ICIC Express Letters ICIC International ©2017 ISSN 1881-803X
Volume **11**, Number **9**, September **2017** pp. 1461–1467

MULTI-MODEL APPROACH FOR NONLINEAR SYSTEM IDENTIFICATION BY EM ALGORITHM

Jinjin Wei[1], Yanyan Yin[1,2] and Fei Liu[1]

[1]Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education)
Institute of Automation
Jiangnan University
No. 1800, Lihu Ave., Wuxi 214122, P. R. China
yinyanyan_2016@126.com

[2]Department of Mathematics and Statistics
Curtin University
Perth 6102, Western Australia

Received March 2017; accepted June 2017

Abstract. *The problem of system identification for nonlinear system is studied in this paper by using EM algorithm, and a stochastic scheduling parameter which follows a Markov jump process is considered. First, multi-model approach is addressed to describe the nonlinear process, where each linear parameter system is represented by an auto regressive exogenous model, and then, EM algorithm is used to do estimation with the help of stochastic scheduling parameter. A simulation example is given to illustrate the effectiveness of the approach proposed.*
Keywords: System identification, Multi-model approach, Stochastic scheduling variable, EM algorithm

1. **Introduction.** Multi model approach proposed in 1969 is reasonable to be applied to describing nonlinear systems [1], and in recent years, it has been widely applied to control problem of nonlinear systems, identification and parameters estimation of nonlinear systems [2-6]. It is worth mentioning that as a typical multi-model system, linear parameter varying (LPV) model [7-9] has great potential to model time-varying nonlinear systems, it has been applied to approximating complex time varying systems successfully and much work has been done in this area [9-14]. As well known, EM algorithm has attracted much attention for the research work of estimation [15], and in [16], robust EM-type algorithms are studied, and in [17], estimation of linear composite quantile regression is investigated, and many results have been obtained based on EM algorithm [18-21].

On another research front line, in real world, many chemical systems are nonlinear and complex, such as petrochemical production processes, the basic and necessary work for the system is estimation, and in this system, the operating trajectory can be described by selecting some working points, which is known as scheduling parameter. With the help of scheduling parameter, the system will be estimated using EM algorithm. Some attempt work has been done for system estimation by using known scheduling parameters or parameters with missing data [8,23]. However, all these works are done based on time-invariant scheduling parameters, it is avoidable that there are some sudden changes in practice, or abrupt variations, and then, the system structure will be randomly changed, which leads to time-varying scheduling parameters [24,25]. We take petrochemical production process as an example, it is a high temperature production process, in such system, coolant is added to control temperature of the reactor, and different temperature will lead to different products, and then, the coolant flow can be used as a scheduling variable. For the modeling of scheduling variable, due to different environment conditions, we will have different system structures for this modelling, and this motivates us to use Markov jumping process to describe such changes, and it is much realistic to use stochastic scheduling parameters in system estimation. This motivates us to do this work and contribution of this paper is: random scheduling parameters are addressed here, and petrochemical production process system is estimated under some abrupt changes conditions.

In this paper, we will do estimation on continuous stirred-tank reactor (CSTR) system using EM algorithm; the rest of the paper is organized as follows: Section 2 describes problem statement, in Section 3, the estimation procedure is given and a simulation example is also given to show the effectiveness of our approach in Section 4, and finally, some concluding remarks are given in Section 5.

2. **Problem Statement.** A nonlinear dynamic system is considered in this paper to describe a nonlinear industrial process

$$f\left(\dot{c}_k, c_k, y_k, u_k, z(r_k), u'_k, T_{1:M}, k, \varepsilon_k\right) = 0 \tag{1}$$

where $c_k \in C \subseteq R^r$ and $y_k \in Y \subseteq R^l$ are system states and measured values at time k, $u_k \in U \subseteq R^m$ is input of the nonlinear process, ε_k is system noise, $f(\cdot)$ is a nonlinear function, $T_{1:M} = \{T_1, T_2, \ldots, T_M\}$ represents different operating points, $\{r_k\}$ is a continuous-time discrete-valued Markov chain, which takes values in a finite state space $\Lambda = \{1, 2, \ldots, f\}$, $\mathrm{z}(r_k)$ is a parameter with known initial value z_0, it is a stochastic parameter which represents the dynamics of system (1), and it is a scheduling variable of system (1) which is described by system (2):

