Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Multi-model approach to nonlinear system identification with unknown time delay ⋆

Lei Chen *,** Biao Huang ** Fei Liu *

* *Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Institute of Automation, Jiangnan University, Wuxi, 214122, P.R. China. (e-mail: chenlei_happy@hotmail.com, fliu@jiangnan.edu.cn)*
** *Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB T6G 2G6, Canada. (e-mail: biao.huang@ualberta.ca)*

Abstract: This paper is concerned with identification of nonlinear systems with a noisy scheduling variable, and the measurement of the system has an unknown time delay. Auto regressive exogenous (ARX) models are selected as the local models, and multiple local models are identified along the process operating points. The dynamics of a nonlinear system are represented by associating a normalized exponential function with each of the ARX models; therein, the normalized exponential function is acted as the probability density function. The parameters of the ARX models and the exponential functions as well as the unknown time delay are estimated simultaneously under the expectation maximization (EM) algorithm using the retarded input-output data. A CSTR example is given to verify the proposed identification approach.

Keywords: Nonlinear system identification, Expectation maximization algorithm, Multiple models, Time delay

1. INTRODUCTION

Time delay is a long discussed problem in engineering sciences, and it is well known that, most of real-life systems may be more accurately represented by nonlinear models. Therefore, basic mathematical models of real process phenomena are constructed by nonlinear time delay systems. Since the unique challenges are posed by the complex stochastic nonlinear dynamics, nonlinear time delay systems represent an additional level of complexity. Identification of nonlinear time delay systems is a thorny problem.

Over the last few decades, the methods for identifying linear time delay systems have been wildly developed [Björklund and Ljung, 2003, Richard, 2003]. For the nonlinear time delay systems, Yazdizadeh and Khorasani [2002] proposed four neuro-dynamic architectures for identifying different classes of nonlinear time delay systems. Liu et al. [2009] proposed a method for the identification of unknown network parameters and topological structure simultaneously; wherein, the uncertain general complex networks has time delay. Cao and Frank [2000] presented a stability analysis and design approach for nonlinear time delay systems under Takagi-Sugeno (TS) fuzzy modeling and control approach. Anguelova and Wennberg [2008] applied the non-commutative rings theory to analyze the delay identification problem for nonlinear constant delay systems.

In general, most of the industrial processes are nonlinear time delay systems. Such systems are often operated along certain fixed operating trajectories, and several predetermined operating points constitute these operating trajectories. This paper is an extension of the work of Chen et al. [2013] to a more general scenario when the nonlinear system has an uncertain scheduling variable as well as unknown time delay in the process output measurements. In Chen et al. [2013], we considered identification problem for nonlinear processes without time delay. In practice, however, most processes have time delays which constitute considerable challenges in identification problem, and therefore solving time delay problem is often a necessary step in identification. Using the EM algorithm, a multiple model based identification procedure is developed, auto regressive exogenous (ARX) models are selected as the local models, and multiple local models are identified in different operating regions using the retarded data. For the complete representation of the nonlinear system, a normalized exponential function is then used to combine all the ARX models. The parameters of the nonlinear model and the unknown time delay are simultaneously estimated.

This paper is organized as follows: Section 2 lays out the problem formulation. Section 3 derives the identification procedure under the EM framework. Section 4 shows the

⋆ This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, Alberta Innovates Technology Futures (AITF), the Fundamental Research Funds for the Central Universities (JUDCF10064), Jiangsu Innovation Program for Graduates (CXZZ11_0464), China Scholarship Council (CSC), the Foundation of NSFC61134007 and 111 Project (B12018)

simulation results on a continuous stirred tank reactor (CSTR) example. Conclusions are given in Section 5.

2. PROBLEM FORMULATION

Let the true nonlinear time delay system be represented as

$$h(\dot{c}_k, c_k, y_k, u_k, z_k, w_k, u'_k, T^0_{1:M}, \lambda, k, \epsilon_k) = 0, \quad (1)$$

where $u_k \in \mathcal{U} \subseteq \mathbb{R}^m$ for $k \in \mathbb{N}$ is the process input; $c_k \in \mathcal{C} \subseteq \mathbb{R}^r$ and $y_k \in \mathcal{Y} \subseteq \mathbb{R}$ are the state and measurement, respectively; $u'_k \in \mathcal{U}' \subseteq \mathbb{R}^s$ affects the scheduling variables as an input variable; $z_k \in \mathcal{Z} \subseteq \mathbb{R}$ and $w_k \in \mathcal{W} \subseteq \mathbb{R}^q$ are the hidden and observed scheduling variables, and z_k and w_k are independent; $T^0_{1:M} = \{T^0_1, T^0_2, \ldots, T^0_M\}$ represent M different operating points; λ is an unknown time delay, which is bounded by λ_{min} and λ_{max}; and $\epsilon_k \in \mathbb{R}$ represents the process noise. Here $h(\cdot)$ is a nonlinear mapping function.

