Mem. Fac. Eng., Kyoto Univ., Vol. 57, No. 2 (1995)

A Subspace Identification of δ-Operator State-Space Model

By

Tohru KATAYAMA, Takaya YAMAMOTO and Yang WU

(Received December 26, 1994)

Abstract

This paper derives a subspace identification algorithm for a δ-operator state-space model by using the methods due to Moonen *et al.* [11], [12], [21]. Since the δ-operator model converges to a continuous-time model as the sampling interval goes to zero, the algorithm obtained is applicable to the identification of continuous-time medels. A method of computing the state vector from the block Hankel matrix is developed. Simulation studies show the present algorithm provides good results for the case of a low N/S ratio. Improvement of the algorithm for the case of a higher N/S ratio remains to be done.

1. Introduction

Some thirty years ago, Ho and Kalman [1] developed a basic minimal realization technique of the state-space model based on the block Hankel matrix constructed by Markov parameters, or the impulse responses. Also, Kung [2] derived an algorithm for obtaining a reduced order state-space model by using SVD (singular value decomposition) [3] of the Hankel matrix. To apply the above techniques, we must first estimate Markov parameters based on the input-output data. Since the estimation of the Markov parameters is not a trivial task [4], the techniques of [1] and [2] are not suitable for practical application.

By defining the predictor space based on the CVA (canonical variate analysis), stochastic realization theory was initiated by the pioneering works of Akaike [4], [5], in which the block Hankel matrix is generated by the covariance matrices of input-output data. Also, Larimore [6], [7] has derived a general reduced order identification technique for MIMO linear state-space models by extending the CVA based technique of [4] so that the arbitrary control inputs can be included in the model. The computation associated with the CVA can effectively be performed by the SVD.

More recently, the subspace method has received much interest in system identification and signal processing [8], [9], [10]. In the subspace method, the identification problem is formulated and solved on signal level; the main problem is thus the approximation

* Department of Applied Mathematics and Physics

of a subspace spanned by the column or row vectors in block Hankel matrices formed by the array of input-output data. The most effective technique for solving this approximation problem is due to the SVD. In particular, subspace state-space identification techniques have been developed based on the SVD of the block Hankel matrix by Moonen *et al.* [11], [12]. Verhaegen and Dewilde [13] have derived a subspace output error method for the identification of the state-space model based on QR decomposition. Also, the subspace methods are analyzed from a statistical point of view by Viberg *et al.* [14].

The classical system identification techniques are based on the least-squares (LS) method or iterative nonlinear optimization techniques (see [15], [16]). The drawbacks of this classical approach are the difficulty in model selection and the overparametrization of the model. For example, pole-zero cancellation in a polynomial model makes the model not identifiable, so that the multivariable ARMAX parametrization is inherently ill-conditioned. For the linear time-invariant models, the subspace identification schemes are possible alternatives to the classical approach in that model selection is much simpler and the application to MIMO cases is almost trivial. Thus for the state-space models, the subspace approach has better numerical conditioning than the classical polynomial model identification, although the determination of the model order is not a trivial task for noisy input-output data. In the CVA approach [6], [7], the model order is selected based on the AIC [4].

Another recent interest in this area is the identification of continuous-time models from sampled data in the literature [17], [18], because the analysis and design of a control system are usually carried out by using continuous-time models since most physical systems are continuous-time. The indirect approach is to first estimate a discrete-time model using sampled-data by the classical approach and then convert it to a continuous-time model. It is shown [18] that the continuous-time model obtained using this approach is highly sensitive to the choice of sampling interval, since the discrete-time and continuous-time models in frequency domain are connected by the transcendental relation $z=e^{s\Delta}$, where Δ is the sampling interval. This difficulty may be overcome by using a δ-operator model rather than a standard shift-operator model [19], [20].

The direct approach is to estimate the parameters of a continuous-model based on the sampled data, without computing an intermediate discrete-time model. A basic idea is to obtain an equivalent discrete-time model whose parameters are identical to those of a continuous time model by using a numerical integration based on a digital filter [17], [18]. Also a direct SVD-based subspace identification method for continuous-time state-space models is presented by Moonen *et al.* [21], in which state variable filters are used for approximate computing of higher order derivatives.

In this paper, motivated by the works of [20] and [21], we derive a subspace identification algorithm for a δ-operator state-space model. Since the δ-operator model reduces to the continuous-time model as the sampling interval tends to zero, the present technique may be applied to the identification of continuous-time state-space models [20]. In Section 2, we briefly describe the δ-operator model based on [19]. In Section 3, based on [11], [12], [21], we present relevant block Hankel matrices formed by the input-output data to determine the state vector of the model. A method of computing the state vector is developed by using the two SVDs, namely the SVD of the Hankel matrix formed by the input-output data and the SVD of a submatrix formed by the left singular vectors obtained from the first SVD. The system matrices are then determined by applying the LS method to an overdetermined system of equations. Section 5 considers the case where the input-output data are corrupted by white noise. A prefiltering scheme is developed in order to compute the higher order differences of the input-output data. The LS estimate of the block Hankel matrix is then derived by using the technique due to De Moor [23]. Numerical results are presented in Section 6 to show the feasibility of the present algorithm by using a version of the model from [20]. The conclusions are given in Section 7. Appendix includes a proof of Lemma 3.

2. δ-Operator Model

Consider a continuous-time model

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

where $x(t) \in R^n$ is the state vector, $u(t) \in R^m$ is the input vector, $y(t) \in R^p$ is the output vector, and A, B, C, D are $n \times n$, $n \times m$, $p \times n$, $p \times m$ constant matrices, respectively.