$$\mathrm{z}\left(r_k\right) = A\left(r_k\right)\mathrm{z}\left(r_{k-1}\right) + B\left(r_k\right)u'_{k-1} \tag{2}$$

where $u'_{k-1} \in U' \subseteq R^s$ is the input vector of system (2), and $A(r_k)$, $B(r_k)$ are known time-varying system matrices with appropriate dimensions corresponding to the model at time k. The transition probability from mode i at time k to mode j at time $k+1$ is defined as $\pi_{ij}(k) = P(r_{k+1} = j | r_k = i)$, $i, j \in \Lambda$, and it satisfies $\pi_{ij}(k) \geq 0$ and $\sum_{i=1}^{f} \pi_{ij}(k) = 1$.

It is worth mentioning that in this paper, u_k, y_k and $T_{1:M}$ are known in system (1), while in system (2), $\left\{u'_{k-1}, A\left(r_k\right), B\left(r_k\right), \pi_{ij}\left(k\right)\right\}$ are also given as a priori as well, then, the observed data set of the process (1) can be represented as $C_{obs} = \{y_{1:N}, u_{1:N}, u'_{1:N}, \mathrm{z}\left(r_{1:N}\right)\}$, which will be used later.

3. **Identification of Nonlinear System by EM Algorithm.** The main task of this paper is to do identification of system (1), with the help of multiple model approach to express the nonlinear process, we choose ARX models to describe local models of system (1) which is given below:

$$y_k = \theta_{I_k}^T x_k + e_k \tag{3}$$

where $x_k \in R^n$ represents the regressor of the system, which is expressed as

$$x_k \triangleq \left[y_{k-1}, y_{k-2}, \ldots, y_{k-n_a}, u_{k-1}^T, u_{k-2}^T, \ldots, u_{k-n_b}^T\right]^T \tag{4}$$

where $y_k \in R^1$ and $u_k \in R^m$ are output and input of system (1) respectively, n_a and n_b are the orders of the output and input, and $n = n_a + mn_b$, I_k is introduced to represent the identity of the local model at sampling time k, e_k is a zero mean Gaussian noise with unknown variance σ^2, and T_i $(i = 1, \ldots, M)$ denotes the ith operating point of nonlinear system (1).

In the neighbourhood of small region for each operating point, a linear model is applied to approximating the process dynamics. Given all the past information, probability of the observed process output y_k is calculated as:

$$p\left(y_k | y_{1:k-1}, u_{1:\mathrm{k}-1}, \mathrm{z}\left(r_{1:k}\right), u'_{1:k-1}\right) = \sum_{i=1}^{M} \alpha_{k,i} p\left(y_k | y_{1:k-1}, u_{1:\mathrm{k}-1}, \theta_i\right) \tag{5}$$

where θ_i represents the parameter of the ith local linear model, $\alpha_{k,i} = p\big(\theta_i | y_{1:k-1}, u_{1:\mathrm{k}-1}, \mathrm{z}\,(r_{1:k})\,, u'_{1:k-1}\big)$ in (5) is a normalized exponential function and it represents the probability of the ith local model at sampling time k, which is written as below:

$$\alpha_{k,i} = \frac{\exp\left(-\dfrac{\left(\mathrm{z}\,(r_k) - T_i\right)^2}{2(o_i)^2}\right)}{\sum\limits_{i=1}^{M} \exp\left(-\dfrac{\left(\mathrm{z}\,(r_k) - T_i\right)^2}{2(o_i)^2}\right)} \tag{6}$$

where $o_i \subseteq R$ is the validity width of the ith local model, which is unknown and bounded, and let $o_{i,\min} \leq o_i \leq o_{i,\max}$, where $o_{i,\min}$ and $o_{i,\max}$ are the lower and upper bounders for o_i. The missing data set is denoted as $C_{mis} = \{I_{1:N}\}$, and then the complete data is written as $\{C_{obs}, C_{mis}\}$ where $I_{1:N}$ is a hidden variable. The parameters which are necessary to be estimated are $\Theta = \{\theta_{1:M}, \sigma, o_{1:M}\}$. From [13], the steps of the EM algorithm are as follows.