The unknown time delay λ is integer-valued and its uncertainty can follow any discrete distribution. In this work, we consider an uniform distribution, i.e.,

$$Pr(\lambda = j) = \frac{1}{\lambda_{max} - \lambda_{min} + 1}, \; j = \lambda_{min}, \ldots, \lambda_{max}. \quad (2)$$

Although the scheduling variable z_k in Eq. (1) is a hidden variable, through another process variable w_k, z_k can be observed; then the dynamics of the scheduling variables can be represented using the following state-space model:

$$\eta_k = f(\eta_{k-1}, u'_{k-1}, \gamma_{k-1}), \quad (3a)$$
$$w_k = g(\eta_k, u'_k, v_k), \quad (3b)$$
$$z_k = \psi(\eta_k), \quad (3c)$$

where $\gamma_k \in \mathbb{R}^p$ and $v_k \in \mathbb{R}^q$ are Gaussian noise represented by $\mathcal{N}(0, R_\gamma)$ and $\mathcal{N}(0, R_v)$, respectively; and $\eta_k \in \mathcal{H} \subseteq \mathbb{R}^p$ is the state of the scheduling variable. $f(\cdot)$ and $g(\cdot)$ are p-dimensional and q-dimensional mapping functions, and each can be nonlinear or linear. $\psi(\cdot)$ generates the scheduling variable, and is a 1-dimensional mapping function[Chen et al., 2013].

Eq. (3) can be represented by

$$Z_0 = z_0 \sim p(z_0), \quad (4)$$
$$Z_k|(Z_{k-1} = z_{k-1}) \sim p(\cdot|z_{k-1}), \quad (5)$$
$$W_k|(Z_k = z_k) \sim p(\cdot|z_k). \quad (6)$$

For any generic sequence χ_k, let $\chi_{i:j} = \{\chi_i, \chi_{i+1}, \ldots, \chi_j\}$, with $\chi_{i:j} = 0$ for $i > j$. It is assumed that the Eq. (3) is known, the data set $\{u'_{1:N}, w_{1:N}\}$ and $T^0_{1:M}$ are given, and the retarded data $\{u_{1:N}, y_{1:N}\}$ can be accessed. Therefore, the observed data set is $C_{obs} = \{y_{1:N}, u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M}\}$ [Chen et al., 2013].

Nonlinear dynamics in different operating conditions may be described by multiple local models. The local dynamics of Eq. (1) may be approximated by ARX models. Ljung [1987] has justified that ARX models can be used in approximating any linear dynamics. Due to unknown time delay, the ARX model is given as

$$y_k = \theta^{\mathrm{T}}_{I_k} x_{k-\lambda} + e_k, \quad (7)$$

where $x_{k-\lambda} \in \mathcal{X} \subseteq \mathbb{R}^n$ is the regressor, which can be expressed as

$$x_{k-\lambda} \triangleq [y_{k-1}, \ldots, y_{k-n_b}, u^{\mathrm{T}}_{k-1-\lambda}, \ldots, u^{\mathrm{T}}_{k-n_a-\lambda}]^{\mathrm{T}}. \quad (8)$$

Here n_a and n_b are the orders of the input and output polynomial, respectively, and $n = mn_a + n_b$; $e_k \in \mathbb{R}$ is Gaussian noise with zero mean and unknown variance σ^2. It is assumed that $u_k = 0$, $y_k = 0$ for $k \leq 0$. I_k in Eq. (7) represents the local model identity at sampling time k, and is a hidden variable. Therefore the parameters of the local model are $\Theta_{I_k} = \{\theta_{I_k}, \sigma\}$. We assume that M local ARX models represent the dynamics around the M operating points of Eq. (1), such that

$$Y_k|(X_k = x_{k-\lambda}, I_k = i) \sim p_{\Theta_i}(\cdot|x_{k-\lambda}), \quad 1 \leq i \leq M. \quad (9)$$

