Automatica 66 (2016) 3–14

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Blind multivariable ARMA subspace identification[☆]

Chengpu Yu, Michel Verhaegen

Delft Center for Systems and Control, Delft University of Technology, Delft 2628CD, Netherlands

ARTICLE INFO

Article history:
Received 12 June 2014
Received in revised form
7 September 2015
Accepted 22 November 2015
Available online 14 January 2016

Keywords:
State-space representation
Blind identification
Cross-relation equation

ABSTRACT

In this paper, we study the deterministic blind identification of multiple channel state-space models having a common unknown input using measured output signals that are perturbed by additive white noise sequences. Different from traditional blind identification problems, the considered system is an autoregressive system rather than an FIR system; hence, the concerned identification problem is more challenging but possibly having a wider scope of application. Two blind identification methods are presented for multi-channel autoregressive systems. A cross-relation identification method is developed by exploiting the mutual references among different channels. It requires at least three channel systems with square and stably invertible transfer matrices. Moreover, a general subspace identification method is developed for which two channel systems are sufficient for the blind identification; however, it requires the additive noises to have identical variances and the transfer matrices having no transmission zeros. Finally, numerical simulations are carried out to demonstrate the performance of the proposed identification algorithms.

1. Introduction

Blind system identification is to estimate transfer functions using only output observations with some a priori knowledge of the system model and system input (Giannakis, Hua, Stoica, & Tong, 2000; Tong & Perreau, 1998; Yu, Zhang, & Xie, 2013a,b). Since only system outputs are required for the identification problems, it has a broad range of potential applications. For instance, the identification of networked systems with unknown disturbed signals (Li, 2005; Yu, Xie, & Soh, 2014) and the blind deblurring for the biomedical or optical imaging (Chen et al., 2013; Yu, Zhang, & Xie, 2012a).

This paper deals with the blind identification of multivariate or multi-input–multi-output (MIMO) autoregressive systems for which the system inputs are deterministic but unknown. Conventional blind MIMO system identification requires to identify the transfer matrix and separate multiple sources. Here, we only concern the transfer matrix estimation part, so there may exist a matrix ambiguity for the identification result (Abed-Meraim, Loubaton, & Moulines, 1997; Huang, Benesty, & Chen, 2006). To date, most of the existing deterministic blind MIMO system identification studies are based on FIR settings (Abed-Meraim et al., 1997; Huang et al., 2006; Moulines, Duhamel, Cardoso, & Mayrargue, 1995), and limited work has been done on ARMA systems (Routtenberg & Tabrikian, 2010; Zhang & Cichocki, 2000). In the present paper, we shall investigate the blind identification of state-space represented MIMO systems.

The blind deconvolution of dynamical systems using state-space approaches was reviewed in Zhang and Cichocki (2000), where the involved cost functions, such as maximum entropy, minimum mutual information and maximum high-order cumulant, are non-convex. Due to the non-convex property, the gradient-descent type of optimization algorithms may get stuck in local optima. In this paper, we use ideas from subspace identification (Ljung, 1999; Verhaegen & Verdult, 2007) to avoid the above mentioned non-convex optimization problems, since the subspace methods do not need to parameterize the model but using reliable linear algebraic calculations such as QR and SVD decompositions.

In this paper, we present two blind identification methods for multi-channel systems in state-space forms. A cross-relation identification method is developed for systems with square and stably invertible transfer matrices. The blind identification problem for this type of systems can be recast into a classic errors-in-variables identification problem, which can then be solved using classic subspace identification methods (Ljung, 1999; Verhaegen & Verdult, 2007). One advantage of the cross-relation

☆ The work of Dr. C. Yu and Prof. M. Verhaegen is sponsored by the European Research Council, Advanced Grant Agreement No. 339681. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Brett Ninness under the direction of Editor Torsten Söderström.

E-mail addresses: c.yu-4@tudelft.nl (C. Yu), M.Verhaegen@tudelft.nl (M. Verhaegen).

http://dx.doi.org/10.1016/j.automatica.2015.12.005

method is that the involved measurement noises can be spatially correlated and may not have identical variances. Moreover, a subspace-based identification algorithm is developed for the multi-channel systems with tall (possibly not stably invertible) transfer matrices by exploiting zero and pole diversities of the multiple channels. For these two blind identification methods, their associated blind identifiability conditions are provided.

The numerator and denominator polynomial matrices of an ARMA system are usually coupled together, so the associated identification problem is challenging. To deal with this identification problem, prior knowledge of the system input has been adopted in traditional identification methods, such as piecewise-smooth input (Bai, Li, & Dasgupta, 2002; Yu, Xie, & Zhang, 2014), white-noise input (Vanbeylen, Pintelon, & Schoukens, 2008; Yu, Zhang, & Xie, 2012b; Yu et al., 2013a), periodically modulated input (Giannakis & Serpedin, 1998) and finite-alphabetic input (Routtenberg & Tabrikian, 2010). In the proposed subspace identification methods, we do not use any prior knowledge of the system input except its persistent excitation property. The proposed identification methods are derived based on the fact that: a square transfer matrix generically possesses transmission zeros, while an augmented transfer matrix constructed by stacking two square transfer matrices generically possesses no transmission zeros.

The rest of the paper is organized as follows. Section 2 describes the multivariate blind system identification problem. Section 3 gives some preliminaries on the identifiability of two-channel autoregressive systems. Section 4 provides a method for blindly estimating the characteristic polynomials of multiple channel systems. This characteristic-polynomial estimation method will be used in the two blind identification algorithms that are developed in Section 5. Section 6 shows two simulation examples, followed by the conclusions in Section 7.

The following notation is adopted throughout the paper. $\mathcal{E}(\cdot)$ denotes the mathematical expectation. $\delta(\cdot)$ stands for the Dirac delta function. $H(q)$ represents the transfer function of a system with impulse response $h(k)$ in time domain, and q^{-1} is the backward shift operator. The upper case letter A denotes a matrix, and $\mathrm{vec}(A)$ represents the vectorization of A. The superscripts ^{T} and $^{-1}$ stand for the matrix transpose and inverse, respectively. $\|A\|_F$ denotes the Frobenius norm of A. $\det(A)$ and $\mathrm{adj}(A)$ represent the determinant and adjoint matrices of A, respectively. I is the identity matrix of appropriate dimension.

2. Problem formulation

We consider the multiple channel systems in state-space forms as follows:

$$\begin{aligned} x_i(k+1) &= A_i x_i(k) + B_i s(k) \\ y_i(k) &= C_i x_i(k) + D_i s(k) + w_i(k), \quad i = 1, \ldots, L, \end{aligned} \tag{1}$$

where $s(k) \in \mathbb{R}^m$ is a common source signal, $w_i(k) \in \mathbb{R}^p$, $x_i(k) \in \mathbb{R}^n$ and $y_i(k) \in \mathbb{R}^p$ are respectively the measurement noise, system state and output of the i-th channel system, and A_i, B_i, C_i, D_i are real system matrices of appropriate dimensions.

In stating the assumptions that will be used in solving the blind identification problem in this paper, use will be made of the following definition.

Definition 1. The input sequence $s(k)$ is persistently exciting of order n_s if and only if for any positive integer k, there exists an integer N such that the matrix

$$\begin{bmatrix} s(k+1) & s(k+2) & \cdots & s(k+N) \\ s(k+2) & s(k+3) & \cdots & s(k+N+1) \\ \vdots & \vdots & \iddots & \vdots \\ s(k+n_s) & s(k+n_s+1) & \cdots & s(k+N+n_s-1) \end{bmatrix}$$

has full row rank.

For the systems in (1), the following standard assumptions are made.

A1. The system input $s(k)$ is persistently exciting of any finite order.

A2. The matrix pair (A_i, B_i) is controllable and (C_i, A_i) is observable for $i \in \{1, \ldots, L\}$.

A3. A_i is stable and D_i has full column rank for $i \in \{1, \ldots, L\}$.

A4. The additive white noise $w_i(k)$ is independent of $s(k)$ and $x(0)$, and satisfies that

$$\mathcal{E}(w_i(k) w_j^T(k-\tau)) = \sigma^2 \delta(i-j)\delta(\tau) \cdot I,$$

where $i, j \in \{1, \ldots, L\}$.

Assumption A1 assures that the concerned systems can be fully excited. Assumption A2 indicates that the concerned systems in state-space forms are minimal realizations. It is assumed in Assumption A3 that all eigenvalues of A_i have amplitudes less than one, indicating that the effect of the current state on the future outputs decays along with the increase of time. In addition, the dimension of the system output is larger than or equal to that of the system input, namely $p \geq m$.

The *problem of interest* is to blindly identify the system matrices $\{A_i, B_i, C_i, D_i\}$ up to a matrix ambiguity based on only the system observations $\{y_i(k)\}_{i=1}^L$. It is noted that the "up to a matrix ambiguity" is different from "up to a similarity transformation" which is commonly used in state-space system identification. Suppose that the true transfer matrix of the i-th channel is $H_i(q) = C_i(qI - A_i)^{-1}B_i + D_i$. Determination up to a matrix ambiguity means that the estimated transfer matrix has the form $\hat{H}_i(q) = H_i(q)\Gamma$ with Γ a non-singular matrix, while determination up to a similarity transformation indicates that the estimated transfer matrix still equals the true transfer matrix.

Generally, there are many matrix fraction description (MFD) forms for each single-channel system. If $H_i(q) = N_i(q)R_i^{-1}(q)$ with $\deg[\det(R_i(q))] = n$ is one MFD form of the i-th transfer matrix, then $H_i(q) = (N_i(q)C(q))\left(C^{-1}(q)R_i^{-1}(q)\right)$ for any unimodular matrix $C(q)$ is another MFD form with $\deg[\det(R_i(q)C(q))] = n$. Since $C(q)$ can be any unimodular matrix, the numerator and denominator polynomial matrices of $H_i(q)$ cannot be determined up to a constant matrix ambiguity. To cope with this problem, in this paper, the i-th transfer matrix is represented in the form $H_i(q) = \frac{Q_i(q)}{p_i(q)}$, where $p_i(q) = \det(qI - A_i)$ is a monic characteristic polynomial and $Q_i(q)$ is a polynomial matrix having the same size as $H_i(q)$.

3. Preliminaries on identifiability of two channel systems

The two-channel system can be considered as a basic element of the multi-channel system, so its identifiability will be investigated in this section.

Definition 2. Consider the following two-channel system without noise effect:

$$\begin{aligned} y_1(k) &= H_1(q)s(k) \\ y_2(k) &= H_2(q)s(k), \end{aligned} \tag{2}$$

where $H_i(q) = C_i(qI - A_i)^{-1}B_i + D_i$ for $i = 1, 2$. Given the system outputs $y_1(k)$ and $y_2(k)$, the above two-channel system is *blindly identifiable* if any solution $\{\hat{H}_1(q), \hat{H}_2(q), \hat{s}(k)\}$ to Eq. (2), with $\hat{H}_1(q)$ and $\hat{H}_2(q)$ having minimal realizations of order n, satisfies

$$\hat{H}_1(q) = H_1(q)\Gamma$$
$$\hat{H}_2(q) = H_2(q)\Gamma$$
$$\hat{s}(k) = \Gamma^{-1}s(k)$$

with $\Gamma \in \mathbb{R}^{m\times m}$ being a non-singular ambiguity matrix.

