ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Laboratoire d'Automatique

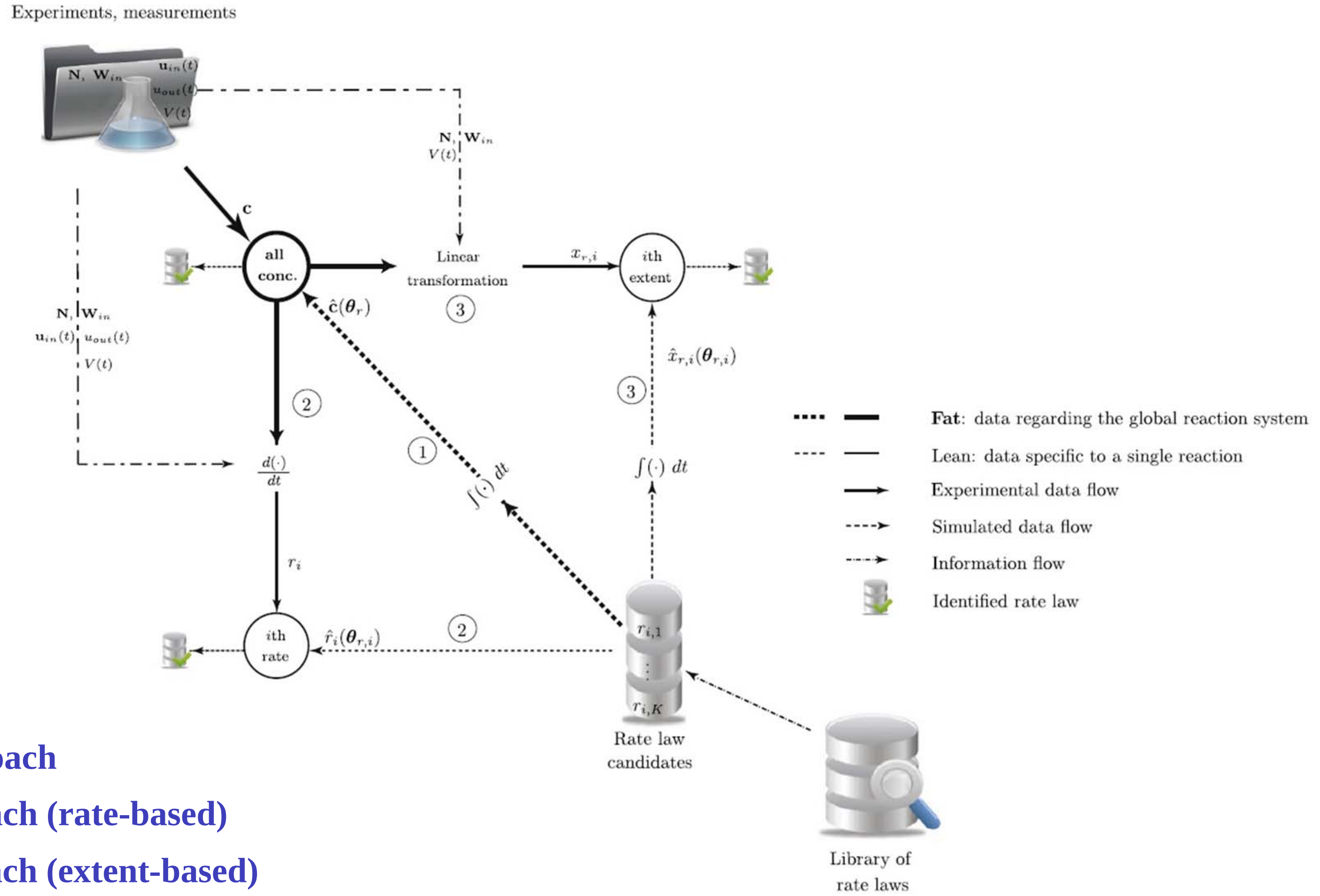
Simultaneous or incremental identification of reaction systems ?