Through a local linear model, the process dynamics within the relatively small region of an operating point can be approximated. We can compute the probability of y_k given all the past information following the approach of Chen et al. [2013]:

$$\begin{aligned} & p_\Theta(y_k|y_{1:k-1}, c_{1:k}, u_{1:k}, z_{0:k}, w_{1:k}, u'_{1:k}, T^0_{1:M}, \lambda) \\ & = \sum_{i=1}^{M} p_\Theta(y_k, I_k = i|y_{1:k-1}, c_{1:k}, u_{1:k}, z_{0:k}, w_{1:k}, \\ & \quad u'_{1:k}, T^0_{1:M}, \lambda) \quad (10a) \\ & = \sum_{i=1}^{M} \alpha_{k,i} p_{\Theta_i}(y_k|x_{k-\lambda}). \quad (10b) \end{aligned}$$

where Θ is a set of model parameters. $\alpha_{k,i} = Pr_\Theta(I_k = i|y_{1:k-1}, c_{1:k}, u_{1:k}, z_{0:k}, w_{1:k}, u'_{1:k}, T^0_{1:M}, \lambda)$ is the probability of $I_k = i$th local model taking effect at sampling time k. Given I_k, y_k is independent of $c_{1:k}$, $z_{0:k}$, $w_{1:k}$, $u'_{1:k}$, $T^0_{1:M}$ (see Eq. (9)). As a result, $p_\Theta(y_k|y_{1:k-1}, c_{1:k}, u_{1:k}, z_{0:k}, w_{1:k}, u'_{1:k}, T^0_{1:M}, I_k = i, \lambda)$ can be simplified, and written as $p_{\Theta_i}(y_k|x_{k-\lambda})$.

Generally, I_k can be inferred from z_k and $T^0_{1:M}$. As a result, in Eq. (10b), $\alpha_{k,i}$ can be simplified as $Pr_\Theta(I_k = i|z_k, T^0_{1:M})$. Here a normalized exponential function is used to model $\alpha_{k,i}$, which is proposed in Jin et al. [2011] and Zhao et al. [2012], such that

$$\alpha_{k,i} = \frac{\exp(-\frac{(z_k - T^0_i)^2}{2(o_i)^2})}{\sum_{i=1}^{M} \exp(-\frac{(z_k - T^0_i)^2}{2(o_i)^2})}, \quad (11)$$

where $o_i \in \mathcal{O} \subseteq \mathbb{R}$ need to be estimated, it denotes the validity width of the $I_k = i$th local model. In Eq. (11), o_i is bounded by $o_{i,max}$ and $o_{i,min}$, such that $o_{i,min} \leq o_i \leq o_{i,max}$. $O = \{o_1, o_2, \ldots, o_M\}$ denote a set of validity widths for the M local models. Therefore, Eq. (11) can be alternatively represented as

$$I_k|(Z_k = z_k, T^0_{1:M}) \sim Pr_O(\cdot|z_k, T^0_{1:M}). \quad (12)$$

Since I_k and z_k are hidden, and the time delay λ is unknown, the missing data set is $C_{mis} = \{I_{1:N}, z_{0:N}, \lambda\}$, and a complete data set can be denoted as $\{C_{obs}, C_{mis}\}$. $\Theta = \{\Theta_{I_k}, O\} = \bigcup_{i=1}^{M} \{\theta_i, \sigma, o_i\}$ are the parameters to be estimated.

In this paper, the main problem addressed is to estimate the parameter Θ and the unknown time delay λ, given C_{obs} and Eq. (3). The EM algorithm is adopted to solve the above problems.

3. EM ALGORITHM BASED MULTIPLE MODEL APPROACH

The EM algorithm is used for the maximum-likelihood estimation from incomplete data, which was first proposed by Dempster et.al. [1977]. The EM algorithm consists of two iterative steps: the expectation step (E-step) and the maximization step (M-step). In the E-step, we calculate the expectation of the log-likelihood function, Q-function, for the complete data set $\{C_{obs}, C_{mis}\}$ with respect to the missing data C_{mis} based on the current estimated parameter set Θ' and the observed data C_{obs}. In the M-step, the parameters are re-estimated through maximizing the Q-function.

The mathematical formulation of EM algorithm is [McLachlan and Krishnan, 2008]:

(1) *Initialization*: Set the initial guess of the parameters to Θ'.
(2) *E-step*: Use the current parameter Θ' to calculate the approximate Q-function as
$$Q(\Theta|\Theta') = \mathrm{E}_{C_{mis}|(C_{obs},\Theta')}\{\log p_\Theta(C_{obs}, C_{mis})\}. \tag{13}$$
(3) *M-step*: Estimate the new parameter through maximizing the Q-function as
$$\Theta = \arg\max_{\Theta} Q(\Theta|\Theta'). \tag{14}$$
Then set $\Theta' = \Theta$.
(4) *Iterate*: Evaluate the relative change of the estimated parameters,
$$\delta = \| \frac{\Theta - \Theta'}{\Theta'} \|. \tag{15}$$
If δ is larger than a pre-determined tolerance, then repeat steps 2 and 3.

