Nonlinear State Estimation Methods – Overview and Application to PET Polymerization

Paul Appelhaus, Ralf Gesthuisen, Stefan Krämer, Sebastian Engell

The first author is now CEO of Appelhaus Miniplants, Lindewerra, the second and the third authors are with INEOS, Köln, Germany

Outline

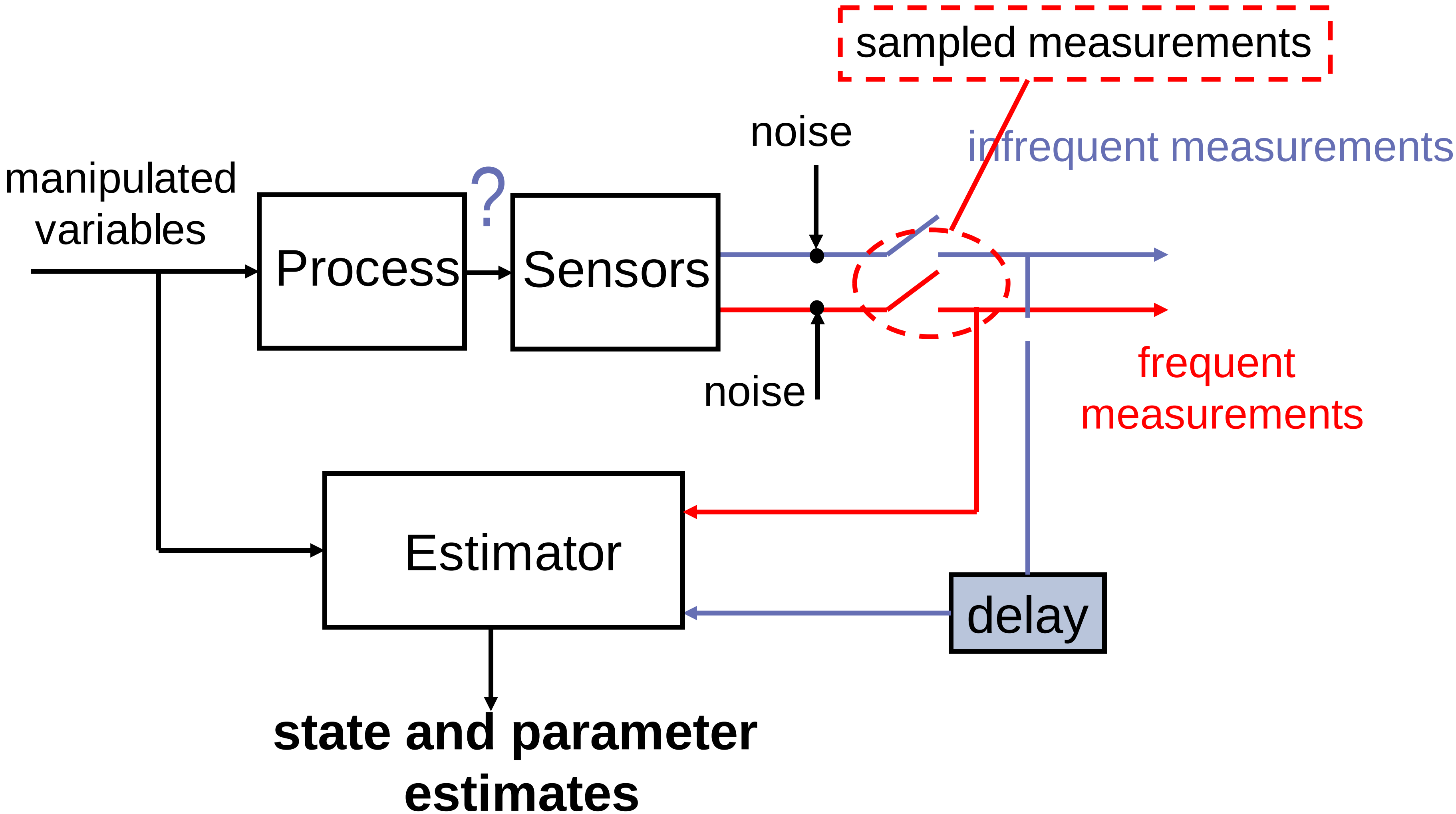
- Introduction
- Approaches to nonlinear state estimation
- Application to PET polycondensation
- Multirate estimation
- Conclusions

Introduction

- Important product qualities are not measureable online but they are needed for control purposes.
- Measurements are often available at different sampling intervals.
- Analytic measurements cause time delays.

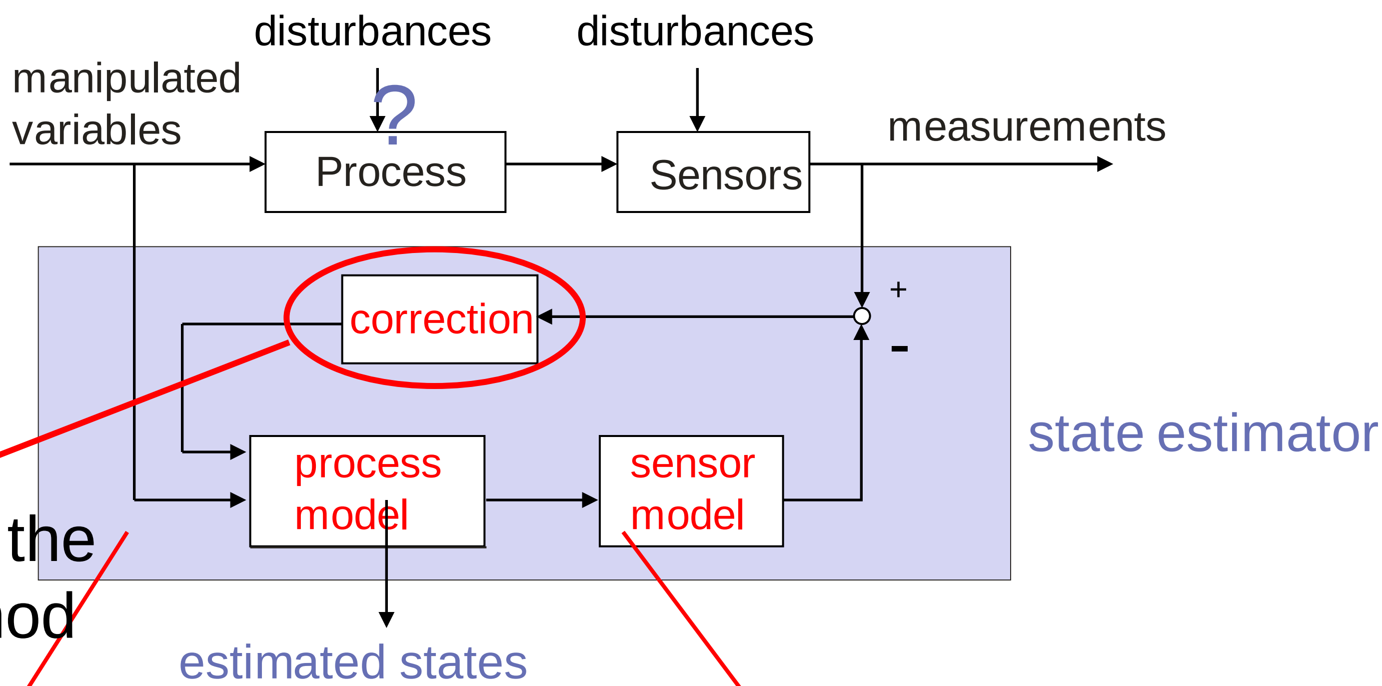
  ⇒ State estimation can provide additional online information.
- Numerous approaches are available: Which one should be applied?
- How can offline analyses be incorporated in state estimation?
  - Black Box Modelling (Neural Nets, …)
  - Extended approaches are necessary to consider measurements with different sampling intervals.

