J. Math. Anal. Appl. 409 (2014) 494–508

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

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

Hankel matrices for system identification[☆]

Bi-Qiang Mu, Han-Fu Chen [*]

Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China

ARTICLE INFO

Article history:
Received 30 January 2013
Available online 19 July 2013
Submitted by Xu Zhang

Keywords:
Hankel matrix
Row-full-rank
Impulse response
Correlation function
Multi-variable linear systems

ABSTRACT

The coefficients of a linear system, even if it is a part of a block-oriented nonlinear system, normally satisfy some linear algebraic equations via Hankel matrices composed of impulse responses or correlation functions. In order to determine or to estimate the coefficients of a linear system it is important to require the associated Hankel matrix be of row-full-rank. The paper first discusses the equivalent conditions for identifiability of the system. Then, it is shown that the row-full-rank of the Hankel matrix composed of impulse responses is equivalent to identifiability of the system. Finally, for the row-full-rank of the Hankel matrix composed of correlation functions, the necessary and sufficient conditions are presented, which appear slightly stronger than the identifiability condition. In comparison with existing results, here the minimum phase condition is no longer required for the case where the dimension of the system input and output is the same, though the paper does not make such a dimensional restriction.

1. Introduction

A block-oriented nonlinear system often includes some linear parts as its subsystems, for example, the Hammerstein system is composed of a nonlinear static block followed by a linear subsystem, and the Wiener–Hammerstein system is a nonlinear static block sandwiched by two linear subsystems. When identifying such kind of systems, one has to estimate not only the nonlinearities but also their linear subsystems. From the existing papers, e.g., [3,11,14,8,9] among others, it is seen that the Hankel matrices composed of impulse responses of the linear subsystem as well as composed of the correlation functions of its output are of crucial importance for estimating the unknown coefficients of the system. Let us explain this more clearly.

Consider the following linear model

$$A(z)y_k = B(z)u_k, \tag{1}$$

where

$$A(z) = I + A_1 z + \cdots + A_p z^p \quad \text{with } A_p \neq 0 \tag{2}$$

$$B(z) = B_0 + B_1 z + \cdots + B_q z^q \quad \text{with } B_q \neq 0 \tag{3}$$

are matrix polynomials in the backward-shift operator $z\ :\ zy_k = y_{k-1}$. The system output y_k and input u_k are of n- and m-dimensions, respectively.

☆ This work was supported by NSFC under Grants 61273193, 61120106011, 61134013, and the National Center for Mathematics and Interdisciplinary Sciences, Chinese Academy of Sciences.

* Corresponding author.
E-mail address: hfchen@iss.ac.cn (H.-F. Chen).

0022-247X/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmaa.2013.07.025

Assume $A(z)$ is stable, i.e., $\det A(z) \neq 0, \forall |z| \leq 1$.

In time series analysis, it is required to estimate the systems orders (p, q), the matrix coefficients $(A_1, \ldots, A_p, B_1 \ldots, B_q)$, and the covariance matrix $\Sigma_u = Eu_k u_k^T$ of the innovation process on the basis of output data $\{y_0, y_1, y_2, \ldots\}$ for the case where $n = m$, $B_0 = I$, and $\{u_k\}$ is a sequence of zero-mean iid (independent and identically distributed) or uncorrelated random vectors.

In system and control, as a rule $n \geq m$, and it is also required to estimate the systems orders (p, q) and the matrix coefficients $(A_1, \ldots, A_p, B_0, B_1 \ldots, B_q)$ with B_0 included. If the linear model is a part of nonlinear systems, e.g., the Hammerstein system, the Wiener system, etc., then the available information for identification may be the noisy or estimated inputs and outputs of the model.

Stability of $A(z)$ gives possibility to define the transfer function:

$$H(z) \triangleq A^{-1}(z)B(z) = \sum_{i=0}^{\infty} H_i z^i, \tag{4}$$

where $H_0 = B_0$, $\|H_i\| = O(e^{-ri})$, $r > 0$, $i > 1$. Then, y_k in (1) can be connected with the input $\{u_k\}$ via impulse responses:

$$y_k = \sum_{i=0}^{\infty} H_i u_{k-i}. \tag{5}$$

Let us first derive the linear equations connecting $\{A_1, \ldots, A_p, B_0, B_1, \ldots, B_q\}$ with $\{H_i\}$.

From (4), it follows that

$$B_0 + B_1 z + \cdots + B_q z^q = (I + A_1 z + \cdots + A_p z^p)(H_0 + H_1 z + \cdots + H_i z^i + \cdots). \tag{6}$$

Identifying coefficients for the same degrees of z at both sides implies

$$B_i = \sum_{j=0}^{i\wedge p} A_j H_{i-j} \quad \forall 0 \leq i \leq q, \tag{7}$$

$$H_i = -\sum_{j=1}^{i\wedge p} A_j H_{i-j} \quad \forall i \geq q+1, \tag{8}$$

where $A_0 = I$ and $a \wedge b$ denotes $\min(a, b)$.

For H_i, $q+1 \leq i \leq q+np$, by (8) we obtain the following linear algebraic equation

$$[A_1, A_2, \ldots, A_p]L = -[H_{q+1}, H_{q+2}, \ldots, H_{q+np}], \tag{9}$$

where

$$L \triangleq \begin{pmatrix} H_q & H_{q+1} & \cdots & H_{q+np-1} \\ H_{q-1} & H_q & \cdots & H_{q+np-2} \\ \vdots & \vdots & \ddots & \vdots \\ H_{q-p+1} & H_{q-p+2} & \cdots & H_{q+(n-1)p} \end{pmatrix}, \tag{10}$$

where $H_i \triangleq 0$ for $i < 0$.

Define

$$\theta_A^T \triangleq [A_1, \ldots, A_p], \qquad W^T \triangleq -[H_{q+1}, H_{q+2}, \ldots, H_{q+np}]. \tag{11}$$

Then, from (9) it follows that

$$\theta_A = (LL^T)^{-1} LW, \tag{12}$$

if L is of row-full-rank.

In this case, if we can obtain estimates for $\{H_i\}$, then replacing H_i, $i = 0, 1, 2, \ldots$ in (9) with their estimates, we derive the estimate for θ_A. Finally, with the help of (7) the estimates for B_i, $i = 0, 1, \ldots, q$ can also be obtained.

From here we see that the row-full-rank of the Hankel matrix L composed of impulse responses is important for estimating the system indeed.

The well-known Yule–Walker equation connects θ_A with the Hankel matrix composed of correlation functions of the system output $\{y_k\}$.

Under the stability assumption on $A(z)$, $\{y_k\}$ is a stationary process with correlation function $R_i \triangleq Ey_k y_{k-i}^T$, if $\{u_k\}$ is a sequence of zero-mean uncorrelated random vectors with the same second moment.

Multiplying y_{k-t}^T, $t \geq q+1$ on the both sides of (1) from right and taking expectation, we obtain

$$E(y_k + A_1 y_{k-1} + \cdots + A_p y_{k-p}) y_{k-t}^T = E(B_0 u_k + B_1 u_{k-1} + \cdots + B_q u_{k-q}) y_{k-t}^T = 0 \quad \forall t \geq q+1,$$

which yields

$$\sum_{i=0}^{p} A_i R_{q-i+l} = 0 \quad \forall l \geq 1. \tag{13}$$

For $R_i,\ q+1 \leq i \leq q+np$, by (13) we have the following linear algebraic equation called the Yule–Walker equation:

$$[A_1, A_2, \ldots, A_p]\Gamma = -[R_{q+1}, R_{q+2}, \ldots, R_{q+np}], \tag{14}$$

where

$$\Gamma \triangleq \begin{pmatrix} R_q & R_{q+1} & \cdots & R_{q+np-1} \\ R_{q-1} & R_q & \cdots & R_{q+np-2} \\ \vdots & \vdots & \ddots & \vdots \\ R_{q-p+1} & R_{q-p+2} & \cdots & R_{q+(n-1)p} \end{pmatrix}. \tag{15}$$

Similar to (12) we can rewrite (14) as

$$\theta_A = (\Gamma\Gamma^T)^{-1}\Gamma U, \tag{16}$$

whenever Γ is of row-full-rank, where $U^T \triangleq -[R_{q+1}, R_{q+2}, \ldots, R_{q+np}]$.

From here it is seen that the row-full-rank of the Hankel matrix Γ composed of correlation functions has the similar importance as L does.

The row-full-rank of L and Γ is closely related with identifiability (see, e.g., [12,5,2] among others) of (1)–(3).

The purpose of this paper is to give the necessary and sufficient conditions for the row-full-rank of L and Γ for the general case $n \geq m$ and with B_0 unknown. In Section 3 it is shown that the row-full-rank of L is equivalent to identifiability of (1)–(3), while the row-full-rank of Γ is equivalent to a condition slightly stronger than identifiability as shown in Section 4. Prior to discussing the row-full-rank of L and Γ, the identifiability issue is addressed in detail in Section 2. A brief conclusion is given in Section 5, while some auxiliary results from matrix polynomials are provided in Appendix.

2. Identifiability

As explained in Introduction, stability of $A(z)$ guarantees stationarity of y_k of the linear model (1) if $\{u_k\}$ is a sequence of zero mean uncorrelated random vectors with the same second order moment, and also allows to define the transfer function (4).

For the transfer function $H(z) = \sum_{i=0}^{\infty} H_i z^i$, $A^{-1}(z)B(z)$ is called its matrix fraction description (MFD) form. It is natural to consider the uniqueness issue of the description [12,5,2].