The EM algorithm guarantees to converge to a stationary point under certain regularity conditions[Wu, 1983].

3.1 Formulation of the identification approach based on the EM algorithm

$p_\Theta(C_{obs}, C_{mis})$ is the complete likelihood function. It can be decomposed using the probability chain rule as

$$\begin{aligned}
&p_\Theta(C_{obs}, C_{mis}) \\
&= p_\Theta(y_{1:N}, u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M}, I_{1:N}, z_{0:N}, \lambda) && (16a)\\
&= p_\Theta(y_{1:N}|u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M}, I_{1:N}, z_{0:N}, \lambda)\\
&\quad \cdot Pr_\Theta(I_{1:N}|u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M}, z_{0:N}, \lambda)\\
&\quad \cdot p_\Theta(z_{0:N}|u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M}, \lambda)\\
&\quad \cdot p_\Theta(\lambda|u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M})\\
&\quad \cdot p_\Theta(u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M}). && (16b)
\end{aligned}$$

Following the similar derivations as Chen et al. [2013], we get

$$\begin{aligned}
&p_\Theta(C_{obs}, C_{mis})\\
&= \prod_{k=1}^{N} p_{\Theta_{I_k}}(y_k|x_{k-\lambda}) \cdot \prod_{k=1}^{N} Pr_O(I_k|z_k, T^0_{1:M}) \cdot C_1,
\end{aligned} \tag{17}$$

where $C_1 = p_\Theta(z_{0:N}|u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M}, \lambda) \cdot p_\Theta(\lambda|u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M}) \cdot p_\Theta(u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M})$, is independent of Θ.

Then by computing the E-step shown in Eq. (13), we get Eq. (18) as shown in the next page.

Since the integrand is a function of I_k, z_k and λ, the multidimensional integral with respect to $p_{\Theta'}(I_{1:N}, z_{0:N}, \lambda|C_{obs})$ can be simplified, and written as Eq. (19a), where $C_2 = \int_{I_{1:N}, z_{0:N}, \lambda} \log C_1 \cdot p_{\Theta'}(I_{1:N}, z_{0:N}, \lambda|C_{obs}) dz_{0:N} dI_{1:N} d\lambda$. Since the local model identity I_k and time delay λ are discrete random variables, we get Eq. (19b) as shown in the next page.

$p_{\Theta'}(I_k = i, z_k, \lambda = j|C_{obs})$ can be decomposed as $Pr_{\Theta'}(I_k = i|z_k, \lambda = j, C_{obs}) p_{\Theta'}(z_k|\lambda = j, C_{obs}) Pr_{\Theta'}(\lambda = j|C_{obs})$. Now to determine the Q-function in Eq. (19b), we need to compute the following probability functions:

(1) $p_{\Theta_i}(y_k|x_{k-\lambda}, \lambda = j)$
(2) $Pr_{o_i}(I_k = i|z_k, T^0_{1:M})$
(3) $Pr_{\Theta'}(I_k = i|z_k, \lambda = j, C_{obs})$
(4) $p_{\Theta'}(z_k|\lambda = j, C_{obs})$
(5) $Pr_{\Theta'}(\lambda = j|C_{obs})$

As Gaussian white noise has been assumed for the ARX model in Eq. (7), $p_{\Theta_i}(y_k|x_{k-\lambda}, \lambda = j)$ is a Gaussian probability density function, and $Pr_{o_i}(I_k = i|z_k, T^0_{1:M})$ can be calculated using Eq. (11).

Using the Bayes' rule, $Pr_{\Theta'}(I_k = i|z_k, \lambda = j, C_{obs})$ and $Pr_{\Theta'}(\lambda = j|C_{obs})$ can be derived as Eq. (20) and Eq. (21), respectively.

$$\begin{aligned}
&Pr_{\Theta'}(I_k = i|z_k, \lambda = j, y_k, x_{k-j}, T^0_{1:M})\\
&= \frac{p_{\Theta'_i}(y_k|x_{k-\lambda}, \lambda = j) Pr_{O'}(I_k = i|z_k, T^0_{1:M}) Pr(\lambda = j)}{\sum_{i=1}^{M} p_{\Theta'_i}(y_k|x_{k-\lambda}, \lambda = j) Pr_{O'}(I_k = i|z_k, T^0_{1:M}) Pr(\lambda = j)}.
\end{aligned} \tag{20}$$

$$\begin{aligned}
&Pr_{\Theta'}(\lambda = j|y_{1:N}, u_{1:N}, u'_{1:N}, w_{1:N}, T^0_{1:M})\\
&= \frac{\prod_{k=1}^{N} p_{\Theta'_i}(y_k|x_{k-\lambda}, \lambda = j) Pr(\lambda = j)}{\sum_{j=\lambda_{min}}^{\lambda_{max}} \left(\prod_{k=1}^{N} p_{\Theta'_i}(y_k|x_{k-\lambda}, \lambda = j) Pr(\lambda = j) \right)}.
\end{aligned} \tag{21}$$