Denote

$$H_i(q) = C_i(qI - A_i)^{-1}B_i + D_i = \frac{Q_i(q)}{p_i(q)} \quad \text{for } i = 1, 2.$$

The two-channel system without additive noise effect can also be written as

$$\begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix} = \begin{bmatrix} \dfrac{Q_1(q)}{p_1(q)} \\ \dfrac{Q_2(q)}{p_2(q)} \end{bmatrix} s(k) = \underbrace{\begin{bmatrix} Q_1(q)p_2(q) \\ Q_2(q)p_1(q) \end{bmatrix}}_{G(q)} \underbrace{\frac{s(k)}{p_1(q)p_2(q)}}_{u(k)}. \tag{3}$$

Let $G_1(q) = Q_1(q)p_2(q)$ and $G_2(q) = Q_2(q)p_1(q)$. The signal $u(k) = \frac{s(k)}{p_1(q)p_2(q)}$ is a pseudo common source signal. As shown in Eq. (3), the two autoregressive systems can be transformed to two FIR systems. To ensure the blind identifiability of $G(q)$ in (3), it is necessary that the polynomial matrix $G(q)$ is irreducible (Giannakis et al., 2000, Chapter 3), i.e. the matrix $G(q)$ has full column rank for any $q \in \mathbb{C}$. The next theorem shows that this is not the case for $G(q)$ in (3).

Theorem 1. *Suppose that A_1 and A_2 are cyclic matrices (Gopal, 1993). Let the dimension parameters $m, p \geq 2$. Then, the polynomial matrix $G(q)$ defined in (3) is reducible even if Assumptions A1–A3 hold and $\{p_1(q), p_2(q)\}$ are coprime.*

Proof. By the following relations for any $i \in \{1, 2\}$:

$$(qI - A_i)^{-1} = \frac{\mathrm{adj}(qI - A_i)}{\det(qI - A_i)}$$

$$p_i(q) = \det(qI - A_i),$$

$Q_i(q)$ can be represented as

$$Q_i(q) = C_i \mathrm{adj}(qI - A_i)B_i + D_i p_i(q). \tag{4}$$

The polynomial matrix $G(q)$ can be written as

$$G(q) = \begin{bmatrix} C_1\mathrm{adj}(qI - A_1)B_1p_2(q) + D_1p_1(q)p_2(q) \\ C_2\mathrm{adj}(qI - A_2)B_2p_1(q) + D_2p_2(q)p_1(q) \end{bmatrix}. \tag{5}$$

Let z_0 be a zero of $p_1(q)$, namely $p_1(z_0) = 0$. It follows that

$$G(z_0) = \begin{bmatrix} C_1\mathrm{adj}(z_0I - A_1)B_1p_2(z_0) \\ 0 \end{bmatrix}.$$

In addition, it can be established that

$$[\mathrm{adj}(z_0I - A_1)]\,(z_0I - A_1) = \det(z_0I - A_1)I = 0. \tag{6}$$

Since A_1 is a cyclic matrix, $z_0I - A_1$ is rank deficient by one (Gopal, 1993). It can then be derived that $\mathrm{adj}(z_0I - A_1)$ has rank less than or equal to one, so does the matrix $C_1\mathrm{adj}(z_0I - A_1)B_1p_2(z_0)$. When the dimension parameters $m, p \geq 2$, it is obvious that $G(z_0)$ is rank deficient, namely $G(q)$ is reducible.

Theorem 1 shows that the FIR transfer matrix $G(q)$ in (3) cannot be identified up to a matrix ambiguity. By taking an insight in the structure of $G(q)$, we can find that the system poles of $H_1(q)$ and $H_2(q)$ are exactly the latent roots (Kailath, 1980) of $G(q)$. Since the unavailability of system poles causes the two FIR systems in (3) to be unidentifiable, we shall investigate the identification of the characteristic polynomials $p_i(q)$ in Section 4.

Before proceeding to the characteristic-polynomial identification, we would like to investigate the persistent excitation properties of the system outputs $\{y_i(k)\}_{i=1}^2$ in (3), since they are essential for analyzing the system identifiability in the sequel.

The matrix form of (3) can be written as

$$\underbrace{\begin{bmatrix} Y^1_{2n+1,r,N} \\ Y^2_{2n+1,r,N} \end{bmatrix}}_{Y_{2n+1,r,N}} = \underbrace{\begin{bmatrix} \mathcal{G}_r^1 \\ \mathcal{G}_r^2 \end{bmatrix}}_{\mathcal{G}_r} U_{1,2n+r,N}, \tag{7}$$

where

$$Y^i_{2n+1,r,N} = \begin{bmatrix} y_i(2n+1) & \cdots & y_i(2n+N) \\ \vdots & \iddots & \vdots \\ y_i(2n+r) & \cdots & y_i(2n+r+N-1) \end{bmatrix}$$

with the superscript ^{i} denoting the channel index, the first subscript $_{2n+1}$ indicating the time index of the top-left entry, the second and third subscripts $_{r,N}$ representing the numbers of block rows and columns, respectively. The matrix $\mathcal{G}_r^i \in \mathbb{R}^{rp\times(2n+r)m}$ is defined by

$$\mathcal{G}_r^i = \begin{bmatrix} G^i_{2n} & \cdots & G^i_0 & & \\ & \ddots & \vdots & \ddots & \\ & & G^i_{2n} & \cdots & G^i_0 \end{bmatrix}$$

with the superscript ^{i} denoting the channel index, the subscript $_r$ indicating the number of block rows, and $\{G^i_j\}_{j=0}^{2n}$ being the matrix coefficients of $G_i(q)$. In the sequel, we assume that $N \gg r$, namely $Y_{2n+1,r,N}$ in (7) is a flat matrix.

The rank property of the augmented block Toeplitz matrix $\mathcal{G}_r$ is shown in the next lemma. Here, we make use of the following definition.

Definition 3. The $m \times m$ minors of the transfer function $H_1(q)$ are the determinants of all $m \times m$ sub-matrices of $H_1(q)$.

Lemma 1. *Suppose that the following assumptions hold:*

(1) *Assumptions A1–A3 hold;*
(2) *$H_1(q)$ and $H_2(q)$ do not share common zeros and poles;*
(3) *The poles of $H_i(q)$ for any $i \in \{1, 2\}$ do not coincide with its zeros;*
(4) *The greatest degree of all $m\times m$ minors of $Q_1(q)$ (or $Q_2(q)$) equals that of $\begin{bmatrix} Q_1(q) \\ Q_2(q) \end{bmatrix}$.*

Then, for any $r > 2n$, the rank of $\mathcal{G}_r$ satisfies

$$\mathrm{rank}(\mathcal{G}_r) = rm + 2n.$$

Furthermore, the rank of $Y_{2n+1,r,N}$ in (7) satisfies

$$\mathrm{rank}\left(Y_{2n+1,r,N}\right) = rm + 2n.$$

Proof of this lemma can be found in Appendix A.

Following the above analysis of two channel systems, the rank properties of single-channel systems will be derived analogously. For the i-th channel with $i \in \{1, 2\}$, the associated system model without measurement noise can be written as

$$y_i(k) = Q_i(q)\frac{s(k)}{p_i(q)}, \tag{8}$$

where $\frac{s(k)}{p_i(q)}$ is regarded as a pseudo source signal.

Corollary 1. *Suppose that Assumptions A1–A3 hold. We have that:*

(1) *when $p = m$, $\mathrm{rank}\left(Y^i_{2n+1,r,N}\right) = rm$.*
(2) *when $p > m$ and $H_i(q)$ has no transmission zeros, $\mathrm{rank}\left(Y^i_{2n+1,r,N}\right) = rm + n$ for any $r > n$.*

The above corollary can be proven by using the same means adopted in the proof of Lemma 1. For the sake of brevity, the proof will not be detailed here.

According to the results of Lemma 1 and Corollary 1, we can see that the two-channel output sequence $\begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix}$ lacks persistent excitation whenever $\{H_i(q)\}_{i=1}^2$ are square or tall. For a single channel with $p = m$, $y_i(k)$ is persistently exciting under some mild conditions; however, $y_i(k)$ lacks persistent excitation when $p > m$.

4. Blind identification of characteristic polynomials

As shown in the previous section, the pole information is crucial for the blind identification of autoregressive systems. In this section, we shall investigate the identification of the characteristic polynomials of two channel systems.

Suppose that $H_1(q) = L_1^{-1}(q)N_1(q)$ with $\deg[\det(L_1(q))] = n$ is an MFD of the first channel system. In order to estimate the input signal from the first channel, we need to find a left inverse of $H_1(q)$, which is denoted by $\tilde{H}_1(q)$. One direct way is to represent the inverse transfer matrix in a state-space form with order n. However, one drawback of such a way is that, when the number of transmission zeros is less than n, unknown poles will be included in $\tilde{H}_1(q)$ apart from those zeros of $H_1(q)$. To this end, the inverse transfer matrix is represented in terms of an MFD form rather than the state-space form.

When D_i is strictly tall, the numerator polynomial matrix $N_i(q)$ is generically right coprime, namely it has no transmission zeros. In such a case, the left inverse of $H_i(q)$ does not possess any poles. To make it more rigorous, we make the following assumption:

A5. $H_i(q)$ has no transmission zeros for $i \in \{1, 2\}$.

Under Assumption A5 and by rational matrix theory (Kailath, 1980), we can always find a polynomial matrix which is the left-inverse of $H_i(q)$.

Lemma 2. *Suppose that Assumptions* A2–A3 *and* A5 *hold. Then, for any positive integer $K \geq 2n$, there exists an $m \times p$ polynomial matrix $E_i(q)$ of degree K such that $E_i(q)H_i(q) = I$.*

Proof. By Assumptions A2–A3, there exists a matrix fraction description $H_i(q) = L_i^{-1}(q)N_i(q)$, where $\deg[\det(L_i(q))] = n$ and $N_i(q)$ is column reduced with its column degrees being summed up to n. Assumption A5 implies that $N_i(q)$ is right coprime. Since $N_i(q)$ is right coprime and column reduced, for any positive integer $K_0 \geq n$, there exists a polynomial matrix $\tilde{N}_i(q)$ of degree K_0 such that $\tilde{N}_i(q)N_i(q) = I$ (Gorokhov & Loubaton, 1997, Lemma 1). Therefore, the polynomial matrix $E_i(q) = \tilde{N}_i(q)L_i(q)$, with its degree being larger than or equal to $2n$, is a left inverse of $H_i(q)$.