Denote by $\mathcal{M}$ the totality of the matrix pairs $[X(z)\ Y(z)]$ satisfying $X^{-1}(z)Y(z) = H(z)$, where $X(z) \in \mathbb{R}^{n\times n}$ is stable and monic with order less than or equal to p and $Y(z) \in \mathbb{R}^{n\times m}$ is with order less than or equal to q.

By (4) $[A(z)\ B(z)] \in \mathcal{M}$. We are interested in conditions guaranteeing the uniqueness of MFD. To clarify this, we first prove a lemma concerning the orders of factors in a matrix polynomial factorization. In the one-dimensional case, the orders of factors of a polynomial are certainly less than the order of the polynomial. However, in the multi-dimensional case the picture is different. Let the $n \times m$-matrix polynomial $B(z) = B_0 + B_1 z + \cdots + B_q z^q$ be factorized as a product of two matrix polynomials $C(z)$ and $D(z)$: $B(z) = C(z)D(z)$. Since $B(z) = C(z)U(z)U^{-1}(z)D(z)$ and $U^{-1}(z)D(z)$ remains to be a matrix polynomial for any unimodular matrix $U(z)$, the factors $C'(z) \triangleq C(z)U(z)$ and $D'(z) \triangleq U^{-1}(z)D(z)$ in the factorization $B(z) = C'(z)D'(z)$ may be with arbitrarily high orders.

The following lemma shows that the order of $D'(z)$ in the factorization $B(z) = C'(z)D'(z)$ can be made no higher than that of $B(z)$ by appropriately choosing the unimodular matrix $U(z)$.

Lemma 1. *Assume an $n \times t$-matrix polynomial $G(z)$ of order r is factorized as $G(z) = C(z)D(z)$, where $C(z)$ and $D(z)$ are matrix polynomials of $n \times n$ and $n \times t$ dimensions, respectively. Then, an $n \times n$ unimodular matrix $U(z)$ can be chosen such that in the factorization $G(z) = C'(z)D'(z)$ with $C'(z) \triangleq C(z)U(z)$ and $D'(z) \triangleq U^{-1}(z)D(z)$ where $\deg D'(z) \leq \deg G(z) = r$.*

Proof. It is well known from [4] that the elementary column transformations, i.e., multiplying $C(z)$ from right by the matrices corresponding to exchanging the places of its ith column with the jth column, multiplying the ith column of $C(z)$ by a constant, and adding its ith column with its jth column multiplied by a polynomial, may lead the matrix polynomial $C(z)$ to a lower-triangular matrix, for which at each row the highest degree appears at its diagonal element. Denoting by $U(z)$ the unimodular matrix resulting from all the elementary transformations yielding $C(z)$ to the lower-triangular form, we have

$$C(z)U(z) = \begin{pmatrix} c'_{1,1}(z) & 0 & \cdots & 0 \\ c'_{2,1}(z) & c'_{2,2}(z) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c'_{n,1}(z) & c'_{n,2}(z) & \cdots & c'_{n,n}(z) \end{pmatrix}, \tag{17}$$

where $\deg c'_{i,s}(z) \leq \deg c'_{i,i}(z)\ \forall s \leq i\ \forall i:\ i = 1, \ldots, n$.

We now show that this $U(z)$ is the one required by the lemma.

Let $D'(z) \triangleq U^{-1}(z)D(z) = \{d'_{i,j}(z)\}_{i=1,\dots,n}^{j=1,\dots,t}$ and $G(z) = \{g_{i,j}(z)\}_{i=1,\dots,n}^{j=1,\dots,t}$.

For the lemma it suffices to show that for any fixed $j:\ j = 1, \dots, t$

$$\deg d'_{i,j}(z) \le \max_{1\le l\le n} \deg g_{l,j}(z) \quad \forall i : i = 1, \dots, n. \tag{18}$$

For $i = 1$, we have $g_{1,j}(z) = c'_{1,1}(z)d'_{1,j}(z)$, which obviously implies

$$\deg d'_{1,j}(z) \le \max_{1\le l\le n} \deg g_{l,j}(z).$$

Thus, (18) holds for $i = 1$.

Assume (18) is true for $i = 1, \dots, s-1$. We want to show that (18) also holds for $i = s$.

Assume the converse: $\deg d'_{s,j}(z) > \max_{1\le l\le n} \deg g_{l,j}(z)$.

The inductive assumption incorporated with the converse assumption implies that

$$\deg d'_{s,j}(z) > \deg d'_{i,j}(z) \quad \forall i:\ i = 1, \dots, s-1. \tag{19}$$

By noticing $\deg c'_{s,l}(z) \le \deg c'_{s,s}(z)\ \forall l \le s\ \forall s:\ s = 1, \dots, n$, by (19) from the equality $g_{s,j}(z) = c'_{s,1}(z)d'_{1,j}(z) + \cdots + c'_{s,s-1}(z)d'_{s-1,j}(z) + c'_{s,s}(z)d'_{s,j}(z)$ we derive $\deg g_{s,j}(z) = \deg c'_{s,s}(z)d'_{s,j}(z)$. From here the converse assumption leads to a contradictory inequality:

$\deg g_{s,j}(z) \ge \deg d'_{s,j}(z) > \max_{1\le l\le n} \deg g_{l,j}(z)$.

The obtained contradiction proves the lemma. □

The following lemma is based on [5,2], but its proof is presented in detail below, because it plays an important role for the main results of the paper and also because some assertions in [5] are not easily understandable without Lemma 1.

Lemma 2. *The following conditions are equivalent.*

H1 *The set $\mathcal{M}$ is composed of the unique pair $[A(z)\ B(z)]$.*

H2 *$A(z)$ and $B(z)$ have no common left factor and $[A_p\ B_q]$ is of row-full-rank.*

H3 *There are no n-vector polynomial $d(z)$ and m-vector polynomial $c(z)$ (not both zero) with orders strictly less than p and q, respectively, such that $d^T(z)H(z) + c^T(z) = 0$.*

Proof. H2 $\Rightarrow$ H1

Assume H2 holds. Take any matrix polynomial pair $[\overline{A}(z)\ \overline{B}(z)] \in \mathcal{M}$, where

$$\overline{A}(z) = I + \overline{A}_1 z + \cdots + \overline{A}_{\overline{p}} z^{\overline{p}} \quad \text{with } \overline{p} \le p,$$
$$\overline{B}(z) = \overline{B}_0 + \overline{B}_1 z + \cdots + \overline{B}_{\overline{q}} z^{\overline{q}} \quad \text{with } \overline{q} \le q.$$

We have to show that $[\overline{A}(z)\ \overline{B}(z)] = [A(z)\ B(z)]$. Set $C(z) \triangleq \overline{A}(z)A^{-1}(z)$. Then, we have

$$\overline{A}(z) = C(z)A(z), \tag{20}$$
$$\overline{B}(z) = \overline{A}(z)H(z) = \overline{A}(z)A^{-1}(z)B(z) = C(z)B(z). \tag{21}$$

Since both $A(z)$ and $\overline{A}(z)$ are stable matrices, $\det C(z)$ is not identically equal to zero. So, the rank of $C(z)$ is n.

By Lemma 3 given in Appendix $C(z)$ can be presented in the Smith–McMillan canonical form:

$$\begin{aligned} C(z) &= U(z)\mathrm{diag}\left[\frac{q_1(z)}{p_1(z)}, \frac{q_2(z)}{p_2(z)}, \dots, \frac{q_n(z)}{p_n(z)}\right]V(z) \\ &= U(z)P^{-1}(z)Q(z)V(z), \end{aligned} \tag{22}$$

where $U(z)$ and $V(z)$ are $n \times n$ unimodular matrices,

$$P(z) = \mathrm{diag}\,[p_1(z), p_2(z), \dots, p_n(z)]$$
$$Q(z) = \mathrm{diag}\,[q_1(z), q_2(z), \dots, q_n(z)]$$

with $p_i(z)$ and $q_i(z)$ being coprime $\forall i = 1, \dots, n$.

Putting the expression of $C(z)$ given by (22) into (20) and (21) leads to

$$Q^{-1}(z)P(z)U^{-1}(z)\overline{A}(z) = V(z)A(z), \qquad Q^{-1}(z)P(z)U^{-1}(z)\overline{B}(z) = V(z)B(z). \tag{23}$$

Noticing that the right-hand sides of both equalities in (23) are matrix polynomials, we find that the ith rows of both $P(z)U^{-1}(z)\overline{A}(z)$ and $P(z)U^{-1}(z)\overline{B}(z)$ must be divided by q_i $\forall i = 1, \dots, n$.

Noticing that q_i and p_i are coprime $\forall i = 1, \ldots, n$, we find that $Q(z)$ must be a common left factor of $U^{-1}(z)\bar{A}(z)$ and $U^{-1}(z)\bar{B}(z)$. In other words, both $Q^{-1}(z)U^{-1}(z)\bar{A}(z)$ and $Q^{-1}(z)U^{-1}(z)\bar{B}(z)$ are matrix polynomials. Noticing that $Q^{-1}(z)$ and $P(z)$ in (20) are commutative, we find that $P(z)$ is a left-common factor of $V(z)A(z)$ and $V(z)B(z)$.

Since $A(z)$ and $B(z)$ have no common left factor, there are matrix polynomials $M(z)$ and $N(z)$ such that $A(z)M(z) + B(z)N(z) = I$, and hence $V(z)A(z)M(z)V^{-1}(z) + V(z)B(z)N(z)V^{-1}(z) = I$. This means that $V(z)A(z)$ and $V(z)B(z)$ have neither a common left factor. Consequently, $P(z)$ is unimodular. Then, from (22) it is seen that $C(z)$ is a matrix polynomial: $C(z) \triangleq C_0 + C_1 z + \cdots + C_r z^r$.