Suppose that the input $u(t)$ is a staircase function of the form

$$u(t) = u(k\Delta), \quad k\Delta \le t < (k+1)\Delta, \quad k = 0, 1, 2, \cdots \tag{2}$$

where Δ is the sampling interval. It then follows that

$$x((k+1)\Delta) = e^{A\Delta}x(k\Delta) + \left(\int_0^\Delta e^{A\tau}d\tau\right) Bu(k\Delta)$$

Thus we have

$$x(t+\Delta) = A_q x(t) + B_q u(t), \quad t = 0, \Delta, 2\Delta, \cdots \tag{3}$$

where

$$A_q=e^{A\Delta}, \quad B_q=\left(\int_0^{\Delta} e^{A\tau}d\tau\right)B \tag{4}$$

By using the shift operator q, we have the following discrete-time model relating the sampled input to the sampled output

$$\begin{aligned} qx(t)&=A_qx(t)+B_qu(t) \\ y(t)&=Cx(t)+Du(t), \quad t=0,\ \Delta,\ 2\Delta,\ \cdots \end{aligned} \tag{5}$$

It follows from (4) that for $\Delta\to 0$, we have $A_q\to I$, $B_q\to 0$, so that the discrete-time model degenerates. Hence, in order to derive a model that has a better correspondence with the continuous-time model, we define the delta operator ([17], [19])

$$\delta x(t)=\frac{x(t+\Delta)-x(t)}{\Delta} \tag{6}$$

Note that this is the forward difference with $\delta:=(q-1)/\Delta$. Since $q=1+\Delta\delta$, it follows that (5) is reduced to

$$\begin{aligned} \delta x(t)&=A_\delta x(t)+B_\delta u(t) \\ y(t)&=Cx(t)+Du(t), \quad t=0,\ \Delta,\ 2\Delta,\ \cdots \end{aligned} \tag{7}$$

where

$$A_\delta=\frac{A_q-I}{\Delta}, \quad B_\delta=\frac{B_q}{\Delta}$$

Since $A_\delta\to A$, $B_\delta\to B$ as $\Delta\to 0$, we see that the delta operator model of (7) reduces to the continuous-time model where the sampling interval is very small. This fact shows that the identification algorithm for the δ-operator model is applicable to the identification of a continuous-time model of (1) (see [20]).

3. State Vector and Block Hankel Matrix

Consider the δ-operator state-space model of (7), where A_δ, B_δ, C, D are $n\times n$, $n\times m$, $p\times n$, $p\times m$ matrices, respectively. We define the augmented controllability and observability matrices

$$\mathcal{C}_k=[A_\delta^{k-1}B_\delta \cdots A_\delta B_\delta \; B_\delta], \quad \mathcal{O}_k=\begin{bmatrix} C \\ CA_\delta \\ \vdots \\ CA_\delta^{k-1} \end{bmatrix}$$

where k is assumed to be larger than n. In the following, we assume that the state-space model is controllable and observable, so that we have rank $\mathcal{C}_k=n$, rank $\mathcal{O}_k=n$. It is to be noted that (A_δ, B_δ, C) is minimal if and only if (A_q, B_q, C) is minimal.

By using higher order differences of the input-output variables, we now define two $k\times L$ block Hankel matrices

$$U_{t,j}=\begin{bmatrix} \delta^j u(t) & \delta^j u(t+\Delta) & \cdots & \delta^j u(t+(L-1)\Delta) \\ \delta^{j+1}u(t) & \delta^{j+1}u(t+\Delta) & \cdots & \delta^{j+1}u(t+(L-1)\Delta) \\ \vdots & \vdots & & \vdots \\ \delta^{j+k-1}u(t) & \delta^{j+k-1}u(t+\Delta) & \cdots & \delta^{j+k-1}u(t+(L-1)\Delta) \end{bmatrix} \tag{8}$$

and

$$Y_{t,j}=\begin{bmatrix} \delta^j y(t) & \delta^j y(t+\Delta) & \cdots & \delta^j y(t+(L-1)\Delta) \\ \delta^{j+1}y(t) & \delta^{j+1}y(t+\Delta) & \cdots & \delta^{j+1}y(t+(L-1)\Delta) \\ \vdots & \vdots & & \vdots \\ \delta^{j+k-1}y(t) & \delta^{j+k-1}y(t+\Delta) & \cdots & \delta^{j+k-1}y(t+(L-1)\Delta) \end{bmatrix} \tag{9}$$

where both $U_{t,j}$ and $Y_{t,j}$ have k block rows and L columns, although k, L do not appear as indices of them. We also define the augmented state vector with L columns as

$$X_{t,j}=[\delta^j x(t) \; \delta^j x(t+\Delta) \; \cdots \; \delta^j x(t+(L-1)\Delta)] \tag{10}$$

It follows from (8)–(10) that

$$Y_{t,0}=\mathcal{O}_k X_{t,0}+\Gamma_k U_{t,0} \tag{11}$$

$$Y_{t,k}=\mathcal{O}_k X_{t,k}+\Gamma_k U_{t,k} \tag{12}$$

where Γ_k is the block Toeplitz matrix defined by

$$\Gamma_k=\begin{bmatrix} D & 0 & 0 & \cdots & 0 & 0 \\ CB_\delta & D & 0 & \cdots & 0 & 0 \\ CA_\delta B_\delta & CB_\delta & D & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ CA_\delta^{k-2}B_\delta & CA_\delta^{k-3}B_\delta & CA_\delta^{k-4}B_\delta & \cdots & CB_\delta & D \end{bmatrix}$$

The above input-output matrix relations are used for defining the state vector by using external variables $U_{t,j}$ and $Y_{t,j}$. In the following, we assume that $L \gg \max(km, kp)$, namely both $U_{t,j}$ and $Y_{t,j}$ are rectangular.

We define two block Hankel matrices H_1 and H_2 as

$$H_1=\begin{bmatrix} U_{t,0} \\ Y_{t,0} \end{bmatrix}, \quad H_2=\begin{bmatrix} U_{t,k} \\ Y_{t,k} \end{bmatrix}$$

Let W_1 and W_2 be subspaces spanned by the row vectors of H_1 and H_2, respectively. Then we have

$$W_1=\mathrm{Im}\,(H_1^T), \quad W_2=\mathrm{Im}\,(H_2^T)$$

Lemma 1 ([11]) Suppose that the following three conditions hold.

1) $\mathrm{rank} X_{t,0}=n$
2) $\mathrm{Im}\,(X_{t,0}^T) \cap \mathrm{Im}\,(U_{t,0}^T)=\phi$
3) $\mathrm{rank} U_{t,0}=km$

Then it follows that

$$\mathrm{rank} H_1=km+n \tag{13}$$

Proof: See Moonen *et al.* [11]. □

The following lemma gives a fundamental relation between the state vector and the subspaces defined by external variables.