By Lemma 2, there exists a polynomial matrix $E_1(q)$ of degree $2n$ such that

$$s(k) = E_1(q)\left(y_1(k) - w_1(k)\right).$$

Substituting the above equation into the second channel yields that

$$y_2(k) = H_2(q)E_1(q)\left(y_1(k) - w_1(k)\right) + w_2(k). \tag{9}$$

Note that $H_2(q)E_1(q)$ in the above equation may not be a proper transfer matrix, so it may not be able to be represented in a regular (non-descriptor) state-space form. Substituting $H_2(q) = \frac{Q_2(q)}{p_2(q)}$ into Eq. (9) yields that

$$\begin{aligned} y_2(k) &= \frac{Q_2(q)}{p_2(q)}E_1(q)\left(y_1(k) - w_1(k)\right) + w_2(k), \\ \text{or} \quad & p_2(q)\left(y_2(k) - w_2(k)\right) = Q_2(q)E_1(q)\left(y_1(k) - w_1(k)\right), \end{aligned} \tag{10}$$

where $p_2(q)$ has degree n and $Q_2(q)E_1(q)$ has degree $3n$. Let $\mathbf{p}_2$ be the coefficient vector of $p_2(q)$ and $\bar{E}_1$ the stacked matrix coefficients of $Q_2(q)E_1(q)$. The matrix form of (10) can then be written as

$$\begin{bmatrix} -Y^1_{\tau+1,3n+1,N} + W^1_{\tau+1,3n+1,N} \\ Y^2_{\tau+1,n+1,N} - W^2_{\tau+1,n+1,N} \end{bmatrix}^T \begin{bmatrix} \bar{E}_1 \\ \mathbf{p}_2 \otimes I \end{bmatrix} = 0, \tag{11}$$

where τ is a positive time index. In order to remove the noise effect in the above equation, we apply the instrumental-variable method (Ljung, 1999) as follows:

$$\frac{1}{N}\begin{bmatrix} Y^1_{1,\tau,N} \\ Y^2_{1,\tau,N} \end{bmatrix}\begin{bmatrix} -Y^1_{\tau+1,3n+1,N} + W^1_{\tau+1,3n+1,N} \\ Y^2_{\tau+1,n+1,N} - W^2_{\tau+1,n+1,N} \end{bmatrix}^T \begin{bmatrix} \bar{E}_1 \\ \mathbf{p}_2 \otimes I \end{bmatrix} = 0, \tag{12}$$

where the instrumental variable $\begin{bmatrix} Y^1_{1,\tau,N} \\ Y^2_{1,\tau,N} \end{bmatrix}$ and the measurement noise $\begin{bmatrix} W^1_{\tau+1,3n+1,N} \\ W^2_{\tau+1,n+1,N} \end{bmatrix}$ are uncorrelated, i.e.,

$$\lim_{N\to\infty}\frac{1}{N}\begin{bmatrix} Y^1_{1,\tau,N} \\ Y^2_{1,\tau,N} \end{bmatrix}\begin{bmatrix} W^1_{\tau+1,3n+1,N} \\ W^2_{\tau+1,n+1,N} \end{bmatrix}^T = 0.$$

Then, we have that

$$\lim_{N\to\infty}\frac{1}{N}\begin{bmatrix} Y^1_{1,\tau,N} \\ Y^2_{1,\tau,N} \end{bmatrix}\begin{bmatrix} -Y^1_{\tau+1,3n+1,N} \\ Y^2_{\tau+1,n+1,N} \end{bmatrix}^T \begin{bmatrix} \bar{E}_1 \\ \mathbf{p}_2 \otimes I \end{bmatrix} = 0. \tag{13}$$

Let the QR factorization be given:

$$\frac{1}{N}\begin{bmatrix} Y^1_{1,\tau,N} \\ Y^2_{1,\tau,N} \end{bmatrix}\begin{bmatrix} -Y^1_{\tau+1,3n+1,N} \\ Y^2_{\tau+1,n+1,N} \end{bmatrix}^T = U\begin{bmatrix} R_{11,N} & R_{12,N} \\ 0 & R_{22,N} \end{bmatrix}, \tag{14}$$

where U is the matrix consisting of $(4n+2)p$ orthonormal column vectors, $R_{11,N} \in \mathbb{R}^{(3n+1)p\times(3n+1)p}$, $R_{22,N} \in \mathbb{R}^{(n+1)p\times(n+1)p}$. Substituting Eq. (14) into (13) yields that

$$\lim_{N\to\infty} R_{22,N}\left(\mathbf{p}_2 \otimes I\right) = 0. \tag{15}$$

Since $p_2(q)$ is a monic polynomial of degree n, the coefficient vector $\mathbf{p}_2$ contains n unknown variables.

Next, we shall analyze the identifiability of $p_2(q)$ in the errors-in-variables model (10). The concept of (dual) minimal basis of a polynomial matrix (David Forney, 1975) will be used in the following lemma.

Lemma 3. *Consider the two-channel system model in* (3) *without measurement noise. Suppose that the following assumptions hold:*

(1) *Assumptions* A1–A3 *and* A5 *hold;*
(2) *There exists an integer index $i \in \{1, 2, \ldots, p\}$ such that the degrees of the minimal polynomial basis of $\begin{bmatrix} G_1(q) \\ G_{2,i}(q) \end{bmatrix}$, with $G_{2,i}(q)$ being the i-th row vector of $G_2(q)$, are summed up to $2n$, and the greatest degree of the dual minimal basis of $\begin{bmatrix} G_1(q) \\ G_{2,i}(q) \end{bmatrix}$ is smaller than or equal to n.*

Let $S_i \in \mathbb{R}^{(n+1)\times(n+1)p}$ be a selection matrix defined as

$$S_i = I_{(n+1)\times(n+1)} \otimes \mathbf{e}_i,$$

where $\mathbf{e}_i$ is the i-th row of a $p \times p$ identity matrix. Then we have that

$$\operatorname{rank}\begin{bmatrix} Y^1_{\tau+1,3n+1,N} \\ S_iY^2_{\tau+1,n+1,N} \end{bmatrix} = \operatorname{rank}\left[Y^1_{\tau+1,3n+1,N}\right] + n.$$

Proof of this lemma is given in Appendix B.

Remark 1. The result of Lemma 3 indicates that there exist n linearly independent equations for estimating $p_2(q)$. Let $\left(Y^2_{\tau+1,n+1,N}/Y^1_{\tau+1,3n+1,N}\right)$ denote the projection of $Y^2_{\tau+1,n+1,N}$ on the orthogonal complement to the row space of $Y^1_{\tau+1,3n+1,N}$. Under noise-free measurements, it can be derived from (11) that

$$\left(\mathbf{p}_2^T \otimes I\right)\left(Y^2_{\tau+1,n+1,N}/Y^1_{\tau+1,3n+1,N}\right) = 0.$$

It then follows that

$$\mathbf{p}_2^T S_i\left(Y^2_{\tau+1,n+1,N}/Y^1_{\tau+1,3n+1,N}\right) = 0 \quad \text{for } i = 1, \ldots, p.$$

Based on the result of Lemma 3, there exists an integer $i \in \{1, \ldots, p\}$ such that the matrix $S_i\left(Y^2_{\tau+1,n+1,N}/Y^1_{\tau+1,3n+1,N}\right)$ has rank n. Since $\mathbf{p}_2$ contains only n variables, it can be uniquely determined from (11) without the noise effect. However, in the presence of measurement noise, by properly choosing the value of τ, the instrumental variable $\begin{bmatrix} Y^1_{1,\tau,N} \\ Y^2_{1,\tau,N} \end{bmatrix}$ in (13) can be of high rank such that it does not destroy the rank properties of $\begin{bmatrix} Y^1_{\tau+1,3n+1,N} \\ S_i Y^2_{\tau+1,n+1,N} \end{bmatrix}$ shown in Lemma 3.

Remark 2. In the proof of Lemma 1, we can find that the polynomial matrix, which consists of the first $p+1$ rows of the coprime part of $\begin{bmatrix} G_1(q) \\ G_2(q) \end{bmatrix}$ in (A.3) in Appendix A, is likely to be irreducible with the degrees of its minimal polynomial basis being summed up to $2n$. In addition, according to the properties of the dual minimal basis of a polynomial matrix that are described (David Forney, 1975; Kailath, 1980), the dimension of the dual minimal basis of $\begin{bmatrix} G_1(q) \\ G_{2,i}(q) \end{bmatrix}$ is $(p+1-m)$, and the associated basis degrees are summed up to $2n$. As a result, the condition in Lemma 3 that *the greatest degree of the dual minimal basis is smaller than or equal to n* can easily be satisfied in practical scenarios.

As shown above, in order to identify the first characteristic polynomial $p_1(q)$ of the two channel systems, the second channel system $H_2(q)$ should not possess any zeros, and vice versa. When the concerned two channel systems have square transfer matrices, they are very likely to have transmission zeros, so their corresponding characteristic polynomials cannot be identified using the proposed method above. To cope with this problem, we adopt one more channel and combine it with the original two channel systems for identifying characteristic polynomials. In order to identify the characteristic polynomial of the first channel, the three-channel system is rewritten as

$$\begin{aligned} y_1(k) &= H_1(q)s(k) + w_1(k) \\ \underbrace{\begin{bmatrix} y_2(k) \\ y_3(k) \end{bmatrix}}_{\bar{y}_2(k)} &= \underbrace{\begin{bmatrix} H_2(q) \\ H_3(q) \end{bmatrix}}_{\bar{H}_2(q)} s(k) + \underbrace{\begin{bmatrix} w_2(k) \\ w_3(k) \end{bmatrix}}_{\bar{w}_2(k)}. \end{aligned} \tag{16}$$

According to the above discussion, if the tall transfer matrix $\bar{H}_2(q)$ in the above system does not have any zeros, then the characteristic polynomial of the first channel can be identified. Similarly, the other characteristic polynomials can be estimated using the same approach.

After obtaining all the characteristic polynomials $p_i(q)$, we shall investigate the identification of the numerator polynomial matrices $Q_i(q)$ in next section.

5. Blind identification of numerator polynomial matrices

In this section, two methods for identifying the numerator polynomial matrices $\{Q_i(q)\}_{i=1}^2$ will be developed. The first method requires the associated transfer matrices $\{H_i(q)\}_{i=1}^2$ to be square and stably invertible. The second one relaxes this requirement, but requires all the measurement noises to satisfy Assumption A4 and the transfer matrices $\{H_i(q)\}_{i=1}^2$ having no transmission zeros.

5.1. Blind identification of square and stably invertible transfer matrices

For two scalar FIR systems with a common source signal, the associated cross-relation equation can be easily constructed (Xu, Liu, Tong, & Kailath, 1995). However, for the two channel multivariable systems, due to the fact that the product of two polynomial matrices is not commutable, the so called cross-relation equation cannot be directly derived. In this section, we shall develop a new cross-relation identification method for the autoregressive MIMO systems.

Denote the state-space representation of the i-th channel model (Zhou, 1996):

$$H_i(q) := \left[\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right]. \tag{17}$$

The corresponding transfer matrix is $H_i(q) = C_i(qI - A_i)^{-1}B_i + D_i$. By Assumption A3, D_i is a square and regular matrix. Then, the inverse of $H_i(q)$ can be expressed as:

$$\tilde{H}_i(q) := \left[\begin{array}{c|c} \tilde{A}_i & \tilde{B}_i \\ \hline \tilde{C}_i & \tilde{D}_i \end{array}\right] = \left[\begin{array}{c|c} A_i - B_i D_i^{-1} C_i & -B_i D_i^{-1} \\ \hline D_i^{-1} C_i & D_i^{-1} \end{array}\right]. \tag{18}$$

It follows that

$$H_i(q)\tilde{H}_i(q) = \tilde{H}_i(q)H_i(q) = I.$$

To ensure the stability of $\tilde{H}_i(q)$, we need the following assumption:

A6. $H_i(q)$ is stably invertible for $i = 1, 2$, namely $A_i - B_i D_i^{-1} C_i$ is a stable matrix.