4th EuCheMS Chemistry Congress
26 – 30 August 2012, Prague – Czech Republic

Julien Billeter, Sriniketh Srinivasan and Dominique Bonvin

Ecole Polytechnique Fédérale de Lausanne
Laboratoire d'Automatique
Switzerland

Kinetic investigation
From measurements to rate expressions

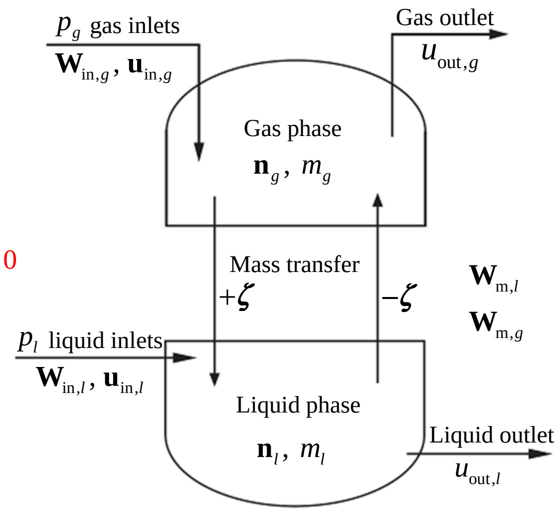


① **Simultaneous approach**

② **Incremental approach (rate-based)**

③ **Incremental approach (extent-based)**

Differential mole balance equations



Gas phase

S_g species, p_{m} mass transfers, p_g inlets and 1 outlet

$$\dot{\mathbf{n}}_g(t) = -\mathbf{W}_{\text{m},g}\zeta(t) + \mathbf{W}_{\text{in},g}\mathbf{u}_{\text{in},g}(t) - \frac{u_{\text{out},g}(t)}{m_g(t)}\mathbf{n}_g(t), \qquad \mathbf{n}_g(0) = \mathbf{n}_{g0}$$

Liquid phase

S_l species, R reactions, p_{m} mass transfers, p_l inlets and 1 outlet

$$\underset{(S_l\times 1)}{\dot{\mathbf{n}}_l(t)} = \underset{(R\times S_l)}{\mathbf{N}^{\mathrm{T}}} V_l(t)\underset{(R\times 1)}{\mathbf{r}(t)} + \underset{(S_l\times p_{\text{m}})(p_{\text{m}}\times 1)}{\mathbf{W}_{\text{m},l}\zeta(t)} + \underset{(S_l\times p_l)(p_l\times 1)}{\mathbf{W}_{\text{in},l}\mathbf{u}_{\text{in},l}(t)} - \frac{u_{\text{out},l}(t)}{m_l(t)}\mathbf{n}_l(t), \qquad \mathbf{n}_l(0) = \underset{(S_l\times 1)}{\mathbf{n}_{l0}}$$

Assumptions: G and L phases are homogeneous, reactions take place in the L phase only,
mass transfers occur with no accumulation in the film,
mass transfer rates are positive from G to L phase.

Remark: For a subset of measured concentrations, $S_l = S_{l,a} + S_{l,u}$, dimensions are adapted…

Simultaneous model identification

The simultaneous model identification proceeds in **one step**:

Model identification

A kinetic model comprising **all reaction and mass transfer** rate laws is postulated and a **coupled** regression problem is solved using the **integral method** of parameters estimation:

$$\min_{\boldsymbol{\theta}_{\mathrm{rm}}} \left\| \mathbf{n}_{l,a}(t) - \hat{\mathbf{n}}_{l,a}(t, \boldsymbol{\theta}_{\mathrm{rm}}) \right\|^2$$

$$\text{s.t.} \quad \dot{\hat{\mathbf{n}}}_l(t, \boldsymbol{\theta}_{\mathrm{rm}}) = \mathbf{N}^{\mathrm{T}} V_l(t)\, \mathbf{r}(t, \boldsymbol{\theta}_{\mathrm{r}}) + \mathbf{W}_{\mathrm{m},l}\, \boldsymbol{\zeta}(t, \boldsymbol{\theta}_{\mathrm{m}}) + \mathbf{W}_{\mathrm{in},l}\, \mathbf{u}_{\mathrm{in},l}(t) - \frac{u_{\mathrm{out},l}(t)}{m_l(t)} \hat{\mathbf{n}}_l(t), \qquad \hat{\mathbf{n}}_l(0) = \mathbf{n}_{l0}$$

$$\boldsymbol{\theta}_{\mathrm{r}}^{L} \leq \boldsymbol{\theta}_{\mathrm{r}} \leq \boldsymbol{\theta}_{\mathrm{r}}^{U}$$

$$\boldsymbol{\theta}_{\mathrm{m}}^{L} \leq \boldsymbol{\theta}_{\mathrm{m}} \leq \boldsymbol{\theta}_{\mathrm{m}}^{U}$$

Incremental model identification

The kinetic problem is decomposed into sub-problems of lower complexity.

