A Data-driven Subspace Identification Algorithm for Industrial 4-stage Evaporator

Xueshan Lin, Xiaosuo Luo

***Abstract*—**For the problem that it is difficult in building the accurate mechanism models of the industrial 4-stage evaporator, A new data-driven subspace identification algorithm is proposed. Firstly, a basic procedure of subspace identification method is introduced. Then, the state-space model, obtained from the PO-MOESP (Past Output MOESP, MOESP is one form of the subspace identification methods) algorithm, is regarded as the system model. Lastly, it is applied to the process simulation on the industrial 4-stage evaporator. Through comparisons of performance with a traditional subspace identification algorithm, the superiority of the proposed algorithm is illustrated.

***Keywords*—**Data-driven method, Subspace identification, State-space model, Industrial 4-stage evaporator

I. Introduction

With the development of industrial technology, the industrial processes become more complex than before and it's more difficult in building the accurate mechanism models of these processes. But large amounts of data are produced and stored in these processes every day. Hence, the data-driven approach has been obtained widespread attention since it emerged [1]. And there are many industrial control methods to control the processes. But one drawback of the traditional industrial control methods is based on input-output model, including parametric and nonparametric ones. In order to improve the control performance, a state-space model should be adopted, so the modern filter theory and the design method of controller developed in recent years can play a role [2]. Subspace identification is one of system identification algorithms for state-space modeling [3-5]. The control workers may relieve completely from the tedious mechanism modeling and the accurate state-space model can be obtained when there is enough process input-output data which it's an excellent data-driven method [6]. The comparison of subspace and classical identification methods can be seen in Fig. 1 [7-9].

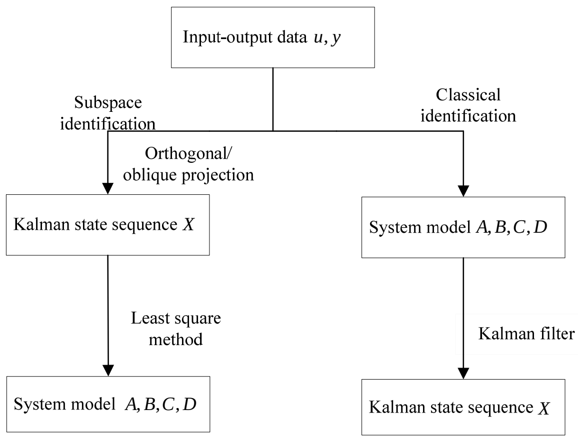

Fig. 1 The frame of subspace and classical identification methods

The PO-MOESP subspace identification algorithm is used to solve the problem when one wants to jointly model the deterministic and stochastic part of the system. Therefore it is important in control and prediction [10-11].

The evaporator is a nonlinear industrial process control system, and considering the complexity of system, it's difficult in building the accurate mechanism models [12]. The conventional model-free control methods, such as PID control, will result in poor control performance. So we should adopt the model-based control method and the premise is to obtain the system accurate model. The PO-MOESP subspace identification algorithm can achieve this goal.

The main contribution of this paper is that we propose a subspace-based data-driven identification method for the industrial 4-stage evaporator and it employs the PO-MOESP algorithm which is more suitable for industrial processes. Firstly, the basic procedure of subspace identification is given. Then, through PO-MOESP algorithm, the state-space model can be obtained from the data and it's used in the industrial 4-stage Evaporator.

The paper is arranged as follows. We start in Section 2 with the basic procedure of subspace identification. In Section 3, we give the PO-MOESP subspace identification algorithm. In Section 4 the simulation example is presented that show the potential of the proposed method. The conclusion is drawn in Section 5.

This work was supported by the Major State Basic Research Development Program 973 (No.2012CB215202), the National Natural Science Foundation of China (No.61134001, 61773081), Key Laboratory of Dependable Service Computing in Cyber Physical Society (Chongqing University), Ministry of Education, Chongqing Research Program of Basic Research and Frontier Exploration (No.cstc2018jcyjAX0082), Young project of science and technology research program of Chongqing Education Commission of China (No.KJQN201803112) and Research Center of Intelligent Building Engineering Technology (No.ZBPT201901).

Xueshan Lin is with the Chongqing College of Electronic Engineering, Chongqing 401331, China (e-mail: 826561356@qq.com).

Xiaosuo Luo is with the School of Automation, Chongqing University, Chongqing 400044, China and Chongqing College of Electronic Engineering, Chongqing 401331, China (e-mail: luoxiaosuo@cqu.edu.cn).

