Preprints of the 8th IFAC Symposium on Advanced Control of Chemical Processes
The International Federation of Automatic Control
Furama Riverfront, Singapore, July 10-13, 2012

A Bias-Eliminated Subspace Identification Method for Errors-in-Variables Systems

Tao Liu*, **

*AVT-Process Systems Engineering, RWTH Aachen University, Turmstrasse 46, Aachen 52064, Germany
** Institute of Advanced Control Technology, Dalian University of Technology, Dalian 116024, P.R. China
(e-mail: liurouter@ieee.org)

Abstract: For model identification of industrial operating systems subject to noisy input-output observations, known as the error-in-variables (EIV) problem, a subspace identification method is proposed in this paper by developing an orthogonal projection approach to guarantee consistent estimation of the deterministic part of such a system. The rank condition for such orthogonal projection is analyzed in terms of the nominal state-space model structure. Using the principal component analysis (PCA), the extended observability matrix and low triangular block-Toeliptz matrix of the state-space model are analytically derived. Accordingly, the system state-space matrices can be retrieved in a transparent manner from the above matrices through linear algebra or an ordinary least-squares (LS) algorithm. A benchmark example used in the existing references is adopted to demonstrate the effectiveness and merit of the proposed subspace identification method.

Keywords: Subspace identification, Errors-in-variables system, Extended observability matrix, Orthogonal projection, Least-squares algorithm.

1. INTRODUCTION

In industrial engineering applications, many operating systems are subject to noisy input-output observations, known as the error-in-variables (EIV) problem. Different solutions have been developed in the past years for model identification of linear dynamic systems from noise-corrupted input and output measurements. As surveyed in the recent literature (Söderström, 2007), most of existing references have been devoted to bias-eliminated (BE) identification methods in terms of the discrete-time transfer function models or frequency domain models. Owing to that state-space models have been widely used for model predictive control (MPC) in industry, subspace identification methods have been rapidly developed in the recent years (Gustafsson, 2002; Chiuso and Picci, 2005; Qin, 2006; Katayama and Tanaka, 2007; Pour, Huang and Shah, 2009), but most of them have been focused on BE identification of open-loop or closed-loop systems with only output measurement noise, since the early works occurred in the last two decades (Larimore, 1990; Van Overschee and De Moor, 1994; Verhaegen, 1994; Viberg, 1995; Jansson and Wahlberg, 1996; Ljung and McKelvey, 1996). For robust identification of EIV systems against both the input and output measurement noises, a QR factorization approach was proposed in the reference (Chou and Verhaegen, 1997) to eliminate the influence from these noises; a dynamic principle component analysis (PCA) based identification algorithm was developed in the paper (Li and Qin, 2001), which was further improved by introducing a more suitable instrumental variable to remove the noise influence in the later paper (Wang and Qin, 2002). For closed-loop systems subject to the input and output measurement noises, a subspace identification algorithm was derived by using an orthogonal projection approach on the closed-loop identification data (Huang, Ding and Qin, 2005).

An extended subspace identification method is proposed in this paper compared with the above subspace identification methods for EIV systems. Based on an analysis on the rank condition for using the orthogonal projection to eliminate the noise influence, a consistent estimation is obtained for the extended observability matrix and low triangular block-Toeliptz matrix of the system state-space model. Accordingly, the system state-space matrices can be retrieved directly from the above matrices or using an ordinary least-squares (LS) algorithm.

