State estimation of two-time scale multiple models with unmeasurable premise variables. Application to biological reactors

A. M. Nagy Kiss, B. Marx, G. Mourot, G. Schutz, J. Ragot

Centre de Recherche en Automatique de Nancy, Nancy-Université, France
Centre de Recherche Public Henri Tudor, Luxembourg

CRAN

A. M. Nagy Kiss, B. Marx, G. Mourot, G. Schutz, J. Ragot

Objectives

1. Propose a state estimation method for two-time scale nonlinear systems
2. Application to the model of an activated sludge bioreactor of a Waste Water Treatment Plant
3. Need for state estimation of environmental plant with limited sensors

Objectives and context

Objectives

1. Propose a state estimation method for two-time scale nonlinear systems
2. Application to the model of an activated sludge bioreactor of a Waste Water Treatment Plant
3. Need for state estimation of environmental plant with limited sensors

Context and tools

1. Difficulty to deal with the **modeling complexity** of nonlinear systems
   $\rightarrow$ Multiple Model approach
2. Existence of multiple **time scale dynamics**
   $\rightarrow$ descriptor approach
3. State estimation based on $\mathcal{L}_2$-gain minimization

Outline of the presentation

What is a Multiple Model ?

- A multiple model (or Takagi-Sugeno model) is defined by

$$\dot{x}(t) = \sum_{i=1}^{r} \mu_i(z(t))\,(A_i x(t) + B_i u(t))$$

$$y(t) = \sum_{i=1}^{r} \mu_i(z(t))\,(C_i x(t) + D_i u(t))$$

where $x(t)$ is the state, $u(t)$ is the input and $y(t)$ is the output.

- A multiple model (or Takagi-Sugeno model) is defined by

$$\dot{x}(t) = \sum_{i=1}^{r} \mu_i(z(t)) \left(A_i x(t) + B_i u(t)\right)$$

$$y(t) = \sum_{i=1}^{r} \mu_i(z(t)) \left(C_i x(t) + D_i u(t)\right)$$

  where $x(t)$ is the state, $u(t)$ is the input and $y(t)$ is the output.
- The activating functions $\mu_i(.)$ depend on the premise variable $z(t)$ and satisfy

$$\sum_{i=1}^{r} \mu_i(z(t)) = 1 \quad \text{and} \quad 0 \leq \mu_i(z(t)) \leq 1$$

- The premise variable can be
  - measurable (e.g. u or y)
  - unmeasurable (e.g. x) : more general, more difficult, less studied.

Interests to use a Multiple Model

CRAN

A. M. Nagy Kiss, B. Marx, G. Mourot, G. Schutz, J. Ragot

- Any nonlinear system can be equivalently written as a Multiple Model on a compact set of the state space (sector nonlinearity approach)

$$\begin{cases} \dot{x} = f(x,u) \\ y = g(x,u) \\ nonlinear \end{cases} \Rightarrow \begin{cases} \dot{x} = A(x,u)x + B(x,u)u \\ y = C(x,u)x + D(x,u)u \\ Quasi - LPV \end{cases} \Rightarrow \begin{cases} \dot{x} = \sum_{i=1}^{r} \mu_i(x,u)(A_i x + B_i u) \\ y = \sum_{i=1}^{r} \mu_i(x,u)(C_i x + D_i u) \\ Multiple\ Model \end{cases}$$

$\rightarrow$ MM with unmeasurable premise variable are generally obtained

1. Nagy, Mourot, Marx, Ragot, Schutz, Systematic multi-modeling methodology applied to an activated sludge reactor model, Industrial & Engineering Chemistry Research, Vol. 46(6), pp. 2790-2799, 2010