Given the inverse transfer matrix $\tilde{H}_1(q)$, it can be derived from the first channel model that $s(k) = \tilde{H}_1(q)\,(y_1(k) - w_1(k))$. Substituting it into the second channel model yields that

$$y_2(k) = H_2(q)\tilde{H}_1(q)\,(y_1(k) - w_1(k)) + w_2(k). \tag{19}$$

The above system can be regarded as a multivariable errors-in-variables system with a noisy input and a noisy output. Let $H_{21}(q) = H_2(q)\tilde{H}_1(q)$ and

$$H_{21}(q) := \left[\begin{array}{c|c} A_{21} & B_{21} \\ \hline C_{21} & D_{21} \end{array}\right] = \left[\begin{array}{cc|c} A_2 & B_2\tilde{C}_1 & B_2\tilde{D}_1 \\ 0 & \tilde{A}_1 & \tilde{B}_1 \\ \hline C_2 & D_2\tilde{C}_1 & D_2\tilde{D}_1 \end{array}\right]. \tag{20}$$

The state-space representation of (19) can be defined as

$$\begin{aligned} x_{21}(k+1) &= A_{21}x_{21}(k) + B_{21}(y_1(k) - w_1(k)) \\ y_2(k) &= C_{21}x_{21}(k) + D_{21}(y_1(k) - w_1(k)) + w_2(k), \end{aligned} \tag{21}$$

where $A_{21} \in \mathbb{R}^{2n\times 2n}$, $B_{21} \in \mathbb{R}^{2n\times p}$, $C_{21} \in \mathbb{R}^{p\times 2n}$, $D_{21} \in \mathbb{R}^{p\times p}$. The associated data equation for the above system model is written as

$$Y^2_{1,r,N} = \mathcal{O}_r X_{1,N} + \mathcal{T}_r Y^1_{1,r,N} - \mathcal{T}_r W^1_{1,r,N} + W^2_{1,r,N}, \tag{22}$$

where

$$X_{1,N} = \begin{bmatrix} x_{21}(1) & x_{21}(2) & \cdots & x_{21}(N) \end{bmatrix} \in \mathbb{R}^{n\times N},$$

$$\mathcal{O}_r = \begin{bmatrix} C_{21} \\ C_{21}A_{21} \\ \vdots \\ C_{21}A_{21}^{r-1} \end{bmatrix},$$

and

$$\mathcal{T}_r = \begin{bmatrix} D_{21} & & & \\ C_{21}B_{21} & D_{21} & & \\ \vdots & \ddots & \ddots & \\ C_{21}A_{21}^{r-2}B_{21} & \cdots & C_{21}B_{21} & D_{21} \end{bmatrix}.$$

Remark 3. By Corollary 1, in the absence of measurement noise, $Y^1_{1,r,N}$ in (22) has full row rank when the transfer matrices $\{H_i(q)\}_{i=1}^2$ are square, while it is rank deficient under tall transfer matrices. In other words, the input signal $y_1(k)$ in (21) lacks persistent excitation when $\{H_i(q)\}_{i=1}^2$ are tall transfer matrices. This is the reason why the presented identification method in this subsection cannot be applied to the multiple channel systems with tall transfer matrices.

Next, we shall apply the classical instrumental-variable technique to deal with the identification of the errors-in-variables system in (21). Eq. (22) can be extended to the following form:

$$\begin{bmatrix} Y^2_{1,r,N} \\ Y^2_{r+1,r,N} \end{bmatrix} = \begin{bmatrix} \mathcal{O}_r X_{1,N} \\ \mathcal{O}_r X_{r+1,N} \end{bmatrix} + \begin{bmatrix} \mathcal{T}_r \left(Y^1_{1,r,N} - W^1_{1,r,N}\right) \\ \mathcal{T}_r \left(Y^1_{r+1,r,N} - W^1_{r+1,r,N}\right) \end{bmatrix} + \begin{bmatrix} W^2_{1,r,N} \\ W^2_{r+1,r,N} \end{bmatrix}.$$

Post-multiplying the second equation above by $\begin{bmatrix} Y^{1,T}_{1,r,N} & Y^{2,T}_{1,r,N} \end{bmatrix}$ yields that

$$\begin{aligned} \frac{1}{N} Y^2_{r+1,r,N} \begin{bmatrix} Y^{1,T}_{1,r,N} & Y^{2,T}_{1,r,N} \end{bmatrix} &= \frac{1}{N} \mathcal{O}_r X_{r+1,N} \begin{bmatrix} Y^{1,T}_{1,r,N} & Y^{2,T}_{1,r,N} \end{bmatrix} \\ &+ \frac{1}{N} \mathcal{T}_r Y^1_{r+1,r,N} \begin{bmatrix} Y^{1,T}_{1,r,N} & Y^{2,T}_{1,r,N} \end{bmatrix} \\ &- \frac{1}{N} \mathcal{T}_r W^1_{r+1,r,N} \begin{bmatrix} Y^{1,T}_{1,r,N} & Y^{2,T}_{1,r,N} \end{bmatrix} \\ &+ \frac{1}{N} W^2_{r+1,r,N} \begin{bmatrix} Y^{1,T}_{1,r,N} & Y^{2,T}_{1,r,N} \end{bmatrix}. \end{aligned} \tag{23}$$

For any $\tau > 0$, $w_1(k+\tau)$ and $w_2(k+\tau)$ are independent of $y_2(k)$ and $y_1(k)$. Thus, the last two terms in the above equation approach zero as $N \to \infty$. Let the following QR factorization be given:

$$\begin{bmatrix} Y^1_{r+1,r,N} Y^{1,T}_{1,r,N} & Y^1_{r+1,r,N} Y^{2,T}_{1,r,N} \\ Y^2_{r+1,r,N} Y^{1,T}_{1,r,N} & Y^2_{r+1,r,N} Y^{2,T}_{1,r,N} \end{bmatrix} = \begin{bmatrix} L_{11,N} & 0 \\ L_{21,N} & L_{22,N} \end{bmatrix} \begin{bmatrix} V_{1,N} \\ V_{2,N} \end{bmatrix}. \tag{24}$$

It follows that

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} L_{22,N} = \lim_{N\to\infty} \frac{1}{\sqrt{N}} \mathcal{O}_r X_{r+1,N} \begin{bmatrix} Y^{1,T}_{1,r,N} & Y^{2,T}_{1,r,N} \end{bmatrix} V^T_{2,N}. \tag{25}$$

The following result can be derived subsequently.

Theorem 2. *Assume that Assumptions* A1–A3 *and* A6 *hold. In view of the QR factorization in* (24)*, if the matrix* $\lim_{N\to\infty} \frac{1}{\sqrt{N}} X_{r+1,N} \begin{bmatrix} Y^{1,T}_{1,r,N} & Y^{2,T}_{1,r,N} \end{bmatrix} V^T_{2,N}$ *has full row rank, then*

$$\operatorname{range}\left(\lim_{N\to\infty} \frac{1}{\sqrt{N}} L_{22,N}\right) = \operatorname{range}\left(\mathcal{O}_r\right). \tag{26}$$

Remark 4. The regularity of the matrix

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} X_{r+1,N} \begin{bmatrix} Y^{1,T}_{1,r,N} & Y^{2,T}_{1,r,N} \end{bmatrix} V^T_{2,N}$$

was discussed in Chou and Verhaegen (1997) and Verhaegen and Verdult (2007), which shows that the regularity condition is easy to be satisfied when system model in (21) is minimal and the corresponding system input is persistently exciting. By Lemma 1 and Corollary 1, it can be established that rank $\left(L_{22,N}\right) = 2n$ under square transfer matrices, while rank $\left(L_{22,N}\right) = n$ under tall transfer matrices. In Eq. (26), $\mathcal{O}_r$ is supposed to have rank $2n$. Therefore, the range of $\mathcal{O}_r$ can be determined under square transfer matrices rather than tall transfer matrices.

From Theorem 2, we can see that the column space of $\mathcal{O}_r$ can be numerically computed, so the matrices A_{21} and C_{21} can be estimated using the classic subspace identification method (Verhaegen & Verdult, 2007). Next, we shall estimate the matrices B_{21} and D_{21} based on the system model (21) with available A_{21} and C_{21}. In view of the data equation in (22), the following equation can be derived:

$$\begin{aligned} y_2(k+l) &= A^l_{21} x_{21}(k) + D_{21}\left(y_1(k+l) - w_1(k+l)\right) \\ &+ \sum_{i=1}^{l} C_{21} A^{i-1}_{21} B_{21}\left(y_1(k+l-i) - w_1(k+l-i)\right) \\ &+ w_2(k+l). \end{aligned} \tag{27}$$

By Assumptions A3 and A6, the matrix A_{21} is stable; therefore, the first term $A^l_{21} x_{21}(k)$ is negligible if l is large enough. Then, Eq. (27) can be accurately approximated as

$$y_2(k+l) - w_2(k+l) \cong \Gamma_l \begin{bmatrix} y_1(k) - w_1(k) \\ \vdots \\ y_1(k+l) - w_1(k+l) \end{bmatrix}, \tag{28}$$

where $\Gamma_l = \begin{bmatrix} C_{21} A^{l-1}_{21} B_{21} & \cdots & C_{21} B_{21} & D_{21} \end{bmatrix}$. The matrix form of the above equation is written as

$$Y^2_{l+1,1,N} - W^2_{l+1,1,N} \cong \Gamma_l \left(Y^1_{1,l+1,N} - W^1_{1,l+1,N}\right). \tag{29}$$

Applying the classic instrumental-variable technique yields that

$$\underbrace{\lim_{N\to\infty} \frac{1}{N} Y^2_{2l+2,1,N} Y^{1,T}_{1,l+1,N}}_{R_{Y_2 Y_1}} \cong \Gamma_l \underbrace{\lim_{N\to\infty} \frac{1}{N} Y^1_{l+2,l+1,N} Y^{1,T}_{1,l+1,N}}_{R_{Y_1 Y_1}}. \tag{30}$$

By Corollary 1, without noise effect, $Y^1_{l+2,l+1,N}$ and $Y^1_{1,l+1,N}$ are of full row rank under square transfer matrices; hence, $R_{Y_1 Y_1}$ in the above equation is a regular matrix. As a result, Γ_l can be determined from (30). Partition the matrix $R_{Y_1 Y_1}$ into block rows as

$$R_{Y_1 Y_1} = \begin{bmatrix} \bar{R}_1 \\ \vdots \\ \bar{R}_{l+1} \end{bmatrix}.$$

The vectorization form of Eq. (30) is written as

$$\begin{aligned} \operatorname{vec}\left(R_{Y_2 Y_1}\right) &= \begin{bmatrix} \sum_{i=1}^{l} \bar{R}^T_i \otimes \left(C_{21} A^{l-i}_{21}\right) & \bar{R}^T_{l+1} \otimes I \end{bmatrix} \\ &\times \begin{bmatrix} \operatorname{vec}\left(B_{21}\right) \\ \operatorname{vec}\left(D_{21}\right) \end{bmatrix}. \end{aligned} \tag{31}$$

Then, B_{21} and D_{21} can be estimated accordingly by solving the above equation.

Using the above developed algorithm, the system matrices $\{A_{21}, B_{21}, C_{21}, D_{21}\}$ corresponding to the transfer matrix $H_{21}(q) = H_2(q)\tilde{H}_1(q)$ can be estimated accordingly. Similarly, the system matrices $\{A_{12}, B_{12}, C_{12}, D_{12}\}$ corresponding to the transfer matrix $G_{12}(q) = H_1(q)\tilde{H}_2(q)$ can be estimated as well. Next, we will try to recover the transfer matrices $H_1(q)$ and $H_2(q)$ from the estimated $H_{12}(q)$ and $H_{21}(q)$.