The incremental model identification proceeds in **two steps**:

1. **Transformation**
   Computation of the contribution of each reaction and
   each mass transfer as *rates* or *extents* (+ *state reconstruction* if necessary)

2. **Model identification**
   Individual identification of each reaction rate law and
   each mass-transfer rate expression from *rates* or *extents*

Rate-based incremental identification

1. Transformation

Computation of rates via **differentiation** of the measured concentrations

$$\left[V_l(t)\mathbf{r}(t) \quad \boldsymbol{\zeta}(t)\right]^{\mathrm{T}} = \underbrace{\left[\mathbf{N}_a^{\mathrm{T}} \quad \mathbf{W}_{\mathrm{m},l,a}\right]^{+}}_{\text{rank} = R + p_{\mathrm{m}}} \left(\frac{d}{dt}\left[\mathbf{n}_{l,a}\right](t) - \mathbf{W}_{\mathrm{in},l,a}\mathbf{u}_{\mathrm{in},l,a}(t) + \frac{u_{\mathrm{out},l}(t)}{m_l(t)}\mathbf{n}_{l,a}(t)\right)$$

State reconstruction (by integration):

$$\dot{\mathbf{n}}_{l,u}(t) = \mathbf{N}_u^{\mathrm{T}} V_l(t)\mathbf{r}(t) + \mathbf{W}_{\mathrm{m},l,u}\boldsymbol{\zeta}(t) + \mathbf{W}_{\mathrm{in},l,u}\mathbf{u}_{\mathrm{in},l}(t) - \frac{u_{\mathrm{out},l}(t)}{m_l(t)}\mathbf{n}_{l,u}(t), \qquad \mathbf{n}_{l,u}(0) = \mathbf{n}_{l0,\mathrm{u}}$$

2. Model identification

A rate law for **each rate of reaction** and **each rate of mass transfer** is postulated and $R+p_{\mathrm{m}}$ regression problems are solved **individually**:

$$\min_{\boldsymbol{\theta}_{\mathrm{r},i}} \left\| r_i(t) - \hat{r}_i(t, \boldsymbol{\theta}_{\mathrm{r},i}) \right\|^2 \qquad \boldsymbol{\theta}_{\mathrm{r,i}}^L \le \boldsymbol{\theta}_{\mathrm{r,i}} \le \boldsymbol{\theta}_{\mathrm{r,i}}^U, \quad i = 1, \ldots, R$$

$$\min_{\boldsymbol{\theta}_{\mathrm{m},j}} \left\| \zeta_j(t) - \hat{\zeta}_j(t, \boldsymbol{\theta}_{\mathrm{m},j}) \right\|^2 \qquad \boldsymbol{\theta}_{\mathrm{m},j}^L \le \boldsymbol{\theta}_{\mathrm{m},j} \le \boldsymbol{\theta}_{\mathrm{m},j}^U, \quad j = 1, \ldots, p_{\mathrm{m}}$$

Extent-based incremental identification (Transformation)

1. **Transformation**

1a: Computation of $R+p_{\mathrm{m}}+p_l+1$ extents

$$\begin{bmatrix} \mathbf{x}_{\mathrm{r}}(t) \\ \mathbf{x}_{\mathrm{m},l}(t) \\ \mathbf{x}_{\mathrm{in},l}(t) \\ \lambda_l(t) \end{bmatrix} = \mathscr{L}_a \mathbf{n}_{l,a}(t)$$