Throughout this paper, the following notations are used: $\Re^{n\times m}$ denotes a $n\times m$ real matrix space. For any matrix $P\in\Re^{m\times m}$, $P>0$ (or $P\geq 0$) means P is a positive (or semipositive) definite symmetric matrix, in which the symmetric elements are indicated by '*'. Denote by P^T the transpose of P. For $P\in\Re^{m\times m}$ of full rank, P^{-1} denotes the inverse of P; for $P\in\Re^{m\times n}$ of full row (or column) rank, $P^{\dagger}$ denotes the Moore-Penrose pseudo-inverse of P. The identity vector/matrix and the zero vector/matrix with appropriate dimensions are denoted by $\boldsymbol{I}$ and $\mathbf{0}$, respectively. Note that $\boldsymbol{I}_{m\times m}$ (or $\mathbf{0}_{m\times m}$) means $\boldsymbol{I}_{m\times m}\in\Re^{m\times m}$ ($\mathbf{0}_{m\times m}\in\Re^{m\times m}$). Denote by $\mathrm{rank}(P)$ the rank number of P. Denote by $E\{\bullet\}$ the statistical expectation operator. Denote by Λ/V an orthogonal projection of the row space of $\Lambda\in\Re^{m\times j}$ onto the row space of $\mathrm{V}\in\Re^{n\times j}$, which may be computed through

$\Lambda/\mathrm{V}=\Lambda\mathrm{V}^T(\mathrm{V}\mathrm{V}^T)^{-1}\mathrm{V}$. The orthogonal complement of the row space Λ is denoted by $\Lambda^{\perp}\in\Re^{(j-m)\times j}$, and correspondingly, there is $\Lambda/\mathrm{V}^{\perp}=\Lambda-\Lambda/\mathrm{V}$. Denote by $\delta_{i,j}$ the Kronecker delta function, $\delta_{i,j}=1$ for $i=j$ and $\delta_{i,j}=0$ for $i\neq j$. Denote by $PE(n)$ the persistent excitation order n of the input excitation signal for identification.

For the rank analysis of matrix multiplication, the following lemma is needed here:

Lemma 1 *(Dennis, 2009): Let $A\in\Re^{m\times n}$ and $B\in\Re^{n\times l}$. Then, the following statements hold:*

(i) $\mathrm{rank}(A)\le\min\{m,\ n\}$;

(ii) $\mathrm{rank}(A)=\mathrm{rank}(A^T)=\mathrm{rank}(A^TA)=\mathrm{rank}(AA^T)$;

(iii) $\mathrm{rank}(AB)\le\min\{\mathrm{rank}(A),\ \mathrm{rank}(B)\}$;

(iv) $\mathrm{rank}(AB)\ge\mathrm{rank}(A)+\mathrm{rank}(B)-n$;

(v) $\mathrm{rank}(AB)=n$ if $\mathrm{rank}(A)=\mathrm{rank}(B)=n$.

2. PROBLEM FORMULATION

Consider an EIV system that is generally described by

$$\begin{cases} x(t+1)=Ax(t)+Bu(t)+G\zeta_x(t) \\ y(t)=Cx(t)+Du(t)+\zeta_y(t) \\ u(t)=\hat{u}(t)+\zeta_u(t) \end{cases} \tag{1}$$

where t denotes the time step in the discrete time domain. $x(t)\in\Re^{n_x}$, $y(t)\in\Re^{n_y}$, $u(t)\in\Re^{n_u}$, $\zeta_x(t)\in\Re^{n_x}$, $\zeta_u(t)\in\Re^{n_u}$, and $\zeta_y(t)\in\Re^{n_y}$.

Without loss of generality, the following assumptions are considered herein for study:

A1: The system described by (1) is asymptotically stable, i.e., all eigenvalues of A lie inside the unit circle.

A2: The system pair $(A,\ C)$ is observable.

A3: All these noises are stationary zero-mean Gaussian white noise processes, and their variances and covariances are given by the following unknown matrix:

$$E\left\{\begin{bmatrix}\zeta_x(t)\\ \zeta_y(t)\\ \zeta_u(t)\end{bmatrix}\begin{bmatrix}\zeta_x(k)\\ \zeta_y(k)\\ \zeta_u(k)\end{bmatrix}^T\right\}=\begin{bmatrix} r_{xx} & r_{xy} & r_{xu}\\ * & r_{yy} & r_{yu}\\ * & * & r_{uu}\end{bmatrix}\delta_{t,k} \tag{2}$$

A4: All these noises are independent of the system state sequence $x(t)$ including the true output $y(t)$, and the input excitation sequence $u(t)$ used in an identification test, i.e.