From (20) and (21) we have $[\bar{A}(z)\ \bar{B}(z)] = C(z)[A(z)\ B(z)]$. Comparing the matrix coefficients of the highest order at both sides of $\bar{A}(z) = C(z)A(z)$ and $\bar{B}(z) = C(z)B(z)$, respectively, gives us $C_r[A_p\ B_q] = 0$. By H2 $[A_p\ B_q]$ is of row-full-rank, so $C_r = 0$.

Similarly, we derive $C_i = 0,\ i \geq 1$. Therefore, $C(z)$ is a constant matrix: $C(z) = C_0$. Setting $z = 0$ in (20), we find $C(z) \equiv I$.

Then, from (20) and (21) we conclude that $\bar{A}(z) \equiv A(z)$ and $\bar{B}(z) \equiv B(z)$. Thus, H1 holds.

H1 $\Rightarrow$ H2

Let $[A(z)\ \ B(z)]$ be the unique pair in $\mathcal{M}$. Assume the converse: either $[A_p\ B_q]$ is not of row-of-rank or $A(z)$ and $B(z)$ have a common-left factor.

In the case $[A_p\ B_q]$ is not of row-of-rank, there exists a nonzero vector $\alpha \in \mathbb{R}^n$ such that

$$\alpha^T[A_p\ B_q] = 0.$$

Set $K(z) \triangleq I + \beta\alpha^T z$, where β is a nonzero vector $\beta \in \mathbb{R}^n$ with sufficiently small $\|\beta\|$ so that $K(z)$ is stable. Then, $K(z)A(z)$ and $K(z)B(z)$ are of orders less than or equal to p and q, respectively. Therefore, $[K(z)A(z)\ K(z)B(z)] \in \mathcal{M}$. This contradicts the uniqueness of $[A(z)\ \ B(z)]$ in $\mathcal{M}$, and proves the row-full-rank of $[A_p\ B_q]$.

In the case $A(z)$ and $B(z)$ have a common-left factor $C(z):\ A(z) = C(z)\bar{A}(z),\ \ B(z) = C(z)\bar{B}(z)$.

Let $U(z)$ be the unimodular matrix defined in Lemma 1. Then, $C'(z) \triangleq C(z)U(z)$ is of the form (17) with $\deg c'_{i,s}(z) \leq \deg c'_{i,i}(z)\ \forall s \leq i\ \forall i:\ \ i = 1, \ldots, n$. Define $A'(z) \triangleq U^{-1}(z)\bar{A}(z)$ and $B'(z) \triangleq U^{-1}(z)\bar{B}(z)$. Then $A(z) = C'(z)A'(z)$ and $B(z) = C'(z)B'(z)$. By Lemma 1, we have $\deg A'(z) \leq p,\ \ \deg B'(z) \leq q$. Since $I = C'(0)A'(0)$, we may assume that both $C'(z)$ and $A'(z)$ are monic and stable. Therefore, $[A'(z)\ B'(z)] \in \mathcal{M}$. However, this contradicts the uniqueness of $[A(z)\ B(z)]$ in $\mathcal{M}$. So, H2 holds.

H1 $\Rightarrow$ H3

Let H1 hold and let $[A(z)\ B(z)] \in \mathcal{M}$ with orders p and q, respectively.

We now show H3. Assume the converse that there exist n-vector polynomial $d(z)$ and m-vector polynomial $c(z)$ (not both zero) with orders strictly less than p and q, respectively, such that $d^T(z)H(z) + c^T(z) = 0$.

Let $\xi \in \mathbb{R}^n \neq 0$ and define $\widetilde{A}(z) \triangleq A(z) + z\xi d^T(z) = I + \widetilde{A}_1 z + \cdots + \widetilde{A}_p z^p,\ \ \widetilde{B}(z) \triangleq B(z) - z\xi c^T(z) = \widetilde{B}_0 + \widetilde{B}_1 z + \cdots + \widetilde{B}_q z^q$. Then, we have

$$\widetilde{A}(z)H(z) = (A(z) + z\xi d^T(z))H(z) = B(z) - z\xi c^T(z) = \widetilde{B}(z). \tag{24}$$

It is clear that $\widetilde{A}(z)$ remains stable if $\|\xi\| > 0$ is small enough. Therefore, $[\widetilde{A}(z)\ \ \widetilde{B}(z)] \in \mathcal{M}$. This contradicts H1, and hence H3 holds.

H3 $\Rightarrow$ H1

Let H3 hold. We now show H1. Assume the converse: there are two different matrix polynomials $[A(z)\ B(z)] \in \mathcal{M}$ and $[\bar{A}(z)\ \bar{B}(z)] \in \mathcal{M}$ with orders less than or equal to p and q, respectively.

Set $X(z) \triangleq A(z) - \bar{A}(z) = X_1 z + \cdots + X_p z^p$ and $Y(z) \triangleq B(z) - \bar{B}(z) = Y_0 + Y_1 z + \cdots + Y_q z^q$. From here it follows that

$$X(z)H(z) = Y(z). \tag{25}$$

Setting $z = 0$ in (25) we find that $Y_0 = 0$. By assumption there exists at least one nonzero row in (25). Take any nonzero row in $[X(z)\ Y(z)]$ and write it as $z[d^T(z)\ \ -c^T(z)]$. It is clear that $[d^T(z)\ \ -c^T(z)]$ is a row polynomial with orders strictly less than p and q, respectively. By (25) we have $d^T(z)H(z) + c^T(z) = 0$, which contradicts H3.

The proof of the lemma is completed. $\square$

3. Row-full-rank of Hankel matrix *L*

In Introduction it has been pointed out that for identification of linear models the row-full-rank of certain Hankel matrices is of crucial importance. We now present the necessary and sufficient conditions for the row-full-rank of L.

Theorem 1. *Assume $[A(z)\ B(z)] \in \mathcal{M}$. Then the following condition* H4 *is equivalent to* H1*, or* H2*, or* H3 *defined in Lemma 2.*

H4 *The matrix L defined by (10) is of row-full-rank.*

Proof. By Lemma 2, H1–H3 are equivalent. So, for the theorem it suffices to show that H3 and H4 are equivalent.

H3 $\Rightarrow$ H4

Assume $[A(z)\ \ B(z)]$ is the unique pair in $\mathcal{M}$ with orders p and q, respectively. Then, we have that

$$H(z) = A^{-1}(z)B(z) = \frac{A^*(z)B(z)}{a(z)} = \frac{B^*(z)}{a(z)}, \tag{26}$$

where $a(z) \triangleq \det(A(z)) = \sum_{i=0}^{np} a_i z^i$, $A^*(z)$ is the adjoint matrix of $A(z)$, and $B^*(z) \triangleq A^*(z)B(z) = \sum_{j=0}^{(n-1)p+q} B_j^* z^j$.

From (26) it follows that

$$(1 + a_1 z + \cdots + a_{np} z^{np})(H_0 + H_1 z + \cdots + H_i z^i + \cdots) = B_0^* + B_1^* z + \cdots + B_{(n-1)p+q}^* z^{(n-1)p+q}. \tag{27}$$

Identifying coefficients for the same degrees of z at both sides of (27), we obtain

$$H_t = -\sum_{i=1}^{np} a_i H_{t-i} \quad \forall t > q + (n-1)p. \tag{28}$$

If the matrix L were not of row-full-rank, then there would exist a vector $x = [x_1^T, \ldots, x_p^T]^T \neq 0$ with $x_i \in \mathbb{R}^n$ such that $x^T L = 0$, i.e.,

$$\sum_{j=1}^{p} x_j^T H_{q-j+l} = 0 \quad \forall 1 \le l \le np. \tag{29}$$

In this case we show that (29) holds $\forall l \ge 1$.

Noticing (28) and (29), for $l = np + 1$ we have

$$\begin{aligned} \sum_{j=1}^{p} x_j^T H_{q-j+np+1} &= -\sum_{j=1}^{p} x_j^T \sum_{i=1}^{np} a_i H_{q-j+np+1-i} \\ &= -\sum_{i=1}^{np} a_i \sum_{j=1}^{p} x_j^T H_{q-j+np+1-i} = 0. \end{aligned} \tag{30}$$

Hence (29) holds for $l = np + 1$. Carrying out a similar treatment as that done in (30), we find

$$\sum_{j=1}^{p} x_j^T H_{q-j+l} = 0 \quad \forall l \ge 1. \tag{31}$$

Defining $d(z) \triangleq \sum_{i=1}^{p} x_i z^{i-1}$, we have

$$\begin{aligned} d^T(z)H(z) &= \left(\sum_{i=1}^{p} x_i^T z^{i-1}\right) \cdot \left(\sum_{j=0}^{\infty} H_j z^j\right) \\ &= \sum_{i=1}^{p} \left(x_i^T z^{i-1} \left(\sum_{j=0}^{q-i} H_j z^j + \sum_{j=q-i+1}^{\infty} H_j z^j \right)\right) \\ &= \sum_{i=1}^{p} \sum_{j=0}^{q-i} x_i^T H_j z^{i+j-1} + \sum_{k=1}^{\infty} \left(\sum_{i=1}^{p} x_i^T H_{q-i+k} \right) z^{q+k-1} \\ &= \sum_{i=1}^{p} \sum_{j=0}^{q-i} x_i^T H_j z^{i+j-1} \triangleq -c^T(z). \end{aligned} \tag{32}$$

Consequently, $d^T(z)H(z) + c^T(z) = 0$ and the orders of $d(z)$ and $c(z)$ are strictly less than p and q, respectively. This contradicts H3, and hence H4 holds.