Lemma 2 ([11]) Suppose that the conditions in Lemma 1 hold. Then the subspace spanned by the row vectors of $X_{t,k}$ coincides with the intersection of subspaces W_1 and W_2, namely, under the assumption that rank $\begin{bmatrix} U_{t,0} \\ U_{t,k} \end{bmatrix}=2km$,

$$\mathrm{Im}\,(X_{t,k}^T)=W_1 \cap W_2 \tag{14}$$

Proof: See Moonen *et al.* [11]. □

4. Determination of State Vector and System Matrices

For convenience, we redefine H_1 and H_2 as

$$H_1=\begin{bmatrix} u(t) & u(t+\Delta) & \cdots & u(t+(L-1)\Delta) \\ y(t) & y(t+\Delta) & \cdots & y(t+(L-1)\Delta) \\ \delta u(t) & \delta u(t+\Delta) & \cdots & \delta u(t+(L-1)\Delta) \\ \delta y(t) & \delta y(t+\Delta) & \cdots & \delta y(t+(L-1)\Delta) \\ \vdots & \vdots & & \vdots \\ \delta^{k-1}u(t) & \delta^{k-1}u(t+\Delta) & \cdots & \delta^{k-1}u(t+(L-1)\Delta) \\ \delta^{k-1}y(t) & \delta^{k-1}y(t+\Delta) & \cdots & \delta^{k-1}y(t+(L-1)\Delta) \end{bmatrix} \tag{15}$$

and

$$H_2=\begin{bmatrix} \delta^k u(t) & \delta^k u(t+\Delta) & \cdots & \delta^k u(t+(L-1)\Delta) \\ \delta^k y(t) & \delta^k y(t+\Delta) & \cdots & \delta^k y(t+(L-1)\Delta) \\ \delta^{k+1} u(t) & \delta^{k+1} u(t+\Delta) & \cdots & \delta^{k+1} u(t+(L-1)\Delta) \\ \delta^{k+1} y(t) & \delta^{k+1} y(t+\Delta) & \cdots & \delta^{k+1} y(t+(L-1)\Delta) \\ \vdots & \vdots & & \vdots \\ \delta^{2k-1} u(t) & \delta^{2k-1} u(t+\Delta) & \cdots & \delta^{2k-1} u(t+(L-1)\Delta) \\ \delta^{2k-1} y(t) & \delta^{2k-1} y(t+\Delta) & \cdots & \delta^{2k-1} y(t+(L-1)\Delta) \end{bmatrix} \tag{16}$$

It is clear that Lemma 2 also holds for these block Hankel matrices.

Define $H=\begin{bmatrix} H_1 \\ H_2 \end{bmatrix}$. Let the SVD of H be given by

$$H=U_H S_H V_H^T,\quad \begin{bmatrix} H_1 \\ H_2 \end{bmatrix}=\begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}\begin{bmatrix} S_{11} & 0 \\ 0 & 0 \end{bmatrix} V_H^T \tag{17}$$

where

$$\begin{aligned} &U_H\text{: } 2k(m+p)\times 2k(m+p); && U_{11}, U_{21}\text{: } k(m+p)\times(2km+n) \\ &S_H\text{: } 2k(m+p)\times L; && U_{12}, U_{22}\text{: } k(m+p)\times(2kp-n) \\ &V_H\text{: } L\times L; && S_{11}\text{: } (2km+n)\times(2km+n) \end{aligned}$$

Lemma 3 ([12]) The SVD of U_{12} of (17) is given by

$$U_{12}=[Q_1\ Q_2\ Q_3]\ \Sigma_{12} W^T \tag{18}$$

where

$$\Sigma_{12}=\begin{bmatrix} I_{kp-n} & & \\ & C_n & \\ & & 0_{km\times(kp-n)} \end{bmatrix},\quad C_n=\mathrm{diag}\,(c_1, \cdots, c_n)$$

and

$$\begin{aligned} &Q_1\text{: } k(m+p)\times(kp-n) \\ &Q_2\text{: } k(m+p)\times n \\ &Q_3\text{: } k(m+p)\times km \\ &W\text{: } (2kp-n)\times(2kp-n) \end{aligned}$$

Proof: A proof is given in Appendix based on [22], since no proof is provided in [12]. □

4.1 Determination of State Vector

Lemma 4 ([11]) The subspace spanned by the row vectors of $U_{12}^T H_1$ is included in $W_1 \cap W_2$, namely

$$\mathrm{Im}\ \{(U_{12}^T H_1)^T\} \subset W_1 \cap W_2 \tag{19}$$

Proof: Since U_H is orthogonal,

$$[U_{12}^T\ \ U_{22}^T] \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} = [U_{12}^T\ \ U_{22}^T] \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} S_{11} & 0 \\ 0 & 0 \end{bmatrix} V_H^T = 0$$

so that

$$U_{12}^T H_1 = -U_{22}^T H_2 \tag{20}$$

we see that the row spaces of both sides of (20) are included in W_1 and W_2. □

Lemma 5 The row space of $U_{12}^T H_1$ coincides with that of $X_{t,k}$, namely

$$\mathrm{Im}\ \{(U_{12}^T H_1)^T\} = \mathrm{Im}\ (X_{t,k}^T) \tag{21}$$

Proof: It follows from Lemmas 3 and 4 that $\mathrm{Im}\ \{(U_{12}^T H_1)^T\} \subset \mathrm{Im}\ (X_{t,k}^T)$. Moreover, from Appendix, we have rank $(U_{12}^T H_1) = n$. This implies (21). □

There exist n independent rows among $2kp - n$ rows of $U_{12}^T H$, so that any n independent row basis vectors form the state vector. The SVD of U_{12} gives n independent bases of the state space.

Theorem 1 ([12]) Suppose that the SVD of U_{12} is given by (18). Then a state vector is given by

$$X_{t,k} = Q_2^T H_1 \tag{22}$$

Proof: This follows from Lemma 5 and (A7) in Appendix. □

4.2 Determination of System Matrices

We introduce the "colon" notation [3]. Let $A(p{:}q,\ r{:}s)$ be the submatrix of A at the intersection of rows $p, p+1, \cdots, q$ and columns $r, r+1, \cdots, s$. For example,

$$A(3{:}4,\ 2{:}5) = \begin{bmatrix} a_{32} & a_{33} & a_{34} & a_{35} \\ a_{42} & a_{43} & a_{44} & a_{45} \end{bmatrix}$$

Moreover, $A(p{:}q, :)$ and $A(:, r{:}s)$ denote submatrices of A consisting of rows $p, p+1, \cdots, q$ and columns $r, r+1, \cdots, s$, respectively.

Theorem 2 Suppose that SVDs of H and U_{12} are given by (17) and (18), respectively.