Introduction: Application of State Estimation



Nonlinear State Estimation

- Basic principle of state estimation



depends on the design method

$$\dot{x} = f(x, u, t) \qquad y = h(x)$$

Approaches to Nonlinear State Estimation

- Process model is given as ODE-system or system of difference equations

$$\dot{\hat{x}} = f\left(\hat{x}, u, t, y\right) \qquad \hat{x}_{k+1} = F\left(\hat{x}_k, u_k, t_k, y_k\right)$$

$$\hat{y} = h\left(\hat{x}\right) \qquad \hat{y}_k = H\left(\hat{x}_k\right)$$

- The estimator must fulfill:
  - Simulation condition:
    if $x\left(t = t_0\right) = \hat{x}\left(t = t_0\right)$ then $x\left(t\right) = \hat{x}\left(t\right) \forall\, t > t_0$
  - Convergence condition:

$$\lim_{t \to \infty}\left(x\left(t\right) - \hat{x}\left(t\right)\right) = 0$$

for all inputs *u(t)* and for all inital errors.

Observability

- **Theory:**
  - Assume a linear system: $\dot{x} = Ax + Bu\,,\; y = Cx$ with $x \in \Re^n, y \in \Re^p$
  - The state variables of the system (A, C) can be divided into two groups:
    - **Observable states**:
      The error dynamics can be prescribed arbitrarily by choice of the observer gains and are independent of the states and the known inputs.
    - **Unobservable states**:
      the error dynamics are given by the (unobservable) eigenvalues of the system.
  - **Definition:** A linear system is said to be observable, if the matrix $Q = [C^T, A^T C^T, (A^T)^2 C^T, \dots (A^T)^{n\text{-}1} C^T]$ has rank *n*. In this case, all states are observable.

Observability of Nonlinear Systems

- **Theory:**
  - Assume a nonlinear system: $\dot{x} = f(x,u),\ y = h(x)$
  - **Definition:**
    A nonlinear system is said to be globally observable, if all initial states $x_0 \in X_0$ can be computed from observations of the output y(t) over an arbitrary interval of time. The states are then called observable.
  - **Definition:**
    A nonlinear system is said to be locally observable at $x_1 \in X_0$, if initial states $x_0 \in X_0$ in the neighborhood of x_1 are observable.

  - To check global observability a unique solution for the nonlinear observability map has to be found.
  - Local observability can be checked using the linearized model (see conditions for linear models).

State and Parameter Estimation

- Observers can be extended to the **estimation** of **unknown process parameters**.
- The parameters are assumed to be additional state variables with dummy dynamics.

$$\dot{\hat{x}} = f\left(\hat{x}, u, \hat{p}\right)$$

$$\dot{\hat{p}} = 0$$

- Observability in general becomes worse if more parameters must be estimated.
- The number of unknown parameters which can be estimated is restricted by the number of measurements!

Approaches to Nonlinear State Estimation

- **Gain Scheduled Observer (Luenberger observer with gains that depend on the estimated states and the inputs)**
  - **Observer equations:**

$$\dot{x} - \dot{\hat{x}} = \dot{\tilde{x}} = f(x,u) - f(\hat{x},u) - K(\hat{x},u)\left(y - h(\hat{x})\right)$$

$K(\hat{x},u)$ calculated by eigenvalue placement for linearized error dynamics

$$\lambda_i\left(I - \left(\left.\frac{\partial f}{\partial x}\right|_{\hat{x},u} - K\left.\frac{\partial h}{\partial x}\right|_{\hat{x},u}\right)\right) = \lambda_{i,\text{desired}}$$

  - Low effort for design
  - Performance and stability only are guaranteed close to a fixed (slowly varying) operating point.

Approaches to Nonlinear State Estimation

- **Sliding Mode Observer:**
  - For nonlinear systems of the form:

Measured nonlinearity

Bounded nonlinearities not measured

$$\dot{x} = Ax + \alpha(y,u) - E\Psi(x,u), \quad x(0) = x_0$$

$$\Psi(x,u) \in \Re^w \text{ with } \|\Psi(x,u)\| \leq \rho(u) \quad \forall \quad x \in \Re^n$$