H4 $\Rightarrow$ H3

Assume the converse: there exist $\widetilde{d}(z) = \sum_{i=1}^{p} \widetilde{x}_i z^{i-1}$ and $\widetilde{c}(z)$ (not both zero) with orders strictly less than p and q, respectively, such that

$$\widetilde{d}^T(z)H(z) = \widetilde{c}^T(z). \tag{33}$$

Noticing

$$\widetilde{d}^T(z)H(z) = \sum_{i=1}^{p}\sum_{j=0}^{\infty}\widetilde{x}_i^T H_j z^{i+j-1} = \sum_{j=0}^{\infty}\left(\sum_{i=1}^{p}\widetilde{x}_i^T H_{j-i+1}\right)z^j,$$

we have

$$\sum_{j=0}^{\infty}\sum_{i=1}^{p}\widetilde{x}_i^T H_{j-i+1}z^j = \widetilde{c}^T(z). \tag{34}$$

Noting the order of $\widetilde{c}^T(z)$ is less than q, from (34) we must have

$$\sum_{j=q}^{\infty}\sum_{i=1}^{p}\widetilde{x}_i^T H_{j-i+1}z^j = 0,$$

which implies

$$\sum_{i=1}^{p}\widetilde{x}_i^T H_{q-i+l} = 0 \quad \forall l \geq 1.$$

This means that the rows of the matrix L are linearly dependent, which contradicts H4. Consequently, H3 must be held. □

Remark 1. For L to be of row-full-rank, the sufficient condition was discussed in [11,14,8,9]. In contrast to this, Theorem 1 states the necessary and sufficient conditions for the row-full-rank of L. It is worth noting that θ_A and $\theta_B \triangleq [B_0, B_1, \ldots, B_q]$ can be estimated by Theorem 1 with the help of (12) and (7), once the estimates for $\{H_i\}$ have been derived (see, e.g., [14,8,9]).

4. Row-full-rank of Hankel matrix Γ

We now consider the row-full-rank of the Hankel matrix Γ composed of correlation functions. Assume that $\{u_k\}$ is a sequence of zero mean uncorrelated random vectors with $Eu_k u_k^T = I$.

If Γ is of row-full-rank and $\{R_i\}$ can be estimated, then by (16), θ_A can also be estimated. As concerns the coefficients θ_B, let us set $\chi_k \triangleq B(z)u_k$.

The spectral density of χ_k is

$$\Phi^{\chi}(z) = B(z)B^T(z^{-1}),$$

while the spectral density of $\{y_k\}$ given by (1)–(3) is

$$\Phi(z) \triangleq \sum_{j=-\infty}^{\infty} R_j z^j = A^{-1}(z)B(z)B^T(z^{-1})A^{-T}(z^{-1}), \tag{35}$$

which implies

$$\Phi^{\chi}(z) = B(z)B^T(z^{-1}) = A(z)\Phi(z)A^T(z^{-1}). \tag{36}$$

Since the right-hand side of (36) is equal to

$$\begin{aligned} A(z)\Phi(z)A^T(z^{-1}) &= \sum_{i=0}^{p}A_i z^i \sum_{k=-\infty}^{\infty}R_k z^k \sum_{j=0}^{p}A_j^T z^{-j} \\ &= \sum_{i=0}^{p}\sum_{k=-\infty}^{\infty}\sum_{j=0}^{p}A_i R_k A_j^T z^{i+k-j} = \sum_{k=-\infty}^{\infty}\left(\sum_{i=0}^{p}\sum_{j=0}^{p}A_i R_{k+j-i}A_j^T\right)z^k, \end{aligned} \tag{37}$$

we have

$$B(z)B^T(z^{-1}) = \sum_{k=-q}^{q}\left(\sum_{i=0}^{p}\sum_{j=0}^{p}A_i R_{k+j-i}A_j^T\right)z^k. \tag{38}$$

Therefore, to derive θ_B it is the matter of factorizing the right-hand-side of (38).

The following theorem tells us that the row-full-rank of Γ is slightly stronger than that of L.

It is worth noting that the theorem requires no stability-like condition on $B(z)$, and $B(z)$ is even allowed to not be a square matrix.

Theorem 2. *Assume $[A(z)\ \ B(z)] \in \mathcal{M}$ and B_0 is of column-full-rank. Then, the following* H5 *and* H6 *are equivalent.*

H5 *The matrix Γ defined by* (15) *is of row-full-rank.*

H6 *The matrix $[A_p\ B_q]$ is of row-full-rank and the matrix polynomials $A(z)$ and $B(z)B^T(z^{-1})z^q$ have no common left factor.*

Proof. H6 $\Rightarrow$ H5

We first note that H6 implies H2, and hence by Lemma 2, $[A(z)\ B(z)]$ is the unique pair in $\mathcal{M}$ under H6. So, it suffices to show that $[A(z)\ B(z)]$ cannot be unique if H5 is not true. In other words, we intend to show that there exists a pair $[\widehat{A}(z)\ \widehat{B}(z)] \in \mathcal{M}$ such that $[\widehat{A}(z)\ \widehat{B}(z)] \neq [A(z)\ B(z)]$, if Γ is not of row-full-rank.

The proof of H6 $\Rightarrow$ H5 is completed by four steps.

Step 1. We first show that the converse assumption leads to

$$\sum_{j=1}^{p} \eta_j^T R_{q-j+l} = 0 \quad \forall l \geq 1, \tag{39}$$

where the column vector $\eta = [\eta_1^T, \ldots, \eta_p^T]^T \neq 0$ with $\eta_i \in \mathbb{R}^n$ is such that $\eta^T \Gamma = 0$, i.e.,

$$\sum_{j=1}^{p} \eta_j^T R_{q-j+l} = 0, \quad 1 \leq l \leq np. \tag{40}$$

The existence of such an η is a consequence of the converse assumption that Γ is not of row-full-rank.

By (26), we have

$$y_k + a_1 y_{k-1} + \cdots + a_{np} y_{k-np} = B_0^* u_k + B_1^* u_{k-1} + \cdots + B^*_{(n-1)p+q} u_{k-((n-1)p+q)}. \tag{41}$$

Multiplying both sides of (41) by y_{k-t}^T from right and taking expectation, we have

$$\begin{aligned} E(y_k + a_1 y_{k-1} + \cdots + a_{np} y_{k-np}) y_{k-t}^T &= E(B_0^* u_k + B_1^* u_{k-1} + \cdots + B^*_{(n-1)p+q} u_{k-((n-1)p+q)}) y_{k-t}^T \\ &= 0, \quad \forall\, t > q + (n-1)p, \end{aligned}$$

which yields

$$R_t = -\sum_{i=1}^{np} a_i R_{t-i}, \quad \forall t > q + (n-1)p. \tag{42}$$

Noticing (42), we have

$$\begin{aligned} \sum_{j=1}^{p} \eta_j^T R_{q-j+np+1} &= -\sum_{j=1}^{p} \eta_j^T \sum_{i=1}^{np} a_i R_{q-j+np+1-i} \\ &= -\sum_{i=1}^{np} a_i \sum_{j=1}^{p} \eta_j^T R_{q-j+np+1-i} = 0, \end{aligned} \tag{43}$$

where the last equality follows from (40). Therefore, (39) holds for $l = np + 1$. Carrying out a similar treatment as that done in (43), we arrive at (39).

Step 2. We now define $[\widehat{A}(z)\ \widehat{B}(z)]$, and show $[\widehat{A}(z)\ \ \widehat{B}(z)] \in \mathcal{M}$.

Set

$$\widehat{A}(z) \triangleq A(z) + \beta \widehat{d}^T(z), \tag{44}$$

$$\widehat{B}(z) \triangleq \widehat{A}(z) A^{-1}(z) B(z) = \widehat{A}(z) H(z), \tag{45}$$

where $\widehat{d}(z) \triangleq \sum_{i=1}^{p} \eta_i z^i$ and $\beta \in \mathbb{R}^n$ is an arbitrary nonzero column vector with small enough $\|\beta\|$ such that $\widehat{A}(z)$ is stable.

It is clear that $\widehat{A}(z) \neq A(z)$ and $\deg \widehat{A}(z) \leq p$. From (45) it is seen that $[A(z)\ B(z)]$ and $[\widehat{A}(z)\ \widehat{B}(z)]$ share the same transfer function $H(z)$.

Therefore, to show $[\widehat{A}(z)\ \widehat{B}(z)] \in \mathcal{M}$ it suffices to prove that $\widehat{B}(z)$ is a matrix polynomial and $\deg \widehat{B}(z) \leq q$. If this is done, then the uniqueness is violated. This means that H5 must be held.

Step 3. Before doing this we first prove that $C(z) \triangleq \widehat{A}(z)A^{-1}(z)$ is a matrix polynomial.

For this, we explicitly express each terms in $\widehat{A}(z)\Phi(z)\widehat{A}^T(z^{-1})$.