$$\begin{cases} \dot{\mathbf{x}}_{\mathrm{r}}(t) = V_l(t)\mathbf{r}(t) - \tau_{\mathrm{out},l}^{-1}(t)\mathbf{x}_{\mathrm{r}}(t), & \mathbf{x}_{\mathrm{r}}(0) = \mathbf{0}_R \\ \dot{\mathbf{x}}_{\mathrm{m},l}(t) = \boldsymbol{\zeta}(t) - \tau_{\mathrm{out},l}^{-1}(t)\mathbf{x}_{\mathrm{m},l}(t), & \mathbf{x}_{\mathrm{m},l}(0) = \mathbf{0}_{p_{\mathrm{m}}} \\ \dot{\mathbf{x}}_{\mathrm{in},l}(t) = \mathbf{u}_{\mathrm{in},l}(t) - \tau_{\mathrm{out},l}^{-1}(t)\mathbf{x}_{\mathrm{in},l}(t), & \mathbf{x}_{\mathrm{in},l}(0) = \mathbf{0}_{p_l} \\ \dot{\lambda}_l(t) = -\tau_{\mathrm{out},l}^{-1}(t)\lambda_l(t), & \lambda_l(0) = 1 \end{cases}$$

with $\tau_{\mathrm{out},l}^{-1}(t) = u_{\mathrm{out},l}(t)m_l^{-1}(t)$

$$\mathrm{rank}\left(\begin{bmatrix} \mathbf{N}_a^{\mathrm{T}} & \mathbf{W}_{\mathrm{m},l,a} & \mathbf{W}_{\mathrm{in},l,a} & \mathbf{n}_{l0,a} \end{bmatrix}\right) = R + p_{\mathrm{m}} + p_l + 1$$

1b: Computation of $R+p_{\mathrm{m}}$ extents (rank $< R + p_{\mathrm{m}} + p_l + 1$)

$$\begin{bmatrix} \mathbf{x}_{\mathrm{r}}(t) \\ \mathbf{x}_{\mathrm{m},l}(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{N}_a^{\mathrm{T}} & \mathbf{W}_{\mathrm{m},l,a} \end{bmatrix}^{+}}_{\text{rank } = R + p_{\mathrm{m}}} \left(\mathbf{n}_{l,a}(t) - \mathbf{W}_{\mathrm{in},l,a}\mathbf{x}_{\mathrm{in},l}(t) - \mathbf{n}_{l0,a}\lambda_l(t)\right)$$

State reconstruction: $\mathbf{n}_{l,u}(t) = \mathbf{N}_u^{\mathrm{T}}\mathbf{x}_{\mathrm{r}}(t) + \mathbf{W}_{\mathrm{m},l,u}\mathbf{x}_{\mathrm{m},l}(t) + \mathbf{W}_{\mathrm{in},l,u}\mathbf{x}_{\mathrm{in},l}(t) + \mathbf{n}_{l0,u}\lambda_l(t)$

(no integration)

Extent-based incremental identification (Identification)

2. Model identification

A rate law for **each extent of reaction** and **each extent of mass transfer** is postulated and $R+p_{\mathrm{m}}$ regression problems are solved **individually** using the **integral method** of parameters estimation:

$$\min_{\boldsymbol{\theta}_{\mathrm{r},i}} \left\| \mathbf{x}_{\mathrm{r},i}(t) - \hat{\mathbf{x}}_{\mathrm{r},i}(t,\boldsymbol{\theta}_{\mathrm{r},i}) \right\|^2 \qquad i = 1,\ldots,R$$

$$\text{s.t.} \quad \dot{\hat{x}}_{\mathrm{r},i}(t,\boldsymbol{\theta}_{\mathrm{r},i}) = V_l(t)\, r_i(t,\boldsymbol{\theta}_{\mathrm{r},i}) - u_{\mathrm{out},l}(t)\, m_l^{-1}(t)\, \hat{x}_{\mathrm{r},i}(t), \qquad \hat{x}_{\mathrm{r},i}(0) = 0$$

$$\boldsymbol{\theta}_{\mathrm{r},i}^{L} \le \boldsymbol{\theta}_{\mathrm{r},i} \le \boldsymbol{\theta}_{\mathrm{r},i}^{U}$$