Since z_k is a continuous random variable, an approximation method is necessary for computing the integral in Eq. (19b). For nonlinear dynamics of the scheduling variable, we use the Sequential Monte-Carlo (SMC) methods to get numerical solutions; for linear case, we use the Kalman smoother. Then the Q-function can be calculated as Eq. (22), where $\delta(\cdot)$ is a dirac-delta function, z^l_k is chosen from $p(z_k|\lambda = j, w_{1:N})$, and L is number of samples.

For the M-step shown in Eq. (14), we take derivatives of the Q-function with respect to θ_i, σ and o_i. Differentiating Eq. (22) with respect to θ_i,σ and o_i and equating it to zero, we get the updated parameters shown in Eqs. (23)-(25).

$$Q(\Theta|\Theta') = \mathrm{E}_{p_{\Theta'}(I_{1:N},z_{0:N},\lambda|C_{obs})}\left\{\log\prod_{k=1}^{N} p_{\Theta_{I_k}}(y_k|x_{k-\lambda}) + \log\prod_{k=1}^{N} Pr_O(I_k|z_k,T^0_{1:M}) + \log C_1\right\} \tag{18a}$$

$$= \int_{I_{1:N},z_{0:N},\lambda}\left(\sum_{k=1}^{N}\log p_{\Theta_{I_k}}(y_k|x_{k-\lambda}) + \sum_{k=1}^{N}\log Pr_O(I_k|z_k,T^0_{1:M}) + \log C_1\right) p_{\Theta'}(I_{1:N},z_{0:N},\lambda|C_{obs})dz_{0:N}dI_{1:N}d\lambda. \tag{18b}$$

$$\sum_{k=1}^{N}\int_{I_k,z_k,\lambda}\log p_{\Theta_{I_k}}(y_k|x_{k-\lambda})p_{\Theta'}(I_k,z_k,\lambda|C_{obs})dz_kdI_kd\lambda + \sum_{k=1}^{N}\int_{I_k,z_k,\lambda}\log Pr_O(I_k|z_k,T^0_{1:M})p_{\Theta'}(I_k,z_k,\lambda|C_{obs})dz_kdI_kd\lambda + C_2 \tag{19a}$$

$$= \sum_{k=1}^{N}\sum_{j=\lambda_{min}}^{\lambda_{max}}\sum_{i=1}^{M}\int_{z_k}\log p_{\Theta_i}(y_k|x_{k-\lambda},\lambda=j)p_{\Theta'}(I_k=i,z_k,\lambda=j|C_{obs})dz_k + \sum_{k=1}^{N}\sum_{j=\lambda_{min}}^{\lambda_{max}}\sum_{i=1}^{M}\int_{z_k}\log Pr_{o_i}(I_k=i|z_k,T^0_{1:M})p_{\Theta'}(I_k=i,z_k,\lambda=j|C_{obs})dz_k + C_2. \tag{19b}$$

$$Q(\Theta|\Theta') = \frac{1}{L}\sum_{k=1}^{N}\sum_{j=\lambda_{min}}^{\lambda_{max}}\sum_{i=1}^{M}\sum_{l=1}^{L}\log p_{\Theta_i}(y_k|x_{k-j})Pr_{\Theta'}(I_k=i|z_k^l,\lambda=j,C_{obs})Pr_{\Theta'}(\lambda=j|C_{obs}) + \frac{1}{L}\sum_{k=1}^{N}\sum_{j=\lambda_{min}}^{\lambda_{max}}\sum_{i=1}^{M}\sum_{l=1}^{L}\log Pr_{o_i}(I_k=i|z_k^l,T^0_{1:M})Pr_{\Theta'}(I_k=i|z_k^l,\lambda=j,C_{obs})Pr_{\Theta'}(\lambda=j|C_{obs}) + C_2. \tag{22}$$