$$y = Cx \text{ and } \quad y \in \Re^m$$

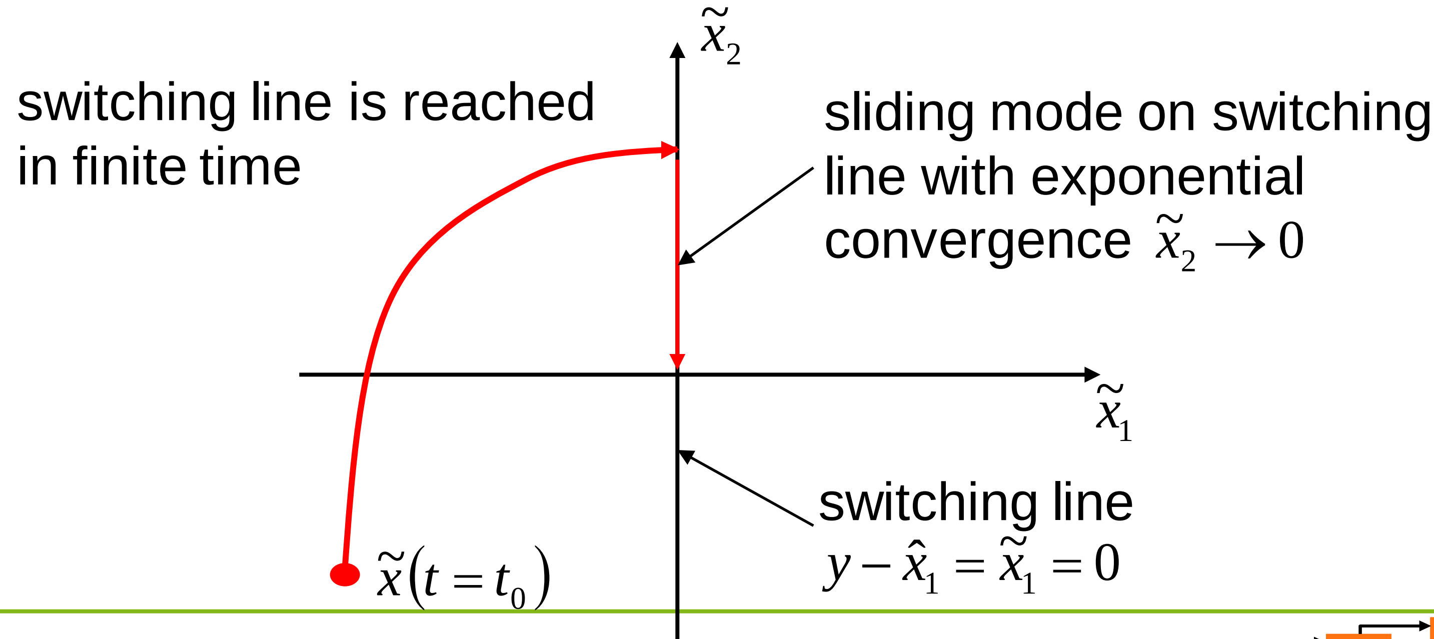
  - Number of measurements ≥ number of unmeasured nonlinearities $\Psi(x,u)$
  - Observer equation:

$$\dot{\hat{x}} = A\hat{x} + \alpha(y,u) + \rho(u)E\frac{D(y - C\hat{x})}{D\|y - C\hat{x}\|} + G(y - C\hat{x}), \quad \hat{x}(0) = x_0$$

  - P,D,G are weighting matrices which have to be determined in an iterative manner (an implicit Ljapunov equation has to be solved).

Sliding Mode Observer

- **Properties of the SMO:**
  - Large effort for design
  - Guaranteed global stability
  - Convergence in 2 steps:
    1. Movement to the switching plane: $y - C\hat{x} = 0$
    2. Sliding on the switching plane until the desired state is reached.



Extended Kalman Filter

- Extension of the model equations by stochastic errors (zero mean and given variance)

$$\dot{x} = f(x,u,t) + \xi$$
$$y = h(x) + \omega$$

- Filter minimizes the variance of the estimation error:

$$E\left\{(x-\hat{x})^T (x-\hat{x})\right\} = E\left\{\tilde{x}^T \tilde{x}\right\} \rightarrow \text{Min}$$

- For nonlinear time discrete (sampled) systems the solution is a 2-step-algorithm:
  - Correction of the predicted states and estimated covariance matrix $P = E\left\{(x_i - \hat{x}_i)(x_i - \hat{x}_i)^T\right\} = E\left\{\tilde{x}_i \tilde{x}_i^T\right\}$ after a new measurement was obtained
  - Prediction of the states and the covariance matrix of the estimation error up to the next time step by nonlinear forward simulation

EKF Algorithm

- Notation: Estimated state at $t=t_{k+1}$ based on measurements up to $t=t_k$

  $\hat{x}_{k+1,k}$
- Correction:

$$K_k = P_{k,k-1} H_{k,k-1}^T \left(H_{k,k-1} P_{k,k-1} H_{k,k-1}^T + R \right)^{-1}$$

$$\hat{x}_{k,k} = \hat{x}_{k,k-1} + K_k \left(y_k - h\left(\hat{x}_{k,k-1}\right)\right)$$

$$P_{k,k} = \left(I - K_k H_{k,k-1} \right) P_{k,k-1}$$

- Prediction:

$$\hat{x}_{k,k+1} = F\left(\hat{x}_{k,k}, u_k\right)$$

$$P_{k,k+1} = A_{k,k} P_{k,k} A_{k,k}^T + Q \quad \text{with} \quad A_{k,k} = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{k,k}} \quad \text{and} \quad H_{k,k-1} = \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_{k,k-1}}$$

- Constraints on states and disturbances cannot be considered.
- Only (at best) local stability!
- The critical part is the update of the covariance matrix based on the linearized state equations.

Batch Estimation

- Suppose k values of the process inputs and of the measured variables are available. Then the estimation of the corresponding states can be formulated as an optimization problem:

$$\min_{\hat{\xi}_j, \hat{\varphi}_j} \left(\Psi_k = \hat{\xi}_{0,k}^T P_0^{-1} \hat{\xi}_{0,k} + \sum_{j=1}^{k-1} \hat{\xi}_{j,k}^T Q^{-1} \hat{\xi}_{j,k} + \sum_{j=1}^{k} \hat{\varphi}_{j,k}^T R^{-1} \hat{\varphi}_{j,k} \right)$$

Subject to

$$\hat{x}_{0,k} = \hat{x}_{0,0} + \hat{\xi}_{0,k}$$

$$\hat{x}_{j,k} = F\left(\hat{x}_{j-1,k}, u_{j-1}\right) + \hat{\xi}_{j,k} \qquad j = 1, \ldots, k-1$$

$$y_j = h\left(\hat{x}_{j,k}\right) + \hat{\varphi}_{j,k} \qquad j = 1, \ldots, k$$

- In this formulation, there are no assumptions on the statistics of the errors.
- No linearization is performed!
- Nonlinear optimization problem!
- Computational effort increases with k → reduce to a finite horizon!