$$\min_{\boldsymbol{\theta}_{\mathrm{m},j}} \left\| \mathbf{x}_{\mathrm{m},l,j}(t) - \hat{\mathbf{x}}_{\mathrm{m},l,j}(t,\boldsymbol{\theta}_{\mathrm{m},j}) \right\|^2 \qquad j = 1,\ldots,p_{\mathrm{m}}$$

$$\text{s.t.} \quad \dot{\hat{x}}_{\mathrm{m},l,j}(t,\boldsymbol{\theta}_{\mathrm{m},j}) = \zeta_j(t,\boldsymbol{\theta}_{\mathrm{m},j}) - u_{\mathrm{out},l}(t)\, m_l^{-1}(t)\, \hat{x}_{\mathrm{m},l,j}(t), \qquad \hat{x}_{\mathrm{m},l,j}(0) = 0$$

$$\boldsymbol{\theta}_{\mathrm{m},j}^{L} \le \boldsymbol{\theta}_{\mathrm{m},j} \le \boldsymbol{\theta}_{\mathrm{m},j}^{U}$$

Case study
Acetoacetylation of pyrrole

The acetoacetylation of pyrrole (A) with diketene (B) in toluene (T) is a **homogeneous** reaction system catalyzed by pyridine (G).

This reaction system involves S_l = 8 species (including the solvent) and R = 4 reactions.

The main reaction (**R1**) between pyrrole and diketene produces 2-acetoacetyl pyrrole (C). The side reactions include the dimerization (**R2**) of diketene to dehydroacetic acid (D), the oligomerization (**R3**) of diketene to oligomers (E) and the reaction (**R4**) of diketene and acetoacetyl pyrrole to py-product (F).

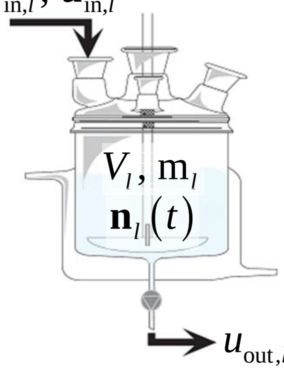
$$\begin{aligned}
&\mathbf{R1}: && \mathrm{A}+\mathrm{B} \xrightarrow{\mathrm{G}} \mathrm{C} && r_1 = k_1 c_{l,\mathrm{A}} c_{l,\mathrm{B}} c_{l,\mathrm{G}} \\
&\mathbf{R2}: && \mathrm{B}+\mathrm{B} \xrightarrow{\mathrm{G}} \mathrm{D} && r_2 = k_2 c_{l,\mathrm{B}}^2 c_{l,\mathrm{G}} \\
&\mathbf{R3}: && \mathrm{B} \longrightarrow \mathrm{E} && r_3 = k_3 c_{l,\mathrm{B}} \\
&\mathbf{R4}: && \mathrm{C}+\mathrm{B} \xrightarrow{\mathrm{G}} \mathrm{F} && r_4 = k_4 c_{l,\mathrm{C}} c_{l,\mathrm{B}} c_{l,\mathrm{G}}
\end{aligned}$$

$$\begin{array}{cc}
 & \begin{array}{cccccccc} \mathrm{A} & \mathrm{B} & \mathrm{C} & \mathrm{D} & \mathrm{E} & \mathrm{F} & \mathrm{G} & \mathrm{T} \end{array} \\
\mathbf{N} = & \begin{bmatrix} -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}
\end{array}$$

Case study
Experimental conditions

The experiment is performed in a CSTR, assuming a constant density, with an inlet of pure diketene B (p_l = 1) and one outlet. All the terms of mass transfer vanish…

$\mathbf{u}_{\mathrm{in},l} = u_{\mathrm{out},l} = 151.34\ \mathrm{g\ min^{-1}}$
$V_l = 1\ \mathrm{L},\ \mathrm{m}_l = 1.022\ \mathrm{kg}$
$M_{w,\mathrm{B}} = 84.08\ \mathrm{g\ mol^{-1}},\ w_{\mathrm{in,B}} = 1$