II. The basic procedure of subspace identification

Consider a typical industrial state-space system of order n in stochastic form:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + w(k) \\ y(k) &= Cx(k) + Du(k) + v(k) \end{aligned} \tag{1}$$

where $u(k) \in \mathbb{R}^m$ is input, $y(k) \in \mathbb{R}^l$ is output, $x(k) \in \mathbb{R}^n$ is state, $w(k) \in \mathbb{R}^n$ is process noise and $v(k) \in \mathbb{R}^l$ is measured noise. (A,B,C,D) are system state-space matrices. The stochastic model identification problem is stated with Fig. 2.

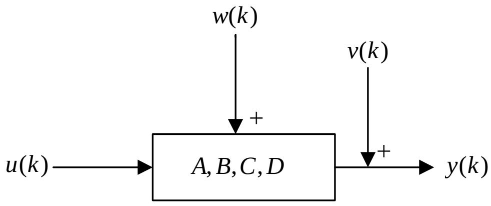

Fig. 2. The scheme of the stochastic model

Assuming that k is the current time, f is the length of the future time, and (1) can be transformed by iteration:

$$y_f = \Gamma_f x(k) + H_f u_f + G_f w_f + v_f \tag{2}$$

where system vectors y_f, u_f, w_f, v_f, extended observability matrix Γ_f and Toeplitz matrices H_f, G_f are:

$$y_f = \begin{bmatrix} y(k) \\ y(k+1) \\ \vdots \\ y(k+f-1) \end{bmatrix},\ \Gamma_f = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{f-1} \end{bmatrix},$$

$$H_f = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ CA^{f-2}B & \cdots & CB & D \end{bmatrix},\ G_f = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ C & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ CA^{f-2} & \cdots & CA & 0 \end{bmatrix},$$

$$u_f = \begin{bmatrix} u(k) \\ u(k+1) \\ \vdots \\ u(k+f-1) \end{bmatrix},\ w_f = \begin{bmatrix} w(k) \\ w(k+1) \\ \vdots \\ w(k+f-1) \end{bmatrix},$$

$$v_f = \begin{bmatrix} v(k) \\ v(k+1) \\ \vdots \\ v(k+f-1) \end{bmatrix}.$$

The (2) can be written in the form of Hankel matrices:

$$Y_f = \Gamma_f X_f + H_f U_f + G_f W_f + V_f \tag{3}$$

and the form of Hankel matrices at the past time p can be presented as

$$Y_p = \Gamma_p X_p + H_p U_p + G_p W_p + V_p \tag{4}$$

The input Hankel matrix is expressed as

$$U_f = \begin{bmatrix} u(k) & u(k+1) & \cdots & u(N-f+1) \\ u(k+1) & u(k+2) & \cdots & u(N-f+2) \\ \vdots & \vdots & \ddots & \vdots \\ u(k+f-1) & u(k+f) & \cdots & u(N) \end{bmatrix},$$

$$U_p = \begin{bmatrix} u(k-p) & u(k-p+1) & \cdots & u(N-f-p+1) \\ u(k-p+1) & u(k-p+2) & \cdots & u(N-f-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ u(k-1) & u(k) & \cdots & u(N-f) \end{bmatrix}$$

.

where N is the sampling time. Similarly, Y_f, Y_p, W_f, V_f, W_p, and V_p have similar definitions.

The state vector matrices are

$$X_f = \begin{bmatrix} x(k) & x(k+1) & \cdots & x(N-f+1) \end{bmatrix},$$
$$X_p = \begin{bmatrix} x(k-p) & x(k-p+1) & \cdots & x(N-f-p+1) \end{bmatrix}.$$

In order to obtain the system matrices, the subspace identification method is generally composed of two steps: (1) Determine the extended observability matrix Γ_f or estimate the system's state sequence $\widehat{X}_f$; (2) Compute system matrices.

III. The PO-MOESP subspace identification algorithm

The stochastic model identification problem is of interest when one wants to jointly model the deterministic and stochastic part of the system. With the PO-MOESP subspace

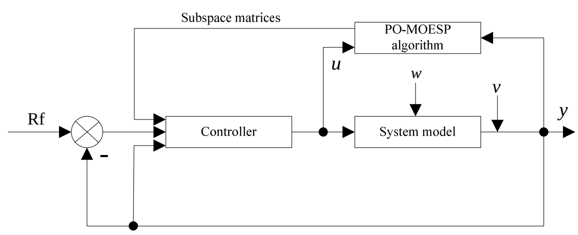

Fig. 3 The structure of the controller design based on PO-MOESP algorithm

identification algorithm, a controller can be implemented as shown in Fig. 3.