According to the definitions of $H_{12}(q)$ and $H_{21}(q)$, the following relations can be derived:

$$\begin{bmatrix} I & -H_{12}(q) \\ -H_{21}(q) & I \end{bmatrix} \begin{bmatrix} H_1(q) \\ H_2(q) \end{bmatrix} = 0. \tag{32}$$

Pre-multiplying the left-hand side of (32) by the matrix $\begin{bmatrix} I & 0 \\ H_{21}(q) & I \end{bmatrix}$ yields that

$$\begin{aligned} &\begin{bmatrix} I & 0 \\ H_{21}(q) & I \end{bmatrix} \begin{bmatrix} I & -H_{12}(q) \\ -H_{21}(q) & I \end{bmatrix} \begin{bmatrix} H_1(q) \\ H_2(q) \end{bmatrix} \\ &= \begin{bmatrix} I & -H_{12}(q) \\ 0 & I - H_{21}(q) H_{12}(q) \end{bmatrix} \begin{bmatrix} H_1(q) \\ H_2(q) \end{bmatrix} = 0. \end{aligned} \tag{33}$$

When the transfer matrix $H_i(q)$ is square, we have that

$$H_{21}(q) H_{12}(q) = H_2(q)\tilde{H}_1(q) H_1(q)\tilde{H}_2(q) = I.$$

Therefore, the second equation of (33) is redundant. It indicates that it is unnecessary to estimate $H_{12}(q)$. Then, Eq. (32) can be simplified into

$$\begin{bmatrix} -H_{21}(q) & I \end{bmatrix} \begin{bmatrix} H_1(q) \\ H_2(q) \end{bmatrix} = 0. \tag{34}$$

Substituting $H_i(q) = \frac{Q_i(q)}{p_i(q)}$ into (34) yields that

$$\begin{bmatrix} -H_{21}(q) & I \end{bmatrix} \begin{bmatrix} Q_1(q)p_2(q) \\ Q_2(q)p_1(q) \end{bmatrix} = 0. \tag{35}$$

By Theorem 1, the augmented polynomial matrix $\begin{bmatrix} Q_1(q)p_2(q) \\ Q_2(q)p_1(q) \end{bmatrix}$ in the above equation is not right coprime; hence, it is not blindly identifiable.

However, once $\{p_i(q)\}_{i=1}^2$ have been estimated using the method in Section 4, Eq. (35) can be recasted as

$$\begin{bmatrix} -H_{ji}(q)p_j(q) & p_i(q)I \end{bmatrix} \begin{bmatrix} Q_i(q) \\ Q_j(q) \end{bmatrix} = 0, \tag{36}$$

where $Q_i(q)$ and $Q_j(q)$ are unknown polynomial matrices which are to be estimated. In the above equation, in order to determine $\begin{bmatrix} Q_i(q) \\ Q_j(q) \end{bmatrix}$ up to a matrix ambiguity, it requires $\begin{bmatrix} Q_i(q) \\ Q_j(q) \end{bmatrix}$ to be right coprime and column reduced and all its column degrees should be identical (Giannakis et al., 2000, Chapter 3). By Assumption A3, i.e. D_i and D_j have full column rank, and in view of the expression of $Q_i(q)$ in (4), we can obtain that $\begin{bmatrix} Q_i(q) \\ Q_j(q) \end{bmatrix}$ is column reduced. In addition, by the assumption that the set consisting of all the zeros and poles of $H_i(q)$ does not intersect with those of $H_j(q)$, it can be verified that $\begin{bmatrix} Q_i(q) \\ Q_j(q) \end{bmatrix}$ is right coprime. Therefore, if $\begin{bmatrix} Q_i(q) \\ Q_j(q) \end{bmatrix}$ has identical column degrees, it can be blindly identified from Eq. (36).

The matrix form of (36) can be written as

$$\begin{bmatrix} -\mathcal{T}^{ji}_{2n+1}\mathcal{P}^{j}_{n+1} & \mathcal{P}^{i}_{n+1} \end{bmatrix} \begin{bmatrix} \bar{Q}^i \\ \bar{Q}^j \end{bmatrix} = 0, \tag{37}$$

where

$$\mathcal{T}^{ji}_{2n+1} = \begin{bmatrix} D_{ji} & & & \\ C_{ji}B_{ji} & D_{ji} & & \\ \vdots & \ddots & \ddots & \\ C_{ji}A_{ji}^{2n-1}B_{ji} & \cdots & C_{ji}B_{ji} & D_{ji} \end{bmatrix},$$

$$\mathcal{P}^{i}_{n+1} = \underbrace{\begin{bmatrix} I & & \\ p_1^i I & \ddots & \\ \vdots & \ddots & I \\ p_n^i I & & p_1^i I \\ & \ddots & \vdots \\ & & p_n^i I \end{bmatrix}}_{n+1 \text{ block columns}}$$

with $\{p_j^i\}_{j=1}^n$ being the coefficients of $p_i(q)$, and $\bar{Q}^i = \begin{bmatrix} Q_0^i \\ \vdots \\ Q_n^i \end{bmatrix}$ with $\{Q_j^i\}_{j=0}^n$ being the matrix coefficients of $Q_i(q)$. Then, a nontrivial solution of $\begin{bmatrix} \bar{Q}^i \\ \bar{Q}^j \end{bmatrix}$ can be obtained by taking the singular value decomposition of $\begin{bmatrix} -\mathcal{T}^{ji}_{2n+1}\mathcal{P}^{j}_{2n+1} & \mathcal{P}^{i}_{2n+1} \end{bmatrix}$.

For ease of reference, the blind identification of three channel systems with square and stably invertible transfer matrices is summarized in Algorithm 1. The objective of Algorithm 1 is to estimate the system matrices $\{A_i, B_i, C_i, D_i\}_{i=1}^3$ using only the system outputs $\{y_i(k)\}_{i=1}^3$. The first four steps are carried out for estimating the transfer functions $\{H_i(q) = \frac{Q_i(q)}{p_i(q)}\}_{i=1}^3$: the first two steps are devoted to estimating the characteristic polynomials $\{p_i(q)\}_{i=1}^3$ while the third and fourth steps are designed for estimating the numerator polynomial matrices $\{Q_i(q)\}_{i=1}^3$. In the third step, the hybrid transfer function $H_{i,j}(q)$ is estimated through identifying its corresponding system matrices of the state-space system model in (21). To identify the system matrices in (21), we use the subspace identification method described in Eqs. (22)–(31). In the last step, the system matrices $\{A_i, B_i, C_i, D_i\}_{i=1}^3$ are estimated using the classic deterministic realization method "Ho–Kalman's method" described in Katayama (2006) and Verhaegen and Verdult (2007):

(1) Expand the transfer function of the i-th channel system as

$$\frac{Q_i(q)}{p_i(q)} = M_0^i + M_1^i q^{-1} + M_2^i q^{-2} + \cdots;$$

(2) Form the block Hankel matrix

$$\mathbf{M} = \begin{bmatrix} M_1^i & M_2^i & M_3^i & \cdots \\ M_2^i & M_3^i & M_4^i & \cdots \\ M_3^i & M_4^i & M_5^i & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

(3) Compute the extended observability and controllability matrices by taking the singular value decomposition of the above block Hankel matrix;
(4) Estimate the system matrices of the i-th channel system by exploiting the shifting structures of the extended observability/controllability matrices.

Algorithm 1 Blind identification of three channel systems with square transfer matrices

(1) Recast the three channel systems into the form of (16).
(2) Estimate characteristic polynomials $\{p_i(q)\}_{i=1}^3$ using Eq. (15) derived in Section 4.
(3) Identify the hybrid transfer matrices $H_{ij}(q)$ of the system in (21) for $i \neq j \in \{1, 2, 3\}$.
(4) Form Eq. (37) and estimate matrix coefficients of $\{Q_i(q)\}_{i=1}^3$.
(5) Estimate the system matrices $\{A_i, B_i, C_i, D_i\}_{i=1}^3$ using the standard Ho–Kalman method.

The implementation of Algorithm 1 requires at least three channel systems. The main reason is that, based on only two channel outputs, we are not able to estimate the denominator parts of their square transfer matrices. It is noteworthy that Algorithm 1 can be applied to the case that $w_i(k)$ is spatially correlated but temporally uncorrelated.

5.2. Blind identification of tall transfer matrices having no transmission zeros

Under tall transfer matrices, the characteristic polynomials can be identified by the method in Section 4. Herein, we only consider the estimation of the numerator polynomial matrices $\{Q_i(q)\}$. The presented algorithm in this subsection is called the generalized subspace identification method.

Analogous to (3), the two channel systems with measurement noise can be written as

$$\begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix} = \underbrace{\begin{bmatrix} Q_1(q)p_2(q) \\ Q_2(q)p_1(q) \end{bmatrix}}_{G(q)} \underbrace{\frac{s(k)}{p_1(q)p_2(q)}}_{u(k)} + \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix}. \tag{38}$$

The matrix form of the above equation is written as

$$\underbrace{\begin{bmatrix} Y^1_{2n+1,r,N} \\ Y^2_{2n+1,r,N} \end{bmatrix}}_{Y_{2n+1,r,N}} = \underbrace{\begin{bmatrix} \mathcal{G}^1_r \\ \mathcal{G}^2_r \end{bmatrix}}_{\mathcal{G}_r} U_{1,r,N} + \underbrace{\begin{bmatrix} W^1_{2n+1,r,N} \\ W^2_{2n+1,r,N} \end{bmatrix}}_{W_{2n+1,r,N}}, \tag{39}$$

where $\mathcal{G}^i_r$ for $i \in \{1, 2\}$ is defined in (7).

By Assumption A4, it can be derived that

$$\underbrace{\lim_{N\to\infty} \frac{Y_{2n+1,r,N} Y^T_{2n+1,r,N}}{N}}_{R_{YY}} = \mathcal{G}_r \underbrace{\lim_{N\to\infty} \frac{U_{1,r,N} U^T_{1,r,N}}{N}}_{R_{UU}} \mathcal{G}^T_r + \sigma^2 I. \tag{40}$$

By Assumption A1, it can be established that R_{UU} is a regular matrix. Furthermore, it can be derived that

$$\text{range}\left(R_{YY} - \sigma^2 I\right) = \text{range}\left(\mathcal{G}_r\right).$$

According to Lemma 1, the matrix $\mathcal{G}_r$ is rank deficient, so is $R_{YY} - \sigma^2 I$. Let the eigenvalue decomposition of R_{YY} be given:

$$R_{YY} = \begin{bmatrix} U_s & U_n \end{bmatrix} \begin{bmatrix} \Sigma + \sigma^2 I & \\ & \sigma^2 I \end{bmatrix} \begin{bmatrix} U^T_s \\ U^T_n \end{bmatrix}, \tag{41}$$

where U_s and U_n denote the signal and noise subspace of R_{YY}, respectively. By Lemma 1, the matrix U_n consists of $n_l = 2rp - (rm + 2n)$ independent orthonormal column vectors. Note that n_l is always non-negative since $\mathcal{G}_r \in \mathbb{R}^{2rp\times(2n+r)m}$ is a tall matrix. Let $U_n = \begin{bmatrix} \phi_1 \cdots \phi_{n_l} \end{bmatrix}$ with $\phi_i \in \mathbb{R}^{2rp}$ the i-th column vector. Then it holds that

$$\phi^T_i \mathcal{G}_r = 0 \quad \text{for } i = 1, \ldots, n_l. \tag{42}$$

Partition ϕ_i as $\phi^T_i = \begin{bmatrix} \phi_{i,1} & \cdots & \phi_{i,2r} \end{bmatrix}$ with $\phi_{i,j} \in \mathbb{R}^{1\times p}$ for $j = 1, \ldots, 2r$. Since $\mathcal{G}_r$ is a stacked block Toeplitz matrix, the above equation is equivalent to the following polynomial equation:

$$\begin{bmatrix} \Phi_{i,1}(q) & \Phi_{i,2}(q) \end{bmatrix} \begin{bmatrix} G_1(q) \\ G_2(q) \end{bmatrix} = 0, \tag{43}$$

where $\Phi_{i,1}(q) = \phi_{i,1} + \phi_{i,2}q + \cdots + \phi_{i,r}q^{r-1}$, $\Phi_{i,2}(q) = \phi_{i,r+1} + \phi_{i,r+2}q + \cdots + \phi_{i,2r}q^{r-1}$, and $G_1(q)$ and $G_2(q)$ are defined in (3). Stacking all equations of (43) yields that

$$\begin{bmatrix} \Phi_1(q) & \Phi_2(q) \end{bmatrix} \begin{bmatrix} G_1(q) \\ G_2(q) \end{bmatrix} = 0, \tag{44}$$

where $\Phi_i(q) = \begin{bmatrix} \Phi^T_{i,1}(q) \cdots \Phi^T_{i,n_l}(q) \end{bmatrix}^T$ for $i = 1, 2$.