$$E\left\{\begin{bmatrix}\zeta_x(t)\\ \zeta_y(t)\\ \zeta_u(t)\end{bmatrix}\begin{bmatrix}x(k)\\ u(k)\end{bmatrix}^T\right\}=\mathbf{0},\ \forall t,\ k \tag{3}$$

Regarding the sample time step, t_0+i, in an identification test, we define the past and future input block-Hankel matrices as below:

$$U_p=\begin{bmatrix} u(t_0) & u(t_0+1) & \cdots & u(t_0+j-1)\\ u(t_0+1) & u(t_0+2) & \cdots & u(t_0+j)\\ \vdots & \vdots & \cdots & \vdots\\ u(t_0+i-1) & u(t_0+i) & \cdots & u(t_0+i+j-2)\end{bmatrix}\in\Re^{n_u i\times j} \tag{4}$$

$$U_f=\begin{bmatrix} u(t_0+i) & u(t_0+i+1) & \cdots & u(t_0+i+j-1)\\ u(t_0+i+1) & u(t_0+i+2) & \cdots & u(t_0+i+j)\\ \vdots & \vdots & \cdots & \vdots\\ u(t_0+2i-1) & u(t_0+2i) & \cdots & u(t_0+2i+j-2)\end{bmatrix}\in\Re^{n_u i\times j} \tag{5}$$

where $j>i>n_x$ is a precondition adopted for data collection and computation.

Similarly, the block-Hankel matrices of the system output, state, and noises are defined by Y_p, $Y_f\in\Re^{n_y i\times j}$, ξ_{x-p}, $\xi_{x-f}\in\Re^{n_x i\times j}$, ξ_{u-p}, $\xi_{u-f}\in\Re^{n_u i\times j}$, and ξ_{y-p}, $\xi_{y-f}\in\Re^{n_y i\times j}$, respectively.

Meanwhile, define the state vectors, $X_p\triangleq[x^T(t),\ x^T(t+1),\ \cdots,x^T(t+j-1)]^T\in\Re^{n_x\times j}$ and $X_f\triangleq[x^T(t+i),\ x^T(t+i+1),\ \cdots,x^T(t+i+j-1)]^T\in\Re^{n_x\times j}$.

By iterating (1) in terms of the past and future time sequences with respect to t_0+i, we obtain

$$Y_p=\Gamma_i X_p+H_iU_p+\Phi_i\xi_{x-p}+\xi_{y-p} \tag{6}$$

$$Y_f=\Gamma_i X_f+H_iU_f+\Phi_i\xi_{x-f}+\xi_{y-f} \tag{7}$$

where

$$\Gamma_i=[C^T,\ A^TC^T,\ \cdots,\ (A^{i-1})^TC^T]^T\in\Re^{n_y i\times n_x} \tag{8}$$

$$H_i=\begin{bmatrix} D & 0 & 0 & \cdots & 0\\ CB & D & 0 & \cdots & 0\\ CAB & CB & D & \cdots & 0\\ \vdots & \vdots & \vdots & D & 0\\ CA^{i-2}B & CA^{i-3}B & \cdots & CB & D\end{bmatrix}\in\Re^{n_y i\times n_u i} \tag{9}$$

$$\Phi_i=\begin{bmatrix} \mathbf{0} & 0 & 0 & \cdots & 0\\ CG & \mathbf{0} & 0 & \cdots & 0\\ CAG & CG & \mathbf{0} & \cdots & 0\\ \vdots & \vdots & \vdots & \mathbf{0} & 0\\ CA^{i-2}G & CA^{i-3}G & \cdots & CG & \mathbf{0}\end{bmatrix}\in\Re^{n_y i\times n_x i} \tag{10}$$

3. CONSISTENT ESTIMATION OF THE EXTENDED OBSERVABILITY MATRIX

It can be seen from (6) or (7) that the state-space model identification is associated with the estimation of the extended observability matrix Γ_i and the block Toeplitz matrix H_i. For consistent estimation of the state-space model against the input and output noises, consistent estimation on Γ_i and H_i must be required in the first place.