- Any nonlinear system can be equivalently written as a Multiple Model on a compact set of the state space (sector nonlinearity approach)

$$\begin{cases} \dot{x} = f(x,u) \\ y = g(x,u) \\ nonlinear \end{cases} \Rightarrow \begin{cases} \dot{x} = A(x,u)x + B(x,u)u \\ y = C(x,u)x + D(x,u)u \\ Quasi - LPV \end{cases} \Rightarrow \begin{cases} \dot{x} = \sum_{i=1}^{r} \mu_i(x,u)(A_i x + B_i u) \\ y = \sum_{i=1}^{r} \mu_i(x,u)(C_i x + D_i u) \\ Multiple\ Model \end{cases}$$

$\rightarrow$ MM with unmeasurable premise variable are generally obtained

- The nonlinearities are rejected in the activating functions
$\rightarrow$ stability/performance analysis and controller/observer design can be carried out with classical tools (Lyapunov functions, LMI conditions, . . .)

1. Nagy, Mourot, Marx, Ragot, Schutz, Systematic multi-modeling methodology applied to an activated sludge reactor model, Industrial & Engineering Chemistry Research, Vol. 46(6), pp. 2790-2799, 2010

Interests to use a Multiple Model

- Any nonlinear system can be equivalently written as a Multiple Model on a compact set of the state space (sector nonlinearity approach)

$$\begin{cases} \dot{x} = f(x,u) \\ y = g(x,u) \\ nonlinear \end{cases} \Rightarrow \begin{cases} \dot{x} = A(x,u)x + B(x,u)u \\ y = C(x,u)x + D(x,u)u \\ Quasi-LPV \end{cases} \Rightarrow \begin{cases} \dot{x} = \sum_{i=1}^{r} \mu_i(x,u)(A_i x + B_i u) \\ y = \sum_{i=1}^{r} \mu_i(x,u)(C_i x + D_i u) \\ Multiple\ Model \end{cases}$$

  $\rightarrow$ MM with unmeasurable premise variable are generally obtained

- The nonlinearities are rejected in the activating functions
  $\rightarrow$ stability/performance analysis and controller/observer design can be carried out with classical tools (Lyapunov functions, LMI conditions, ...)

- Different equivalent re-writing of the original nonlinear system can be obtained
  $\rightarrow$ guidelines in order to select the most suitable one have been given [1]

1. Nagy, Mourot, Marx, Ragot, Schutz, Systematic multi-modeling methodology applied to an activated sludge reactor model, Industrial & Engineering Chemistry Research, Vol. 46(6), pp. 2790-2799, 2010

Two-time scale Multiple Models : a singular approach

- nonlinear system with two time scale are generally dealt with singular perturbation

$$\varepsilon \dot{x}_f(t) = f_f(x_s(t), x_f(t), u(t), \varepsilon)$$
$$\dot{x}_s(t) = f_s(x_s(t), x_f(t), u(t), \varepsilon)$$

where x_s is the slow state, x_f is the fast state and $\varepsilon > 0$ is the singular perturbed parameter

- nonlinear system with two time scale are generally dealt with singular perturbation

$$\varepsilon \dot{x}_f(t) = f_f(x_s(t), x_f(t), u(t), \varepsilon)$$
$$\dot{x}_s(t) = f_s(x_s(t), x_f(t), u(t), \varepsilon)$$

where x_s is the slow state, x_f is the fast state and $\varepsilon > 0$ is the singular perturbed parameter

- In the limit case (i.e. $\varepsilon \to 0$), a differential-algebraic system is obtained

$$0 = f_f(x_s(t), x_f(t), u(t), 0)$$
$$\dot{x}_s(t) = f_s(x_s(t), x_f(t), u(t), 0)$$

- A Q-LPV form can be derived

$$0 = A_{ff}(x,u)x_f(t) + A_{fs}(x,u)x_s(t) + B_f(x,u)u(t)$$
$$\dot{x}_s(t) = A_{sf}(x,u)x_f(t) + A_{ss}(x,u)x_s(t) + B_s(x,u)u(t)$$
$$y(t) = C_f(x,u)x_f(t) + C_s(x,u)x_s(t) + D(x,u)u(t)$$

