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

$$\dot{n}_g(t) = -W_{m,g} \zeta(t) + W_{in,g} u_{in,g}(t) - \frac{u_{out,g}(t)}{m_g(t)} n_g(t), \quad n_g(0) = n_{g0}$$

Liquid phase

$S_{l}$ species, $R$ reactions, $p_{m}$ mass transfers, $p_{l}$ inlets and 1 outlet

$$\dot{n}_l(t) = N^T V_l(t) r(t) + W_{m,l} \zeta(t) + W_{in,l} u_{in,l}(t) - \frac{u_{out,l}(t)}{m_l(t)} n_l(t), \quad n_l(0) = n_{l0}$$

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:

\[
\begin{align*}
\min_{\theta_m} & \quad \| n_{l,a}(t) - \hat{n}_{l,a}(t, \theta_{rm}) \|^2 \\
\text{s.t.} & \quad \hat{n}_l(t, \theta_{rm}) = N^T V_l(t) r(t, \theta_r) + W_{m,l} \xi(t, \theta_m) + W_{in,l} u_{in,l}(t) - \frac{u_{out,l}(t)}{m_l(t)} \hat{n}_l(t), \quad \hat{n}_l(0) = n_{l0} \\
\theta_r^L & \leq \theta_r \leq \theta_r^U \\
\theta_m^L & \leq \theta_m \leq \theta_m^U
\end{align*}
\]
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

\[
\begin{bmatrix}
V_i(t) \mathbf{r}(t) \\
\zeta(t)
\end{bmatrix}^T = 
\begin{bmatrix} N_a^T \quad W_{m,l,a} \end{bmatrix}^+ 
\frac{d}{dt} \begin{bmatrix} n_{l,a} \end{bmatrix}(t) - W_{in,l,a} u_{in,l,a}(t) + \frac{u_{out,l}(t)}{m_l(t)} n_{l,a}(t)
\]

\text{rank} = R + p_m

*State reconstruction* (by integration):

\[
\dot{n}_{l,u}(t) = N_u^T V_i(t) \mathbf{r}(t) + W_{m,l,u} \zeta(t) + W_{in,l,u} u_{in,l}(t) - \frac{u_{out,l}(t)}{m_l(t)} n_{l,u}(t),
\]

\[n_{l,u}(0) = n_{l0,u}\]

2. Model identification
A rate law for **each rate of reaction** and **each rate of mass transfer** is postulated and \(R+p_m\) regression problems are solved **individually**:

\[
\min_{\theta_{r,i}} \| r_i(t) - \hat{r}_i(t, \theta_{r,i}) \|^2 \quad \theta_{r,i}^L \leq \theta_{r,i} \leq \theta_{r,i}^U \quad i = 1, \ldots, R
\]

\[
\min_{\theta_{m,j}} \| \zeta_j(t) - \hat{\zeta}_j(t, \theta_{m,j}) \|^2 \quad \theta_{m,j}^L \leq \theta_{m,j} \leq \theta_{m,j}^U \quad j = 1, \ldots, p_m
\]
Extent-based incremental identification (Transformation)

1. Transformation

1a: Computation of \( R+p_m+p_l+1 \) extents

\[
\begin{bmatrix}
    x_r(t) \\
    x_{m,l}(t) \\
    x_{in,l}(t) \\
    \lambda_l(t)
\end{bmatrix} = \mathcal{L}_a \mathbf{n}_{l,a}(t)
\]

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

1b: Computation of \( R+p_m \) extents (rank < \( R+p_m+p_l+1 \))

\[
\begin{bmatrix}
    x_r(t) \\
    x_{m,l}(t)
\end{bmatrix} = \left[ \mathbf{N}_a^T \mathbf{W}_{m,l,a} \right]^+ \left( \mathbf{n}_{l,a}(t) - \mathbf{W}_{in,l,a} x_{in,l}(t) - \mathbf{n}_{l0,a} \lambda_l(t) \right)
\]

\[
\text{rank} = R + p_m
\]

State reconstruction: \( \mathbf{n}_{l,u}(t) = \mathbf{N}_u x_r(t) + \mathbf{W}_{m,l,u} x_{m,l}(t) + \mathbf{W}_{in,l,u} x_{in,l}(t) + \mathbf{n}_{l0,u} \lambda_l(t) \)
(no integration)

\[
\begin{align*}
\dot{x}_r(t) &= V_l(t) r(t) - \tau_{out,l}^{-1}(t) x_r(t), \\
\dot{x}_{m,l}(t) &= \zeta(t) - \tau_{out,l}^{-1}(t) x_{m,l}(t), \\
\dot{x}_{in,l}(t) &= u_{in,l}(t) - \tau_{out,l}^{-1}(t) x_{in,l}(t), \\
\dot{\lambda}_l(t) &= -\tau_{out,l}^{-1}(t) \lambda_l(t),
\end{align*}
\]

with \( \tau_{out,l}^{-1}(t) = u_{out,l}(t) m_l^{-1}(t) \)
2. Model identification