For the convenience of analysis, denote two short-hands, $W_p=[Y_p^T,\ U_p^T]^T\in\Re^{i(n_u+n_y)\times j}$ and $W_f=[Y_f^T,\ U_f^T]^T\in\Re^{i(n_u+n_y)\times j}$. Define

$$r_u(\tau)=\lim_{N\to\infty}\frac{1}{N}\sum_{t=0}^{N-1}u(t_0+t+\tau)u^T(t_0+t) \tag{11}$$

$$R_u(i)=\begin{bmatrix} r_u(i) & r_u(i+1) & \cdots & r_u(2i-1) \\ r_u(i-1) & r_u(i) & \cdots & r_u(2i-2) \\ \vdots & \vdots & \vdots & \vdots \\ r_u(1) & r_u(2) & \cdots & r_u(i) \end{bmatrix} \tag{12}$$

It follows from (7) that

$$\begin{bmatrix} \boldsymbol{I} & -H_i \end{bmatrix} W_f = \Gamma_i X_f + \Phi_i \xi_{x-f} + \xi_{y-f} \tag{13}$$

In view of that W_p is incorrelated with ξ_{x-f} and ξ_{y-f}, we perform an orthogonal projection of (13) onto the row space of W_p to remove the noise terms, obtaining

$$\begin{bmatrix} \boldsymbol{I} & -H_i \end{bmatrix} W_f / W_p = \Gamma_i X_f / W_p \tag{14}$$

Then, premultiplying both sides of (14) by the orthogonal complement of Γ_i, denoted by $\Gamma_i^{\perp}$, yields

$$(\Gamma_i^{\perp})^T \begin{bmatrix} \boldsymbol{I} & -H_i \end{bmatrix} W_f / W_p = \boldsymbol{0} \tag{15}$$

Given the fact that $PE(u) > i$ in performing an identification test, W_p is guaranteed full row rank and therefore, its SVD can be obtained as

$$W_p = U_{W_p} \begin{bmatrix} \Lambda_{W_p} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} V_{W_{p1}}^T \\ V_{W_{p2}}^T \end{bmatrix} \tag{16}$$

where $U_{W_p} \in \Re^{i(n_u+n_y)\times i(n_u+n_y)}$ is an unitary matrix, $\Lambda_{W_p} \in \Re^{i(n_u+n_y)\times i(n_u+n_y)}$ is a diagonal matrix with nonnegative entries, $V_{W_{p1}} \in \Re^{j\times i(n_u+n_y)}$ and $V_{W_{p2}} \in \Re^{j\times(j-n_u i-n_y i)}$ are of full column rank matrix composed of unitary orthogonal column vectors.

Substituting (16) into the orthogonal projection of W_f / W_p, it can be derived that

$$W_f / W_p = W_f V_{W_{p1}} V_{W_{p1}}^T \tag{17}$$

To solve the extended observability matrix Γ_i from (15), the following theorem is given:

Theorem 1 *Given the input excitation sequence of $PE(u) \geq i$ that guarantees both W_p and W_f full row rank and $R_u(i) > 0$, there exists*

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}} V_{W_{p1}}^T) = n_u i + n_x \tag{18}$$

and for $N \to \infty$,

$$\begin{bmatrix} \Gamma_i^{\perp} \\ -H_i^T \Gamma_i^{\perp} \end{bmatrix} = \tilde{\mathrm{P}} \Xi \tag{19}$$

where $\tilde{\mathrm{P}}$ is obtained from the PCA-based SVD of

$$W_f V_{W_{p1}} V_{W_{p1}}^T = \begin{bmatrix} \mathrm{P} & \tilde{\mathrm{P}} \end{bmatrix} \begin{bmatrix} \Lambda_{\mathrm{P}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} V_{\mathrm{P}}^T \\ V_{\tilde{\mathrm{P}}}^T \end{bmatrix} \tag{20}$$

and $\Xi \in \Re^{(n_y i-n_x)\times(n_y i-n_x)}$ is any matrix of full rank.