Moving Horizon State Estimation

- Further motivation for a deterministic observer based on numerical optimization:
  - Possiblity of imposing constraints on states and disturbances
- **Moving Horizon Estimator (MHE):**
  - Takes measurements over a finite horizon in the past into account
  - No approximation of the nonlinear system within this horizon
  - Model: Discrete-time nonlinear system

$$x_{k+1} = F(x_k, u_k) + \xi_k$$
$$y_k = h(x_k) + \varphi_k$$

Moving Horizon State Estimator

- The Moving Horizon Estimator can be formulated as:

Minimize the error

$$\min_{\hat{\xi}_{k-N-1,k},\ldots,\hat{\xi}_{k-1,k},\hat{\varphi}_{k-N,k},\ldots,\hat{\varphi}_{k,k}} \Psi_k^N = \hat{\xi}_{k-N-1,k}^T \mathbf{P}_{k-N}^{-1} \hat{\xi}_{k-N-1,k} + \sum_{j=k-N}^{k-1} \hat{\xi}_{j,k}^T \mathbf{Q}_k^{-1} \hat{\xi}_{j,k} + \sum_{j=k-N}^{k-1} \hat{\varphi}_{j,k}^T \mathbf{R}_k^{-1} \hat{\varphi}_{j,k}$$

$$s.t.: \quad \hat{\mathbf{x}}_{j+1,k} = \mathbf{F}\left(\hat{\mathbf{x}}_{j,k}, \mathbf{u}_j\right) + \hat{\xi}_{j,k} \qquad j = k-N,\ldots,k-1$$

$$\hat{\mathbf{x}}_{k-N,k} = \hat{\mathbf{x}}_{k-N,k-N-1} + \hat{\xi}_{k-N-1,k}, \qquad \mathbf{y}_j = \mathbf{h}\left(\hat{\mathbf{x}}_{j,k}\right) + \hat{\varphi}_{j,k}$$

Model equations

$$\xi_{\min} \le \hat{\xi}_{k,k} \le \xi_{\max}, \qquad \varphi_{\min} \le \hat{\varphi}_{k,k} \le \varphi_{\max}, \qquad \mathbf{x}_{\min} \le \hat{\mathbf{x}}_{k,k} \le \mathbf{x}_{\max}$$

Physical or other constraints

Constrained Extended Kalman Filter (CEKF)

- For a horizon of N=0 the MHE becomes the Constrained EKF
- CEKF equations:

$$\min_{\hat{\xi}_{k-1,k},\hat{\varphi}_{k,k}} \left(\Psi_k^0 = \hat{\xi}_{k-1,k}^T P_k^{-1} \hat{\xi}_{k-1,k} + \hat{\varphi}_{k,k}^T R^{-1} \hat{\varphi}_{k,k} \right)$$

subject to

$$\left.\begin{aligned} \hat{x}_{k,k} &= \hat{x}_{k,k-1} + \hat{\xi}_{k-1,k} \\ y_k &= h\left(\hat{x}_{k,k}\right) + \hat{\varphi}_{k,k} \end{aligned}\right\} \text{ equal. constr.}$$

$$\hat{x}_{k,k+1} = F\left(\hat{x}_{k,k}, u_k\right)$$

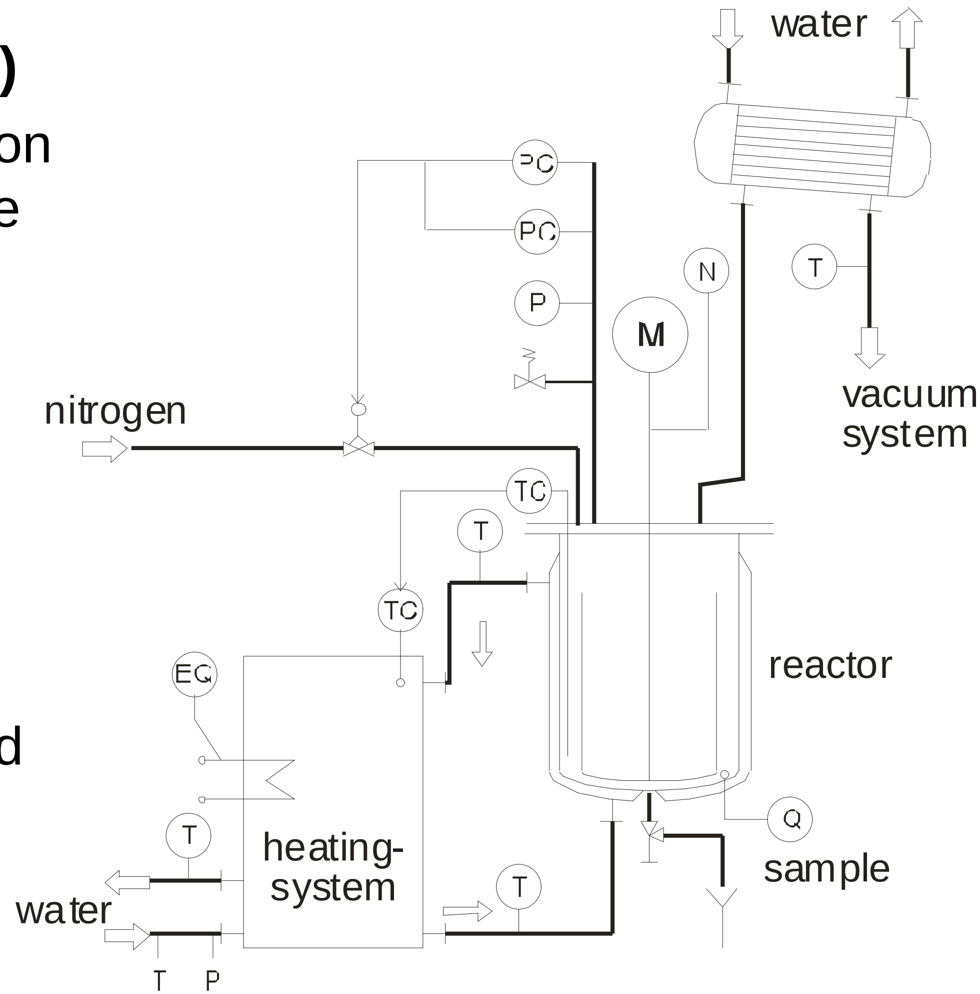
$$P_{k,k+1} = A_{k,k} P_{k,k} A_{k,k}^T + Q \quad \text{with} \quad A_{k,k} = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{k,k}} \quad \text{and } H_{k,k-1} = \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_{k,k-1}}$$