Then the system matrices A_δ, B_δ, C, and D are obtained by solving the following overdetermined equation by the LS technique.

$$\begin{bmatrix} Q_2^T U_H(m+p+1:(k+1)(m+p),\ 1:2km+n)S_{11} \\ U_H(k(m+p)+m+1:(k+1)(m+p),\ 1:2km+n)S_{11} \end{bmatrix} = \begin{bmatrix} A_\delta & B_\delta \\ C & D \end{bmatrix} \begin{bmatrix} Q_2^T U_H(1:k(m+p),\ 1:2km+n)S_{11} \\ U_H(k(m+p)+1:k(m+p)+m,\ 1:2km+n)S_{11} \end{bmatrix} \tag{23}$$

Proof: It follows from Theorem 1 that

$$\begin{aligned} X_{t,k} &= Q_2^T H_1 \\ &= Q_2^T H(1:k(m+p),\ :) \\ &= Q_2^T U_H(1:k(m+p),\ :)S_H V_H^T \end{aligned}$$

and

$$\begin{aligned} X_{t,k+1} &= Q_2^T H(m+p+1:(k+1)(m+p),\ :) \\ &= Q_2^T U_H(m+p+1:(k+1)(m+p),\ :)S_H V_H^T \end{aligned}$$

Also, we have

$$\begin{aligned} &[\delta^k u(t)\ \delta^k u(t+\Delta)\ \cdots\ \delta^k u(t+(L-1)\Delta)] \\ &\quad = H(k(m+p)+1:k(m+p)+m,\ :) \\ &\quad = U_H(k(m+p)+1:k(m+p)+m,\ :)S_H V_H^T \end{aligned}$$

and

$$\begin{aligned} &[\delta^k y(t)\ \delta^k y(t+\Delta)\ \cdots\ \delta^k y(t+(L-1)\Delta)] \\ &\quad = H(k(m+p)+m+1:(k+1)(m+p),\ :) \\ &\quad = U_H(k(m+p)+m+1:(k+1)(m+p),\ :)S_H V_H^T \end{aligned}$$

Substituting the above equations into

$$\begin{bmatrix} \delta^{k+1}x(t) & \delta^{k+1}x(t+\Delta) & \cdots & \delta^{k+1}x(t+(L-1)\Delta) \\ \delta^k y(t) & \delta^k y(t+\Delta) & \cdots & \delta^k y(t+(L-1)\Delta) \end{bmatrix} = \begin{bmatrix} A_\delta & B_\delta \\ C & D \end{bmatrix} \begin{bmatrix} \delta^k x(t) & \delta^k x(t+\Delta) & \cdots & \delta^k x(t+(L-1)\Delta) \\ \delta^k u(t) & \delta^k u(t+\Delta) & \cdots & \delta^k u(t+(L-1)\Delta) \end{bmatrix}$$

yields

$$\begin{bmatrix} Q_2^T U_H(m+p+1:(k+1)(m+p),:)S_H V_H^T \\ U_H(k(m+p)+m+1:(k+1)(m+p),:)S_H V_H^T \end{bmatrix} = \begin{bmatrix} A_\delta & B_\delta \\ C & D \end{bmatrix} \begin{bmatrix} Q_2^T U_H(1:k(m+p),:)S_H V_H^T \\ U_H(k(m+p)+1:k(m+p)+m,:)S_H V_H^T \end{bmatrix}$$

Since the orthogonal matrix V_H^T has no effect on the LS estimate, it can be removed. Also, since S_H has zeros except for S_{11}, we have (23). □

It should be noted that a considerable computational saving is achieved in the SVD of (17), since a large orthogonal matrix V_H is not needed in actual computation.

5. Generation of Block Hankel Matrices

5.1 Prefiltering

We need higher order differences of $u(t)$ and $y(t)$ to form the block Hankel matrices H_1 and H_2 of (15) and (16). But since the raw differences are susceptible to noise, we instead use filtered differences.

Define a stable polynomial with order $2k$ by

$$E(\delta)=\delta^{2k}+e_1\delta^{2k-1}+\cdots+e_{2k-1}\delta+e_{2k} \tag{24}$$

where $e_1, \cdots, e_{2k}$ are constants. Also define

$$x^f(t)=\frac{1}{E(\delta)}x(t), \quad u^f(t)=\frac{1}{E(\delta)}u(t), \quad y^f(t)=\frac{1}{E(\delta)}y(t) \tag{25}$$

Pre-multiplying (7) by a stable filter $1/E(\delta)$ gives

$$\begin{aligned} \delta x^f(t) &= A_\delta x^f(t)+B_\delta u^f(t) \\ y^f(t) &= Cx^f(t)+Du^f(t), \quad t=0, \Delta, 2\Delta, \cdots \end{aligned} \tag{26}$$

Thus we can use the filtered differences $\delta^i u^f(t+i\Delta)$, $\delta^i y^f(t+i\Delta)$ in place of the raw differences $\delta^i u(t+i\Delta)$, $\delta^i y(t+i\Delta)$ in H_1 and H_2. The block Hankel matrix thus obtained will be denoted by $H^f=\begin{bmatrix} H_1^f \\ H_2^f \end{bmatrix}$.

In order to form H^f, we define

$$\Phi_{u_i}(t)=\begin{bmatrix} u_i^f(t) \\ \delta u_i^f(t) \\ \vdots \\ \delta^{2k-1}u_i^f(t) \end{bmatrix}, \quad \Phi_{y_j}(t)=\begin{bmatrix} y_j^f(t) \\ \delta y_j^f(t) \\ \vdots \\ \delta^{2k-1}y_j^f(t) \end{bmatrix}; \quad \begin{aligned} i&=1, 2, \cdots, m \\ j&=1, 2, \cdots, p \end{aligned}$$

It follows from (24) and (25) that

$$\delta^{2k}u^f(t) = -e_1\delta^{2k-1}u^f(t) - \cdots - e_{2k-1}\delta u^f(t) - e_{2k}u^f(t) + u(t) \tag{27}$$

$$\delta^{2k}y^f(t) = -e_1\delta^{2k-1}y^f(t) - \cdots - e_{2k-1}\delta y^f(t) - e_{2k}y^f(t) + y(t) \tag{28}$$

These equations are respectively expressed as

$$\delta\Phi_{u_i^f}(t) = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & & \\ & & & 0 & 1 \\ -e_{2k} & -e_{2k-1} & \cdots & & -e_1 \end{bmatrix} \Phi_{u_i^f}(t) + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u_i(t),\ i=1, 2, \cdots, m$$

$$\delta\Phi_{y_j^f}(t) = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & & \\ & & & 0 & 1 \\ -e_{2k} & -e_{2k-1} & \cdots & & -e_1 \end{bmatrix} \Phi_{y_j^f}(t) + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} y_j(t),\ j=1, 2, \cdots, p$$

Solving the above equations with initial conditions $\Phi_{u_i^f}(0)=0$, $\Phi_{y_i^f}(0)=0$, we get filtered differences to form H^f.

