Published k [b]	k [b]	Published k [b]
Batch conditions	only (1)	**1.58** (± 0.02)	**1.74** (± 0.05) [c]	**1.63** (± 0.02)	**1.77** (± 0.03) [c]
	(1) + (2)	**1.58** (± 0.02)	**1.40** [d]	**1.63** (± 0.02)	**1.51** [d]
	(1) + (4)	**1.57** (± 0.02)		**1.62** (± 0.02)	
Dosing *Aa*	(1) + (3)	**1.60** (± 0.04)		**1.65** (± 0.04)	
Dosing *B*	(1) + (3)	**1.60** (± 0.03)		**1.64** (± 0.03)	
	(1) + (3)	**1.55** (± 0.02)		**1.57** (± 0.02)	
Dosing *B* + *Aa*	only (3)	**1.59** (± 0.06)		**1.64** (± 0.06)	
Dosing *P*	(1) + (3)	**1.62** (± 0.06)		**1.67** (± 0.06)	

a) (1): defining uncoloured species, (2): providing known pure spectra, (3): dosing, (4): second order global analysis
b) in $L^2 \cdot mol^{-2} \cdot s^{-1} \times 10^{-4}$
c) Billeter et al., Chemom. Intell. Lab. Syst., 93 (2008), 120-131
c) Carvalho et al., Talanta, 68 (2006), 1190-1200

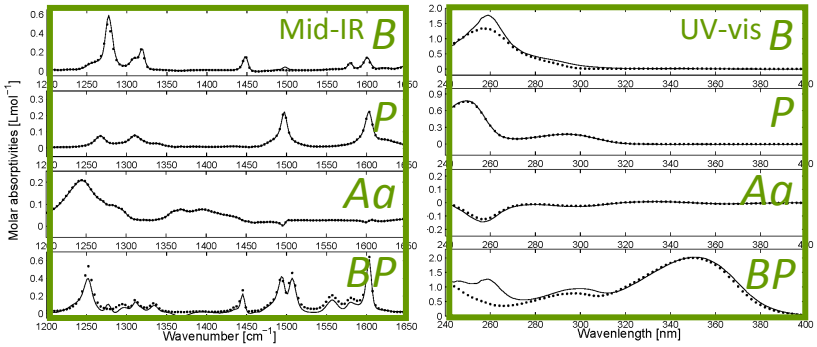
$$B + P + Aa \xrightarrow{k_1} BP + Aa$$