By the definition of $\Phi(z)$ given by (35) we have

$$\begin{aligned}
\widehat{d}^T(z)\Phi(z)A^T(z^{-1}) &= \sum_{i=1}^{p}\sum_{j=-\infty}^{\infty}\sum_{l=0}^{p}\eta_i^T R_j A_l^T z^{i+j-l} \\
&= \sum_{i=1}^{p}\sum_{l=0}^{p}\left(\sum_{j=q-i+1}^{\infty}\eta_i^T R_j A_l^T z^{i+j-l} + \sum_{j=-\infty}^{q-i}\eta_i^T R_j A_l^T z^{i+j-l}\right) \\
&= \sum_{i=1}^{p}\sum_{l=0}^{p}\sum_{j=q-i+1}^{\infty}\eta_i^T R_j A_l^T z^{i+j-l} + \sum_{i=1}^{p}\sum_{l=0}^{p}\sum_{j=-\infty}^{q-i}\eta_i^T R_j A_l^T z^{i+j-l} \qquad (46) \\
&= \sum_{l=0}^{p}\sum_{j=1}^{\infty}\left(\sum_{i=1}^{p}\eta_i^T R_{q-i+j}\right) z^{q+j}A_l^T z^{-l} + \sum_{i=1}^{p}\sum_{l=0}^{p}\sum_{j=-\infty}^{q-i}\eta_i^T R_j A_l^T z^{i+j-l} \\
&= \sum_{i=1}^{p}\sum_{l=0}^{p}\sum_{j=i-q}^{\infty}\eta_i^T R_j^T A_l^T z^{i-j-l} \\
&= \sum_{i=1}^{p}\sum_{l=0}^{p}\sum_{j=i-q}^{q-l}\eta_i^T R_j^T A_l^T z^{i-j-l} + \sum_{i=1}^{p}\sum_{l=0}^{p}\sum_{j=q-l+1}^{\infty}\eta_i^T R_j^T A_l^T z^{i-j-l} \\
&= \sum_{i=1}^{p}\sum_{l=0}^{p}\sum_{j=i-q}^{q-l}\eta_i^T R_j^T A_l^T z^{i-j-l} + \sum_{i=1}^{p}\eta_i^T z^i\sum_{j=1}^{\infty}\left(\sum_{l=0}^{p}R_{q-l+j}^T A_l^T\right) z^{-q-j} \\
&= \sum_{i=1}^{p}\sum_{l=0}^{p}\sum_{j=i-q}^{q-l}\eta_i^T R_j^T A_l^T z^{i-j-l}, \qquad (47)
\end{aligned}$$

where for the fifth equality (39) is invoked, while for the last equality (13) is used.

Similarly, we obtain

$$A(z)\Phi(z)\widehat{d}(z^{-1}) = \sum_{l=0}^{p}\sum_{i=1}^{p}\sum_{j=l-q}^{q-i}A_l R_j^T \eta_i z^{l-i-j} \qquad (48)$$

and

$$\begin{aligned}
\widehat{d}^T(z)\Phi(z)\widehat{d}(z^{-1}) &= \sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{j=-\infty}^{\infty}\eta_i^T R_j \eta_l z^{i+j-l} \\
&= \sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{j=q-i+1}^{\infty}\eta_i^T R_j \eta_l z^{i+j-l} + \sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{j=-\infty}^{q-i}\eta_i^T R_j \eta_l z^{i+j-l} \\
&= \sum_{l=1}^{p}\sum_{j=1}^{\infty}\left(\sum_{i=1}^{p}\eta_i^T R_{q-i+j}\right) z^{q+j}\eta_l z^{-l} + \sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{j=-\infty}^{q-i}\eta_i^T R_j \eta_l z^{i+j-l} \\
&= \sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{j=i-q}^{\infty}\eta_i^T R_j^T \eta_l z^{i-j-l} \\
&= \sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{j=i-q}^{q-l}\eta_i^T R_j^T \eta_l z^{i-j-l} + \sum_{i=1}^{p}\eta_i^T z^i\sum_{j=1}^{\infty}\left(\sum_{l=1}^{p}R_{q-l+j}^T \eta_l\right) z^{-q-j} \\
&= \sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{j=i-q}^{q-l}\eta_i^T R_j^T \eta_l z^{i-j-l}. \qquad (49)
\end{aligned}$$

From (35) and (47)–(49) it follows that

$$\begin{aligned}
\widehat{A}(z)\Phi(z)\widehat{A}^T(z^{-1}) &= (A(z) + \beta\widehat{d}^T(z))\Phi(z)(A(z^{-1}) + \beta\widehat{d}^T(z^{-1}))^T \\
&= A(z)\Phi(z)A^T(z^{-1}) + A(z)\Phi(z)\widehat{d}(z^{-1})\beta^T + \beta\widehat{d}^T(z)\Phi(z)A^T(z^{-1}) + \beta\widehat{d}^T(z)\Phi(z)\widehat{d}(z^{-1})\beta^T
\end{aligned}$$

$$= B(z)B^T(z^{-1}) + \sum_{l=0}^{p}\sum_{i=1}^{p}\sum_{j=l-q}^{q-i} A_l R_j^T y_i z^{l-i-j}\beta^T + \sum_{i=1}^{p}\sum_{l=0}^{p}\sum_{j=i-q}^{q-l} \beta y_i^T R_j^T A_l^T z^{i-j-l}$$
$$+ \sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{j=i-q}^{q-l} \beta y_i^T R_j^T y_l \beta^T z^{i-j-l} \triangleq F(z). \tag{50}$$

The degrees of z in $F(z)$ are between $-q$ and q. So, it may diverge to infinity only at $z = 0$ and at z equal to infinity, and hence all its nonzero finite poles should be canceled with its zeros.

Noticing

$$F(z) = \widehat{A}(z)\Phi(z)\widehat{A}^T(z^{-1}) = \widehat{A}(z)A^{-1}(z)B(z)B^T(z^{-1})A^{-T}(z^{-1})\widehat{A}^T(z^{-1}), \tag{51}$$

we see that all poles of $A^{-1}(z)$ should be canceled with zeros of $F(z)$. However, H6 requires that $A(z)$ and $B(z)B^T(z^{-1})z^q$ have no common left factor. This means that any pole of $A^{-1}(z)$ cannot be canceled with zeros of $B(z)B^T(z^{-1})z^q$. By stability of $A(z)$ the poles of $A^{-1}(z)$, being outside the closed unit disk, can neither be canceled with zeros of $A^{-T}(z^{-1})$ and $\widehat{A}^T(z^{-1})$, since their zeros are inside the unit disk. Therefore, all poles of $A^{-1}(z)$ must be canceled with zeros of $\widehat{A}(z)$. In other words, $C(z) = \widehat{A}(z)A^{-1}(z)$ must be a matrix polynomial:

$$C(z) = C_0 + C_1 z + \cdots + C_r z^r.$$

Step 4. We now complete the proof of H6 $\Rightarrow$ H5.

From definition, we have $\widehat{A}(z) = C(z)A(z)$, which leads to $C_0 = I$ by setting $z = 0$.

Further, $\widehat{B}(z) = C(z)B(z)$ is also a matrix polynomial denoted by $\widehat{B}(z) = \widehat{B}_0 + \widehat{B}_1 z + \cdots + \widehat{B}_{\widehat{q}} z^{\widehat{q}}$, and we have $\widehat{B}_0 = B_0$ since $C_0 = I$. By (45) and (51) we have

$$\widehat{B}(z)\widehat{B}^T(z^{-1}) = F(z). \tag{52}$$

It remains to show that $\deg \widehat{B}(z) \le q$. If $\widehat{q} > q$ and $\widehat{B}_{\widehat{q}} \ne 0$, then comparing the matrix coefficients of the degree $\widehat{q}$ at both sides of (52) we obtain $\widehat{B}_{\widehat{q}} B_0^T = 0$, since the maximal degree of z in $F(z)$ defined by (50) is q. Since B_0 is of column-full-rank, we find that $\widehat{B}_{\widehat{q}} = 0$. A similar treatment for $q+1 \le s \le \widehat{q}-1$ leads to $\widehat{B}_s = 0$, $q+1 \le s \le \widehat{q}$ in $\widehat{B}(z)$. Thus, we have proved that $\deg \widehat{B}(z) \le q$, and at the same time $[\widehat{A}(z)\ \ \widehat{B}(z)] \in \mathcal{M}$. The violation of uniqueness implies that Γ is of row-full-rank.

H5 $\Rightarrow$ H6

The proof of H5 $\Rightarrow$ H6 is completed by three steps.

Step 1. We first show that $[A_p\ B_q]$ is of row-full-rank.

If $[A_p\ B_q]$ is not of row-full-rank, then $[A(z)\ \ B(z)]$ is not unique in $\mathcal{M}$ by Lemma 2. Then, the matrix L is not of row-full-rank by Theorem 1. This means that there exists a nonzero column vector $\widetilde{x} = [\widetilde{x}_1^T, \ldots, \widetilde{x}_p^T]^T$ such that

$$\sum_{i=1}^{p} \widetilde{x}_i^T H_{q-i+l} = 0 \quad \forall 1 \le l \le np.$$

From here as shown in (29)–(31), we have

$$\sum_{i=1}^{p} \widetilde{x}_i^T H_{q-i+l} = 0 \quad \forall l \ge 1.$$

Therefore, for any $l \ge 1$ we have

$$\sum_{i=1}^{p} \widetilde{x}_i^T R_{q-i+l} = \sum_{i=1}^{p} \widetilde{x}_i^T \sum_{j=0}^{\infty} H_{q-i+l+j} H_j^T = \sum_{j=0}^{\infty} \left(\sum_{i=1}^{p} \widetilde{x}_i^T H_{q-i+l+j} \right) H_j^T = 0,$$

which means that the rows of the matrix Γ are linearly dependent. This contradicts H5.

Consequently, $[A_p\ B_q]$ is of row-full-rank.

Step 2. We explain how to prove that $A(z)$ and $B(z)B^T(z^{-1})z^q$ have no common left factor.

Assume the converse: $A(z)$ and $B(z)B^T(z^{-1})z^q$ have a common left factor, i.e., there exists $C(z)$ being not unimodular such that

$$[A(z)\ B(z)B^T(z^{-1})z^q] = C(z)[\overline{A}(z)\ \overline{D}(z)]. \tag{53}$$

Applying Lemma 1 to (53) we may assume that $\deg[\overline{A}(z)] \le p$, and the matrix $C(z) = \{c_{i,j}(z)\}_{1 \le i,j \le n}$ is lower triangular with $c_{ii}(0) = 1$ and the degree of $c_{i,i}(z)$ is the greatest among the entries of the ith row $\forall i : 1 \le i \le n$. It is clear that $A(z) \ne \overline{A}(z)$.