Proof: It follows from (7) that

$$W_f = \begin{bmatrix} Y_f \\ U_f \end{bmatrix} = \begin{bmatrix} \Gamma_i & H_i \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} X_f \\ U_f \end{bmatrix} + \begin{bmatrix} \Phi_i & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \xi_{x-f} \\ \xi_{y-f} \end{bmatrix} \tag{21}$$

Postmultiplying both sides of (21) by W_p^T to remove the noise part yields

$$\lim_{N\to\infty} \frac{1}{N} W_f W_p^T = \lim_{N\to\infty} \frac{1}{N} \begin{bmatrix} \Gamma_i & H_i \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} X_f \\ U_f \end{bmatrix} W_p^T \tag{22}$$

Since $j > i > n_x$, there exist

$$\operatorname{rank}(\Gamma_i) = n_x \tag{23}$$

$$\operatorname{rank}(\begin{bmatrix} \Gamma_i & H_i \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}) = n_u i + n_x \tag{24}$$

Owing to the fact that $PE(u) > i$ and $R_u(i) > 0$, we have

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} X_f W_p^T) = n_x \tag{25}$$

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} U_f W_p^T) = n_u i \tag{26}$$

Hence, it follows from the statement (v) in Lemma 1 that

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f W_p^T) = n_u i + n_x \tag{27}$$

Note that

$$\operatorname{rank}(\begin{bmatrix} \Lambda_{W_p} & \boldsymbol{0} \end{bmatrix}^T U_{W_p}^T) = i(n_u + n_y) \tag{28}$$

According to the statements (iii) and (iv) in Lemma 1, we have

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f W_p^T) \leq \operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}}) \tag{29}$$

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f W_p^T) \geq \operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}}) + \operatorname{rank}(\begin{bmatrix} \Lambda_{W_p} & \boldsymbol{0} \end{bmatrix}^T U_{W_p}^T) - i(n_u + n_y) \tag{30}$$

Substituting (28) into (30), one can see that

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f W_p^T) = \operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}}) \tag{31}$$

Similarly, it follows that

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}} V_{W_{p1}}^T) \leq \operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}}) \tag{32}$$

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}} V_{W_{p1}}^T) \geq \operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}}) + \operatorname{rank}(V_{W_{p1}}^T) - i(n_u + n_y) \tag{33}$$

In view of $\operatorname{rank}(V_{W_{p1}}^T) = i(n_u + n_y)$, there exists

$$\operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}} V_{W_{p1}}^T) = \operatorname{rank}(\lim_{N\to\infty} \frac{1}{N} W_f V_{W_{p1}}) \tag{34}$$

Therefore, the rank condition of (18) is true.

Accordingly, in the SVD of $W_f V_{W_{p1}} V_{W_{p1}}^T$ as shown in (20), it follows

$$\operatorname{rank}(\mathrm{P}) = n_u i + n_x, \ \mathrm{P} \in \Re^{i(n_u+n_y)\times(n_u i+n_x)} \tag{35}$$

$$\operatorname{rank}(\tilde{\mathrm{P}}) = n_y i - n_x, \ \tilde{\mathrm{P}} \in \Re^{i(n_u+n_y)\times(n_y i-n_x)} \tag{36}$$

Therefore, the orthogonal column space of $W_f V_{W_{p1}} V_{W_{p1}}^T$ can be spanned by $\tilde{\mathrm{P}}$. Meanwhile, it can be seen from (15) that the orthogonal column space of $W_f V_{W_{p1}} V_{W_{p1}}^T$ is equal to the column space of $\left((\Gamma_i^{\perp})^T \begin{bmatrix} \boldsymbol{I} & -H_i \end{bmatrix}\right)^T$. Hence, it follows that

$$\left((\Gamma_i^{\perp})^T \begin{bmatrix} \boldsymbol{I} & -H_i \end{bmatrix}\right)^T = \tilde{\mathrm{P}} \Xi \tag{37}$$

which is exactly the result of (19) in Theorem 1. This completes the proof. □

For the convenience of computation, it is suggested to take $\Xi = \boldsymbol{I}$ and correspondingly, it can be obtained that