$$\theta_i = \frac{\sum_{k=1}^{N}\sum_{j=\lambda_{min}}^{\lambda_{max}}\sum_{l=1}^{L}Pr_{\Theta'}(I_k=i|z_k^l,\lambda=j,C_{obs})Pr_{\Theta'}(\lambda=j|C_{obs})x_{k-j}^{\mathrm{T}}y_k}{\sum_{k=1}^{N}\sum_{j=\lambda_{min}}^{\lambda_{max}}\sum_{l=1}^{L}Pr_{\Theta'}(I_k=i|z_k^l,\lambda=j,C_{obs})Pr_{\Theta'}(\lambda=j|C_{obs})x_{k-j}^{\mathrm{T}}x_{k-j}}. \tag{23}$$

$$\sigma^2 = \frac{\sum_{k=1}^{N}\sum_{j=\lambda_{min}}^{\lambda_{max}}\sum_{i=1}^{M}\sum_{l=1}^{L}Pr_{\Theta'}(I_k=i|z_k^l,\lambda=j,C_{obs})Pr_{\Theta'}(\lambda=j|C_{obs})(y_k-\theta_i{}^{\mathrm{T}}x_{k-j})^{\mathrm{T}}(y_k-\theta_i{}^{\mathrm{T}}x_{k-j})}{\sum_{k=1}^{N}\sum_{j=\lambda_{min}}^{\lambda_{max}}\sum_{i=1}^{M}\sum_{l=1}^{L}Pr_{\Theta'}(I_k=i|z_k^l,\lambda=j,C_{obs})Pr_{\Theta'}(\lambda=j|C_{obs})}. \tag{24}$$

$$\max_{o_i,i=1,2,\cdots,M}\sum_{k=1}^{N}\sum_{j=\lambda_{min}}^{\lambda_{max}}\sum_{i=1}^{M}\sum_{l=1}^{L}\log Pr_{o_i}(I_k=i|z_k^l,T^0_{1:M})Pr_{\Theta'}(I_k=i|z_k^l,\lambda=j,C_{obs}). \tag{25}$$
$$S.t.\ \ o_{i,min}\le o_i, i=1,2,\dots,M\le o_{i,max}$$

The unknown time delay can be estimated as

$$\hat{\lambda} = \arg\max_j Pr_{\Theta'}(\lambda=j|C_{obs}),\ j=\lambda_{min},\dots,\lambda_{max}. \tag{26}$$

3.2 Summary of the proposed identification algorithm

The proposed multiple model approach for nonlinear system identification with a single uncertain scheduling variable and unknown measurement time delay using the EM algorithm is as follows:

(1) *Initialization*: Set the initial value to Θ'.
(2) *E-step*: According to Eq. (22), the approximate Q-function can be evaluated using current parameter Θ'.
(3) *M-step*: Maximize Eq. (22) to re-estimate the parameter, and then set $\Theta' = \Theta$.
(4) *Iterate*: Evaluate the relative change shown in Eq. (15), and according to the pre-determined tolerance, decide whether to repeat step 2 and 3 or terminate.

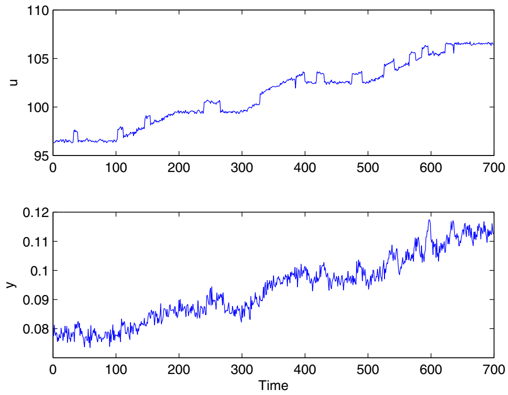

Fig. 1. The input-output data of CSTR

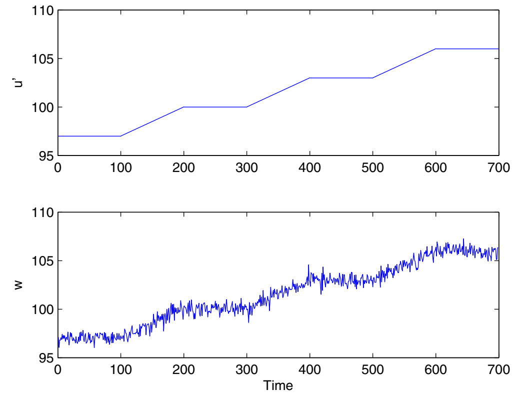

Fig. 2. The input-output data of the scheduling variable

4. SIMULATION EXAMPLE

Continuous stirred tank reactor (CSTR) process is an exothermic; irreversible reaction. The mathematical models are given as follows [Zhao et al., 2012, Chen et al., 2013]:

$$\frac{dC_A(t)}{dt} = \frac{q(t)}{V}(C_{A0} - C_A(t)) - k_0 C_A(t) \exp(\frac{-E}{RT(t)}), \quad (27a)$$

$$\frac{dT(t)}{dt} = \frac{q(t)}{V}(T_0(t) - T(t)) - \frac{(-\Delta H) k_0 C_A(t)}{\rho C_p} \exp(\frac{-E}{RT(t)}) + \frac{\rho_c C_{pc}}{\rho C_p V} q_c(t) \{1 - \exp(\frac{-hA}{q_c(t) \rho C_p})\} (T_{c0}(t) - T(t)). \quad (27b)$$