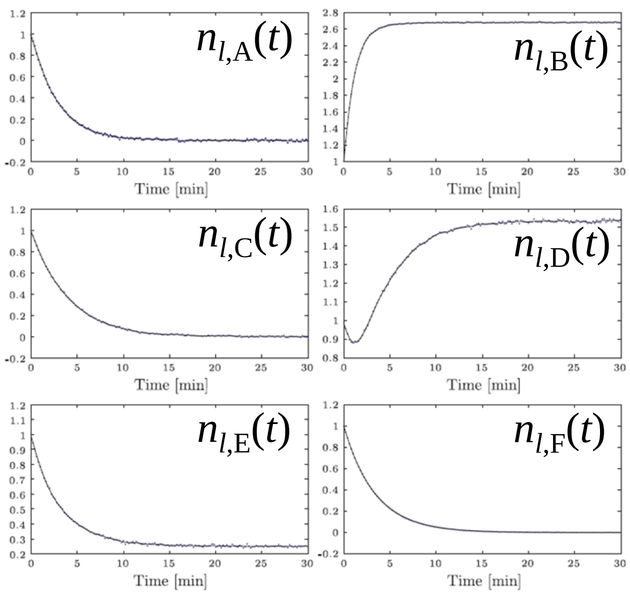
$$\mathbf{n}_{l0} = \begin{array}{c} \left[\begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0.5 \\ 2.1 \end{array}\right] \\ \mathrm{mol} \end{array} \begin{array}{c} \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \\ \mathrm{D} \\ \mathrm{E} \\ \mathrm{F} \\ \mathrm{G} \\ \mathrm{T} \\ \end{array} \qquad \mathbf{W}_{\mathrm{in},l} = \begin{array}{c} \left[\begin{array}{c} 0 \\ \mathrm{M}_{w,\mathrm{B}}^{-1} w_{\mathrm{in,B}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right] \\ \mathrm{mol\ g^{-1}} \end{array} \begin{array}{c} \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \\ \mathrm{D} \\ \mathrm{E} \\ \mathrm{F} \\ \mathrm{G} \\ \mathrm{T} \\ \end{array}$$

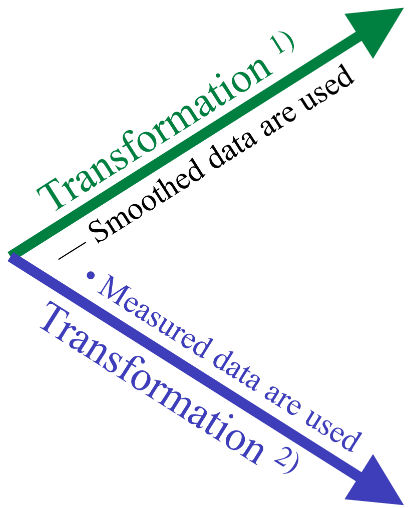
	Scenario 1 Base case	**Scenario 2** High noise	**Scenario 3** Fewer meas. conc.	**Scenario 4** Fewer time points
Noise level	1%	10%	1%	1%
Measured species concentrations	A – F	A – F	B – F	A – F
Measured time points over 30 min	150 (0.2 min)	150 (0.2 min)	150 (0.2 min)	20 (1.5 min)

Case study
Transformation into *rates* / *extents*

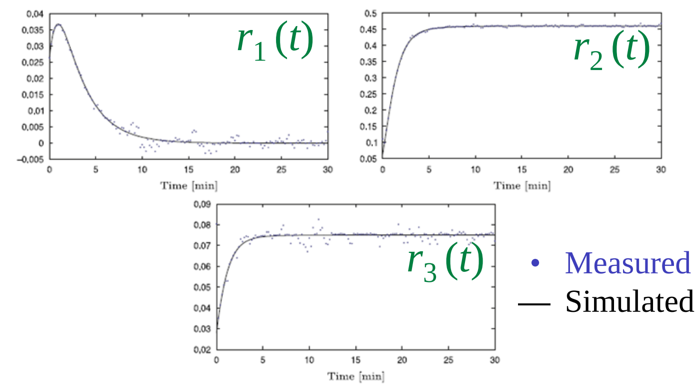
1) $\mathbf{r}(t) = V_l^{-1}(t)\,\mathbf{N}_a^{\mathrm{T}+}\left(\frac{d}{dt}\left[\mathbf{n}_{l,a}\right](t) - \mathbf{W}_{\mathrm{in},l,a}\mathbf{u}_{\mathrm{in},l,a}(t) + \frac{u_{\mathrm{out},l}(t)}{m_l(t)}\mathbf{n}_{l,a}(t)\right)$