$$\begin{bmatrix} \Gamma_i^{\perp} \\ -H_i^T \Gamma_i^{\perp} \end{bmatrix} = \tilde{\mathrm{P}} = \begin{bmatrix} \tilde{\mathrm{P}}_y \\ \tilde{\mathrm{P}}_u \end{bmatrix} \tag{38}$$

where $\tilde{\mathrm{P}}_y \in \Re^{n_y i \times (n_y i - n_x)}$ and $\tilde{\mathrm{P}}_u \in \Re^{n_u i \times (n_y i - n_x)}$.
Correspondingly, there are

$$\Gamma_i = \tilde{\mathrm{P}}_y^{\perp} \tag{39}$$

$$-\tilde{\mathrm{P}}_y^T H_i = \tilde{\mathrm{P}}_u^T \tag{40}$$

To obtain H_i from (40), denote

$$-\tilde{\mathrm{P}}_y^T = \alpha = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_i] \tag{41}$$

$$\tilde{\mathrm{P}}_u^T = \beta = [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_i] \tag{42}$$

where $\alpha_k \in \Re^{(n_y i - n_x) \times n_y}$ and $\beta_k \in \Re^{(n_y i - n_x) \times n_u}$, $k = 1, 2, \ldots, i$.
Substituting (9), (41) and (42) into (40) yields

$$[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_i] \begin{bmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & D & 0 \\ CA^{i-2}B & CA^{i-3}B & \cdots & CB & D \end{bmatrix} = [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_i] \tag{43}$$

Let

$$H_{i-1} = \begin{bmatrix} D \\ CB \\ CAB \\ \vdots \\ CA^{i-2}B \end{bmatrix} \tag{44}$$

It can be derived that

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{i-1} & \alpha_i \\ \alpha_2 & \alpha_3 & \cdots & \alpha_i & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_i & 0 & \cdots & 0 & 0 \end{bmatrix} H_{i-1} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_i \end{bmatrix} \tag{45}$$

It is obvious that the coefficient matrix of H_{i-1} is full column rank. Therefore, H_{i-1} can be derived from the standard least-squares algorithm. Consequently, the block Toeplitz matrix H_i can be obtained in terms of the exact triangular structure.

4. DETERMINATION OF THE STATE-SPACE TRANSFER MATRICES

With the above estimation of Γ_i and H_i, it can directly be obtained that

$$C = \Gamma_i(1 : n_y, :) \tag{46}$$

In the sequel, an LS solution of A can be obtained from

$$\Gamma_i(1 : n_y(i-1), :) A = \Gamma_i(n_y + 1 : n_y i, :) \tag{47}$$

It follows from (44) that

$$H_{i-1} = \begin{bmatrix} \boldsymbol{I}_{n_y \times n_y} & \mathbf{0}_{n_y \times n_x} \\ \mathbf{0}_{n_y(i-1) \times n_y} & \Gamma_i(1 : n_y(i-1), :) \end{bmatrix} \begin{bmatrix} D \\ B \end{bmatrix} \tag{48}$$

It is obvious that the coefficient matrix of B and D in (48) is full column rank. Therefore, B and D can be derived from

$$\begin{bmatrix} D \\ B \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_{n_y \times n_y} & \mathbf{0}_{n_y \times n_x} \\ \mathbf{0}_{n_y(i-1) \times n_y} & \Gamma_i(1 : n_y(i-1), :) \end{bmatrix}^{\dagger} H_{i-1} \tag{49}$$

In case of $D = \mathbf{0}$, let

$$\bar{H}_{i-1} = \begin{bmatrix} CB \\ CAB \\ \vdots \\ CA^{i-2}B \end{bmatrix} \tag{50}$$

In view of

$$\bar{H}_{i-1} = \Gamma_i(1 : n_y(i-1), :) B \tag{51}$$

where $\Gamma_i(1 : n_y(i-1), :) \in \Re^{n_y(i-1) \times n_x}$ is full column rank, it can be obtained that

$$B = \Gamma_i(1 : n_y(i-1), :)^{\dagger} \bar{H}_{i-1} \tag{52}$$

Note that the above algorithm gives an estimation of the system matrices (A, B, C, D) up to a similarity transformation. Model validation should be performed by verifying the eignvalues of A that equal the system poles, or the system transfer matrix composed of an algebraic product of (A, B, C, D).