plus additional constraints on errors and states

- Quadratic problem for linear measurement function

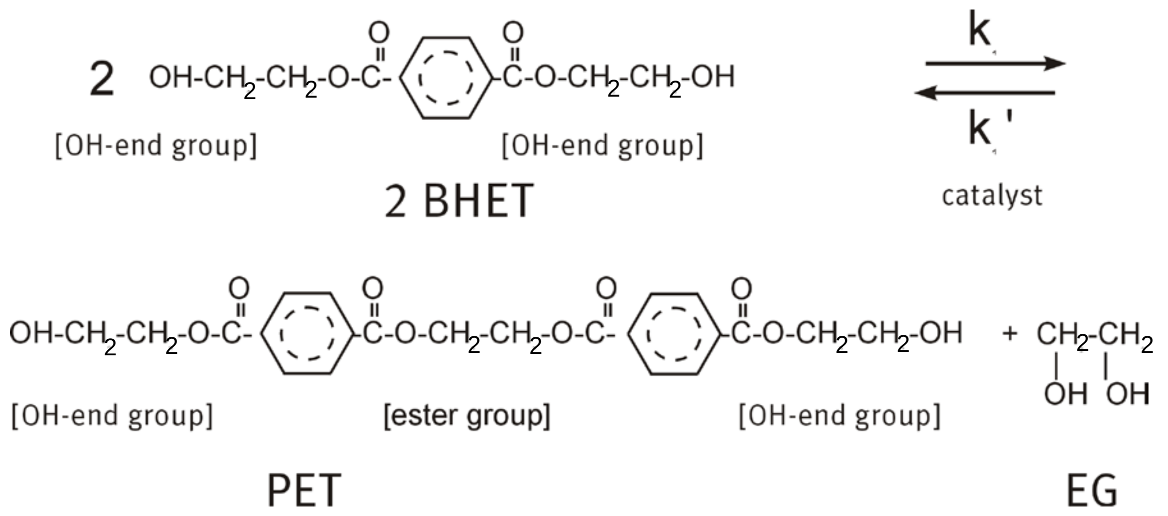
Case-Study: Polycondensation of PET

- **Polycondensation of polyethylenterephtalate (PET)**
  - The mass transfer of a reaction byproduct from the melt to the gas phase determines the progress of the poly-condensation process.
  - No measurements of mass transfer coefficients under industrial process conditions available
  - Process realized in a jacketed 10 l stainless steel reactor controlled by a DCS (Contronic-P, ABB)



PET Polycondensation

- Batch process, polymerization in the melt
- Main chemical reaction:



- The reverse reaction is 8 times faster than the forward reaction
- The removal of ethyleneglycol (EG) determines the speed of the polymerization

PET Polycondensation

- Main and side reactions:
  - -[OH] + -[OH] $\rightleftarrows$ -[E]- + EG (polymerisation)
  - -[E]- $\longrightarrow$ -[COOH] + -[$CO_2C_2H_3$] (thermal degradation)
  - -[COOH] + EG $\rightleftarrows$ -[OH] + H_2O
  - -[COOH] + -[OH] $\rightleftarrows$ -[E]- + H_2O
  - -[OH] $\longrightarrow$ -[COOH] + CH_3CHO
  - -[$CO_2C_2H_3$] + -[OH] $\longrightarrow$ -[E]- + CH_3CHO

- Mathematical modelling yields a system of differential equations for the concentrations of hydroxy group -[OH], ethylene glycol EG, carboxyl group -[COOH], acetaldehyde CH_3CHO, vinyl group -[$CO_2C_2H_3$], ester group -[E] and water H_2O

Model Reduction

- Model of 7^{th} order too complex, several unknown mass transfer parameters
- Simplification by:
  - Neglecting the side reactions
  - Estimation of the thermal degradation by steady state assumption as this reaction has slow dynamics compared to the main reaction
- Resulting model of 3^{rd} order includes dynamics of
  - Concentrations of hydroxygroup **-[OH**] and ethylene glycol **EG** and
  - A parameter which characterizes the mass transfer of EG: βa

Estimation Problem

- **Available data:**

  Measurements of

  - temperature
  - stirrer torque
  - stirrer speed
  - pressure

- **Problems:**
  - Online measurement of the degree of polymerisation not possible
  - Removal of the volatile reaction-product ethylene glycol crucial, but no infomation on the mass transfer coefficient is available

- **Solution**
  - Computation of the degree of polymerisation based on available process data
  - Design of an estimator based on a simple reaction model for
    - concentration of the most important end groups
    - mass transfer coefficient for the removal of ethylene glycol

Example: PET-Poly-Condensation

- Computation of the degree of polymerization
  Inputs: Temperature, stirrer torque, stirrer speed

  a) Semi-empirical model

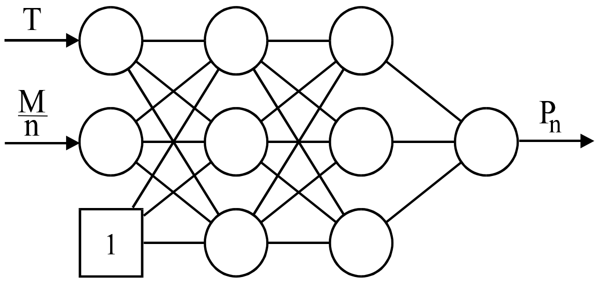
$$P_n = C_1 \cdot \left(\frac{M}{n^a} \exp\left(\frac{-7000}{T} \right) \right)^{C_2}$$