When $\{p_i(q)\}^2_{i=1}$ are available, substituting $G(q)$ shown in (38) into Eq. (44) yields that

$$\begin{bmatrix} \Phi_1(q)p_2(q) & \Phi_2(q)p_1(q) \end{bmatrix} \begin{bmatrix} Q_1(q) \\ Q_2(q) \end{bmatrix} = 0. \tag{45}$$

Analogous to Eq. (36), $\begin{bmatrix} Q_1(q) \\ Q_2(q) \end{bmatrix}$ in the above equation can be identified up to a matrix ambiguity if the following assumptions hold:

(1) Assumption A3 holds;
(2) The set consisting of all the zeros and poles of $H_1(q)$ does not intersect with those of $H_2(q)$;
(3) $\begin{bmatrix} Q_1(q) \\ Q_2(q) \end{bmatrix}$ has identical column degrees.

The matrix form of (45) can be written as

$$\begin{bmatrix} \bar{\Phi}^1_{2n+1}\mathcal{P}^2_{n+1} & \bar{\Phi}^2_{2n+1}\mathcal{P}^1_{n+1} \end{bmatrix} \begin{bmatrix} \bar{Q}^1 \\ \bar{Q}^2 \end{bmatrix} = 0, \tag{46}$$

where

$$\bar{\Phi}^i_{2n+1} = \underbrace{\begin{bmatrix} \Phi^i_1 & & \\ \Phi^i_2 & \ddots & \\ \vdots & \ddots & \Phi^i_1 \\ \Phi^i_r & & \Phi^i_2 \\ & \ddots & \vdots \\ & & \Phi^i_r \end{bmatrix}}_{2n+1 \text{ block columns}}$$

with $\{\Phi^i_j\}^r_{j=1}$ being the matrix coefficients of $\Phi_i(q)$, and $\mathcal{P}^i_{n+1}$ and $\bar{Q}^i$ are the same as in (37). From (46), we can numerically obtain a nontrivial solution of $\begin{bmatrix} \bar{Q}^1 \\ \bar{Q}^2 \end{bmatrix}$ by taking the singular value decomposition of $\begin{bmatrix} \bar{\Phi}^1_{2n+1}\mathcal{P}^2_{n+1} & \bar{\Phi}^2_{2n+1}\mathcal{P}^1_{n+1} \end{bmatrix}$.

For ease of reference, the blind identification of two channel systems with tall transfer matrices is summarized in Algorithm 2. We estimate the associated transfer matrices $\{H_i(q)\}^2_{i=1}$, followed by estimating the system matrices $\{A_i, B_i, C_i, D_i\}^2_{i=1}$. The first step aims to estimate the denominator parts of the associated transfer matrices, while the second and third steps are devoted to estimating the numerator parts. The last step refers to the realization of the state-space system models from their transfer matrices, which is accomplished using the "Ho–Kalman's method" described in Katayama (2006) and Verhaegen and Verdult (2007).

Algorithm 2 Blind identification of two channel systems with tall transfer matrices

(1) Estimate characteristic polynomials $\{p_i(q)\}^2_{i=1}$ using Eq. (15) derived in Section 4.
(2) Derive Eq. (43) using the method shown in Eqs. (39)–(42).
(3) Form Eq. (46) and estimate coefficient matrices of $\{Q_i(q)\}^2_{i=1}$.
(4) Estimate the system matrices $\{A_i, B_i, C_i, D_i\}^2_{i=1}$ using the standard Ho–Kalman method.

Comparing Algorithm 1 with Algorithm 2, we can find that Eqs. (43) and (35) have similar forms, and Eqs. (45) and (36) have similar forms as well. It is remarked that Algorithm 2 is developed based on Assumption A4, i.e. the covariance matrix of $w_i(k)$ is a scaled identity matrix. Under Assumption A4, steps 2–3 in Algorithm 2 can also be applied to identify two square transfer matrices.

6. Numerical simulations

In this section, two numerical simulation examples are carried out to validate the proposed blind identification algorithms for multiple channel systems sharing a common source signal.

The proposed algorithms aim to identify the system matrices of the concerned state-space models *up to a matrix ambiguity*. Instead of directly measuring the estimation error of the system matrices, we assess the estimation performance of the associated transfer matrices. Denote by $\{\hat{A}_i, \hat{B}_i, \hat{C}_i, \hat{D}_i\}$ the estimated coefficient matrices of the i-th channel and $\hat{H}_i(q) = \frac{\hat{Q}_i(q)}{\hat{p}_i(q)} = \hat{C}_i(qI - \hat{A}_i)^{-1}\hat{B}_i + \hat{D}_i$ the corresponding transfer matrix. Here, we adopt the following normalized mean-square error to assess the identification performance of the numerator parts $\{Q_i(q)\}^L_{i=1}$:

$$\text{nMSE}_N = \frac{1}{K} \sum_{j=1}^{K} \frac{\min_{\Gamma} \|\bar{Q} - \hat{Q}^j \Gamma\|^2_F}{\|\bar{Q}\|^2_F}, \tag{47}$$

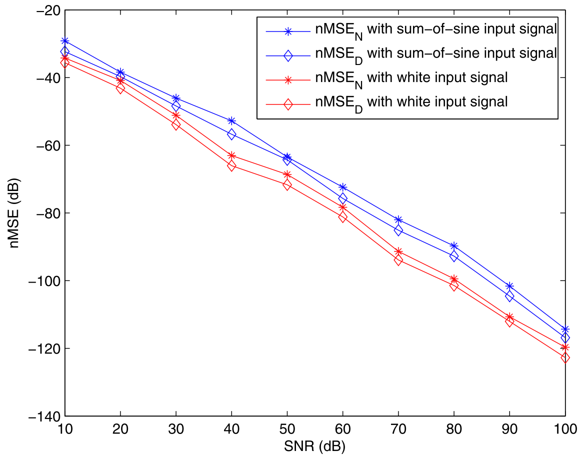


Fig. 1. Example 1: identification performance of Algorithm 2 against SNR. Blue curves correspond to the sum-of-sine input signal, while red ones correspond to the white noise input. Solid-star curves correspond to numerator matrices, while solid-diamond ones correspond to characteristic polynomials. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

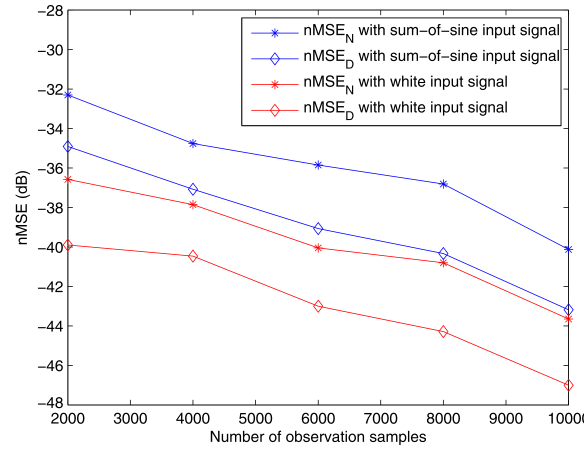


Fig. 2. Example 1: identification performance of Algorithm 2 against the number of observation samples. Blue curves correspond to the sum-of-sine input signal, while red ones correspond to the white noise input. Solid-star curves correspond to numerator matrices, while solid-diamond ones correspond to characteristic polynomials. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where Γ denotes the matrix ambiguity, $\bar{Q}$ stands for a block vector stacked by the matrix coefficients of $\{Q_i(q)\}_{i=1}^{L}$, $\hat{Q}^j$ represents the estimate of $\bar{Q}$ in the j-th experimental trial, and K is the total number of Monte-Carlo trials. Similarly, the identification performance of characteristic polynomials is evaluated by

$$\text{nMSE}_D = \frac{1}{K}\sum_{j=1}^{K}\frac{\|\bar{\mathbf{p}} - \hat{\mathbf{p}}^j\|_2^2}{\|\bar{\mathbf{p}}\|_2^2}, \tag{48}$$

where $\bar{\mathbf{p}}$ denotes a vector stacked by the coefficients of all the characteristic polynomials and $\hat{\mathbf{p}}^j$ is the estimate in the j-th experimental trial. In the simulation, the common system input is generated as a white-noise signal or a sum-of-sine signal. To show the identification performance against noise effect, the signal-to-noise ratio (SNR) is defined as

$$\text{SNR} = 10\log\left(\frac{\sum_{i=1,k=1}^{L,N}\|y_i(k) - w_i(k)\|_2^2}{\sum_{i=1,k=1}^{L,N}\|w_i(k)\|_2^2}\right). \tag{49}$$

Example 1. Two channel systems with tall transfer matrices are considered. Their system matrices are shown as follows:

$$\begin{aligned}
&A_1 = \begin{bmatrix}-0.6537 & 0.1005\\ 1.0000 & 0\end{bmatrix}, && B_1 = \begin{bmatrix}1.4525 & 0.6578\\ 0.6726 & 0.2015\end{bmatrix},\\
&C_1 = \begin{bmatrix}-0.0082 & 0.0885\\ -0.9403 & -1.0258\\ 0.1616 & -2.0666\end{bmatrix}, && D_1 = \begin{bmatrix}-0.0810 & -0.3231\\ -0.1936 & 1.7654\\ -1.0544 & 0.3209\end{bmatrix},\\
&A_2 = \begin{bmatrix}-1.0060 & -0.2162\\ 1.0000 & 0\end{bmatrix}, && B_2 = \begin{bmatrix}-0.1088 & -0.4299\\ -1.0141 & -0.9198\end{bmatrix},\\
&C_2 = \begin{bmatrix}-0.2591 & 0.7364\\ -1.7646 & 0.0137\\ -0.3099 & 0.6287\end{bmatrix}, && D_2 = \begin{bmatrix}-1.2215 & -0.3806\\ -0.5776 & 2.0738\\ -2.9127 & -1.2171\end{bmatrix}.
\end{aligned} \tag{50}$$

It can be verified that the above two transfer matrices possess no transmission zeros and have no common poles.