- From the Q-LPV form can be derived a singular multiple model form

$$0 = \sum_{i=1}^{r} \mu_i(x,u)\left(A^i_{ff}x_f(t) + A^i_{fs}x_s(t) + B^i_f u(t)\right)$$

$$\dot{x}_s(t) = \sum_{i=1}^{r} \mu_i(x,u)\left(A^i_{sf}x_f(t) + A^i_{ss}x_s(t) + B^i_s u(t)\right)$$

$$y(t) = \sum_{i=1}^{r} \mu_i(x,u)\left(C^i_f x_f(t) + C^i_s x_s(t) + D^i u(t)\right)$$

- From the Q-LPV form can be derived a singular multiple model form

$$0 = \sum_{i=1}^{r} \mu_i(x,u) \left(A_{ff}^i x_f(t) + A_{fs}^i x_s(t) + B_f^i u(t) \right)$$

$$\dot{x}_s(t) = \sum_{i=1}^{r} \mu_i(x,u) \left(A_{sf}^i x_f(t) + A_{ss}^i x_s(t) + B_s^i u(t) \right)$$

$$y(t) = \sum_{i=1}^{r} \mu_i(x,u) \left(C_f^i x_f(t) + C_s^i x_s(t) + D^i u(t) \right)$$

- It is assumed that the output is linear in x_f and x_s and that $B_f(x,u)$ (in the Q-LPV form) only depend on u. Then it follows

$$0 = \sum_{i=1}^{r} \mu_i(x,u) \left(A_{ff}^i x_f(t) + A_{fs}^i x_s(t) \right) + \sum_{i=1}^{\tilde{r}} \tilde{\mu}_i(u) B_f^i u(t)$$

$$\dot{x}_s(t) = \sum_{i=1}^{r} \mu_i(x,u) \left(A_{sf}^i x_f(t) + A_{ss}^i x_s(t) + B_s^i u(t) \right)$$

$$y(t) = C_f x_f(t) + C_s x_s(t)$$

State estimation

- As classically done with descriptor systems
  $\rightarrow$ the algebraic equation is injected into an augmented output $y_a(t)$
  $\rightarrow$ the algebraic state (i.e. x_f) is considered as unknown input $d(t)$

$$\dot{x}_s(t) = \sum_{i=1}^{r} \mu_i(x_s, d, u) \left(A^i_{sf} d(t) + A^i_{ss} x_s(t) + B^i_s u(t) \right)$$

$$y_a(t) = \begin{bmatrix} \sum_{i=1}^{\tilde{r}} \tilde{\mu}_i(u) B^i_f u(t) \\ y(t) \end{bmatrix} = \sum_{i=1}^{\tilde{r}} \tilde{\mu}_i(u) \left(\begin{bmatrix} A^i_{fs} \\ C_s \end{bmatrix} x_s(t) + \begin{bmatrix} A^i_{ff} \\ C_f \end{bmatrix} d(t) \right)$$

- As classically done with descriptor systems
$\rightarrow$ the algebraic equation is injected into an augmented output $y_a(t)$
$\rightarrow$ the algebraic state (i.e. x_f) is considered as unknown input $d(t)$

$$\dot{x}_s(t) = \sum_{i=1}^{r} \mu_i(x_s, d, u) \left(A_{sf}^i d(t) + A_{ss}^i x_s(t) + B_s^i u(t) \right)$$

$$y_a(t) = \begin{bmatrix} \sum_{i=1}^{\tilde{r}} \tilde{\mu}_i(u) B_f^i u(t) \\ y(t) \end{bmatrix} = \sum_{i=1}^{\tilde{r}} \tilde{\mu}_i(u) \left(\begin{bmatrix} A_{fs}^i \\ C_s \end{bmatrix} x_s(t) + \begin{bmatrix} A_{ff}^i \\ C_f \end{bmatrix} d(t) \right)$$