5.2 Least-Squares Estimate of H^f

If the input-output data $u(t)$ and $y(t)$ are disturbed by noises, then H^f is corrupted by noise. In order to reduce the effect of noise, we consider the LS estimate of H^f under the assumption that H^f is perturbed by noise, namely

$$H^f = H + N \tag{29}$$

We assume that the unperturbed H has the SVD of (17), namely

$$H = [U_1\ U_2] \begin{bmatrix} S_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \tag{30}$$

where

U_1: $2k(m+p)\times(2km+n)$; V_1: $L\times(2km+n)$

U_2: $2k(m+p)\times(2kp-n)$; V_2: $L\times(L-2km-n)$

S_{11}: $(2km+n)\times(2km+n)$

In the following, we assume that

(A1) $NN^T = \sigma^2 I$

(A2) $HN^T = 0$

Assumption (A1) implies that N is an orthogonal matrix where the norm of each row vector is σ. Also, (A2) shows that the row spaces of N and H are orthogonal. It should be noted that (A1) is not very realistic, since if the input-output data $u(t)$ and $y(t)$ are disturbed by white noise, then elements of H^f are perturbed by colored noise whose statistical characteristics are determined by the prefilterings. To cope with the colored noise, we can apply the techniques of [12], [21].

Since $[U_1\ U_2]$ is orthogonal,

$$\begin{aligned} H^f &= H + N \\ &= U_1 S_{11} V_1^T + (U_1 U_1^T + U_2 U_2^T) N \\ &= [U_1\ U_2] \begin{bmatrix} (S_{11} + \sigma^2 I)^{\frac{1}{2}} & 0 \\ 0 & \sigma I \end{bmatrix} \begin{bmatrix} (S_{11} + \sigma^2 I)^{-\frac{1}{2}} (S_{11} V_1^T + U_1^T N) \\ \sigma^{-1} U_2^T N \end{bmatrix} \end{aligned} \tag{31}$$

We see that this is the SVD of H^f, so that we may write

$$H^f = [U_{f1}\ U_{f2}] \begin{bmatrix} S_{f1} & 0 \\ 0 & S_{f2} \end{bmatrix} \begin{bmatrix} V_{f1}^T \\ V_{f2}^T \end{bmatrix} \tag{32}$$

where

$$S_{f1} = \sqrt{S_{11} + \sigma^2 I_{2km+n}} \tag{33}$$

$$S_{f2} = \sigma I_{2kp-n} \tag{34}$$

Lemma 6 Let the singular values of H^f be $\mu_1, \cdots, \mu_{2k(m+p)}$. Suppose that $\operatorname{rank} H = 2km + n$. Then the estimate of σ^2 is given by

$$\hat{\sigma}^2 = \frac{\mu_{2km+n+1}^2 + \cdots + \mu_{2k(m+p)}^2}{2kp - n} \tag{35}$$

Proof: A proof is immediate from (34). □

It follows from (33) that the estimate of S_{11} is given by

$$\hat{S}_{11} = \sqrt{S_{f1}^2 - \hat{\sigma}^2 I_{2km+n}} \tag{36}$$

We now wish to derive the LS estimate of H based on H^f. Since the LS estimate $\hat{H}$ is given by the orthogonal projection of H on the row space of H^f, this problem is equivalent to finding X such that

$$\min_{X \in R^{q \times q}} \|XH^f - H\|_F^2, \quad q = 2k(m+p)$$

where $\|\cdot\|_F^2$ denotes the Frobenius norm. Note that X is optimal if and only if $XH^f - H$ is orthogonal to H^f, namely $(XH^f - H)(H^f)^T = 0$. Thus we get $X = HH^{fT}[H^f(H^f)^T]^{-1}$.

Lemma 7 The LS estimate of H is given by

$$\hat{H} = U_{f1}\,[\hat{S}_{11}^2(\hat{S}_{11}^2 + \hat{\sigma}^2 I)^{-\frac{1}{2}}]\,V_{f1}^T \tag{37}$$

Proof: It can be shown that

$$\begin{aligned}
\hat{H} &= XH^f \\
&= H(H^f)^T[H^f(H^f)^T]^{-1}H^f \\
&= H[V_{f1}\ V_{f2}] \begin{bmatrix} V_{f1}^T \\ V_{f2}^T \end{bmatrix} \\
&= [U_1\ U_2] \begin{bmatrix} S_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} [(V_1 S_{11} + N^T U_1)(S_{11}^2 + \sigma^2 I)^{-\frac{1}{2}}\ \sigma^{-1} N^T U_2] \begin{bmatrix} V_{f1}^T \\ V_{f2}^T \end{bmatrix} \\
&= [U_{f1}\ U_{f2}] \begin{bmatrix} S_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} S_{11}(S_{11}^2 + \sigma^2 I)^{-\frac{1}{2}} & 0 \\ * & * \end{bmatrix} \begin{bmatrix} V_{f1}^T \\ V_{f2}^T \end{bmatrix}
\end{aligned}$$

By replacing S_{11} by $\hat{S}_{11}$ of (36), we get (37). □

5.3 Identification Algorithm

The identification algorithm is summarized as follows.

Step 0: Set k, and $E(\delta)$.

Step 1: For given input-output data, generate H^f, and compute SVD of (32).

Step 2: Compute $\hat{H}$ from (37), and put $H := \hat{H}$.

Step 3: Compute SVD of H and U_{12}.

Step 4: Solve the overdetermined equation (23) to get A_δ, B_δ, C, D.

6. Example

We consider the system of output error type shown in Fig. 1, where

$$G(s) = \frac{10s + 5}{s^3 + 6s^2 + 21s + 26} \tag{38}$$

The above system is simulated over 15 seconds, where $e(t)$ is the white Gaussian noise with mean zero and variance σ_0^2 and the input $u(t)$ is a composite sine wave

$$u(t)=10\cos 2t+4\sin \pi t+6\cos 1.7t$$

which is used in [20], where the equation error model is employed for numerical examples. Figs. 2 and 3 display the input $u(t)$ and the output $y(t)$ for $\sigma_0^2=(0.1)^2$, respectively.