1) Initialization: Given Θ^{old} as initial values.

2) E-step: Calculate the Q-function by using the current parameter Θ^{old} as below

$$\mathrm{Q}\left(\Theta|\Theta^{old}\right) = E_{C_{mis}|\left(C_{obs},\Theta^{old}\right)\{\log p(C_{mis},C_{obs}|\Theta)\}} \tag{7}$$

3) M-step: Maximize the Q-function,

$$\Theta = \underset{\Theta}{\operatorname{argmax}}\, \mathrm{Q}\left(\Theta|\Theta^{old}\right) \tag{8}$$

and then, set $\Theta^{old} = \Theta$.

4) Iterate: Repeat steps 2) and 3) until it is convergent. By applying the EM algorithm, and using the probability chain rule, the complete likelihood function $p\,(C_{mis}, C_{obs}|\Theta)$ of (7) is decomposed as

$$\begin{aligned} \log p\,(C_{mis}, C_{obs}|\Theta) &= \log p\left(y_{1:N}, u_{1:N}, \mathrm{z}\,(r_{1:N})\,, u'_{1:N}, I_{1:N}|\Theta\right) \\ = \log &\left(\prod_{k=1}^{N} p\left(y_k|x_k, \mathrm{z}\,(r_{1:k})\,, u'_{1:k}, I_{1:k}, \Theta\right) \cdot p\,(I_k|\mathrm{z}\,(r_{1:k})\,, I_{1:k-1}, \Theta)\right. \\ &\left.\times p\left(u_{1:N}, \mathrm{z}\,(r_{1:N})\,, u'_{1:N}|\Theta\right)\right) \\ = &\sum_{k=1}^{N} \left(\log p\,(y_k|x_k, \Theta_{I_k}) + \log p\,(I_k|\mathrm{z}\,(r_k)\,, \Theta_{I_k}) + \log C\right) \end{aligned} \tag{9}$$

and

$$\begin{aligned} &\mathrm{Q}\left(\Theta|\Theta^{old}\right) = E_{C_{mis}|\left(C_{obs},\Theta^{old}\right)\{\log p(C_{mis},C_{obs}|\Theta)\}} \\ = &\int p\left(I_{1:N}|\Theta^{old}, C_{obs}\right) \sum_{k=1}^{N} \left(\log p\,(y_k|x_k, \Theta_{I_k}) + \log p\,(I_k|\mathrm{z}\,(r_k)\,, \Theta_{I_k}) + \log C\right) d_{I_{1:N}} \\ = &\sum_{k=1}^{N}\sum_{i=1}^{M} p\left(I_k = i|\Theta^{old}, C_{obs}\right) \log p\,(I_k = i|\mathrm{z}\,(r_k)\,, o_i) \\ &+ \sum_{k=1}^{N}\sum_{i=1}^{M} p\left(I_k = i|\Theta^{old}, C_{obs}\right) \log p\,(y_k|x_k, \theta_i, \sigma) \\ &+ \sum_{k=1}^{N}\sum_{i=1}^{M} p\left(I_k = i|\Theta^{old}, C_{obs}\right) \times \log C \end{aligned} \tag{10}$$

where $u_{1:N}$, $\mathrm{z}(r_{1:N})$ and $u'_{1:N}$ are independent of Θ, and $C = P\left(u_{1:N}, \mathrm{z}(r_{1:N}), u'_{1:N}|\Theta\right)$ is a constant. To compute the Q-function in (10), the following probability functions will be calculated: 1) $p\left(y_k|y_{1:k-1}, u_{1:k-1}, \theta_i, \sigma\right)$; 2) $p\left(I_k = i|\Theta^{old}, C_{obs}\right)$.