According to the Section 2, firstly, we need to remove the $H_f U_f$, $G_f W_f$ and V_f in (3). The instrumental variable in the PO-MOESP method is chosen as a combination of past inputs and outputs:

$$W_p = \begin{bmatrix} U_p \\ Y_p \end{bmatrix} \tag{5}$$

Multiplying Equation (2) from the right by projection operators $\Pi_{U_f^\perp}$ and Π_{W_p} to remove $H_f U_f$, $G_f W_f$ and V_f terms:

$$\lim_{N\to\infty} \frac{1}{N} Y_f \Pi_{U_f^\perp} \Pi_{W_p} = \lim_{N\to\infty} \frac{1}{N} \Gamma_f X_f \Pi_{U_f^\perp} \Pi_{W_p} \tag{6}$$

It can be solved in the usual way using the RQ decomposition:

$$\begin{bmatrix} U_f \\ W_p \\ Y_f \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & 0 & 0 \\ R_{21} & R_{22} & 0 & 0 \\ R_{31} & R_{32} & R_{33} & 0 \\ R_{41} & R_{42} & R_{43} & R_{44} \end{bmatrix} \begin{bmatrix} Q_1^{\mathrm{T}} \\ Q_2^{\mathrm{T}} \\ Q_3^{\mathrm{T}} \\ Q_4^{\mathrm{T}} \end{bmatrix} \tag{7}$$

It can be deduced that:

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \begin{bmatrix} R_{42} & R_{43} \end{bmatrix} = \lim_{N\to\infty} \frac{1}{\sqrt{N}} \Gamma_f X_f \begin{bmatrix} Q_2 \\ Q_3 \end{bmatrix}^{\mathrm{T}} \tag{8}$$

Through SVD decomposition of $\begin{bmatrix} R_{42} & R_{43} \end{bmatrix}$:

$$\begin{bmatrix} R_{42} & R_{43} \end{bmatrix} = \begin{bmatrix} U_n & U_n^\perp \end{bmatrix} \begin{bmatrix} S_n & 0 \\ 0 & S_n^\perp \end{bmatrix} \begin{bmatrix} V_n^{\mathrm{T}} \\ V_n^{\perp\mathrm{T}} \end{bmatrix} \tag{9}$$

where n is the system model order. U_n can be obtained and we can get:

$$U_{n1} = U_n\left(1:(f-1)l, 1:n\right) \tag{10}$$

$$U_{n2} = U_n\left((1+l): fl, 1:n\right) \tag{11}$$

and system matrices A and C can be derived as

$$A = U_{n1}^{\dagger} U_{n2} \tag{12}$$

$$C = U_{n1}\left(1:l, :\right) \tag{13}$$

Next, carry out the conversion of (1):

$$y(k) = CA^k x(0) + \sum_{\tau=0}^{k-1} CA^{k-1-\tau} Bu(\tau) + Du(k) + v(k) \tag{14}$$

It can be rewritten as the Kronecker product:

$$\begin{aligned} y(k) &= CA^k x(0) + \left[\sum_{\tau=0}^{k-1} u(\tau)^{\mathrm{T}} \otimes CA^{k-1-\tau} \right] \mathrm{vec}(B) \\ &+ \left[u(k)^{\mathrm{T}} \otimes I_l \right] \mathrm{vec}(D) + v(k) \end{aligned} \tag{15}$$

where $\otimes$ expresses Kronecker product. Define the following matrices:

$$Y_{0,N} = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix}, \ \Gamma_N = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{N-1} \end{bmatrix},$$

$$\psi = \begin{bmatrix} 0 \\ u(0)^{\mathrm{T}} \otimes C \\ \vdots \\ \sum_{\tau=0}^{N-2} u(\tau)^{\mathrm{T}} \otimes CA^{N-2-\tau} \end{bmatrix}, \ \upsilon = \begin{bmatrix} u(0)^{\mathrm{T}} \otimes I_l \\ u(1)^{\mathrm{T}} \otimes I_l \\ \vdots \\ u(N-1)^{\mathrm{T}} \otimes I_l \end{bmatrix},$$

$$E = \begin{bmatrix} v(0) \\ v(1) \\ \vdots \\ v(N-1) \end{bmatrix}.$$

and the (15) can be denoted as the matrix form:

$$Y_{0,N} = \begin{bmatrix} \Gamma_N & \psi & \upsilon \end{bmatrix} \begin{bmatrix} x_0 \\ \mathrm{vec}(B) \\ \mathrm{vec}(D) \end{bmatrix} + E \tag{16}$$

Noise $v(k)$ is independent with $u(k)$ and x_0, so

$$\begin{bmatrix} x_0 \\ \mathrm{vec}(B) \\ \mathrm{vec}(D) \end{bmatrix} = \begin{bmatrix} \Gamma_N & \psi & \upsilon \end{bmatrix}^{\dagger} Y_{0,N} \tag{17}$$

The system matrices B and D can be obtained by the least square method of (17).