We assume that $k=3$ and let the filter be given by

$$\frac{1}{E(\delta)}=\frac{1}{(\delta+3)(\delta+5)(\delta^2+2\delta+2)(\delta^2+4\delta+13)}$$

Thus we have a block Henkel matrix of $12\times L$, where L is related to the data length used for identification. Tables 1 and 2 show the identification results for the sampling intervals $\Delta=0.01$ and 0.005, respectively, where

$$d=\sqrt{\sum_{i=1}^{5}\left(\frac{\theta_i-\hat{\theta}_i}{\theta_i}\right)^2}$$

In each case, 20 realizations are generated over 15 seconds. From the estimates obtained in each of 20 independent realizations, the sample mean and standard deviation (s.d.) are evaluated. For the noise variance $\sigma_0^2=(0.05)^2$ and $(0.1)^2$, we see that the parameters so obtained show good agreement with the true parameters. We also observed, although have not presented here, that if the noise variance is getting larger, the identification results are quite unsatisfactory.

7. Conclusions

This paper has developed a subspace identification algorithm for a δ-operator model. As Δ tends to zero, the δ-operator model converges to a continuous-time model, so the present technique can be applied to the identification of a continuous-time model. We show by simulation studies that if the output N/S ratio is low, then the estimated parameters

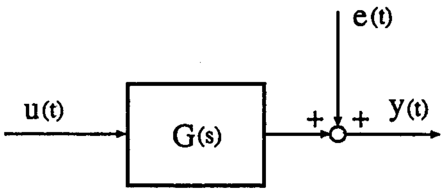

Fig. 1. Continuous-time output-error model

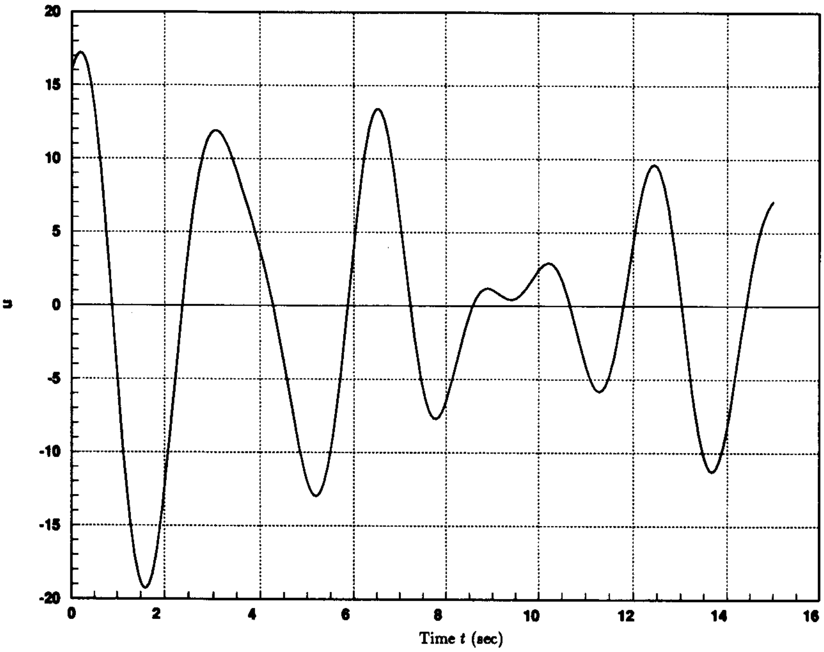

Fig. 2. Input $u(t)$

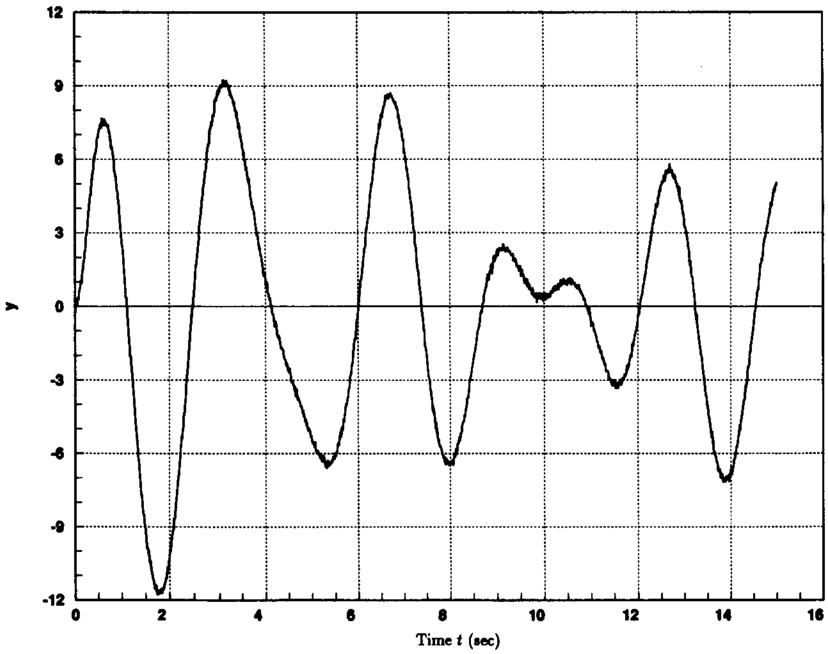

Fig. 3. Output $y(t)$ ($\sigma_0^2=0.01$)

Table 1. Identification results for $\Delta=0.01$

		$\sigma_0^2=(0.05)^2$		$\sigma_0^2=(0.1)^2$	
	true	mean	s.d.	mean	s.d.
a_1	6	6.1006	0.1186	6.3067	0.3180
a_2	21	20.9521	0.5845	21.5290	1.4872
a_3	26	25.8184	1.0489	27.4145	2.6457
b_1	10	9.8501	0.3539	10.1494	0.8863
b_2	5	5.0422	0.2104	5.5250	0.5349
NSR (%)		0.0104		0.0336	
d		0.0251		0.1321	

Table 2. Identification results for $\Delta=0.005$

		$\sigma_0^2=(0.05)^2$		$\sigma_0^2=(0.1)^2$	
	true	mean	s.d.	mean	s.d.
a_1	6	6.0483	0.1146	6.1140	0.2261
a_2	21	20.9930	0.5361	21.2523	1.1520
a_3	26	25.8951	0.9802	26.4961	2.0495
b_1	10	9.9386	0.3239	10.0876	0.7043
b_2	5	5.0051	0.1960	5.1701	0.4095
NSR (%)		0.0069		0.0277	
d		0.0109		0.0459	

s.d.: = standard deviation
NSR: $=\mathrm{var}\{e(t)\}/\mathrm{var}\{y(t)\}$

show good agreement with the true parameters. For noisy cases, the algorithm remains to be improved, e.g., based on the canonical correlation approach.