If we can show that there is $\widetilde{B}(z)$ with $\deg \widetilde{B}(z) \le q$ so that $[\widetilde{A}(z)\ \widetilde{B}(z)]$ and $[A(z)\ B(z)]$ have the same correlation functions $\{R_i\}$, where $\widetilde{A}(z) \triangleq \overline{A}(0)^{-1}\overline{A}(z)$ with $\widetilde{A}(0) = I$, then by (14) this implies

$$[A_1, A_2, \ldots, A_p]\Gamma = -[R_{q+1}, R_{q+2}, \ldots, R_{q+np}], \tag{54}$$

$$[\widetilde{A}_1, \widetilde{A}_2, \ldots, \widetilde{A}_p]\Gamma = -[R_{q+1}, R_{q+2}, \ldots, R_{q+np}], \tag{55}$$

where $\widetilde{A}_i,\ 1 \le i \le p$ are coefficients of $\widetilde{A}(z)$. From here it follows that

$$[A_1 - \widetilde{A}_1, A_2 - \widetilde{A}_2, \ldots, A_p - \widetilde{A}_p]\Gamma = 0. \tag{56}$$

Since the matrix $[A_1 - \widetilde{A}_1, A_2 - \widetilde{A}_2, \ldots, A_p - \widetilde{A}_p]$ is not identically zero, the matrix Γ cannot be of row-full-rank. This contradicts H5, and hence, completes the proof of the theorem.

Thus, the remaining task of the proof is to find $\widetilde{B}(z)$ with property mentioned above.

From (53) we have

$$B(z)B^T(z^{-1})z^q = C(z)\overline{D}(z). \tag{57}$$

Set $\widetilde{D}(z) \triangleq \overline{D}(z)C^{-T}(z^{-1})z^{-q} = \{\widetilde{d}_{i,j}(z), 1 \le i, j \le n\}$. Then, from (57) it follows that

$$\widetilde{D}(z) = C^{-1}(z)B(z)B^T(z^{-1})C^{-T}(z^{-1}), \tag{58}$$

which is equivalent to

$$D(z) \triangleq B(z)B^T(z^{-1}) = C(z)\widetilde{D}(z)C^T(z^{-1}). \tag{59}$$

If we can show that $\widetilde{D}(z) = \sum_{i=-q}^{q} \widetilde{D}_i z^i$ with $\widetilde{D}_{-i} = \widetilde{D}_i^T$ is of rank m $(m \le n)$, and is non-negative definite on the unit circle $|z| = 1$, then by Lemma 5 given in Appendix there exists an $n \times m$ matrix polynomial $\overline{B}(z)$ with $\deg[\overline{B}(z)] \le q$ such that $\widetilde{D}(z) = \overline{B}(z)\overline{B}^T(z^{-1})$.

By (53) and (59) we have

$$A(z) = C(z)\overline{A}(z) \quad \text{and} \quad B(z)B^T(z^{-1}) = C(z)\overline{B}(z)\overline{B}^T(z^{-1})C^T(z^{-1}) \tag{60}$$

with $\deg \overline{A}(z) \le p$ and $\deg \overline{B}(z) \le q$.

Define $\widetilde{C}(z) \triangleq C(z)\overline{A}(0)$ and $\widetilde{B}(z) \triangleq \overline{A}(0)^{-1}\overline{B}(z)$. Then we have

$$A(z) = \widetilde{C}(z)\widetilde{A}(z) \quad \text{and} \quad B(z)B^T(z^{-1}) = \widetilde{C}(z)\widetilde{B}(z)\widetilde{B}^T(z^{-1})\widetilde{C}^T(z^{-1}), \tag{61}$$

which yields

$$\begin{aligned} A^{-1}(z)B(z)B^T(z^{-1})A^{-T}(z^{-1}) &= [\widetilde{C}(z)\widetilde{A}(z)]^{-1}B(z)B^T(z^{-1})[\widetilde{C}(z^{-1})\widetilde{A}(z^{-1})]^{-T} \\ &= \widetilde{A}(z)^{-1}\widetilde{C}^{-1}(z)B(z)B^T(z^{-1})\widetilde{C}^{-T}(z^{-1})\widetilde{A}^{-T}(z^{-1}) \\ &= \widetilde{A}(z)^{-1}\widetilde{B}(z)\widetilde{B}^T(z^{-1})\widetilde{A}^{-T}(z^{-1}). \end{aligned} \tag{62}$$

This means that the two different linear systems $[A(z)\ B(z)]$ and $[\widetilde{A}(z)\ \widetilde{B}(z)]$ share the same spectral density, and hence they have the same correlation functions. This will prove the theorem.

Step 3. To complete the proof we now show that $\widetilde{D}(z) = \sum_{i=-q}^{q} \widetilde{D}_i z^i$ with $\widetilde{D}_{-i} = \widetilde{D}_i^T$ is of rank m $(m \le n)$, and is non-negative definite on the unit circle $|z| = 1$.

For any scalar rational polynomial $g(z) = g_{-a}z^{-a} + \cdots + g_0 + \cdots + g_b z^b$ with real coefficients we introduce the operators $[\cdot]^+$ and $[\cdot]^-$ such that

$$[g(z)]^+ = g_0 + \cdots + g_b z^b \quad \text{and} \quad [g(z)]^- = g_0 + g_{-1}z + \cdots + g_{-a}z^a.$$

The essential step is to show that $\deg[\widetilde{d}_{i,j}(z)]^+ \le q,\ \deg[\widetilde{d}_{i,j}(z)]^- \le q$ for $1 \le i, j \le n$. This is done by a treatment similar to but more complicated than that used in the proof of Lemma 1. We prove this inductively starting from the first column, and in each column the proof is also carried out inductively.

Noticing that $C(z)$ is lower triangular, from (59) we have

$$D(z) = \{d_{ij}(z)\}, \qquad d_{ij}(z) = \sum_{t=1}^{i}\sum_{s=1}^{j} c_{it}(z)\widetilde{d}_{ts}(z)c_{js}(z^{-1}) \quad \forall i, j: 1 \le i, j \le n.$$

Starting from the first column of $\widetilde{D}(z)$, we show that $\deg[\widetilde{d}_{i1}(z)]^+ \le q$ for $1 \le i \le n$ by induction.

The (1, 1)-element $\widetilde{d}_{11}(z)$ of $\widetilde{D}(z)$ is related to $d_{11}(z)$ as follows

$$d_{11}(z) = c_{11}(z)\widetilde{d}_{11}(z)c_{11}(z^{-1}). \tag{63}$$

Since $\deg[d_{11}(z)]^+ \le q$ and the constant term of $c_{11}(z)$ equals 1, we see that $\deg[\widetilde{d}_{11}(z)]^+ \le q$.

Assume it has been established that $\deg[\widetilde{d}_{i1}(z)]^+ \le q\ \forall i : 1 \le i \le r$. We want to show $\deg[\widetilde{d}_{(r+1)1}(z)]^+ \le q$.

Assume the converse: $\deg[\widetilde{d}_{(r+1)1}(z)]^+ > q$.

Noticing $\deg[\widetilde{d}_{i1}(z)]^+ \le q\ \forall i : 1 \le i \le r$ and $\deg[c_{(r+1)(r+1)}(z)] \ge \deg[c_{(r+1)t}(z)]\ \forall t : 1 \le t \le r$, by the converse assumption and $c_{11}(0) = 1$ we see that

$$\deg\left[c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)1}(z)c_{11}(z^{-1})\right]^+ > \deg\left[\sum_{t=1}^{r} c_{(r+1)t}(z)\widetilde{d}_{t1}(z)c_{11}(z^{-1})\right]^+,$$

and hence

$$\begin{aligned}&\deg\left[\sum_{t=1}^{r} c_{(r+1)t}(z)\widetilde{d}_{t1}(z)c_{11}(z^{-1}) + c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)1}(z)c_{11}(z^{-1})\right]^+\\ &\quad = \deg\left[c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)1}(z)c_{11}(z^{-1})\right]^+ > q.\end{aligned} \tag{64}$$

Since

$$d_{(r+1)1}(z) = \sum_{t=1}^{r+1} c_{(r+1)t}(z)\widetilde{d}_{t1}(z)c_{11}(z^{-1}), \tag{65}$$

by (64) we obtain a contradictory inequality:

$$\begin{aligned}q \ \ge\ \deg[d_{(r+1)1}(z)]^+ &= \deg\left[\sum_{t=1}^{r+1} c_{(r+1)t}(z)\widetilde{d}_{t1}(z)c_{11}(z^{-1})\right]^+\\ &= \deg\left[c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)1}(z)c_{11}(z^{-1})\right]^+ > q.\end{aligned} \tag{66}$$

Thus, we have proved $\deg[\widetilde{d}_{(r+1)1}(z)]^+ \le q$ and inductively $\deg[\widetilde{d}_{i1}(z)]^+ \le q\ \forall i : 1 \le i \le n$.

Similarly, we can show $\deg[\widetilde{d}_{i1}(z)]^- \le q\ \forall i : 1 \le i \le n$. Therefore, the assertion holds for the first column.

We now assume that the assertion is true for the first j columns, i.e.,

$$\deg[\widetilde{d}_{is}(z)]^+ \le q \quad \text{and} \quad \deg[\widetilde{d}_{is}(z)]^- \le q \quad \forall i : 1 \le i \le n\ \forall s : 1 \le s \le j.$$

We want to show that it also holds for the $j+1$ column.

Observing that $\widetilde{d}_{i(j+1)}(z) = \widetilde{d}_{(j+1)i}(z^{-1})$, $1 \le i \le j$, we see $\deg[\widetilde{d}_{i(j+1)}(z)]^+ = \deg[\widetilde{d}_{(j+1)i}(z)]^- \le q\ \forall i : 1 \le i \le j$ by the inductive assumption.