- In order to estimate both the state (i.e. x_s) and the unknown input (i.e. x_f) a proportional integral observer (PIO) is proposed

$$\dot{\hat{x}}_s(t) = \sum_{i=1}^{r} \mu_i(\hat{x}_s, \hat{d}, u) \left(A_{ss}^i \hat{x}_s(t) + A_{sf}^i \hat{d} + B_u^i(t) + K_P^i(y_a(t) - \hat{y}_a(t)) \right)$$

$$\dot{\hat{d}}(t) = \sum_{i=1}^{r} \mu_i(\hat{x}_s, \hat{d}, u) K_I^i (y_a(t) - \hat{y}_a(t))$$

$$\hat{y}_a(t) = \sum_{i=1}^{\tilde{r}} \tilde{\mu}_i(u) \left(\begin{bmatrix} A_{fs}^i \\ C_s \end{bmatrix} \hat{x}_s(t) + \begin{bmatrix} A_{ff}^i \\ C_f \end{bmatrix} \hat{d}(t) \right)$$

where the gains K_P^i and K_I^i are to be determined.

Estimation error derivation

CRAN

A. M. Nagy Kiss, B. Marx, G. Mourot, G. Schutz, J. Ragot

- The estimation error is not easily derived because
  $\rightarrow$ the activating functions in the system depend on x_s and d
  $\rightarrow$ the activating functions in the observer depend on $\hat{x}_s$ and $\hat{d}$

- The estimation error is not easily derived because
  $\rightarrow$ the activating functions in the system depend on x_s and d
  $\rightarrow$ the activating functions in the observer depend on $\hat{x}_s$ and $\hat{d}$
- To overcome this difficulty, the system is written as

$$\dot{x}_s(t) = \sum_{i=1}^{r} \mu_i(\hat{x}_s, \hat{d}, u) \left(A_{sf}^i x_f(t) + A_{ss}^i x_s(t) + B_s^i u(t) \right) + \omega(t)$$

where $\omega(t)$ is defined by

$$\omega(t) = \sum_{i=1}^{r} (\mu_i(x_s, d, u) - \mu_i(\hat{x}_s, \hat{d}, u)) \left(A_{sf}^i x_f(t) + A_{ss}^i x_s(t) + B_s^i u(t) \right)$$

- Observer design is then based on the minimization of the $\mathcal{L}_2$-gain from $\omega(t)$ to the estimation error.
  $\rightarrow$ no Lipschitz assumption is needed on the activating functions.

- Assuming that $\dot{d}(t) = 0$ (technical assumption, that is practically relaxed), the estimation error becomes

$$\dot{e}(t) = \sum_{i=1}^{r} \sum_{j=1}^{\tilde{r}} \mu_i(\hat{x}_s, \hat{d}, u)\tilde{\mu}_j(u) \left(\tilde{A}_i - K_i \tilde{C}_j\right) e(t) + \Gamma\omega(t)$$

where

$$e(t) = \begin{bmatrix} x_s(t) - \hat{x}_s(t) \\ d(t) - \hat{d}(t) \end{bmatrix} \quad \tilde{A}_i = \begin{bmatrix} A^i_{ss} & A^i_{sf} \\ 0 & 0 \end{bmatrix}$$

$$\tilde{C}_j = \begin{bmatrix} A^j_{fs} & A^i_{ff} \\ C_s & C_f \end{bmatrix} \quad K_i = \begin{bmatrix} K^i_P \\ K^i_I \end{bmatrix} \quad \Gamma = \begin{bmatrix} I \\ 0 \end{bmatrix}$$

Estimation error derivation

- Assuming that $\dot{d}(t) = 0$ (technical assumption, that is practically relaxed), the estimation error becomes

$$\dot{e}(t) = \sum_{i=1}^{r} \sum_{j=1}^{\tilde{r}} \mu_i(\hat{x}_s, \hat{d}, u)\tilde{\mu}_j(u) \left(\tilde{A}_i - K_i\tilde{C}_j\right) e(t) + \Gamma\omega(t)$$

where

$$e(t) = \begin{bmatrix} x_s(t) - \hat{x}_s(t) \\ d(t) - \hat{d}(t) \end{bmatrix} \quad \tilde{A}_i = \begin{bmatrix} A_{ss}^i & A_{sf}^i \\ 0 & 0 \end{bmatrix}$$

$$\tilde{C}_j = \begin{bmatrix} A_{fs}^j & A_{ff}^i \\ C_s & C_f \end{bmatrix} \quad K_i = \begin{bmatrix} K_P^i \\ K_I^i \end{bmatrix} \quad \Gamma = \begin{bmatrix} I \\ 0 \end{bmatrix}$$