Figs. 1 and 2 show the identification performance of both numerator matrices and characteristic polynomials. The nMSE curves in Fig. 1 are plotted with respect to SNR. The length of adopted output data is set to 8000 and the number of Monte-Carlo trials is set to $K = 50$. We can see that the nMSE values for both the numerator matrices and characteristic polynomials decrease along with the SNR, indicating that the identification performance can be improved by increasing the SNR. The nMSE values of the characteristic polynomials are slightly smaller than those of numerator matrices, because the identification of the numerator matrices relies on the identification results of the characteristic polynomials. In addition, the nMSE values corresponding to the white-noise input are slightly smaller than those corresponding to the sum-of-sine input signal. This is because the frequency component of the white-noise input is much richer than the sum-of-sine input signal. Fig. 2 shows the identification performance against the number of observation samples, where the SNR is set to 20 dB. We can observe that the nMSE values decrease along with the number of observation samples.

Example 2. Three channel systems with square transfer matrices are considered. Their system matrices are shown as follows:

$$\begin{aligned}
&A_1 = \begin{bmatrix}1.0328 & -0.2000\\ 1.0000 & 0\end{bmatrix}, && B_1 = \begin{bmatrix}0.4692 & -0.4205\\ 0.5282 & 0.1416\end{bmatrix},\\
&C_1 = \begin{bmatrix}0.2957 & -0.1632\\ -0.6861 & 1.0004\end{bmatrix}, && D_1 = \begin{bmatrix}-1.1490 & -0.5662\\ -0.6648 & -1.6503\end{bmatrix},\\
&A_2 = \begin{bmatrix}-1.2533 & -0.3927\\ 1.0000 & 0\end{bmatrix}, && B_2 = \begin{bmatrix}0.7369 & 0.2112\\ 0.7080 & -0.3047\end{bmatrix},\\
&C_2 = \begin{bmatrix}1.8618 & 0.2300\\ 1.9953 & -0.0621\end{bmatrix}, && D_2 = \begin{bmatrix}-1.5296 & -2.9723\\ -2.2695 & 0.8078\end{bmatrix},\\
&A_3 = \begin{bmatrix}1.3981 & -0.4872\\ 1.0000 & 0\end{bmatrix}, && B_3 = \begin{bmatrix}1.4372 & -0.8803\\ -0.5827 & 0.0493\end{bmatrix},\\
&C_3 = \begin{bmatrix}0.3931 & -0.2496\\ 0.4023 & -0.1164\end{bmatrix}, && D_3 = \begin{bmatrix}-1.5417 & -1.4544\\ -0.3472 & -0.0452\end{bmatrix}.
\end{aligned}$$

In the above setting, the first two systems are stably invertible while the third one is not. It can be verified that the above three transfer matrices have no common zeros and poles.

In this example, the characteristic polynomial of each channel can be determined by carrying out steps 1–2 of Algorithm 1. Since the first two systems are stably invertible, their numerator polynomial matrices can then be identified by steps 3–4 of Algorithm 1 or steps 2–3 of Algorithm 2. Due to the fact that the

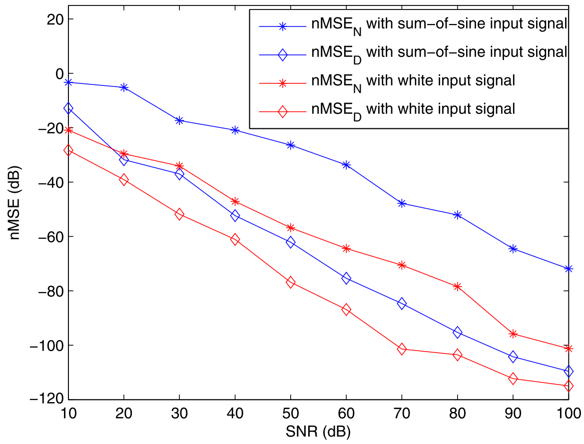


Fig. 3. Example 2: identification performance of the first two channels using Algorithm 1. Blue curves correspond to the sum-of-sine input signal, while red ones correspond to the white noise input. Solid-star curves correspond to numerator matrices, while solid-diamond ones correspond to characteristic polynomials. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

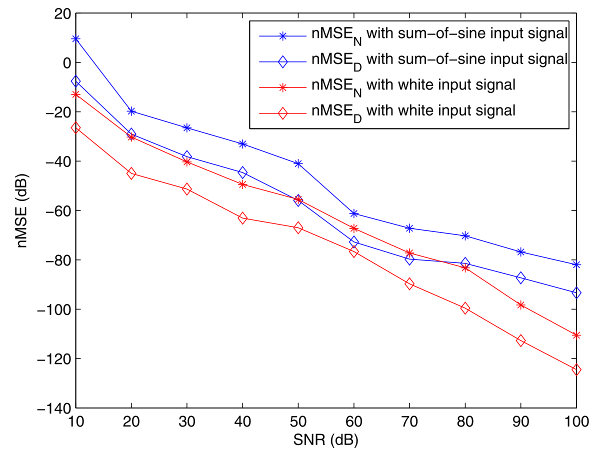


Fig. 4. Example 2: identification performance of the last two channels using Algorithm 2. Blue curves correspond to the sum-of-sine input signal, while red ones correspond to the white noise input. Solid-star curves correspond to numerator matrices, while solid-diamond ones correspond to characteristic polynomials. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

third system is non-invertible, its numerator polynomial matrix $Q_3(q)$ cannot be identified by Algorithm 1. However, since $\begin{bmatrix} Q_2(q) \\ Q_3(q) \end{bmatrix}$ is right coprime, $Q_3(q)$ can then be identified by carrying out steps 2–3 of Algorithm 2.

Fig. 3 shows the identification performance of the first two systems using Algorithm 1, while Fig. 4 shows the identification performance of the last two systems using Algorithm 2. In Algorithm 1, the value of l in (27) is set to 100. Analogous to Example 1, the identification accuracy improves along with the SNR. In addition, the performance associated with a sum-of-sine signal input is slightly worse than that with a white-noise input.

Since the numerator polynomial matrices of the first two systems $\{Q_i(q)\}_{i=1}^2$ can be identified using either Algorithm 1 or Algorithm 2, the identification performances of these two algorithms are compared. From Fig. 5, we can find that the identification performance of Algorithm 1 is slightly worse than that of Algorithm 2, which might be caused by the approximation error introduced by neglecting the first term on the right-hand side of (27).

7. Conclusion

In this paper, we have presented a comprehensive study of the blind identification of multivariable systems in state-space form. Identification algorithms have been developed for systems with invertible or non-invertible, square or tall transfer matrices. The present work is challenging in the following aspects. Different from the blind system identification with scalar transfer functions, the product of two multivariate transfer matrices is non-commutable. Hence, the cross-relation equation between different channels cannot be derived immediately. Unlike the traditional blind identification of FIR systems, the rational transfer matrices of the concerned systems have coupled poles and zeros, which is difficult to deal with. For the proposed identification methods, their blind identifiability conditions have been investigated. In addition, two numerical simulation examples have been provided to validate the presented identification algorithms.

The derived identification results in this paper do not rely on any statistical properties of the input signal. In other words, any persistently exciting deterministic input sequence is acceptable. Due to the fact that both the input and plants are unavailable, the derived identification results possess a wide range of applications, such as the detection of a common fault sequence of multiple plants, the reconstruction of the object image from multiple sensed images, and so on.

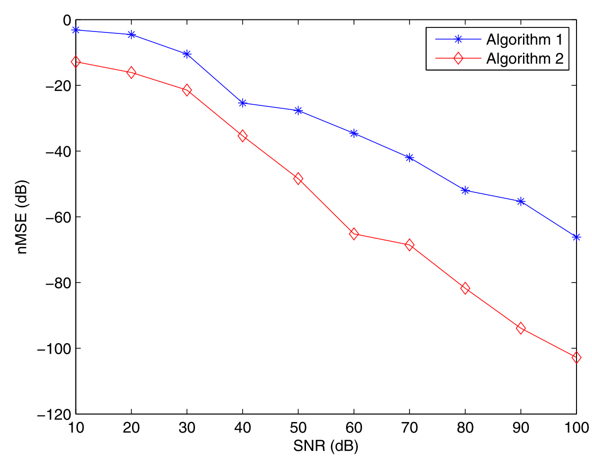


Fig. 5. Example 2: comparison of Algorithm 1 and Algorithm 2 on identifying the first two systems. The star-blue curve corresponds to the performance of Algorithm 1, while the diamond-red curve corresponds to the performance of Algorithm 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Appendix A. Proof of Lemma 1

According to the generalized resultant matrix properties in Kung, Kailath, and Morf (1976) and Lemma 1 in Giannakis et al. (2000, Chapter 3.3.2), it can be established that rank$(\mathcal{G}_r) = rm + \bar{n}$, where $\bar{n}$ denotes the *minimal order* (sum of the degrees of minimal polynomial basis) of $G(q)$. Thus, to obtain the results in the lemma, it is sufficient to prove that $\bar{n} = 2n$.

In this proof, we shall use following facts (David Forney, 1975; Kailath, 1980):

(1) The polynomial matrix $G(q)$ in (3) is irreducible if rank $[G(q)] = m$ for all $q \in \mathbb{C}$;
(2) If $G(q)$ is irreducible, then the *minimal order* of $G(q)$ equals the maximum degree of all $m \times m$ minors of $G(q)$.

By Assumptions A2 and A3, there exists an MFD of the i-th channel system $H_i(q) = N_i(q)R_i^{-1}(q)$ such that deg $[\det(R_i(q))] = n$, $\begin{bmatrix} N_i(q) \\ R_i(q) \end{bmatrix}$ is irreducible and the maximum degree of the $m \times m$ minors of $N_i(q)$ is n.