  Parameters C_1, C_2 and a calculated by nonlinear regression from data of analysis of samples

  b) **Neural net**

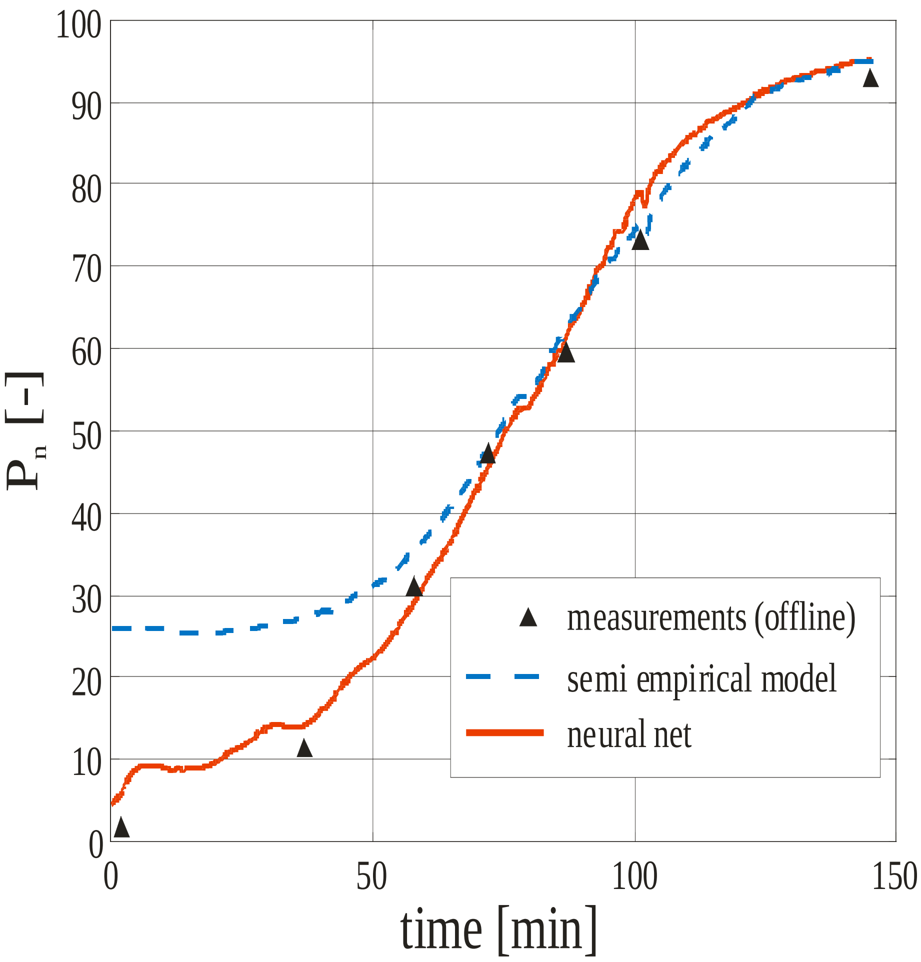
  Inputs: Ratio of stirrer torque and stirrer speed, temperature

  Output: Degree of polymerisation



  Training data: Laboratory analysis of samples

Estimation of the Degree of Polymerization



PET Polycondensation Model

- **Simplified model:**

$$\dot{x}_{EG} = -\beta a(x_{EG} - x_{EG}^{*}) + \frac{1}{2}\left[x_{OH}^{2} - 8x_{EG}\left[x_{E,\max} - \frac{1}{2}x_{OH}\right]\right]$$

$$\dot{x}_{OH} = -k_1 x_{OH}^{2} + 8x_{EG}\left[x_{E,\max} - \frac{1}{2}x_{OH}\right]$$

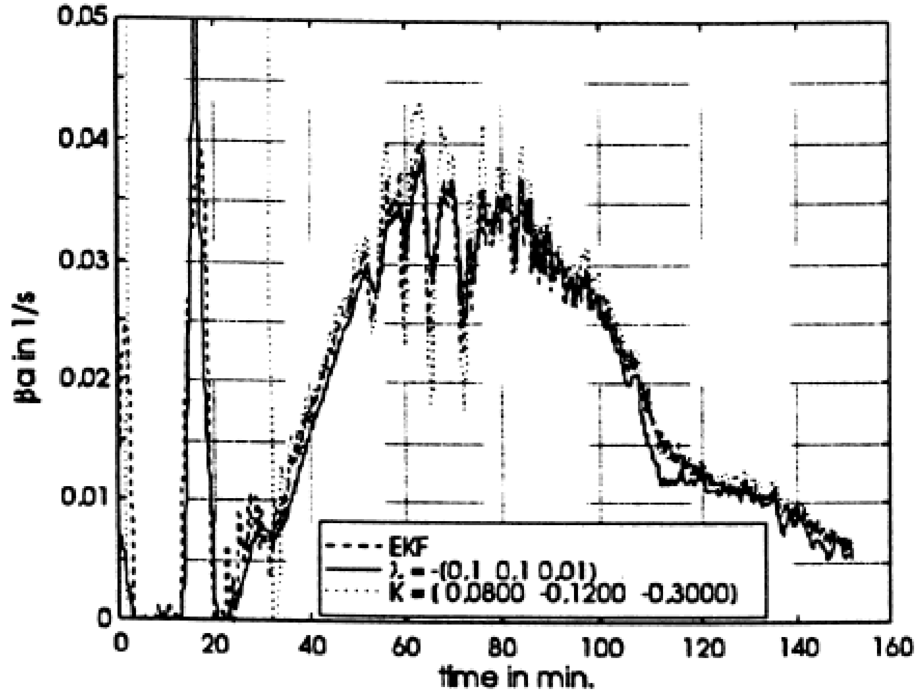
$$\dot{\beta a} = 0$$

Measurement equation: $y = x_{OH}$ where $x_{OH} = f\left(P_n, M_{EG}, M_E\right)$

- **Goal:** Estimation of the mass transfer coefficient in order to find dependencies on process conditions (stirrer speed, temperature)

Thesis by Paul Appelhaus

Results after careful tuning



EKF Parameters:

$$\underline{Q} = \frac{1}{\hat{x}_2}\begin{bmatrix} 0.001 & 0 & 0 \\ 0 & 0.0005 & 0 \\ 0 & -0.01 & 0.06 \end{bmatrix}$$