Measurements

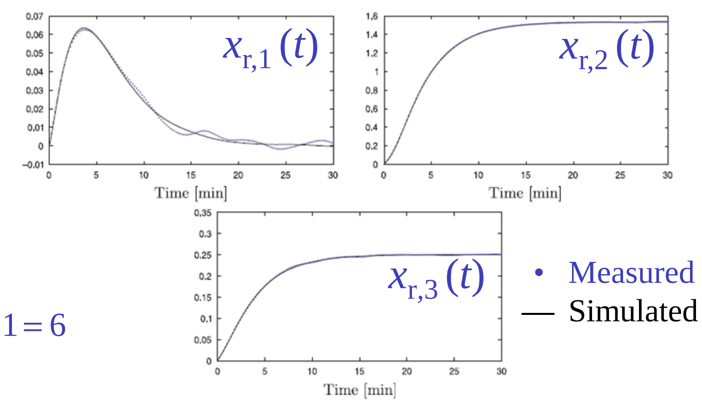




Rates



Extents



2) $\left[\mathbf{x}_\mathrm{r}(t)\ \ \mathbf{x}_{\mathrm{in},l}(t)\ \ \lambda_l(t)\right]^\mathrm{T} = \mathscr{L}_a\,\mathbf{n}_{l,a}(t)$ $\quad$ $\mathrm{rank}\left(\left[\mathbf{N}_a^\mathrm{T}\ \ \mathbf{W}_{\mathrm{in},l,a}\ \ \mathbf{n}_{l0,a}\right]\right) = R + p_l + 1 = 6$

Case study
Model identification

	Rate constant [1)]	Simulated value	Rate-based method Estimate	[95% C.I.]	Extent-based method Estimate	[95% C.I.]	Simultaneous method Estimate	[95% C.I.] [2)]
1 – Base case	k_1	0.0530	0.0501	[0.0446, 0.0556]	0.0533	[0.0528, 0.0538]	0.0526	[0.0519, 0.0533]
	k_2	0.1280	0.1281	[0.1267, 0.1295]	0.1280	[0.1280, 0.1280]	0.1281	[0.1280, 0.1283]
	k_3	0.0280	0.0279	[0.0275, 0.0283]	0.0280	[0.0280, 0.0280]	0.0280	[0.0279, 0.0281]
2 – High noise	k_1	0.0530	0.0723	[0.0328, 0.1118]	0.0461	[0.0418, 0.0504]	0.0553	[0.0479, 0.0626]
	k_2	0.1280	0.1273	[0.1232, 0.1314]	0.1283	[0.1279, 0.1283]	0.1288	[0.1275, 0.1301]
	k_3	0.0280	0.0279	[0.0265, 0.0293]	0.0285	[0.0281, 0.0289]	0.0278	[0.0275, 0.0281]
3 – Fewer measured concentrations	k_1	0.0530	0.0455	[0.0329, 0.0581]	0.0489	[0.0479, 0.0499]	0.0514	[0.0479, 0.0549]
	k_2	0.1280	0.1269	[0.1248, 0.1290]	0.1283	[0.1279, 0.1286]	0.1280	[0.1277, 0.1287]
	k_3	0.0280	0.0272	[0.0263, 0.0281]	0.0280	[0.0279, 0.0280]	0.0280	[0.0278, 0.0281]
4 – Fewer time points	k_1	0.0530	0.0457	[0.0247, 0.0667]	0.0495	[0.0438, 0.0553]	0.0460	[0.0395, 0.0525]
	k_2	0.1280	0.1278	[0.1260, 0.1297]	0.1281	[0.1257, 0.1305]	0.1279	[0.1273, 0.1285]
	k_3	0.0280	0.0275	[0.0270, 0.0280]	0.0282	[0.0278, 0.0285]	0.0280	[0.0275, 0.0285]

1) The 4th reaction is excluded from the analysis due to lack of structural identifiability