Since $N_i(q)R_i^{-1}(q) = Q_i(q)(p_i(q)I)^{-1}$, the polynomial matrix $G(q)$ in (3) can be rewritten as

$$G(q) = \begin{bmatrix} G_1(q) \\ G_2(q) \end{bmatrix} = \begin{bmatrix} N_1(q)R_1^{-1}(q)p_1(q)p_2(q) \\ N_2(q)R_2^{-1}(q)p_1(q)p_2(q) \end{bmatrix}. \tag{A.1}$$

It has that

$$\begin{bmatrix} \det\left(R_1^{-1}(q)p_1(q)p_2(q)\right) \\ \det\left(R_2^{-1}(q)p_1(q)p_2(q)\right) \end{bmatrix} = \begin{bmatrix} p_2(q) \\ p_1(q) \end{bmatrix} p_1^{m-1}(q)p_2^{m-1}(q).$$

By the assumption that $H_1(q)$ and $H_2(q)$ have no common poles, i.e. $p_1(q)$ and $p_2(q)$ have no common zeros, there exist polynomial matrices $\tilde{R}_1(q)$, $\tilde{R}_2(q)$ and $\tilde{C}(q)$ such that

$$\begin{bmatrix} R_1^{-1}(q)p_1(q)p_2(q) \\ R_2^{-1}(q)p_1(q)p_2(q) \end{bmatrix} = \begin{bmatrix} \tilde{R}_2(q) \\ \tilde{R}_1(q) \end{bmatrix} \tilde{C}(q), \tag{A.2}$$

where $\det(\tilde{C}(q)) = p_1^{m-1}(q)p_2^{m-1}(q)$, $\det(\tilde{R}_2(q)) = p_2(q)$, $\det(\tilde{R}_1(q)) = p_1(q)$. Then we have that

$$G(q) = \begin{bmatrix} G_1(q) \\ G_2(q) \end{bmatrix} = \begin{bmatrix} N_1(q)\tilde{R}_2(q) \\ N_2(q)\tilde{R}_1(q) \end{bmatrix} \tilde{C}(q). \tag{A.3}$$

Under the Assumption (2)–(3) of the lemma, it can be verified that $\begin{bmatrix} N_1(q)\tilde{R}_2(q) \\ N_2(q)\tilde{R}_1(q) \end{bmatrix}$ has full column rank for any $q \in \mathbb{C}$, so it is the coprime part of $G(q)$. In addition, by the properties of $N_i(q)$ shown above, we can obtain that the greatest degree of $m \times m$ minors of either $N_1(q)\tilde{R}_2(q)$ or $N_2(q)\tilde{R}_1(q)$ is $2n$. By the Assumption (4) of the lemma, it can be established that the greatest degree of all $m \times m$ minors of $\begin{bmatrix} N_1(q)\tilde{R}_2(q) \\ N_2(q)\tilde{R}_1(q) \end{bmatrix}$ is equal to that of either $N_1(q)\tilde{R}_2(q)$ or $N_2(q)\tilde{R}_1(q)$. As a consequence, the *minimum order* of $\begin{bmatrix} G_1(q) \\ G_2(q) \end{bmatrix}$ is $2n$. So far, it has been proven that $\text{rank}(\mathcal{G}_r) = rm + 2n$.

The pseudo source signal $u(k)$ in (3) is considered as an output of $s(k)$ by linear filtering. By Assumption A1, i.e. $s(k)$ is persistently exciting, it can be established that $u(k)$ is persistently exciting as well (Ljung, 1999); hence, the matrix $U_{1,r,N}$ in (7) has full row rank. We can obtain from Eq. (7) that

$$\text{rank}\left(Y_{2n+1,r,N}\right) = \text{rank}\left(\mathcal{G}_r\right) = rm + 2n.$$

Therefore, the lemma is proven.

Appendix B. Proof to Lemma 3

As shown in the proof of Lemma 1, the pseudo input $u(k)$ is persistently exciting. Without noise effect, the rank of $\begin{bmatrix} Y^1_{\tau+1,3n+1,N} \\ S_iY^2_{\tau+1,n+1,N} \end{bmatrix}$ is therefore equal to that of

$$\mathcal{G}^{1,2}_{3n+1,n+1} = \left[\begin{array}{c|c} \multicolumn{2}{c}{\mathcal{G}^1_{3n+1}} \\ \hline S_i\mathcal{G}^2_{n+1} & 0 \end{array} \right], \tag{B.1}$$

where $\mathcal{G}^1_{3n+1}$ and $\mathcal{G}^2_{n+1}$ are defined in (7). Note that the coefficient matrix $\mathcal{G}^{1,2}_{3n+1,n+1}$ is determined by the polynomial matrix $\begin{bmatrix} G_1(q) \\ G_{2,i}(q) \end{bmatrix}$ with $G_{2,i}(q)$ being the i-th row of $G_2(q)$.

By Assumption (2) of this lemma, we can obtain that (Kung et al., 1976):

$$\text{rank} \begin{bmatrix} \mathcal{G}^1_{n+1} \\ S_i\mathcal{G}^2_{n+1} \end{bmatrix} = (n+1)m + 2n.$$

Furthermore, following the proof procedure in Section III of Kung et al. (1976) or the rank analysis in Chen (1999, Chapter 7.8.2), we can obtain that

$$\begin{aligned} \text{rank}\left(\mathcal{G}^{1,2}_{3n+1,n+1}\right) &= \text{rank} \begin{bmatrix} \mathcal{G}^1_{n+1} \\ S_i\mathcal{G}^2_{n+1} \end{bmatrix} + 2nm \\ &= (3n+1)m + 2n. \end{aligned}$$

It then follows that

$$\begin{aligned} \text{rank} \begin{bmatrix} Y^1_{\tau+1,3n+1,N} \\ S_iY^2_{\tau+1,n+1,N} \end{bmatrix} &= \text{rank}\left(\mathcal{G}^{1,2}_{3n+1,n+1}\right) \\ &= (3n+1)m + 2n. \end{aligned}$$

By Corollary 1, we have that

$$\text{rank}\left(Y^1_{\tau+1,3n+1,N}\right) = (3n+1)m + n.$$

As a consequence, we can obtain that

$$\text{rank} \begin{bmatrix} Y^1_{\tau+1,3n+1,N} \\ S_iY^2_{\tau+1,n+1,N} \end{bmatrix} = \text{rank}\left(Y^1_{\tau+1,3n+1,N}\right) + n.$$

The lemma has been proven.

References

Abed-Meraim, K., Loubaton, P., & Moulines, E. (1997). A subspace algorithm for certain blind identification problems. *IEEE Transactions on Information Theory*, *43*(2), 499–511.

Bai, E. W., Li, Q. Y., & Dasgupta, S. (2002). Blind identifiability of IIR systems. *Automatica*, *38*(1), 181–184.

Chen, C. T. (1999). *Linear system theory and design, Linear system theory and design*. Oxford University Press.

Chen, Jianhua, Lin, Riqiang, Wang, Huina, Meng, Jing, Zheng, Hairong, & Song, Liang (2013). Blind-deconvolution optical-resolution photoacoustic microscopy in vivo. *Optics Express*, *21*(6), 7316–7327.

Chou, C. T., & Verhaegen, M. (1997). Subspace algorithms for the identification of multivariable dynamic errors-in-variables models. *Automatica*, *33*(10), 1857–1869.

David Forney, G., Jr. (1975). Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM Journal on Control*, *13*(3), 493–520.

Giannakis, G. B., Hua, Y., Stoica, P., & Tong, L. (2000). *Signal processing advances in wireless and mobile communications—volume 1*. Prentice Hall PTR.

Giannakis, G. B., & Serpedin, E. (1998). Blind identification of arma channels with periodically modulated inputs. *IEEE Transactions on Signal Processing*, *46*(11), 3099–3104.

Gopal, M. (1993). *Modern control system theory*. Wiley.

Gorokhov, A., & Loubaton, P. (1997). Subspace-based techniques for blind separation of convolutive mixtures with temporally correlated sources. *IEEE Transactions on Circuits and Systems I-Fundamental Theory and Applications*, *44*(9), 813–820.

Huang, Yiteng (Arden), Benesty, Jacob, & Chen, Jingdong (2006). *Acoustic MIMO signal processing*. Springer.

Kailath, Thomas (1980). *Linear systems*. Englewood Cliffs, NJ: Prentice-Hall.

Katayama, T. (2006). *Communications and control engineering, Subspace methods for system identification*. Springer.

Kung, S., Kailath, T., & Morf, M. (1976). A generalized resultant matrix for polynomial matrices. In *Proceedings of the IEEE conference on decision and control, Vol. 15* (pp. 892–895).

Li, Xiaohua (2005). Blind channel estimation and equalization in wireless sensor networks based on correlations among sensors. *IEEE Transactions on Signal Processing*, *53*(4), 1511–1519.

Ljung, Lennart (1999). *System identification: theory for the user* (2nd ed.). Prentice Hall PTR.

Moulines, E., Duhamel, P., Cardoso, J. F., & Mayrargue, S. (1995). Subspace methods for the blind identification of multichannel FIR filters. *IEEE Transactions on Signal Processing*, *43*(2), 516–525.

Routtenberg, T., & Tabrikian, J. (2010). Blind mimo-ar system identification and source separation with finite-alphabet. *IEEE Transactions on Signal Processing*, *58*(3), 990–1000.

Tong, L., & Perreau, S. (1998). Multichannel blind identification: From subspace to maximum likelihood methods. *Proceedings of the IEEE*, *86*(10), 1951–1968.

Vanbeylen, Laurent, Pintelon, Rik, & Schoukens, Johan (2008). Blind maximum likelihood identification of Hammerstein systems. *Automatica*, *44*(12), 3139–3146.

Verhaegen, Michel, & Verdult, Vincent (2007). *Filtering and system identification: a least squares approach*. Cambridge University Press.

Xu, G., Liu, H., Tong, L., & Kailath, T. (1995). A least-squares approach to blind channel identification. *IEEE Transactions on Signal Processing*, *43*(12), 2982–2993.

Yu, Chengpu, Xie, Lihua, & Soh, Yeng Chai (2014). Blind channel and source estimation in networked systems. *IEEE Transactions on Signal Processing*, *62*(17), 4611–4626.

Yu, Chengpu, Xie, Lihua, & Zhang, Cishen (2014). Deterministic blind identification of IIR systems with output-switching operations. *IEEE Transactions on Signal Processing*, *62*(7), 1740–1749.

Yu, Chengpu, Zhang, Cishen, & Xie, Lihua (2012a). A blind deconvolution approach to ultrasound imaging. *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, *59*(2), 271–280.

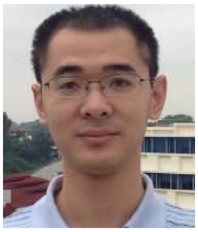
Yu, Chengpu, Zhang, Cishen, & Xie, Lihua (2012b). Blind identification of multichannel arma models based on second-order statistics. *IEEE Transactions on Signal Processing*, *60*(8), 4415–4420.

Yu, Chengpu, Zhang, Cishen, & Xie, Lihua (2013a). Blind identification of non-minimum phase arma systems. *Automatica*, *49*(6), 1846–1854.

Yu, Chengpu, Zhang, Cishen, & Xie, Lihua (2013b). Blind system identification using precise and quantized observations. *Automatica*, *49*(9), 2822–2830.

Zhang, Liqing, & Cichocki, Andrzej (2000). Blind deconvolution of dynamical systems: a state-space approach. *Journal of Signal Processing*, *4*(2), 111–130.

Zhou, Kemin (1996). *Robust and optimal control*. Prentice Hall PTR.



Chengpu Yu received the B.E. and M.E. degrees in electrical engineering from the University of Electronic Science and Technology of China, in 2006 and 2009, respectively, and the Ph.D. degree in electrical engineering from Nanyang Technological University, Singapore, in 2014.

He was with the Internet of Things lab at Nanyang Technological University as a research associate from June 2013 to June 2014. Since June 2014, he has been with Delft Center for Systems and Control as a postdoc. His research interests include system identification, distributed optimization and optical imaging.

Michel Verhaegen received an engineering degree in aeronautics from the Delft University of Technology, The Netherlands, in 1982, and the doctoral degree in applied sciences from the Catholic University Leuven, Belgium, in 1985. From 1985 to 1994 he has been a research fellow of the US National Research Council (NRC) and the Dutch Academy of Arts and Sciences. In the period 1994–1999 he was an Associate Professor of the control laboratory of the Delft University of Technology and became a full professor at the faculty of Applied Physics of the University of Twente in the Netherlands in 1999. From 2001 on he moved back to the University of Delft and joined the Delft Center for Systems and Control. His main research directions include system identification, distributed and fault tolerant control and data driven controller design methodologies. Application areas include control for high resolution imaging and wind energy. In 2014 he was a recipient of an advanced ERC grant.