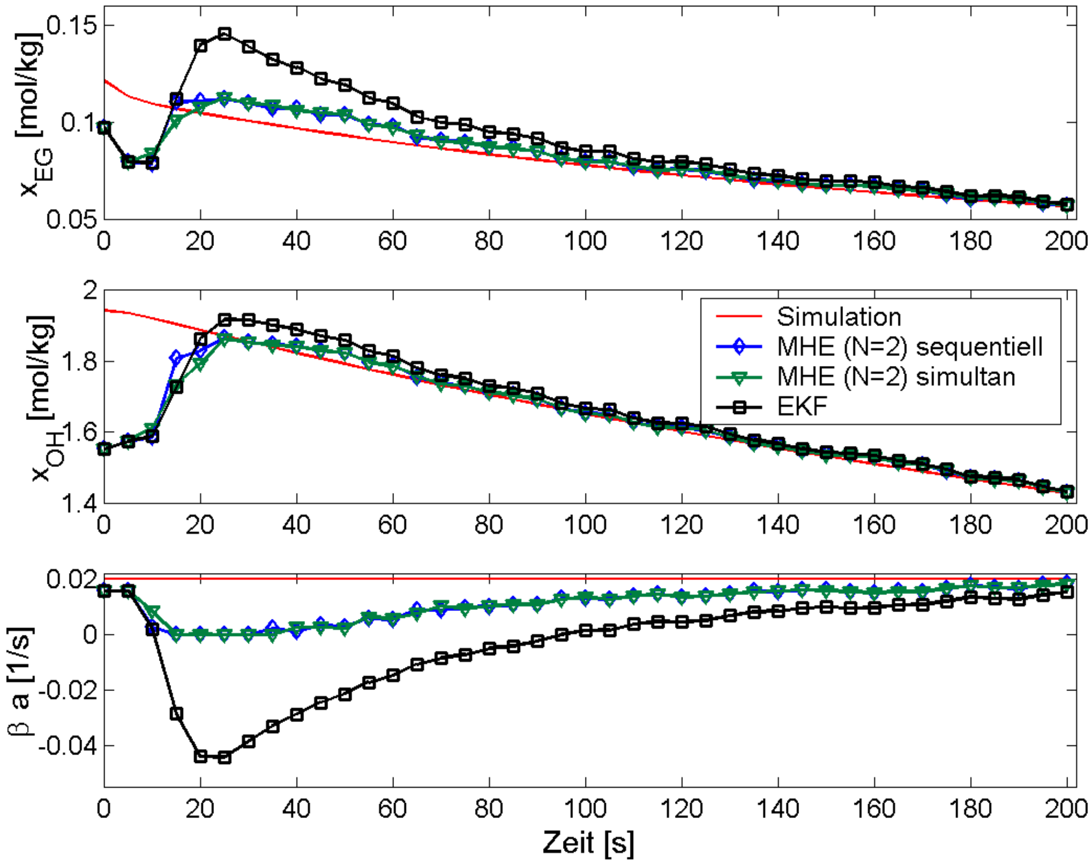
$$\underline{P}_0 = c_p \cdot r(0) \cdot \mathrm{diag}\left(\hat{x}_{10}, \hat{x}_{20}, (\widehat{\beta \cdot a})_0\right)$$

$$r = c_r \cdot \hat{x}_2^{-1}$$

$$c_r = 1$$
$$c_p = 0.5$$

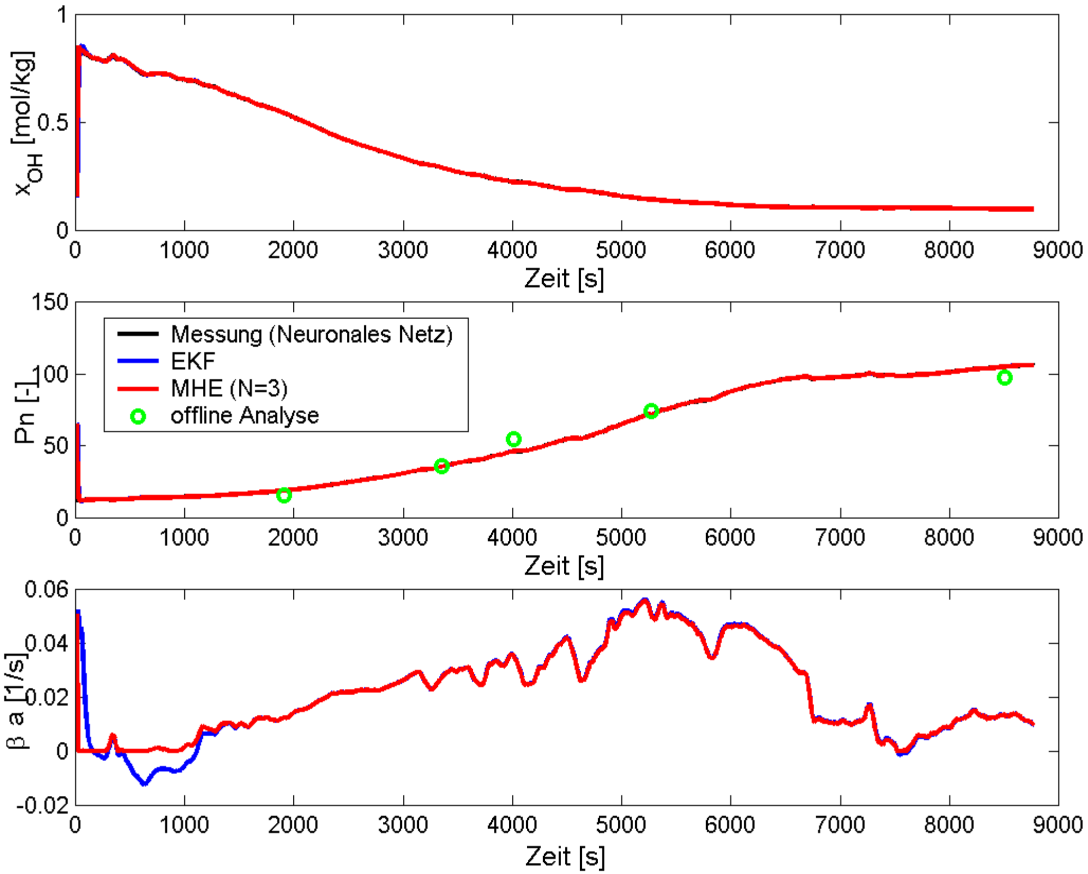
Comparison EKF-MHE, Simulation

- 20 % initial error in all states
- Constraints:
  $x_{EG} \geq 0$, $x_{OH} \geq 0$, $\beta a \geq 0$
- better convergence of the MHE