The model is kinetically validated

Spectral validation of the model

$$B + P + Aa \xrightarrow{k_1} BP + Aa$$

Strategy (3): dosing



Species *B* and *Aa* dosed

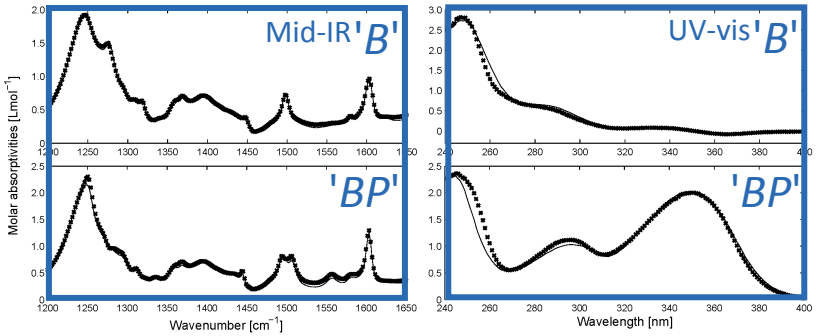
Full spectral resolution

—— Fitted spectra ●●● Measured spectra **XXX** Predicted spectra

Spectral validation of the model

$$B + P + Aa \xrightarrow{k_1} BP + Aa$$

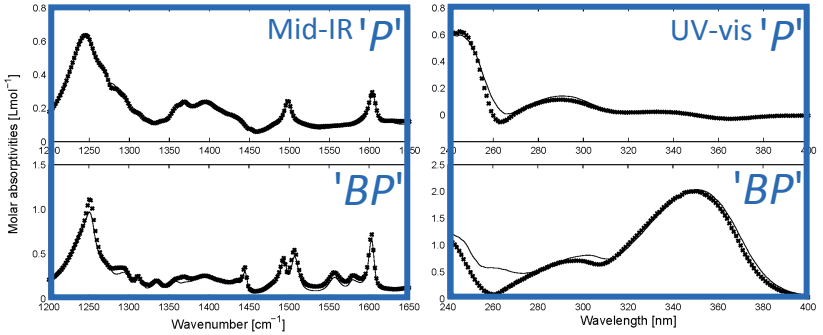
Strategy (1): uncoloured species



Species *P* and *Aa* set uncoloured

Partial spectral resolution

Strategy (1)+(2): provided known spectrum



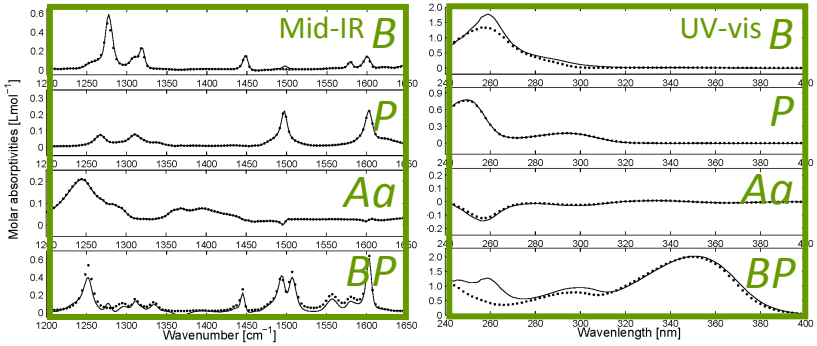
Pure spectrum of *B* provided

Aa set uncoloured

Partial spectral resolution

The model can be spectroscopically validated **with** and **without** rank deficiency!

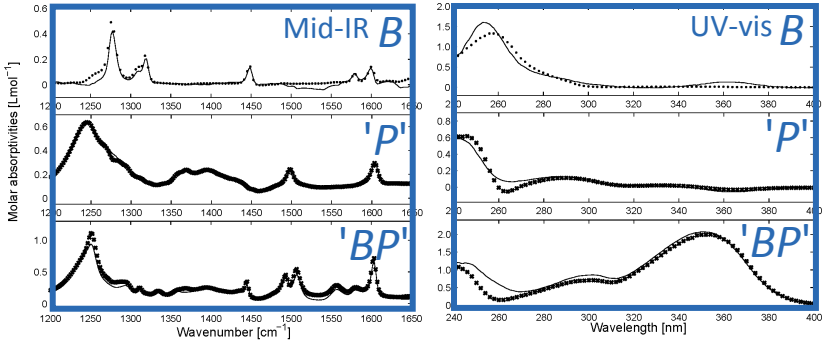
Strategy (3): dosing



Species *B* and *Aa* dosed

Full spectral resolution

Strategy (1)+(4): second order global analysis



Initial concentration of *B* varied, *Aa* set uncoloured

Partial spectral resolution

—— Fitted spectra ●●● Measured spectra **xxx** Predicted spectra

Conclusion: general methods

- **Chemometrics**
  - **Extraction** of information from complex multivariate signals
  - **Identification** of significant contributions (SVD, PCA)
  - **Support** for the elaboration of models
- **Kinetic hard-modelling of spectroscopic data**
  - **Determination** of kinetic parameters (e.g. rate constants)
  - **Assessment** of kinetic models by comparing fitted and independently measured pure component spectra (direct fitting)
  - **Calibration-free** method (implicit calibration)

Conclusion: specific methods in hard-modelling

- **Method of error propagation**
  - **Determination** of reliable uncertainties in the fitted kinetic parameters
  - **Prediction** of the experimental conditions minimising the uncertainties
- **Method for prediction of rank deficiencies and spectral validation**
  - **Design of rank deficient experiments** when Strategies 1 – 4 have to be used
  - **Interpretation** of the fitted pure component spectra when Strategy 1 (defining uncoloured species) is used
  - **Spectral validation** of rank deficient kinetic models

Acknowledgements

Group of Safety and Environmental Technology

Prof. Dr. K. Hungerbühler — **Head of the group**

Subgroup of Reaction Analysis

Dr. Y.-M. Neuhold	**Senior scientist, subgroup leader**
Dr. G. Puxty	**Former subgroup leader** (now at CSIRO)
Dr. G. Richner	Post-doctoral scientist
S. Cap	Doctoral student
T. Godany	Doctoral student
S. Gianoli	Doctorat student