Considering Gaussian noise in ARX model, function $p\left(y_k|y_{1:k-1}, u_{1:k-1}, \theta_i, \sigma\right)$ is expressed as:

$$p\left(y_k|y_{1:k-1}, u_{1:k-1}, \theta_i, \sigma\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{-1}{2\sigma^2}\left(y_k - \theta_i^T x_k\right)^T \left(y_k - \theta_i^T x_k\right) \tag{11}$$

where $p\left(I_k = i|\Theta^{old}, C_{obs}\right)$ is the probability of the $I_k = i$th local model at sampling time k, under known conditions of C_{obs} and the current parameter Θ^{old}, by using Bayes' rule, $p\left(I_k = i|\Theta^{old}, C_{obs}\right)$ is expressed as:

$$p\left(I_k = i|\Theta^{old}, C_{obs}\right) = \frac{p\left(y_k|y_{1:k-1}, u_{1:k-1}, \theta_i^{old}, \sigma^{old}\right) p\left(I_k = i|\mathrm{z}(r_k), o_i^{old}\right)}{\sum_{i=1}^{M} p\left(y_k|y_{1:k-1}, u_{1:k-1}, \theta_i^{old}, \sigma^{old}\right) p\left(I_k = i|\mathrm{z}(r_k), o_i^{old}\right)} \tag{12}$$

and then,

$$p\left(I_k = i|\mathrm{z}(r_k), o_i\right) = \alpha_{k,i} \tag{13}$$

Combining (11), (12) and (13), then, the Q-function is fixed. Do derivative to $\mathrm{Q}\left(\Theta|\Theta^{old}\right)$, the parameters θ_i, σ and o_i in (10) are calculated, and then, we have

$$\frac{\partial}{\partial\theta_i} \sum_{k=1}^{N} \sum_{i=1}^{M} p\left(I_k = i|\Theta^{old}, C_{obs}\right) \log p\left(y_k|y_{1:k-1}, u_{1:k-1}, \theta_i, \sigma\right) = 0 \tag{14}$$

By solving (14), we have θ_i^{New}, and it holds that

$$\theta_i^{New} = \frac{\sum_{k=1}^{N} p\left(I_k = i|\Theta^{old}, C_{obs}\right) x_k^T y_k}{\sum_{k=1}^{N} \sum_{l=1}^{L} Pr_{\Theta^{cu}}\left(I_k = i|\mathrm{z}(r_k), C_{obs}\right) x_k^T x_k} \tag{15}$$

Similarly, we will also get $\left(\sigma^{New}\right)^2$, where

$$\left(\sigma^{New}\right)^2 = \frac{\sum_{k=1}^{N} \sum_{i=1}^{M} p\left(I_k = i|\Theta^{old}, C_{obs}\right) \left(\left(y_k - \left(\theta_i^{New}\right)^T x_k\right)^T \left(y_k - \left(\theta_i^{New}\right)^T x_k\right)\right)}{\sum_{k=1}^{N} \sum_{i=1}^{M} p\left(I_k = i|\Theta^{old}, C_{obs}\right)} \tag{16}$$

Let $I_k = i$th, for the validity width of the ith local model, the mathematical formulation of the optimization problem in searching for the optimal o_i is given in (17):

$$\begin{aligned} &\max_{o_{i,i=1,2,\ldots,M}} \sum_{k=1}^{N} \sum_{i=1}^{M} p\left(I_k = i|\Theta^{old}, C_{obs}\right) \log p\left(I_k = i|\mathrm{z}(r_k), o\right) \\ &\left(o_{\min} \le o_i, i = 1, 2, \ldots, M \le o_{i,\max}\right) \end{aligned} \tag{17}$$

Remark 3.1. *Similar to [12,14], then, the optimal o_i is searched by a constrained nonlinear optimization function which is "fmincon".*

4. **Numerical Example.** In our paper, we assume A → B is an irreversible and exothermic reaction. The concentration of reagent A and the reactor temperature are outputs in this system, coolant flow rate is input, and it is also the scheduling variable. The principle model and parameters of the process are derived in [12]. In this paper, five operating points are pre-determined as 96, 100, 103, 106 and 109, respectively, and the model of stochastic scheduling variable is given as:

$$\mathrm{z}(r_k) = A(r_k)\,\mathrm{z}(r_{k-1}) + B(r_k)\,u'_{k-1}$$

where $A(r_k) = \begin{bmatrix} 0.196 \\ 0.2 \end{bmatrix}$, $B(r_k) = \begin{bmatrix} 0.801 \\ 0.8 \end{bmatrix}$.