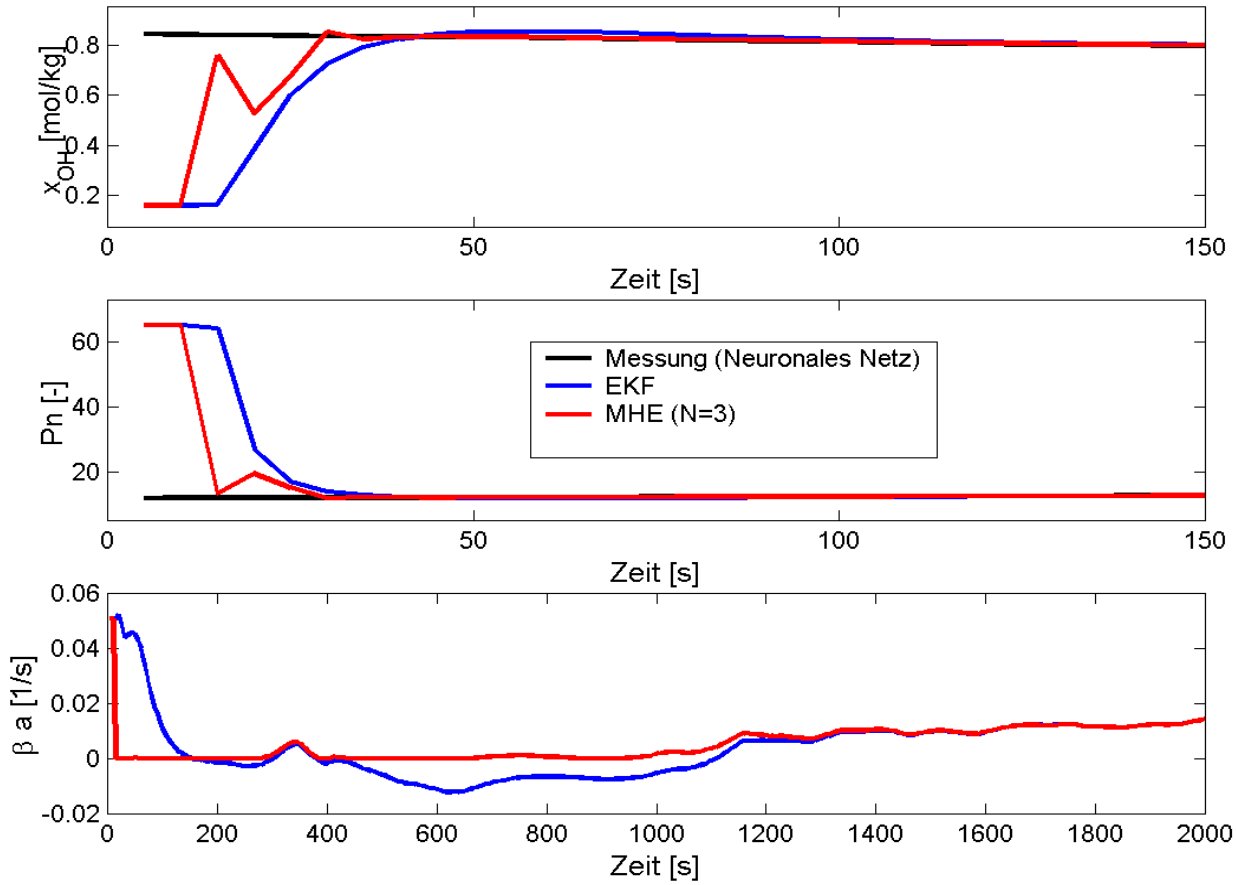
Comparison EKF-MHE

- MHE (N=3), experimental data
  - constraints
    $x_{EG} \geq 0$, $x_{OH} \geq 0$,
    $\beta a \geq 0$



Results PET Polycondensation

- Comparison EKF - MHE (N=3), experimental data

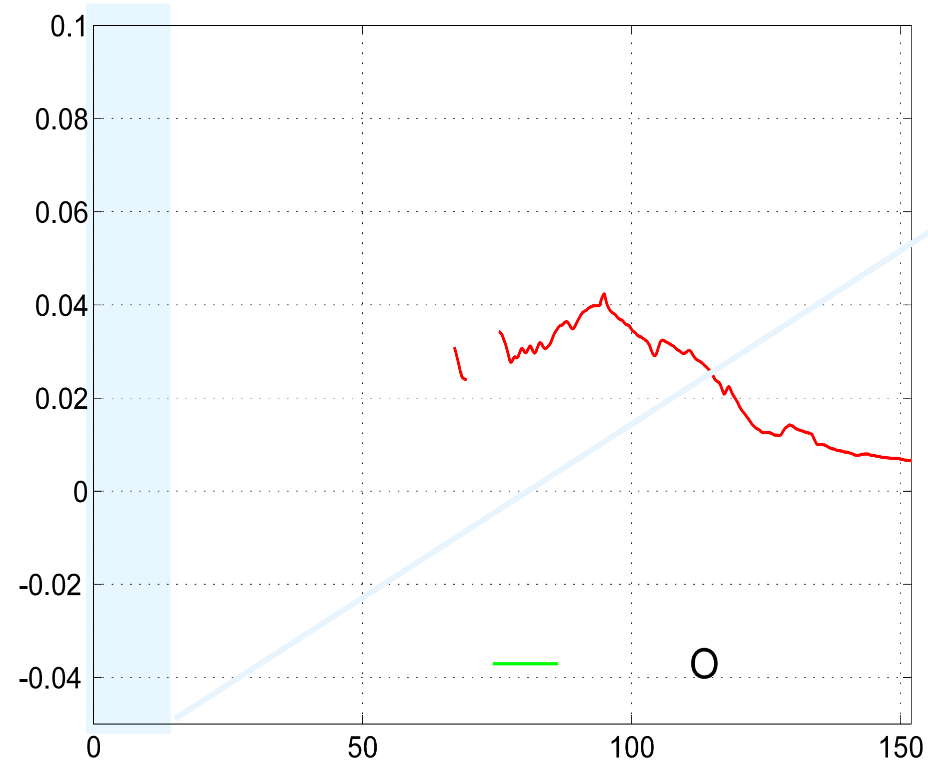


Estimation of the Mass Transfer Coefficient

- Different estimation methods gave similar results.
- Additional information about the process is provided.
- Estimation results enabled to improve process operation.

⇒ Development of improved trajectories of stirrer speed and reaction temperature

- Result: reduction of batch time by 10-15 %



Summary EKF vs. MHE vs. CEKF

- EKF:
  - Simple
  - Often good results
  - Tuning is not straightforward, requires insight and trial-and-error
  - Instability may occur
- MHE:
  - Uses full nonlinear model $\Rightarrow$ better convergence
  - Constraints on states and disturbances can be imposed
  - Numerically demanding
  - Tuning also an issue
- CEKF:
  - Good compromise between performance and effort