The model parameters are shown in Table 1. C_A is the product concentration as the controlled variable, and $q_c(t)$ is the coolant flow rate as the manipulated variable. $q_c(t)$ in the range 96-110 is also the scheduling variable. 97, 100, 103 and 106 are four pre-determined operating points.

Table 1. Parameters of the CSTR process

q	100	L/min
C_{A0}	1	mol/L
T_0	350	K
T_{c0}	350	K
V	100	L
hA	7×10^5	cal/(min K)
K_0	7.2×10^{10}	min^{-1}
E/R	1×10^4	K
$-\Delta H$	-2×10^5	cal/mol
$\rho,\ \rho_c$	1×10^3	g/L
$C_p,\ C_{pc}$	1	cal/(gK)

The dynamic model of the scheduling variable is assumed as

$$z_k = A z_{k-1} + B u'_{k-1} + \gamma_{k-1}, \quad (28a)$$

$$w_k = C z_k + v_k, \quad (28b)$$

where $A = 0.2$, $B = 0.8$, $C = 1$, $R_\gamma = 0.01$ and $R_v = 0.2$. White noise is added to the simulated process output, with variance of about 2% of that of the noise free output. The unknown true time delay is 2, which is bounded from 0 to 5. The simulated data are shown in Figures 1 and 2.

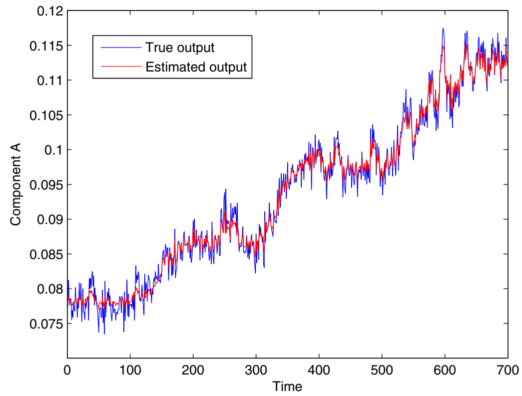

Fig. 3. Comparison of the identified CSTR model (self-validation)

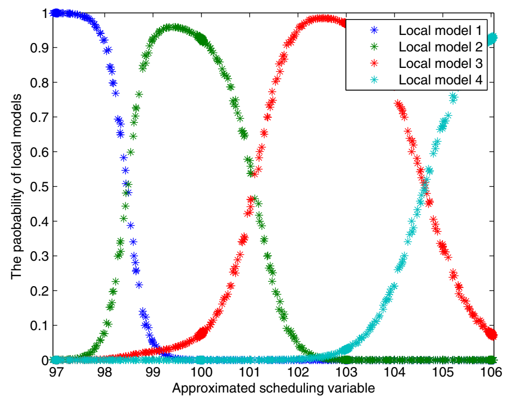

Fig. 4. The probability of each local model (self-validation)

Applying the proposed method, four local models are identified, where the tolerance is 1e-5. The self-validation results are shown in Figures 3 and 4. From Figure 5, we can see that the estimation of the unknown time delay has fast convergence. To verify the identified models, cross validation data are used with operating points at 99 and 106, and the graphic comparisons are shown in Figures 6 and 7.

It can be observed that the predictions obtained by the identified nonlinear process models are in a close agreement