The initial state is given as $z_0 = 97$, and the real transition probability matrix is given as

$$\Pi = \begin{bmatrix} 0.6 & 0.4 \\ 0.6 & 0.4 \end{bmatrix}$$

Here, white noise is given with variance of about 1.5%, and the noise-free output is added to the simulated process output. Without knowing the parameters of each local model a-priori, we apply the proposed algorithm to the input-output data and the self-validation results of CSTR model are obtained as shown in Figure 1, while Figure 2 shows the cross-validation results of CSTR model, and it demonstrates that the estimated curve tracks the real curve effectively. We also use the relative error [26] to measure the model quality,

$$ERR = \frac{\mathrm{var}\,(y - \hat{y})}{\mathrm{var}(y)} * 100\% \tag{18}$$

where y is the true output and $\hat{y}$ is the estimated output. It is observed from Figures 1 and 2 that the relative error of self-validation is 1.14%, the relative error of cross-validation is 4.14%, and the estimation error is small which shows the effectiveness of the methods proposed.

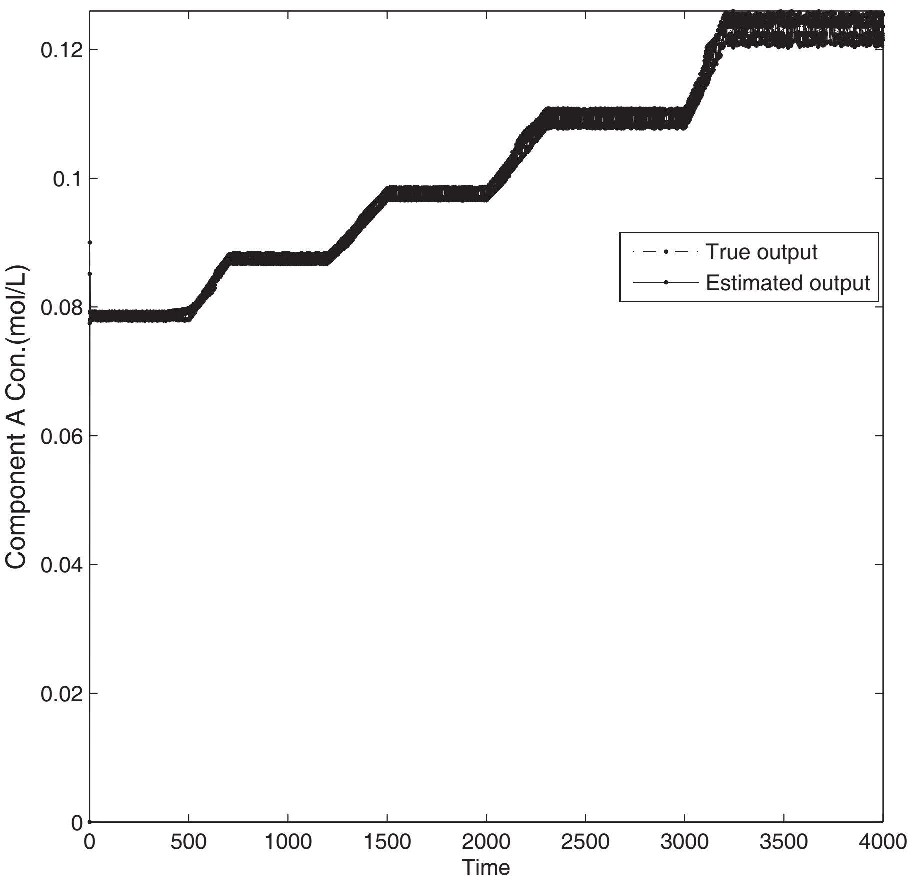

Figure 1. The self-validation results of CSTR model

5. **Conclusions.** The identification of nonlinear industrial system using a stochastic scheduling variable is studied by using EM algorithm and multi-model approach. The LPV model is addressed to express the dynamic of local models. A simulation of CSTR system is shown in this paper, which illustrates that the approach proposed in this paper