Inductively, we now assume that $\deg[\widetilde{d}_{i(j+1)}(z)]^+ \le q\ \forall i : 1 \le i \le r$ for some $r : r \ge j$. We want to prove $\deg[\widetilde{d}_{(r+1)(j+1)}(z)]^+ \le q$.

Assume the converse: $\deg[\widetilde{d}_{(r+1)(j+1)}(z)]^+ > q$.

Noticing $\deg[c_{(r+1)(r+1)}(z)] \le \deg[c_{(r+1)t}(z)]\ \forall t : 1 \le t \le r$ and the inductive assumptions $\deg[\widetilde{d}_{i(j+1)}(z)]^+ \le q\ \forall i : 1 \le i \le r$ and $\deg[\widetilde{d}_{is}(z)]^+ \le q\ \forall i : 1 \le i \le n\ \forall s : 1 \le s \le j$, we find that

$$\begin{aligned}&\deg\left[c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)(j+1)}(z)c_{(j+1)(j+1)}(z^{-1})\right]^+\\ &\quad > \deg\left[\sum_{t=1}^{r}\sum_{s=1}^{j+1} c_{(r+1)t}(z)\widetilde{d}_{ts}(z)c_{(j+1)s}(z^{-1}) + \sum_{s=1}^{j} c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)s}(z)c_{(j+1)s}(z^{-1})\right]^+.\end{aligned}$$

Consequently, we have

$$\begin{aligned}&\deg\Bigg[\sum_{t=1}^{r}\sum_{s=1}^{j+1} c_{(r+1)t}(z)\widetilde{d}_{ts}(z)c_{(j+1)s}(z^{-1})\\ &\qquad + \sum_{s=1}^{j} c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)s}(z)c_{(j+1)s}(z^{-1}) + c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)(j+1)}(z)c_{(j+1)(j+1)}(z^{-1})\Bigg]^+\\ &\quad = \deg\left[c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)(j+1)}(z)c_{(j+1)(j+1)}(z^{-1})\right]^+ > q.\end{aligned} \tag{67}$$

Since

$$d_{(r+1)(j+1)}(z)=\sum_{t=1}^{r+1}\sum_{s=1}^{j+1}c_{(r+1)t}(z)\widetilde{d}_{ts}(z)c_{(j+1)s}(z^{-1}),$$

by (67) we arrive at the following contradictory inequality:

$$\begin{aligned} q &\geq \deg[d_{(r+1)(j+1)}(z)]^{+}=\deg\left[\sum_{t=1}^{r+1}\sum_{s=1}^{j+1}c_{(r+1)t}(z)\widetilde{d}_{ts}(z)c_{(j+1)s}(z^{-1})\right]^{+} \\ &= \deg\left[c_{(r+1)(r+1)}(z)\widetilde{d}_{(r+1)(j+1)}(z)c_{(j+1)(j+1)}(z^{-1})\right]^{+}>q. \end{aligned} \tag{68}$$

This contradiction implies that $\deg[\widetilde{d}_{(r+1)(j+1)}(z)]^{+}\leq q$. As a consequence, we have proved that $\deg[\widetilde{d}_{i(j+1)}(z)]^{+}\leq q$ for $1\leq i\leq n$. Similarly, we can also show that $\deg[\widetilde{d}_{i(j+1)}(z)]^{-}\leq q$ for $1\leq i\leq n$.

Therefore, the assertion holds for the $j+1$ column, i.e., $\deg[\widetilde{d}_{i(j+1)}(z)]^{+}\leq q$, $\deg[\widetilde{d}_{i(j+1)}(z)]^{-}\leq q\ \forall i:1\leq i\leq n$.

As results, $\widetilde{D}(z)$ can be written as $\widetilde{D}(z)=\sum_{i=-q}^{q}\widetilde{D}_i z^i$ with $\widetilde{D}_{-i}=\widetilde{D}_i^T$.

From (58) it follows that $\widetilde{D}(z)$ is of rank m, and is non-negative on the unit circle $|z|=1$. Thus, the proof of the theorem is completed. □

In Theorem 2 the criterion for the row-full-rank of Γ is under the additional assumption that B_0 is of column-full-rank. However, this assumption can be removed with the help of Lemma 4 given in Appendix. As a matter of fact, by Lemma 4, $B(z)$ has the factorization

$$B(z)=\widetilde{B}(z)B_P(z) \tag{69}$$

where $\widetilde{B}(z)$ is an $n\times m$ matrix polynomial such that $\deg[\widetilde{B}(z)]\leq\deg[B(z)]$ and its constant term $\widetilde{B}(0)$ is of column-full-rank, while $B_P(z)$ is an $m\times m$ matrix polynomial satisfying $B_P(z)B_P^T(z^{-1})=I_m$. We write $\widetilde{B}(z)$ as

$$\widetilde{B}(z)=\widetilde{B}_0+\widetilde{B}_1 z+\cdots+\widetilde{B}_q z^q \tag{70}$$

where $\widetilde{B}_0$ is of column-full-rank and $\widetilde{B}_q$ may be equal to 0.

Theorem 2′. *Assume that $[A(z)\ B(z)]\in\mathcal{M}$ and $B(z)$ is of rank m. Then the matrix Γ defined by (15) is of row-full-rank if and only if the matrix $[A_p\ \widetilde{B}_q]$ is of row-full-rank and the matrix polynomials $A(z)$ and $B(z)B^T(z^{-1})z^q$ have no common left factor.*

Proof. We need only to discuss the case where B_0 is not of column-full-rank. By the factorization (69), we see that the two linear systems $\{A(z),B(z)\}$ and $\{A(z),\widetilde{B}(z)\}$ with different impulse responses have the same correlation functions because they have the same spectral density:

$$\begin{aligned} A^{-1}(z)B(z)B^T(z^{-1})A^{-T}(z^{-1}) &= A^{-1}(z)\widetilde{B}(z)B_P(z)B_P^T(z^{-1})\widetilde{B}^T(z^{-1})A^{-T}(z^{-1}) \\ &= A^{-1}(z)\widetilde{B}(z)\widetilde{B}^T(z^{-1})A^{-T}(z^{-1}). \end{aligned} \tag{71}$$

This means that the matrix Γ constructed from the two linear systems are identical. Therefore, we only need to consider the necessary and sufficient conditions that make the matrix Γ derived from $\{A(z),\widetilde{B}(z)\}$ be of row-full-rank. Since $\widetilde{B}_0$ is of column-full-rank, by Theorem 2 the necessary and sufficient conditions for the row-full-rank of Γ is that $[A_p\ \ \widetilde{B}_q]$ is of row-full-rank and $A(z)$ and $\widetilde{B}(z)\widetilde{B}^T(z^{-1})z^q$ have no common left factor.

By noticing

$$B(z)B^T(z^{-1})=\widetilde{B}(z)B_P(z)B_P^T(z^{-1})\widetilde{B}^T(z^{-1})=\widetilde{B}(z)\widetilde{B}^T(z^{-1}), \tag{72}$$

the conclusion of the theorem follows. □

Remark 2. For systems with stable $A(z)$ the well-known result given in [12] is for the special case $n=m$, $B_0=I$, and $\det B(z)\neq 0\ \forall|z|<1$, and it states that the matrix Γ defined by (15) is of row-full-rank if and only if $A(z)$ and $B(z)$ have no common left factor and $[A_p\ B_q]$ is of row-full-rank. We note that Theorems 2 and 2′ are for the general case $n\geq m$ and require neither $B_0=I$ nor $\det B(z)\neq 0\ \ \forall|z|<1$. It is worth noting that "$A(z)$ and $B(z)$ have no common left factor" and "$A(z)$ and $B(z)B^T(z^{-1})z^q$ have no common left factor" are equivalent for systems with stable $A(z)$, whenever $\det B(z)\neq 0$, $\forall|z|<1$. This is because "$\det B(z)\neq 0\ \forall|z|<1$" guarantees that all roots of $\det B^T(z^{-1})$ are inside or on the unit disk.

5. Conclusions

The row-full-rank of the Hankel matrices composed of impulse responses and of correlation functions is of crucial importance for determining or estimating coefficients of the corresponding linear system. Whenever identifiability is concerned, in most of existing papers ([2] may be among a few exceptions) the input and output of the system under consideration usually have the same dimension and the minimum phase condition is normally required [12,5] etc. In this paper, it is shown that such kind of restrictions are not necessary. As a matter of fact, the necessary and sufficient conditions are presented for the row-full-rank of the Hankel matrices. With these new results applied, the corresponding results for identifying a certain kind of block-oriented nonlinear systems containing linear subsystems may compatibly be improved.

Appendix

In the proof of Lemma 2 the Smith–McMillan diagonal decomposition for a square matrix polynomial is used. We formulate the decomposition below. For details we refer to [4,13].

A square matrix polynomial is called unimodular if its determinant is a nonzero constant. From the definition it follows that the inverse of a unimodular matrix is also a matrix polynomial.

The non-negative integer r is called the rank of a rational polynomial matrix if (1) there exists at least one subminor of order r which does not vanish identically, and (2) all subminors of order greater than r vanish identically.