A rate law for each extent of reaction and each extent of mass transfer is postulated and \(R+p_m\) regression problems are solved individually using the integral method of parameters estimation:

\[
\min_{\theta_{r,i}} \left\| x_{r,i}(t) - \hat{x}_{r,i}(t, \theta_{r,i}) \right\|^2 \quad i = 1, \ldots, R
\]

\[
\text{s.t.} \quad \dot{x}_{r,i}(t, \theta_{r,i}) = V_i(t) r_i(t, \theta_{r,i}) - u_{\text{out},i}(t) m_i(t) \hat{x}_{r,i}(t), \quad \hat{x}_{r,i}(0) = 0
\]

\[
\theta_{r,i}^L \leq \theta_{r,i} \leq \theta_{r,i}^U
\]

\[
\min_{\theta_{m,j}} \left\| x_{m,l,j}(t) - \hat{x}_{m,l,j}(t, \theta_{m,j}) \right\|^2 \quad j = 1, \ldots, p_m
\]

\[
\text{s.t.} \quad \dot{x}_{m,l,j}(t, \theta_{m,j}) = \zeta_j(t, \theta_{m,j}) - u_{\text{out},j}(t) m_i(t) \hat{x}_{m,l,j}(t), \quad \hat{x}_{m,l,j}(0) = 0
\]

\[
\theta_{m,j}^L \leq \theta_{m,j} \leq \theta_{m,j}^U
\]
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).

\[
\begin{align*}
\textbf{R1 : } & \quad \text{A} + \text{B} \xrightarrow{G} \text{C} & & r_1 = k_{i,A} c_{i,A} c_{i,B} c_{i,G} \\
\textbf{R2 : } & \quad \text{B} + \text{B} \xrightarrow{G} \text{D} & & r_2 = k_{2} c_{i,B} c_{i,G} \\
\textbf{R3 : } & \quad \text{B} \xrightarrow{} \text{E} & & r_3 = k_{3} c_{i,B} \\
\textbf{R4 : } & \quad \text{C} + \text{B} \xrightarrow{G} \text{F} & & r_4 = k_{4} c_{i,C} c_{i,B} c_{i,G}
\end{align*}
\]

\[
\begin{array}{cccccccc}
A & B & C & D & E & F & G & T \\
\hline
-1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & -2 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & -1 & -1 & 0 & 0 & 1 & 0 & 0
\end{array}
\]
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…

\[
\begin{align*}
\mathbf{u}_{\text{in},l} &= \mathbf{u}_{\text{out},l} = 151.34 \text{ g min}^{-1} \\
V_l &= 1 \text{ L, } m_l = 1.022 \text{ kg} \\
M_{w,B} &= 84.08 \text{ g mol}^{-1}, \ w_{\text{in,B}} = 1
\end{align*}
\]

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
<th>Scenario 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base case</td>
<td>High noise</td>
<td>Fewer meas. conc.</td>
<td>Fewer time points</td>
</tr>
<tr>
<td>Noise level</td>
<td>1%</td>
<td>10%</td>
<td>1%</td>
</tr>
<tr>
<td>Measured species concentrations</td>
<td>A – F</td>
<td>A – F</td>
<td>B – F</td>
</tr>
<tr>
<td>Measured time points over 30 min</td>
<td>150 (0.2 min)</td>
<td>150 (0.2 min)</td>
<td>150 (0.2 min)</td>
</tr>
</tbody>
</table>

Case study
Transformation into *rates / extents*

1) \( r(t) = V_l^{-1}(t)N_a^{T^+} \left( \frac{d}{dt} \left[ n_{l,a} \right](t) - W_{in,l,a} u_{in,l,a}(t) + \frac{u_{out,l}(t)}{m_l(t)} n_{l,a}(t) \right) \)

**Measurements**

\[ n_{l,A}(t) \]
\[ n_{l,B}(t) \]
\[ n_{l,C}(t) \]
\[ n_{l,D}(t) \]
\[ n_{l,E}(t) \]
\[ n_{l,F}(t) \]

**Rates**

\[ r_1(t) \]
\[ r_2(t) \]
\[ r_3(t) \]