IV. Simulation Example

A typical industrial 4-stage evaporator system is shown in Fig. 4. There are three inputs and three outputs in the system. The three inputs are input product flow q_i, vapour flow q_v to first evaporator and cooling water flow q_c to condenser respectively. The three outputs are dry matter content TDS of output product, output product flow q_o and output product temperature T respectively.

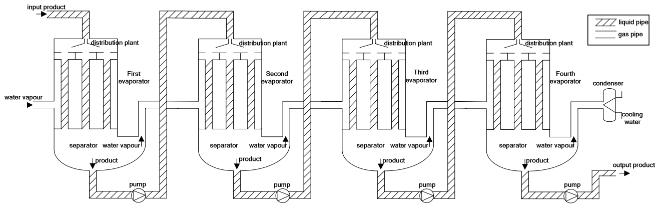

Fig. 4 Diagram of the 4-stage evaporator

In order to validate the algorithm, the data obtained from the evaporator's industrial object identification is used directly, and the first 1000 data are used for model identification and verification. For comparison, the traditional subspace identification algorithm in [13] is conducted. In the two algorithms, a suitably large estimate at the model order was chosen, and the corresponding singular values displayed [14]. By examining the singular values of the model, see Fig. 5, it was shown that most of the system information was stored in the first five singular values, so the model order is chosen as 5. f can be set to three times of model order that $f=15$. The response of the identified model and process output with proposed PO-MOESP is given in Fig. 6 whereas with traditional subspace identification algorithm is shown in Fig. 7. The red solid line is represented as the identified output and the blue dashed line represented as the process output.

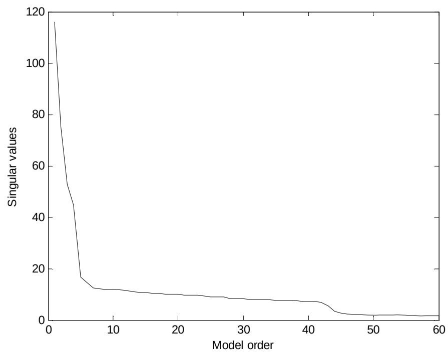

Fig. 5 Singular values of the 4-stage evaporator model

To test the cross validation, a form of prediction error is defined as [15-16]:

$$\text{VAF=max}\{1\text{-}\frac{\text{var}(y_k-\hat{y}_k)}{\text{var}(y_k)},0\}\times 100 \tag{18}$$

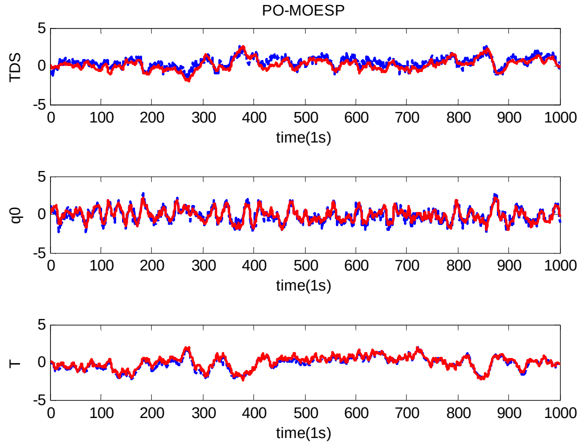

Fig. 6 The response of the identified model and process output with PO-MOESP

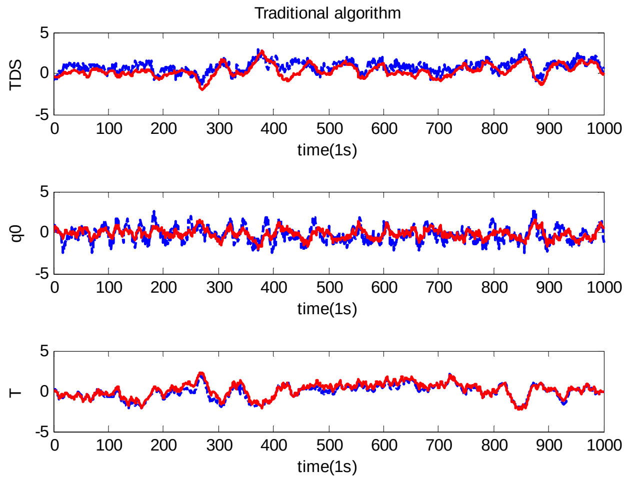

Fig. 7 The response of the identified model and process output with traditional algorithm