References

[1] B. L. Ho and R. E. Kalman: Effective Construction of Linear State-Variable Models from Input/Output Functions. *Regelungstechnik*, vol. 14, no. 12, pp. 545–548, 1966.

[2] S. Y. Kung: A New Identification and Model Reduction Algorithm via Singular Value Decomposition. *Proc. 12th Asilomar Conf. on Circuits, Syst. & Computers*, Pacific Grove, 1978, pp. 715–714.

[3] G. H. Golub and C. R. Van Loan: *Matrix Computations* (2nd ed.). The Johns Hopkins Univ. Press, 1989.

[4] H. Akaike: Canonical Correlation Analysis of Time Series and the Use of an Information Criterion. *System Identification: Advances and Case Studies* (R. Mehra and D. Lainiotis, Eds.), Academic, 1976, pp. 27–96.

[5] H. Akaike: Markovian Representation of Stochastic Processes by Canonical Variables. *SIAM J. Control*, vol. 13, no. 1, pp. 162–173, 1975.

[6] W. E. Larimore: System Identification, Reduced-order Filtering and Modeling via Canonical

Variate Analysis. *Proc. ACC*, San Francisco, 1983, pp. 445–451.
[7] W. E. Larimore: Canonical Variate Analysis in Identification, Filtering, and Adaptive Control. *Proc. 29th CDC*, Honolulu, 1990, pp. 596–604.
[8] B. De Moor, M. Moonen, L. Vandenberghe and J. Vandewalle: Identification of Linear State-Space Models with Singular Value Decomposition using Canonical Correlation Concepts. *SVD and Signal Processing* (Ed. E. F. Deprettere), Elsevier, 1988, pp. 161–169.
[9] J. Vandewalle and B. De Moor: On the Use of the Singular Value Decomposition in Identification and Signal Processing. *Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms* (Eds. G. H. Golub and P. Van Deeren), Springer, 1991, pp. 321–360.
[10] A. J. van der Veen, E. F. Deprettere and A. L. Swindlehurst: Subspace-Based Signal Analysis Using Singular Value Decomposition. *Proc. IEEE*, vol. 81, no. 9, pp. 1277–1308, 1993.
[11] M. Moonen, B. De Moor and J. Vandewalle: On- and Off-Line Identification of Linear State-Space Models. *Int. J. Control*, vol. 49, no. 1, pp. 219–232, 1989.
[12] M. Moonen and J. Vandewalle: QSVD Approach to On- and Off-Line State-Space Identification. *Int. J. Control*, vol. 51, no. 5, pp. 1133–1146, 1990.
[13] M. Verhaegen and P. Dewilde: Subspace Model Identification, Part 1. The Output-Error State-Space Model Identification Class of Algorithms; Part 2. Analysis of the Elementary Output-Error State-Space Model Identification Algorithm. *Int. J. Control*, vol. 56, no. 5, pp. 1187–1210 & pp. 1211–1241, 1992.
[14] M. Viberg, B. Ottersten, B. Wahlberg and L. Ljung: A Statistical Perspective on State-Space Modeling using Subspace Methods. *Proc. 30th CDC*, London, 1991, pp. 1337–1342.
[15] L. Ljung: *System Identification—Theory for the User*. Prentice-Hall, 1987.
[16] T. Söderström and P. Stoica: *System Identification*. Prentice-Hall, 1989.
[17] H. Unbehauen and G. P. Rao: Continuous-Time Approaches to System Identification—A Survey. *Automatica*, vol. 26, no. 1, pp. 23–35, 1990.
[18] N. K. Sinha and G. P. Rao (eds.): *Identification of Continuous-Time Systems*. Kluwer Academic Pub., 1991.
[19] R. H. Middleton and G. C. Goodwin: *Digital Control and Estimation—A Unified Approach*. Prentice-Hall, 1990.
[20] D. L. Stericker and N. K. Sinha: Identification of Continuous-Time Systems from Samples of Input-Output Data Using the δ-Operator. *Control-Theory and Advanced Technology*, vol. 9, no. 1, pp. 113–125, 1993.
[21] M. Moonen, B. De Moor and J. Vandewalle: SVD-Based Subspace Methods for Multivariable Continuous-Time Systems Identification. *Identification of Continuous-Time Systems* (Eds. N. K. Sinha and G. P. Rao), Kluwer, 1991, pp. 473–488.
[22] C. C. Paige and M. A. Saunders: Towards a Generalized Singular Value Decomposition. *SIAM J. Numerical Anal.*, vol. 18, no. 3, pp. 398–405, 1981.
[23] B. De Moor: The Singular Value Decomposition and Long and Short Spaces of Noisy Matrices. *IEEE Trans. Signal Processing*, vol. 41, no. 9, pp. 2826–2838, 1993.

Appendix: Proof of Lemma 3

For simplicity, define $\kappa := k(m+p)$, $\pi := 2kp-n$, $\mu := 2km+n$ $(2\kappa = \pi + \mu)$. Let $U' = \begin{bmatrix} U_{12} & U_{11} \\ U_{22} & U_{21} \end{bmatrix}$, where U' is a $2\kappa \times (\pi+\mu)$ orthogonal matrix. It follows from [22] that there exist four orthogonal matrices Q, $V \in R^{\kappa\times\kappa}$, $W \in R^{\pi\times\pi}$, $Z \in R^{\mu\times\mu}$ such that

$$\begin{bmatrix} Q^T & 0 \\ 0 & V^T \end{bmatrix} \begin{bmatrix} U_{12} & U_{11} \\ U_{22} & U_{21} \end{bmatrix} \begin{bmatrix} W & 0 \\ 0 & Z \end{bmatrix}$$

$$= \left[\begin{array}{ccc|ccc} I_r & & & 0_S^T & & \\ & C_s & & & S_s & \\ & & 0_C & & & I_{\kappa-s-r} \\ \hline 0_S & & & I_{\kappa-\pi+r} & & \\ & S_s & & & -C_s & \\ & & I_{\pi-s-r} & & & 0_C^T \end{array}\right] = \begin{bmatrix} \Sigma_{12} & \Sigma_{11} \\ \Sigma_{22} & \Sigma_{21} \end{bmatrix} \tag{A1}$$

$C_s = \mathrm{diag}\,(\alpha_{r+1}, \cdots, \alpha_{r+s}),\quad 1 > \alpha_{r+1} \geq \cdots \geq \alpha_{r+s} > 0$