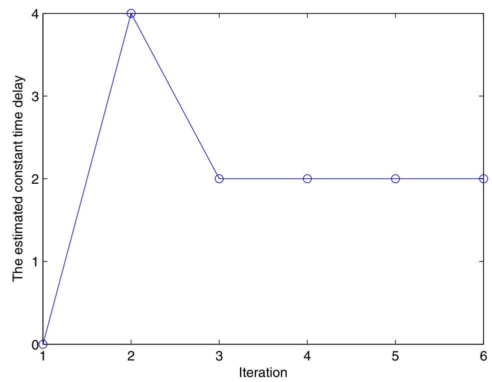

Fig. 5. The estimated time delay

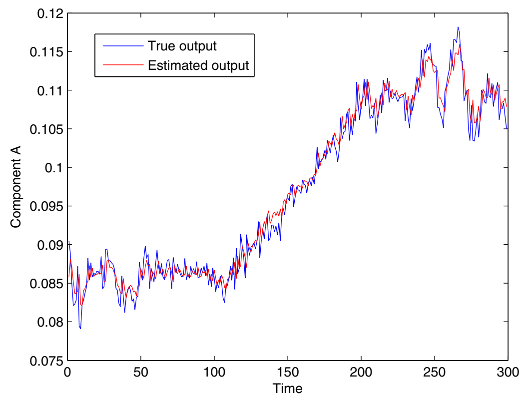

Fig. 6. Comparison of the identified CSTR model (cross-validation)

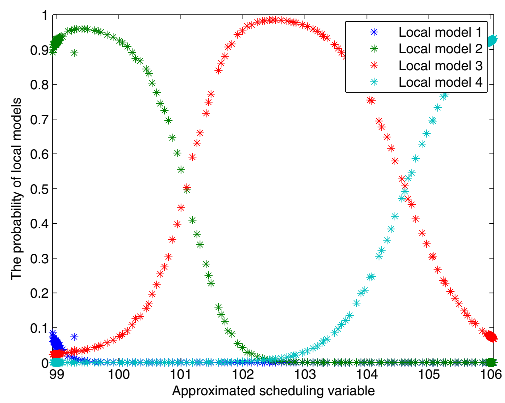

Fig. 7. The probability of each local model (cross-validation)

with the measured outputs, and the unknown time delay can be estimated correctly. These results demonstrate the effectiveness of the proposed approach.

5. CONCLUSIONS

This article considers an identification approach for nonlinear time delay systems with a noisy scheduling variable. The proposed approach follows the framework of the EM algorithm using the multiple model approach. ARX models as the local models are identified at each process operating point; the complete dynamics of the nonlinear system is a combination of all the ARX models associated with normalized exponential function as its probability density function. The parameters of the local models and the exponential functions, as well as the unknown time delay are estimated simultaneously. The validation results of a CSTR example demonstrate that the proposed method can give satisfactory performance for identifying nonlinear time delay systems with an uncertain scheduling variable.

REFERENCES

Björklund, S., Ljung, L. (2003). *A review of time-delay estimation techniques.* In Proceedings of the 42nd IEEE Conference on Decision and Control. Maui, Hawaii USA. 3, 2502-2507.

Richard, J.P. (2003). *Time-delay systems: an overview of some recent advances and open problems.* Automatica, 39(10), 1667-1694.

Yazdizadeh, A., Khorasani, K. (2002). *Adaptive time delay neural network structures for nonlinear system identification.* Neurocumputing, 47(1-4), 207-240.

Liu, H., Lu, J.A., Lü, J., Hill, D. J. (2009). *Structure identification of uncertain general complex dynamical networks with time delay.* Automatica, 45(8), 1799-1807.

Cao, Y., Frank, P. (2000). *Analysis and synthesis of nonlinear time-delay systems via fuzzy control approach.* IEEE Transactions on Fuzzy Systems, 8(2), 200C211.

Anguelova, M., Wennberg, B. (2008). *State elimination and identifiability of the delay parameter for nonlinear time-delay systems.* Automatica, 44(5), 1373-1378.

Chen, L., Tulsyan, A., Huang, B., Liu, F. (2013). *Multiple model approach to nonlinear system identification with uncertain scheduling variables using EM algorithm.* Journal of Process Control, 23(10), 1480-1496.

Ljung, L. (1987). System Identification: Theory for the user, Prentice Hall, New Jersey.

Jin, X., Huang, B., Shook, D. S. (2011). *Multiple model LPV approach to nonlinear process identification with EM algorithm.* Journal of Process Control, 21(1), 182-193.

Zhao, Y., Huang, B., Su, H., Chu, J. (2012). *Prediction error method for identification of LPV models.* Journal of Process Control, 22(1), 180-193.

Dempster, A. P., Laird, N. M., Rubin, D. B. (1977). *Maximum likelihood from incomplete data via the EM algorithm.* Journal of the Royal Statistical Society, Series B, 39(1), 1-38.

McLachlan, G. J., Krishnan, T. (2008). *The EM algorithm and extensions, 2E.* Hoboken, NJ, USA: John Wiley & Sons, Inc.

Wu, J. (1983). *On the convergence properties of the EM algorithm.* The Annals of Statistics. 11(1), 95-103.