where y_k and $\hat{y}_k$ are the values at instant k of process and model output respectively. It is evident that with the increase of VAF, the model is more satisfactory. The prediction error on the validation data set can be seen in Table 1 and the performance using the PO-MOESP is better comparing with the traditional algorithm. This is attributed to the system model incorporating stochastic part can be identified by the PO-MOESP algorithm even better.

Table 1. The prediction error on the validation data set

Outputs	Traditional algorithm VAF	PO-MOESP VAF
TDS	65.5447	80.9608
q_o	45.8329	84.1373
T	73.6743	80.8299

V. Conclusion

In this paper, based on the basic procedure of subspace identification method, the PO-MOESP subspace identification algorithm is addressed. The system matrices are identified through the PO-MOESP subspace identification, then it's applied to the industrial 4-stage evaporator to verify the effectiveness of the proposed algorithm.

REFERENCES

[1] Z. S. Hou and J. X. Xu, "On data-driven control theory: the state of the art and perspective," *ACTA AUTOMATICA SINICA*, vol. 35, no. 6, pp. 650-667, 2009.

[2] P. Chaupa, J. Novak, and P. Januska, "State space MPC Using State Observers," *International Journal of Circuits, Systems and Signal Processing*, vol. 8, pp. 9–14, 2014.

[3] X. S. Luo, "Subspace based model predictive control for linear parameter varying systems," *International Journal of Circuits, Systems and Signal Processing*, vol. 10, pp. 321–325, 2016.

[4] J. W. van Wingerden and M. Verhaegen, "Subspace identification of bilinear and LPV systems for open- and closed-loop data," *Automatica*, vol. 45, no. 2, pp. 372-381, 2009.

[5] M. Bergamasco and M. Lovera, "Continuous-time predictor-based subspace identification using Laguerre filters," *IET Control Theory & Applications*, vol. 5, no. 7, pp. 856-867, 2011.

[6] T. Liu, B. Huang and S. J. Qin, "Bias-eliminated subspace model identification under time-varying deterministic type load disturbance," *Journal of Process Control*, vol. 25, pp. 41-49, 2015.

[7] K. Kameyama and A. Ohsumi, "Subspace-based prediction of linear time-varying stochastic systems," *Automatica*, vol. 43, no. 12, pp. 2009-2021, 2007.

[8] B. Huang, Y. T. Qi and A. K. M. M. Murshed. *Dynamic Modeling and Predictive Control in Solid Oxide Fuel Cells: First Principle and Data-based Approaches*. John Wiley & Sons, New Jersey, 2013.

[9] P. Torres, J. W. van Wingerden and M. Verhaegen. "PO-MOESP subspace identification of Directed Acyclic Graphs with unknown topology," *Automatica*, vol. 53, pp. 60-71, 2015.

[10] R. N. Cabrera, V. M. A. Martinez and M. A. Medina. "Computing the extended observability matrix in the PO-MOESP algorithm: an alternative point of view," *IEEE Latin America Transactions*, vol. 14, no. 1, pp. 314-319, 2016.

[11] M. Bergamasco and M. Lovera. "Continuous-time predictor-based subspace identification using Laguerre filters," *IET Control Theory & Applications*, vol. 5, no. 7, pp. 856-867, 2011.

[12] X. S. Luo and Y. D. Song, "Adaptive predictive control: a data-driven closed-loop subspace identification approach," *Abstract and Applied Analysis*, pp. 1-11, 2014.

[13] W. Favoreel, B. De Moor and P. Van Overschee. "Subspace state space system identification for industrial processes," *Journal of Process Control*, vol. 10, no. 2-3, pp. 149-155, 2000.

[14] N. A. Mardi and L. P. Wang, "Subspace-based model predictive control with data-prefiltering," in Proceedings of the UKACC International Conference on Control, Manchester, UK, 2008.

[15] M. Verhaegen and V. Verdult. *Filtering and system identification: a least squares approach*. Cambridge university press, Cambridge, 2007.

[16] X. S. Luo and Y. D. Song, "Data-driven Predictive Control of Hammerstein-Wiener Systems Based on Subspace Identification," *Information Sciences*, vol. 422, pp. 447–461, 2018.