Subgroup of Reaction Process Design and Optimisation

Dr. S. Papadokonstantakis — Senior scientist, subgroup leader

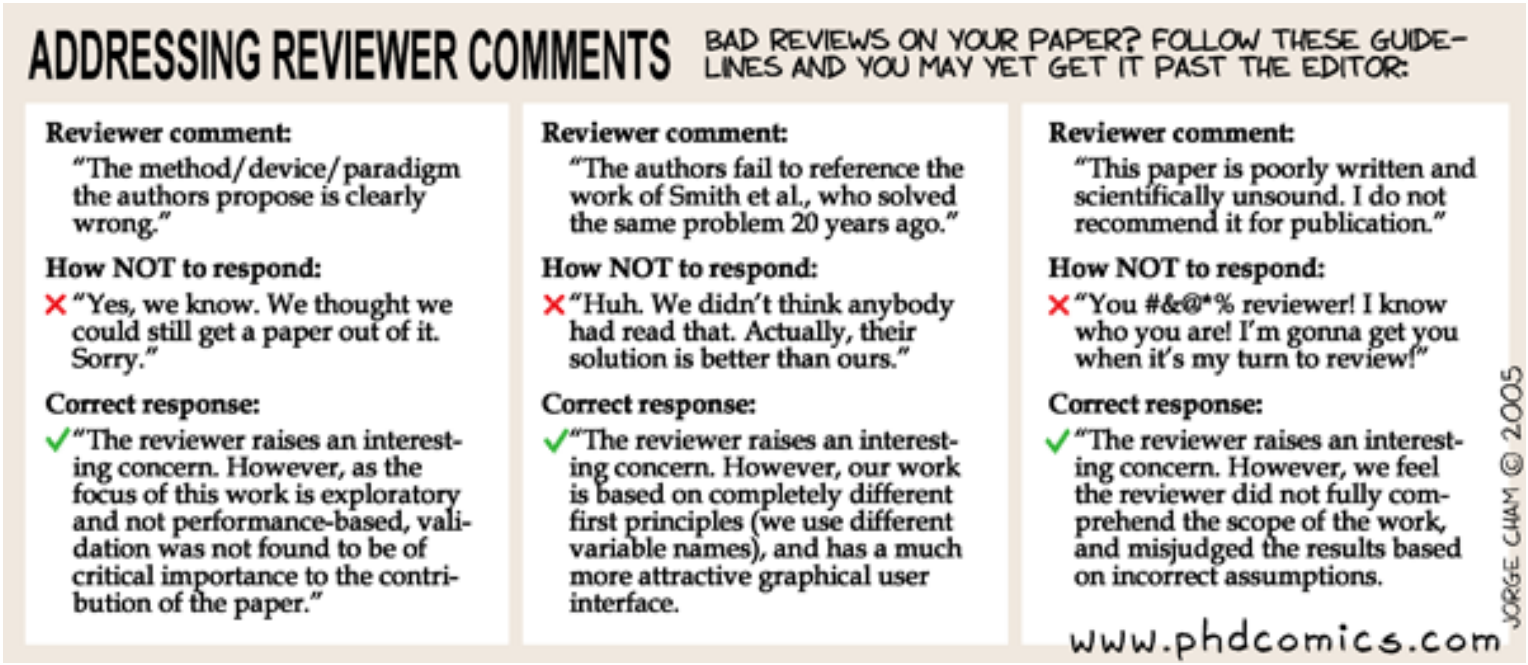
Subgroup of Separation Technologies

Dr. L. Simon — Senior scientist, subgroup leader

and all members of the group of Safety and Environmental Technology

Thank you for your attention

Any question or comment?


ADDRESSING REVIEWER COMMENTS
BAD REVIEWS ON YOUR PAPER? FOLLOW THESE GUIDELINES AND YOU MAY YET GET IT PAST THE EDITOR:
Reviewer comment:
"The method/device/paradigm the authors propose is clearly wrong."
How NOT to respond:
✗ "Yes, we know. We thought we could still get a paper out of it. Sorry."
Correct response:
✓ "The reviewer raises an interesting concern. However, as the focus of this work is exploratory and not performance-based, validation was not found to be of critical importance to the contribution of the paper."
Reviewer comment:
"The authors fail to reference the work of Smith et al., who solved the same problem 20 years ago."
How NOT to respond:
✗ "Huh. We didn't think anybody had read that. Actually, their solution is better than ours."
Correct response:
✓ "The reviewer raises an interesting concern. However, our work is based on completely different first principles (we use different variable names), and has a much more attractive graphical user interface.
Reviewer comment:
"This paper is poorly written and scientifically unsound. I do not recommend it for publication."
How NOT to respond:
✗ "You #&@*% reviewer! I know who you are! I'm gonna get you when it's my turn to review!"
Correct response:
✓ "The reviewer raises an interesting concern. However, we feel the reviewer did not fully comprehend the scope of the work, and misjudged the results based on incorrect assumptions.
JORGE CHAM © 2005
www.phdcomics.com