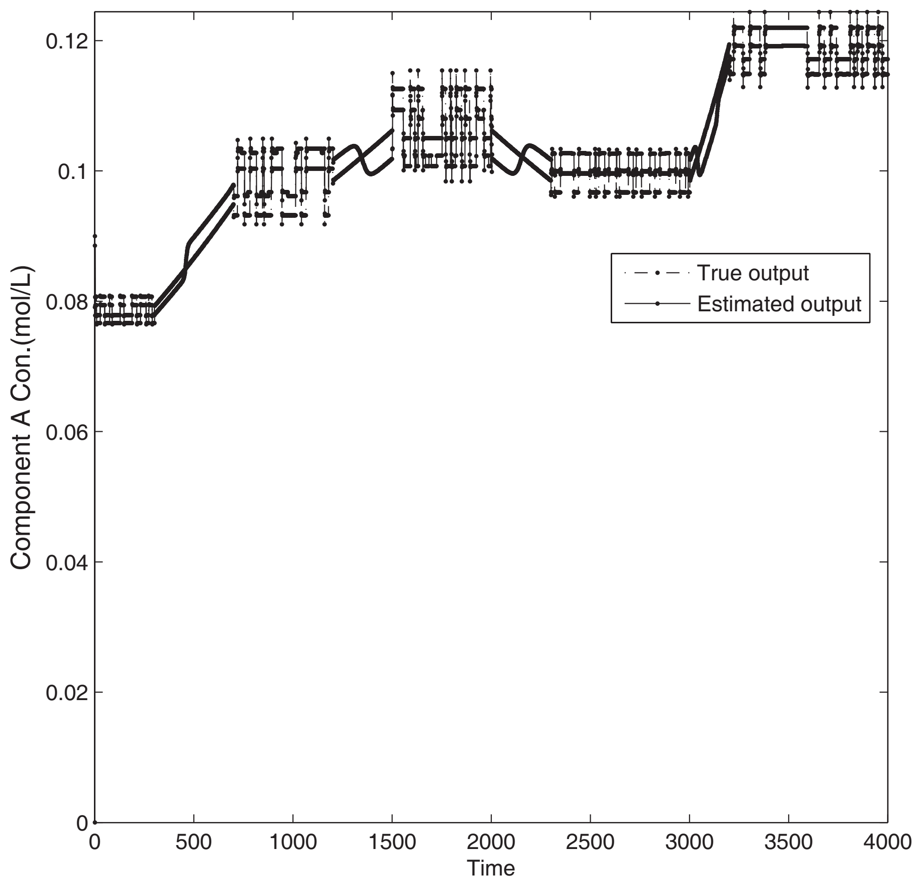

Figure 2. The cross-validation results of CSTR model

is effective. In our future work, control problem will be considered on this system and comparison with other related results will be given.

Acknowledgment. This work was supported in part by the National Natural Science Foundation of China under grant No. 61403169, the Fundamental Research Funds for the Central Universities (NO: JUSRP51407B).

REFERENCES

[1] J. M. Bates and C. W. J. Granger, The combination of forecasts, *Operations Research Quarterly*, vol.20, no.4, pp.451-468, 1969.

[2] N. Nandola and S. Bhartiya, A multiple model approach for predictive control of nonlinear hybrid systems, *Journal of Process Control*, vol.18, no.2, pp.131-148, 2008.

[3] B. V. Patil, S. Bhartiya, P. S. V. Nataraj and N. N. Nandola, Multiple-model based predictive control of nonlinear hybrid systems based on global optimization using the Bernstein polynomial approach, *Journal of Process Control*, vol.22. no.2, pp.423-435, 2012.

[4] M. Pottmann, H. Unbehauen and D. E. Seborg, Application of a general multi-model approach for identification of highly nonlinear processes: A case study, *International Journal of Control*, vol.57, no.1, pp.97-120, 1993.

[5] T. A. Johansen and B. A. Foss, Identification of nonlinear system structure and parameters using regime decomposition, *Automatica*, vol.31, no.2, pp.321-326, 1995.