2) $Corr(k_1,k_2) = -0.03,\ Corr(k_1,k_3) = -0.07,\ Corr(k_2,k_3) = -0.04$

Case study
Model discrimination power

	Right rate law 1)	Wrong rate law 1)	Rate-based method T-criterion 2)	Extent-based method T-criterion 2)	Simultaneous method T-criterion 2)
1 – Base case	1a	1b	29	35	1 476
	2a	2b	210	1 595	16 400
	3a	3b	123	2 568	7 569
2 – High noise	1a	1b	9	8	1 221
	2a	2b	412	1 589	15 010
	3a	3b	76	458	2 876
3 – Fewer measured concentrations	1a	1b	15	26	1 252
	2a	2b	214	1 442	15 690
	3a	3b	342	348	3 784
4 – Fewer time points	1a	1b	0.9	1.2	1.4
	2a	2b	0.3	1.8	13
	3a	3b	2	48	63

1) (1a) $r_1 = k_1 c_{l,A} c_{l,B} c_{l,G}$ (1b) $r_1 = k_1 c_{l,A} c_{l,B}^2 c_{l,G}$ (2a) $r_2 = k_2 c_{l,B}^2 c_{l,G}$ (2b) $r_2 = k_2 c_{l,B}^3 c_{l,G}$ (3a) $r_3 = k_3 c_{l,B}$ (3b) $r_3 = k_3 c_{l,B}^2$

2) $T = \frac{(\hat{\mathbf{c}}_{ib} - \hat{\mathbf{c}}_{ia})^{\mathrm{T}} (\hat{\mathbf{c}}_{ib} - \hat{\mathbf{c}}_{ia})}{2\sigma^2 + \sigma_{ib}^2 + \sigma_{ia}^2}$ according to G. Buzzi-Ferraris and P. Forzatti, *Chem. Eng. Sci.* 38 (**1983**), 225

Case study
Computational time

	Rate-based method [1)] Without MT [2)]	Rate-based method [1)] With MT [2)]	Extent-based method [1)] Without MT [2)]	Extent-based method [1)] With MT [2)]	Simultaneous method [1)]
1 – Base case	3.1	3.6	6.2	6.7	16.2
2 – High noise	3.4	4.2	6.7	7.5	17.0
3 – Fewer measured concentrations	4.6	5.2	8.3	8.9	19.2
4 – Fewer time points	1.2	1.3	3.6	3.7	10.3

1) Computational time in minutes using a PC with 2.2 GHz Intel Core 2 Duo processor, 2 GB RAM

2) MT = Model Tuning by simultaneous method

Conclusion

	Rate-based method	Extent-based method	Simultaneous method
Pros	Computational time	Precision Accuracy (unbiased estimates) Model discrimination power	Precision Accuracy (unbiased estimates) Model discrimination power
Cons	Precision Accuracy (biased estimates) Model discrimination power	Model discrimination power Computational time	Computational time

It is advisable to combine the **extent-based** incremental method with a final adjustment of the rate parameters using the **simultaneous** method of identification…

Thank you for your attention

References

- J. Billeter, S. Srinivasan, D. Bonvin
  Extent-based kinetic identification using spectroscopic measurements and multivariate calibration
  Anal. Chim. Acta., submitted (**2012**)
- N. Bhatt, N. Kerimoglu, M. Amrhein, W. Marquardt, D. Bonvin
  Incremental identification of reaction systems – A comparison between rate-based and extent-based approaches
  Chem. Eng. Sci., in press (**2012**)
- N. Bhatt, M. Amrhein and D. Bonvin
  Incremental identification of reaction and mass-transfer kinetics using the concept of extents
  Ind. Eng. Chem. Res. 50 (**2011**), 12960
- C. Michalik, M. Brendel, W. Marquardt
  Incremental identification of fluid multi-phase reaction systems
  AIChE J. 55 (**2009**), 1009
- M. Brendel, D. Bonvin, W. Marquardt
  Incremental identification of kinetic models for homogeneous reaction systems
  Chem. Eng. Sci. 61 (**2006**), 5404
- A. Bardow, W. Marquardt
  Incremental and simultaneous identification of reaction kinetics: methods and comparison
  Chem. Eng. Sci. 59 (**2004**), 2673