5. ILLUSTRATION

Consider the benchmark example studied by Li and Qin (2001) and Wang and Qin (2002),

$$x(t+1) = \begin{bmatrix} 0.67 & 0.67 & 0 & 0 \\ -0.67 & 0.67 & 0 & 0 \\ 0 & 0 & -0.67 & -0.67 \\ 0 & 0 & 0.67 & -0.67 \end{bmatrix} x(t) + \begin{bmatrix} 0.6598 \\ 1.9698 \\ 4.3171 \\ -2.6436 \end{bmatrix} u(t) + \begin{bmatrix} 0.1762 \\ 0.5278 \\ -0.5532 \\ 0.2983 \end{bmatrix} \zeta_x(t)$$

$$y(t) = \begin{bmatrix} -0.5749 & 1.0751 & -0.5225 & 0.183 \\ 2.4027 & 0.7543 & -0.2159 & 0.0982 \end{bmatrix} x(t) - \begin{bmatrix} 0.7139 \\ 0.5431 \end{bmatrix} u(t)$$

Both the input and output data are corrupted by two independent Gaussian white noise sequences with zero mean and the standard deviations of $\sigma_{\zeta_u} = 0.4$ and $\sigma_{\zeta_y} = 0.4$, respectively. The process noise input is assumed to be another Gaussian white noise sequence with zero mean and standard deviation of $\sigma_{\zeta_x} = 0.1$. To generate the identification data, the input excitation sequence is taken as used in the above references as the combination of 10 sine saves with different frequencies,

$$u(t) = \sum_{k=1}^{10} \sin(0.3898 \pi k t)$$

It is easy to verify that the input sequence satisfies the input condition in Theorem 1. For comparison, 100 Monte-Carlo tests are performed with a data length of $N = 100$ taken in each test for identification. By taking $p = f = 10$, the proposed identification algorithm gives the system poles (i.e., the eigenvalues of A) as listed in Table 1, where the result of each pole is shown by the mean value along with the standard deviation in parentheses. It is seen that the estimation variance is apparently reduced by the proposed algorithm, compared with that of Wang and Qin (2002). Note that under the above noise conditions, the standard MOESP and N4SID algorithms in the MATLAB toolbox give obviously biased estimation and thus is omitted. For illustration, the estimated system poles are plotted in Figure 1 where the true values are denoted by '+' and the estimated results by '×'.

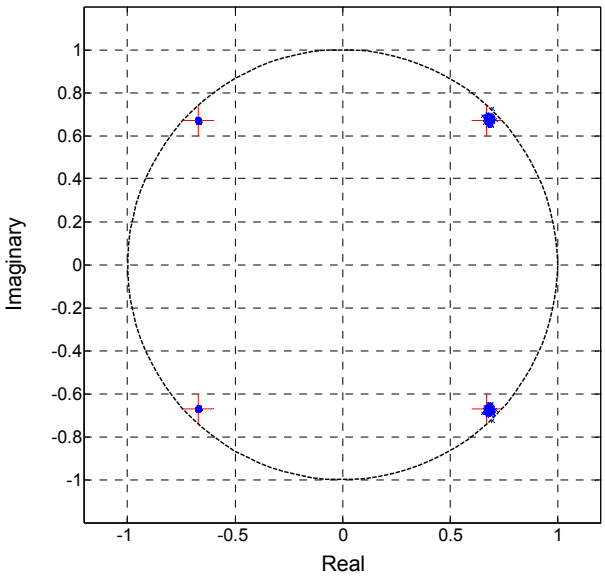

Fig. 1. Plot of the identified system poles ('×') in contrast to the true values ('+')

It is seen that the estimated system poles match well with the true values, thus demonstrating good accuracy of the proposed identification method.