Lemma 3 *([4,13]). Let $G(z)$ be an $n \times n$ rational matrix of rank r. Then there exist two $n \times n$ unimodular matrices $U(z)$ and $V(z)$ such that*

$$\begin{aligned} G(z) &= U(z)\mathrm{diag}\left[\frac{e_1(z)}{\psi_1(z)}, \frac{e_2(z)}{\psi_2(z)}, \ldots, \frac{e_r(z)}{\psi_r(z)}, 0 \ldots, 0\right] V(z) \\ &= U(z)W(z)V(z), \end{aligned} \tag{73}$$

where

(a) *$e_k(z)$ and $\psi_k(z)$ are relatively prime polynomials with unit leading coefficients $\forall k : 1 \le k \le r$;*
(b) *each $e_k(z)$ divides $e_{k+1}(z)$ $\forall k : 1 \le k \le r-1$, and each $\psi_j(z)$ is a factor of $\psi_{j-1}(z)$ $\forall j : 2 \le j \le r$;*
(c) *the diagonal matrix $W(z)$ appearing in (73) satisfies* (a) *and* (b), *uniquely determined by $G(z)$;*
(d) *if $G(z)$ is real, then $U(z)$, $W(z)$ and $V(z)$ may also be chosen to be real.*

Lemma 4 *([6,7]). Let $B(z) = B_0 + B_1 z + \cdots + B_q z^q$ be an $n \times m$ $(n \ge m)$ matrix polynomial with rank m. Then $B(z)$ can be factorized as*

$$B(z) = B_I(z)B_P(z) \tag{74}$$

where $B_I(z)$ is an $n \times m$ matrix polynomial such that $\deg[B_I(z)] \le \deg[B(z)]$ *and its constant term $B_I(0)$ is of column-full-rank, while $B_P(z)$ is an $m \times m$ matrix polynomial satisfying $B_P(z)B_P^T(z^{-1}) = I_m$.*

Proof. Since $B(z)$ is with rank m, any minor of order m being not identically zero must be of the form: $z^x g(z)$, where $x \ge 0$ is an integer and $g(z)$ is a polynomial with a nonzero constant term. Denote the greatest common factor of minors of order m (GCF) by $z^r b(z)$. Without loss of generality, $b(z)$ may be assumed to be monic. To emphasize the degree r in the common factor $z^r b(z)$, we write $B(z)$ as $B_r(z)$.

If $r = 0$, then the GCF of $B_r(z)$ is a monic polynomial $b(z)$. Since the constant term of $b(z)$ is nonzero $(=1)$, B_0 must be of column-full-rank. Then, we may take $B_P(z) = I$ and $B_I(z) = B(z)$, which meet the requirements of the lemma.

If $r > 0$, then the GCF of $B_r(z)$ is zero at $z = 0$. This implies that all minors of order m are zero at $z = 0$. In other words, the columns of B_0 are linearly dependent. Therefore, there exists a nonzero unit m-vector ψ such that $B_0\psi = 0$. This means that

$$B_r(z)\psi = \sum_{i=0}^{q} B_i \psi z^i = z\left(\sum_{i=0}^{q-1} B_{i+1} z^i\right)\psi. \tag{75}$$

Let T_r be an orthogonal matrix with ψ serving as its last column.

Define the matrix polynomial $B_{r-1}(z)$ as follows:

$$B_{r-1}(z) \triangleq B_r(z)T_r\Upsilon(z),$$

where

$$\Upsilon(z) \triangleq \begin{bmatrix} I_{m-1} & 0 \\ 0 & \dfrac{1}{z} \end{bmatrix}.$$

Since T_r is an $m \times m$ orthogonal matrix, the GCF of $B_r(z)T_r$ coincides with that of $B_r(z)$. Further, $B_r(z)T_r\Upsilon(z)$ differs from $B_r(z)T_r$ only at the last column by one degree of z less for the former. Therefore, the GCF of $B_{r-1}(z)$ is $z^{r-1}b(z)$, and $\deg[B_{r-1}(z)] \le \deg[B_r(z)]$.

If $r-1>0$, as before, the columns of the constant term of $B_{r-1}(z)$ are linearly dependent. Proceeding as above for r times, we arrive at

$$B_0(z) \triangleq B_r(z)T_r\Upsilon(z)T_{r-1}\Upsilon(z)\cdots T_1\Upsilon(z).$$

It is clear that $B_0(z)$ is still of rank m with $\deg B_0(z) \le q$, and the GCF of $B_0(z)$ is $b(z)$. So, the constant term of $B_0(z)$ is of column-full-rank.

Define

$$B_I(z) \triangleq B_0(z) \tag{76}$$

$$B_P(z) \triangleq \Upsilon^{-1}(z)T_1^T\Upsilon^{-1}(z)T_2^T\cdots\Upsilon^{-1}(z)T_r^T. \tag{77}$$

It is clear that (74) holds, and all requirements of the lemma are satisfied. □

Lemma 5. *Assume that an $n \times n$ rational polynomial $\widetilde{D}(z) = \sum_{i=-q}^{q}\widetilde{D}_i z^i$ with $\widetilde{D}_{-i} = \widetilde{D}_i^T$ is of rank m ($m \le n$), and is non-negative definite on the unit circle $|z| = 1$. Then there exists an $n \times m$ matrix polynomial $\widetilde{B}(z)$ with $\deg[\widetilde{B}(z)] \le q$ such that $\widetilde{D}(z) = \widetilde{B}(z)\widetilde{B}^T(z^{-1})$.*

Proof. Since $\widetilde{D}(z)$ with $\widetilde{D}_{-i} = \widetilde{D}_i^T$ is of rank m, and is non-negative on the unit circle $|z| = 1$, then, there exists an $n \times m$ real rational spectral factor $\overline{B}(z)$ with the poles being outside the unit circle such that $\widetilde{D}(z) = \overline{B}(z)\overline{B}^T(z^{-1})$ (see, e.g., [13,1,10] among others).

Notice that the poles of $\widetilde{D}(z)$ cannot be anything but 0 and ∞. Thus ∞ is the unique pole of $\overline{B}(z)$, which implies that $\overline{B}(z)$ is a matrix polynomial.

We denote $\overline{B}(z)$ by $\overline{B}(z) = \overline{B}_0 + \overline{B}_1 z + \cdots + \overline{B}_{\overline{q}} z^{\overline{q}}$. By Lemma 4, $\overline{B}(z)$ can be factored as

$$\overline{B}(z) = \widetilde{B}(z)B_P(z) \tag{78}$$

where $\widetilde{B}(z) = \widetilde{B}_0 + \widetilde{B}_1 z + \cdots + \widetilde{B}_{\overline{q}} z^{\overline{q}}$ is an $n \times m$ matrix polynomial with $\deg\widetilde{B}(z) \le \deg\overline{B}(z) = \overline{q}$, $\widetilde{B}_0$ is of column-full-rank, and $B_P(z)$ is an $m \times m$ matrix polynomial satisfying $B_P(z)B_P^T(z^{-1}) = I$. Hence, we obtain a different from $\overline{B}(z)\overline{B}^T(z^{-1})$ real polynomial factorization of $\widetilde{D}(z)$:

$$\widetilde{D}(z) = \overline{B}(z)\overline{B}^T(z^{-1}) = \widetilde{B}(z)B_P(z)B_P^T(z^{-1})\widetilde{B}^T(z^{-1}) = \widetilde{B}(z)\widetilde{B}^T(z^{-1}) \tag{79}$$

and $\widetilde{B}_0$ is of column-full-rank.

We now show that $\widetilde{B}_s = 0\ \forall s : q+1 \le s \le \overline{q}$.

If $\overline{q} > q$, then comparing the matrix coefficients of $z^{\overline{q}}$ on both sides of (79) we obtain $\widetilde{B}_{\overline{q}}\widetilde{B}_0^T = 0$. Since $\widetilde{B}_0$ is of column-full-rank, we have $\widetilde{B}_{\overline{q}} = 0$. By the same argument for $s : q+1 \le s \le \overline{q}-1$, we see that $\widetilde{B}_s = 0\ \forall s : q+1 \le s \le \overline{q}$. Therefore, $\deg[\widetilde{B}(z)] \le q$. □

References

[1] B.D.O. Anderson, M. Deistler, Properties of zero-free spectral matrices, IEEE Trans. Automat. Control 54 (2009) 2365–2375.
[2] P.O. Arambel, G. Tadmor, Identifiability and persistent excitation in full matrix fraction parameter estimation, Automatica 33 (1997) 689–692.
[3] H.F. Chen, New approach to identification for armax systems, IEEE Trans. Automat. Control 55 (2010) 868–879.
[4] F.P. Gantmacher, The Theory of Matrices, Chelsea, New York, 1959.
[5] E.J. Hannan, The identification of vector mixed autoregressive-moving average systems, Biometrika 56 (1969) 223–225.
[6] Y. Inouye, R.W. Liu, A system-theoretic foundation for blind equalization of an fir mimo channel system, IEEE Trans. Circuits Syst. I 49 (2002) 425–436.
[7] T. Kailath, Linear Systems, Prentice-Hall, New York, 1980.
[8] B.Q. Mu, H.F. Chen, Recursive identification of Wiener–Hammerstein systems, SIAM J. Control Optim. 50 (2012) 2621–2658.
[9] B.Q. Mu, H.F. Chen, Recursive identification of mimo Wiener systems, IEEE Trans. Automat. Control 58 (2013) 802–808.
[10] U.A. Rozanov, Stationary Random Processes, Holden-Day, 1967.
[11] Q.J. Song, H.F. Chen, Identification of errors-in-variables systems with arma observation noise, Systems Control Lett. 57 (2008) 420–424.
[12] P. Stoica, Generalized Yule–Walker equations and testing the orders of multivariate time series, Internat. J. Control 37 (1983) 1159–1166.
[13] D.C. Youla, On the factorization of rational matrices, IRE Trans. Inform. Theory 7 (1961) 172–189.
[14] W.X. Zhao, H.F. Chen, Markov chain approach to identifying Wiener systems, Sci. China Inf. Sci. 55 (2012) 1201–1217.