- The bounded real lemma allows to derive sufficient LMI conditions for the $\mathcal{L}_2$-gain from $\omega(t)$ to $e(t)$ to be bounded by a prescribed positive real number.

Theorem :

The optimal PIO for the two-time scale multiple model is obtained if $\exists$ matrix $X = X^T > 0$, matrices M_i and a positive scalar λ, minimizing λ under the following LMI constraints for $i = 1, \ldots, r$ and $j = 1, \ldots, \tilde{r}$:

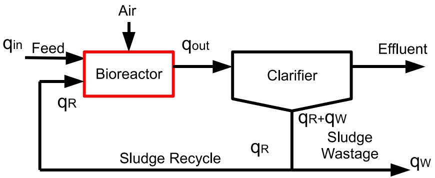
$$\begin{bmatrix} \tilde{A}_i^T X + X\tilde{A}_i - \tilde{C}_j^T M_i^T - M_i\tilde{C}_j + I & X\Gamma \\ \Gamma^T X & -\lambda I \end{bmatrix} < 0$$

The observer gains are given by :

$$\begin{bmatrix} K_P^i \\ K_I^i \end{bmatrix} = X^{-1} M_i$$

Application to a wastewater treatment process

Activated sludge Model of the biological reactor and clarifier (ASM1)



1. Model reduction :
   - only the carbonated ($S_S(t)$) pollution is considered
2. Simplifying assumptions :
   - the bioreactor volume is constant :

$$q_{out}(t) = q_{in}(t) + q_R(t)$$

   - the clarifier is perfect :
     $\rightarrow$ all the biomass is recycled or stored, none is rejected into the effluent

$$(q_{in}(t) + q_R(t))X_{BH} = (q_W(t) + q_R(t))X_{BH,R}(t)$$
$$S_S(t) = S_{S,R}(t)$$

Nonlinear model of the activated sludge bioreactor

The reduced ASM1 model

$$\dot{S}_S = \frac{q_{in}}{V}\left(S_{S,in} - S_S\right) + (1-f)b_H X_{BH} - \frac{\mu_H}{Y_H}\frac{S_S}{K_S + S_S}\frac{S_O}{K_{OH} + S_O}X_{BH}$$

$$\dot{S}_O = -\frac{q_{in}}{V}S_O + K q_a\left(S_{O,sat} - S_O\right) - \frac{1-Y_H}{Y_H}\mu_H\frac{S_S}{K_S + S_S}\frac{S_O}{K_{OH} + S_O}X_{BH}$$

$$\dot{X}_{BH} = \frac{q_{in}}{V}X_{BH,in} - \frac{q_W}{V}\frac{q_{in} + q_R}{q_W + q_R}X_{BH} + \mu_H\frac{S_S}{K_S + S_S}\frac{S_O}{K_{OH} + S_O}X_{BH} - b_H X_{BH}$$

where the state, input and output are defined by

$$x(t) = \begin{pmatrix} S_S(t) \\ S_O(t) \\ X_{BH}(t) \end{pmatrix} \qquad u(t) = \begin{pmatrix} q_{in}(t) \\ q_a(t) \\ S_{S,in}(t) \\ X_{BH,in}(t) \end{pmatrix} \qquad y(t) = \begin{pmatrix} s_S(t) \\ S_O(t) \end{pmatrix} + \delta(t)$$

where $\delta(t)$ is a zero mean measurement noise and $(V, \mu_H, b_H, f, Y_H, S_{O,sat}, K_S, K_{OH}, K)$ are known constant parameters.