$S_s = \mathrm{diag}\,(\beta_{r+1}, \cdots, \beta_{r+s}),\quad 0 < \beta_{r+1} \leq \cdots \leq \beta_{r+s} < 1$

$C_s^2 + S_s^2 = I_s$

$\Sigma_{12},\ \Sigma_{22} \in R^{\kappa\times\pi};\ \Sigma_{11},\ \Sigma_{21} \in R^{\kappa\times\mu}$

$0_C,\ 0_S$: $(\kappa-s-r)\times(\pi-s-r)$ and $(\kappa-\pi+r)\times r$ zero matrices

where s, r are to be determined. From (A1),

$$\begin{bmatrix} U_{12} & U_{11} \\ U_{22} & U_{21} \end{bmatrix} = \begin{bmatrix} Q\Sigma_{12}W^T & Q\Sigma_{11}Z^T \\ V\Sigma_{22}W^T & V\Sigma_{21}Z^T \end{bmatrix} \tag{A2}$$

Since Q, W are orthogonal, the (1,1) block of (A2) gives the SVD of U_{12}. In the following, we prove $s=n$, $r=kp-n$, showing that $Q\Sigma_{12}W^T$ gives the SVD of (18).

(a) Partition $Q=[Q_1\ Q_2\ Q_3]$, where $Q_1 \in R^{\kappa\times r}$, $Q_2 \in R^{\kappa\times s}$, $Q_3 \in R^{\kappa\times(\kappa-s-r)}$. We show that Q_1 is orthogonal to H_1. It follows from (20) and (A2) that

$$\begin{aligned} H_1 &= [U_{11}S_{11}\ \ 0]\ V_H^T \\ &= [Q\Sigma_{11}Z^TS_{11}\ \ 0]\ V_H^T \\ &= \left([Q_1\ Q_2\ Q_3] \begin{bmatrix} 0_S^T & & \\ & S_s & \\ & & I_{\kappa-s-r} \end{bmatrix} Z^TS_{11}\ \ 0\right) V_H^T \end{aligned} \tag{A3}$$

Since $Q_i^TQ_j=0$, $i\neq j$, we get

$$Q_1^TH_1 = ([0\ \ Q_1^TQ_2S_s\ \ Q_1^TQ_3]Z^TS_{11}\ \ 0)\ V_H^T = 0 \tag{A4}$$

This shows that Q_1 is orthogonal to H_1.

(b) We show that $n \geq s$. From (A2), (A3),

$$U_{12}^TH_1 = W\Sigma_{12}^TQ^TH_1 = W\begin{bmatrix} I_r & & \\ & C_s & \\ & & 0_C^T \end{bmatrix}\begin{bmatrix} Q_1^T \\ Q_2^T \\ Q_3^T \end{bmatrix}H_1 = W\begin{bmatrix} 0 \\ C_s \\ 0 \end{bmatrix}Q_2^TH_1 \tag{A5}$$

Also from, (A3)

$$Q_2^T H_1 = ([0 \quad S_s \quad 0]\, Z^T S_{11} \quad 0)\, V_H^T = S_s Z_2^T S_{11} V_{H1}^T \tag{A6}$$

where $Z=[Z_1\ Z_2\ Z_3]$, $Z_1 \in R^{\mu\times(\kappa-\pi+r)}$, $Z_2 \in R^{\mu\times s}$, $Z_3 \in R^{\mu\times(\kappa-s-r)}$, $S_{11}\in R^{\mu\times\mu}$, $V_{H1}^T \in R^{\mu\times L}$. Since each matrix on the right-hand side of (A6) has full rank, we get $\mathrm{rank}(Q_2^T H_1)=s$. It also follows from (A5) that $\mathrm{Im}(U_{12}^T H_1)^T \subset \mathrm{Im}\,(Q_2^T H_1)^T$. Partition $W=[W_1\ W_2\ W_3]$, where $W_1\in R^{\pi\times r}$, $W_2 \in R^{\pi\times s}$, $W_3\in R^{\pi\times(\kappa-s-r)}$. Then, from (A5), we get $C_s^{-1} W_2^T U_{12}^T H_1 = Q_2^T H_1$. This implies that $\mathrm{Im}\,(Q_2^T H_1)^T \subset \mathrm{Im}\,(U_{12}^T H_1)^T$. Thus we have

$$\mathrm{Im}\,\{(U_{12}^T H_1)^T\} = \mathrm{Im}\,\{(Q_2^T H_1)^T\} \tag{A7}$$

Hence, $\mathrm{rank}\,(U_{12}^T H_1) = \mathrm{rank}\,(Q_2^T H_1) = s$ holds. Since, from Lemma 4, $n \geq \mathrm{rank}\,(U_{12}^T H_1)$, it follows that $n\geq s$.

(c) We prove $s\geq n$. Note that H_1 is $\kappa\times L$ and $\mathrm{rank} H_1 = km+n$. Thus we see from (A4) that $\mathrm{rank} Q_1$ is smaller than the rank deficiency of H_1, so that

$$r \leq \kappa-(km+n) = kp-n$$

Similarly to the derivation of (A3), it follows that

$$\begin{aligned} H_2 &= [U_{21} S_{11} \quad 0]\, V_H^T \\ &= [V \Sigma_{21} Z^T S_{11} \quad 0]\, V_H^T \\ &= \left([V_1 \quad V_2 \quad V_3] \begin{bmatrix} I_{\kappa-\pi+r} & & \\ & -C_s & \\ & & 0_C^T \end{bmatrix} Z^T S_{11} \quad 0 \right) V_H^T \end{aligned}$$

where $V_1 \in R^{\kappa\times(\kappa-\pi+r)}$, $V_2\in R^{\kappa\times s}$, $V_3 \in R^{\kappa\times(\pi-s-r)}$. Since $V_i^T V_j = 0$, $i\neq j$, we get

$$V_3^T H_2 = 0 \tag{A8}$$

Since H_2 is a $\kappa\times L$ matrix and $\mathrm{rank} H_2 = km+n$, its rank deficiency is $kp-n$. It therefore follows from (A8) that $\mathrm{rank}\, V_3 = \pi-s-r \leq kp-n$, so that

$$r \geq kp-s \tag{A9}$$

Hence, we see from (A8) and (A9) that $kp-s\leq r\leq kp-n$, or $s\geq n$.

From the above, we see that $n=s$ and $r=kp-n$ hold.