**Extents**

\[ x_{r,1}(t) \]
\[ x_{r,2}(t) \]
\[ x_{r,3}(t) \]

2) \[ \begin{bmatrix} x_r(t) & x_{in,l}(t) & \lambda_l(t) \end{bmatrix}^T = L_a n_{l,a}(t) \]

\( \text{rank} \left( \begin{bmatrix} N_a^T & W_{in,l,a} & n_{f0,a} \end{bmatrix} \right) = R + p_l + 1 = 6 \)

## Case study

### Model identification

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

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

<table>
<thead>
<tr>
<th>Case Description</th>
<th>Right rate law</th>
<th>Wrong rate law</th>
<th>Rate-based method T-criterion</th>
<th>Extent-based method T-criterion</th>
<th>Simultaneous method T-criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – Base case</td>
<td>la</td>
<td>lb</td>
<td>29</td>
<td>35</td>
<td>1476</td>
</tr>
<tr>
<td></td>
<td>2a</td>
<td>2b</td>
<td>210</td>
<td>1595</td>
<td>16400</td>
</tr>
<tr>
<td></td>
<td>3a</td>
<td>3b</td>
<td>123</td>
<td>2568</td>
<td>7569</td>
</tr>
<tr>
<td>2 – High noise</td>
<td>la</td>
<td>lb</td>
<td>9</td>
<td>8</td>
<td>1221</td>
</tr>
<tr>
<td></td>
<td>2a</td>
<td>2b</td>
<td>412</td>
<td>1589</td>
<td>15010</td>
</tr>
<tr>
<td></td>
<td>3a</td>
<td>3b</td>
<td>76</td>
<td>458</td>
<td>2876</td>
</tr>
<tr>
<td>3 – Fewer measured concentrations</td>
<td>la</td>
<td>lb</td>
<td>15</td>
<td>26</td>
<td>1252</td>
</tr>
<tr>
<td></td>
<td>2a</td>
<td>2b</td>
<td>214</td>
<td>1442</td>
<td>15690</td>
</tr>
<tr>
<td></td>
<td>3a</td>
<td>3b</td>
<td>342</td>
<td>348</td>
<td>3784</td>
</tr>
<tr>
<td>4 – Fewer time points</td>
<td>la</td>
<td>lb</td>
<td>0.9</td>
<td>1.2</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>2a</td>
<td>2b</td>
<td>0.3</td>
<td>1.8</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>3a</td>
<td>3b</td>
<td>2</td>
<td>48</td>
<td>63</td>
</tr>
</tbody>
</table>

1) (1a) $r_1 = k_i c_i A c_{1,B} c_{1,G}$  
(1b) $r_1 = k_i c_i A c_{1,B}^2 c_{1,G}$

2) $r_2 = k_2 c_{1,B}^2 c_{1,G}$

$$T = \frac{(\hat{e}_{ib} - \hat{e}_{ia})^T (\hat{e}_{ib} - \hat{e}_{ia})}{2\sigma^2 + \sigma_{ib}^2 + \sigma_{ia}^2}$$


## Case study

### Computational time

<table>
<thead>
<tr>
<th>Case Description</th>
<th>Rate-based method 1)</th>
<th>Extent-based method 1)</th>
<th>Simultaneous method 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without MT 2)</td>
<td>With MT 2)</td>
<td></td>
</tr>
<tr>
<td>1 – Base case</td>
<td>3.1</td>
<td>3.6</td>
<td>16.2</td>
</tr>
<tr>
<td>2 – High noise</td>
<td>3.4</td>
<td>4.2</td>
<td>17.0</td>
</tr>
<tr>
<td>3 – Fewer measured concentrations</td>
<td>4.6</td>
<td>5.2</td>
<td>19.2</td>
</tr>
<tr>
<td>4 – Fewer time points</td>
<td>1.2</td>
<td>1.3</td>
<td>10.3</td>
</tr>
</tbody>
</table>

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

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

- J. Billeter, S. Srinivasan, D. Bonvin
  Extent-based kinetic identification using spectroscopic measurements and multivariate calibration

- N. Bhatt, N. Kerimoglu, M. Amrhein, W. Marquardt, D. Bonvin
  Incremental identification of reaction systems – A comparison between rate-based and extent-based approaches

- N. Bhatt, M. Amrhein and D. Bonvin
  Incremental identification of reaction and mass-transfer kinetics using the concept of extents

- 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

- A. Bardow, W. Marquardt
  Incremental and simultaneous identification of reaction kinetics: methods and comparison