Multiple Model of the ASM1

From the previous nonlinear model, a Multiple Model can be built.

- The decision variables are defined by :

$$z_1(u(t)) = \frac{q_{in}(t)}{V}$$

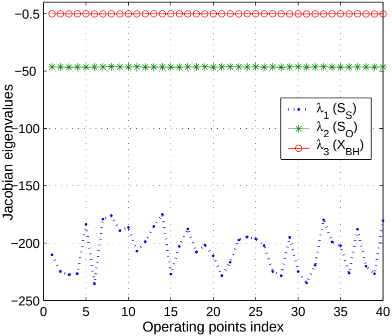
$$z_2(x(t)) = \frac{1}{K_S + S_S(t)} \frac{S_O(t)}{K_{OH} + S_O(t)} X_{BH}(t)$$

$$z_3(u(t)) = q_a(t)$$

- The number of submodels is $r = 8$

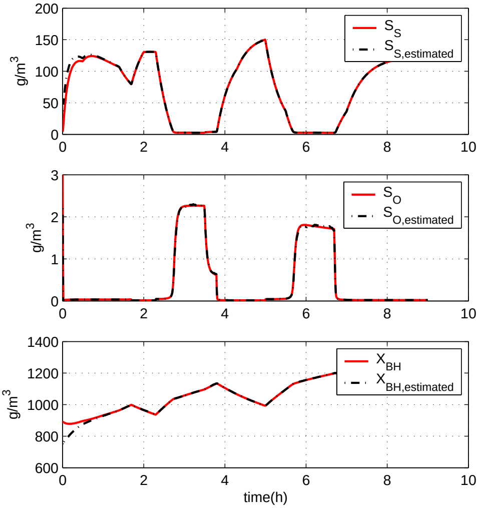
Slow and fast dynamic separation

The mode separation is done by computing the eigenvalues of the Jacobian of the linearized model at 40 operating points :
$\rightarrow$ the separation mode is clearly independent of the operating points.

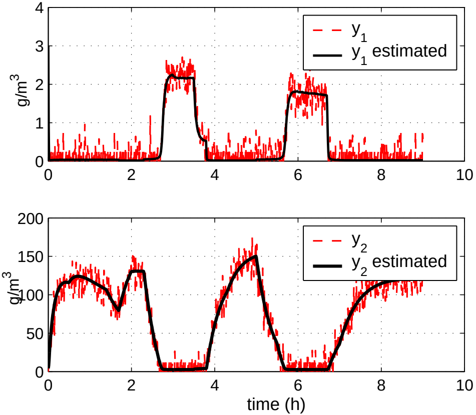


$\rightarrow$ fast variable : carbonated pollution S_S
$\rightarrow$ slow variables : heterotrophic biomass x_{BH} and dissolved oxygen S_0.

State estimation results of the reduced ASM1



Output estimation results of the reduced ASM1



Conclusions and future prospects

CRAN

A. M. Nagy Kiss, B. Marx, G. Mourot, G. Schutz, J. Ragot

Conclusions

1. A nonlinear system is written as a two time scale multiple model with unmeasurable premise variables
2. Observer design of a PI observer by $\mathcal{L}_2$-gain minimization
3. Application to a reduced model of the wastewater treatment plant

CRAN
A. M. Nagy Kiss, B. Marx, G. Mourot, G. Schutz, J. Ragot

Conclusions

1. A nonlinear system is written as a two time scale multiple model with unmeasurable premise variables
2. Observer design of a PI observer by $\mathcal{L}_2$-gain minimization
3. Application to a reduced model of the wastewater treatment plant

Future prospects

1. Application to a less reduced model of the wastewater treatment plant
2. Estimation of singular MM with model uncertainties
3. Extension of the estimation results to system diagnosis (bank of observers for FDI)
4. Apply $\mathcal{L}_2$-based technique for ASM1 model reduction

Thank you for your attention