[6] J. Huang, G. Ji, Y. Zhu and P. V. D. Bosch, Identification of multi-model LPV models with two scheduling variables, *Journal of Process Control*, vol.22, no.7, pp.1198-1208, 2012.

[7] L. Chen, A. Tulsyan, B. Huang and F. Liu, Multiple model approach to nonlinear system identification with an uncertain scheduling variable using EM algorithm, *Journal of Process Control*, vol.3, no.10, pp.1480-1496, 2013.

[8] X. Jin, B. Huang and D. S. Shook, Multiple model LPV approach to nonlinear process identification with EM algorithm, *Journal of Process Control*, vol.21, no.1, pp.182-193, 2011.

[9] B. Bamieh and L. Giarr, Identification of linear parameter varying models, *International Journal of Robust & Nonlinear Control*, vol.12, no.9, pp.841-853, 2002.

[10] A. Banerjee, Y. Arkun, B. Ogunnaike and R. Pearson, Estimation of nonlinear multiple systems using linear models, *AIChE Journal*, vol.43, no.5, pp.1204-1226, 1997.

[11] L. Lee and K. Poolla, Identification of linear parameter-varying systems using nonlinear programming, *Proc. of the 35th IEEE Conference on Decision and Control*, pp.1545-1550, 1996.

[12] R. Murray-Smith and T. A. Johansen, *Multiple Model Approaches to Modelling and Control*, Taylor & Francis, 1997.

[13] Z. Xu, J. Zhao, J. Qian and Y. Zhu, Nonlinear MPC using identified LPV model, *Industrial & Engineering Chemistry Research*, vol.6, no.48, pp.3043-3051, 2009.

[14] V. Laurain, M. Cilson, R. Toth and H. Garnier, Refined instrumental variable methods for identification of LPV Box-Jenkins models, *Automatica*, vol.46, no.6, pp.959-967, 2010.

[15] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, 2nd Edition, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2008.

[16] H. Hu, W. Yao and Y. Wu, The robust EM-type algorithms for log-concave mixtures of regression models, *Computational Statistics & Data Analysis*, vol.111, no.7, pp.14-26, 2017.

[17] W. Huang and Y. Chen, The multiset EM algorithm, *Statistics & Probability Letters*, 2017.

[18] Y. Tian, Q. Zhu and M. Tian, Estimation of linear composite quantile regression using EM algorithm, *Statistics & Probability Letters*, vol.117, no.10, pp.183-191, 2016.

[19] W. Ding and P. Song, EM algorithm in Gaussian copula with missing data, *Computational Statistics & Data Analysis*, vol.101, no.9, pp.1-11, 2016.

[20] R. Ranjan, B. Huang and A. Fatehi, Robust Gaussian process modeling using EM algorithm, *Journal of Process Control*, vol.42, no.9, pp.125-136, 2016.

[21] P. Latouche, P. A. Mattei, C. Bouveyron and J. Chiquet, Combining a relaxed EM algorithm with Occam's razor for Bayesian variable selection in high-dimensional regression, *Journal of Multivariate Analysis*, vol.146, no.4, pp.177-190, 2016.

[22] Y. Zhu and Z. Xu, A method of LPV model identification for control, *The 17th IFAC World Congress*, Seoul, Korea, pp.5018-5023, 2008.

[23] L. Chen, S. Khatibisepehr, B. Huang, F. Liu and Y. Ding, Nonlinear process identification in the presence of multiple correlated hidden scheduling variables with missing data, *AIChE Journal*, vol.61, no.10, pp.3270-3287, 2015.

[24] Y. Yin, Y. Liu, K. L. Teo and S. Wang, Event-triggered probabilistic robust control of linear systems with input constrains: By scenario optimization approach, *International Journal of Robust and Nonlinear Control*, DOI:10.1002/rnc.3858, 2017.

[25] Y. Yin and Z. Lin, Constrained control of uncertain nonhomogeneous Markovian jump systems, *International Journal of Robust and Nonlinear Control*, 2017.

[26] Y. Zhu, H. Telkamp, J. Wang and Q. Fu, System identification using slow and irregular output samples, *Journal of Process Control*, vol.19, no.1, pp.58-67, 2009.