6. CONCLUSIONS

To cope with noisy input and output observations occurring in the so-called EIV operating systems, a bias-eliminated subspace identification method has been proposed in this paper. By analyzing the rank condition for performing the orthogonal projection to eliminate the noise influence, consistent estimation on the extended observability matrix and low triangular block-Toeliptz matrix of the system state-space model is guaranteed through taking the corresponding PCA-based SVD in such orthogonal projection. In the result, the system state-space matrices can be explicitly retrieved from the above matrices, thus facilitating practical applications. An illustrative example has shown well the effectiveness and merit of the proposed subspace identification method.

ACKNOWLEDGEMENT

This work is supported by the Alexander von Humboldt Research Fellowship.

REFERENCES

Bernstein, D. S. (2009). *Matrix Mathematics*, 2nd ed, Princeton: Princeton University Press.

Chiuso, A. and Picci, G. (2005). Consistency analysis of some closed-loop subspace identification methods. *Automatica*, 41, 377-391.

Chou, C. T. and Verhaegen, M. (1997). Subspace algorithms for the identification of multivariable dynamic errors-in-variables models. *Automatica*, 33, 1857–1869.

Gustafsson, T. (2002). Subspace-based system identification: Weighting and pre-filtering of instruments. *Automatica*, 38, 433-443.

Huang, B., Ding, S. and Qin, S. J. (2005). Closed-loop subspace identification: An orthogonal projection approach. *Journal of Process Control*, 15, 53-66.

Jansson, M. and Wahlberg, B. (1996). A linear regression approach to state-space subspace system identification. *Signal Processing*, 52, 103-129.

Katayama, T. and Tanaka, H. (2007). An approach to closed-loop subspace identification by orthogonal decomposition. *Automatica*, 43, 1623-1630.

Larimore, W. E. (1990). Canonical variate analysis in identification, filtering, and adaptive control. *Proceedings of the 29th IEEE Conference on Decision and Control*, *Honolulu, Hawaii*, v1, 596-604.

Li, W. and Qin, S. J. (2001). Consistent dynamic PCA based on errors-in-variables subspace identification. *Journal of Process Control*, 11, 661-678.

Ljung, L. and McKelvey, T. (1996). Subspace identification from closed loop data. *Signal Processing*, 52, 209–215.

Pour, N. D., Huang, B. and Shah, S. L. (2009). Consistency of noise covariance estimation in joint input–output closed-loop subspace identification with application in LQG benchmarking. *Journal of Process Control*, 19, 1649-1657.

Qin, S. J. (2006). An overview of subspace identification. *Computers and Chemical Engineering*, 30, 1502-1513.

Söderström, T. (2007). Errors-in-variables methods in system identification. *Automatica*, 43, 939-958.

Van Overschee, P. and De Moor, B. (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30, 75-93.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica*, 30, 61–74.

Viberg, M. (1995). Subspace-based methods for the identification of linear time-invariant systems. *Automatica*, 31, 1835–1851.

Wang, J. and Qin, S. J. (2002). A new subspace identification approach based on principal component analysis. *Journal of Process Control*, 12, 841–855.

Table 1 Comparison of the identified system poles (the eigenvalues of A in the system state-space model) with the true values

True poles	Proposed	Wang & Qin
$0.67+0.67i$	$0.6849(\pm 0.011)+0.6765(\pm 0.0139)i$	$0.6578(\pm 0.1368)+0.6689(\pm 0.0374)i$
$0.67-0.67i$	$0.6849(\pm 0.011)-0.6765(\pm 0.0139)i$	$0.6578(\pm 0.1368)-0.6689(\pm 0.0374)i$
$-0.67+0.67i$	$-0.6695(\pm 0.0036)+0.6707(\pm 0.004)i$	$-0.6573(\pm 0.1247)+0.6658(\pm 0.0439)i$
$-0.67-0.67i$	$-0.6695(\pm 0.0036)-0.6707(\pm 0.004)i$	$-0.6573(\pm 0.1247)-0.6658(\pm 0.0439)i$