ELSEVIER

Systems & Control Letters 39 (2000) 237–243

SYSTEMS & CONTROL LETTERS

www.elsevier.com/locate/sysconle

Identification of certain noisy MA models: new results

W. Hachem[a], F. Desbouvries[a,*], Ph. Loubaton[b]

[a] Département Signal et Image, Institut National des Télécommunications, 9 rue Charles Fourier, 91011 Évry, France
[b] Laboratoire Systèmes de Communication, Université de Marne la Vallée, Champs sur Marne-5, bd Descartes, 77454 Marne la Vallée Cédex 2, France

Received 6 February 1999; received in revised form 10 August 1999

Abstract

In this paper, we address the identification problem of p-inputs q-outputs MA models, corrupted by a white noise with unknown covariance matrix, in the case where $p < q$. Under certain additional conditions, we show that the generating function of the MA model is identifiable (up to a $p \times p$ constant orthogonal matrix) from the autocovariance function of the observation. Our results extend those already obtained in Desbouvries et al. [5] and Desbouvries and Loubaton [6].

Keywords: Noisy MA models; Stochastic realization; Polynomial bases of rational subspaces; Truncated autocovariance sequence; Wiener–Hopf factorization

1. Introduction

Let $\{y_n\}_{n\in\mathbb{Z}}$ be a q-variate time-series given by

$y_n = [H(z)]v_n + w_n$,

where $H(z) = \sum_{k=0}^{M} H_k z^{-k}$ is a $q \times p$ polynomial of degree M, v_n is a p-dimensional (non-observable) white noise sequence for which $E(v_n v_n^{\mathrm{T}}) = I_p$, and w_n is an additive q-dimensional white noise (i.e. $E(w_n w_m^{\mathrm{T}}) = 0$ if $n \neq m$), independent with v_n. Throughout this paper, it is assumed that (C1) $q > p$, i.e., the dimension of the output y is strictly greater than that of the input v, and that (C2) $H(z)$ is minimum phase, i.e., that

$$\mathrm{Rank}(H(z)) = p \quad \text{for all } z \text{ such that } |z| > 1, \text{ including } z = \infty.$$

If the observations are noiseless, $H(z)$ can of course be identified up to a constant $p \times p$ orthogonal matrix from the autocovariance sequence of y_n. The purpose of this paper is to study the identifiability of $H(z)$ in the noisy case: in other words, if $\Sigma = E(w_n w_n^{\mathrm{T}})$ is non-zero, is it still possible to identify $H(z)$, up to a constant $p \times p$ orthogonal matrix, from the second-order statistics of y_n? The answer is known to be positive if $\Sigma = \sigma^2 I_q$ where σ^2 is an unknown scalar parameter. As a matter of fact, one can identify σ^2 from the covariance matrix $\mathcal{R}_N$ of the vector $Y_N(n) = [y_n^{\mathrm{T}}, \ldots, y_{n-N}^{\mathrm{T}}]^{\mathrm{T}}$, provided the parameter N is chosen "large enough". To show this, let $T_N(H)$ denote the $q(N+1) \times p(M+N+1)$ generalized Sylvester matrix associated with $H(z)$:

$$T_N(H) = \begin{bmatrix} H_0 & \ldots & H_M & & 0 \\ & \ddots & & \ddots & \\ 0 & & H_0 & \ldots & H_M \end{bmatrix}. \tag{1.1}$$

* Corresponding author. Tel.: 33-1-60-76-45-27; fax: 33-1-60-76-44-33.
E-mail addresses: Walid.Hachem@int-evry.fr (W. Hachem), Francois.Desbouvries@int-evry.fr (F. Desbouvries), loubaton@univ-mlv.fr (Ph. Loubaton)

0167-6911/00/$ - see front matter
PII: S0167-6911(99)00101-2

Then, it is clear that $Y_N(n) = T_N(H)V_{M+N}(n) + W_N(n)$, in which $V_{M+N}(n) = [v_n^T, \dots, v_{n-(M+N)}^T]^T$ and $W_N(n) = [w_n^T, \dots, w_{n-N}^T]^T$. The covariance matrix $\mathcal{R}_N$ is thus equal to $\mathcal{R}_N = T_N(H)T_N^T(H) + \sigma^2 I$. Since $p < q$, one can choose N such that $(N+1)q > (M+N+1)p$. If this holds, then $T_N(H)T_N^T(H)$ is a singular matrix, so σ^2 is identified as the smallest eigenvalue of $\mathcal{R}_N$: we are thus essentially led back to the noiseless case.

However, the assumption that $\Sigma = \sigma^2 I$ may be restrictive in certain contexts. Very often, y_n represents the signal sampled behind a sensor array. If w is a thermal noise that stems from the acquisition devices of the sensors, and the physical properties of the sensors are not identical, the noise components are likely to be decorrelated but their variances do not necessarily coincide. As another example, the noise in underwater acoustics is the sum of ambient sea noise, flow noise and traffic noise and may thus be spatially correlated [10,3]. In this paper, we are going to show that in the case where no a priori information on Σ is available, it is still possible, under certain conditions, to identify $H(z)$ (up to a $p \times p$ orthogonal matrix) from the exact second-order statistics of y_n.

Let us denote by $\{R_n\}_{n\in\mathbb{Z}}$ and by $\{R_n^y\}_{n\in\mathbb{Z}}$ the autocovariance lags of the useful signal $[H(z)]v_n$ and of y_n, respectively. As w_n is assumed to be white and $H(z)$ is a degree M polynomial, it is clear that $R_0^y = R_0 + \Sigma$, $R_n^y = R_n$ for $1 \leqslant |n| \leqslant M$, and $R_n^y = R_n = 0$ for $|n| > M$. In particular, R_0^y does not bring any information on R_0. We thus reformulate our problem as follows:

> Let $H(z) = \sum_{k=0}^{M} H_k z^{-k}$ be a $q \times p$ minimum phase polynomial, and let $\{R_n\}$ be the autocovariance function associated to the "spectral density" $S(z) = H(z)H^T(z^{-1})$. Identify $H(z)$ (up to a $p \times p$ orthogonal matrix) from the truncated sequence $\{R_n\}_{1\leqslant n\leqslant M}$.

This problem was first introduced in [7] in the case $p = 1$. It was shown that the unknown $q \times 1$ transfer function $H(z)$ is not necessarily identifiable if $q = 2$. In case of identifiability, an identification procedure based on the stochastic realization theory was proposed; however, it is based on a difficult non-convex optimization problem, for which no satisfying solution was proposed. In the meantime, subspace-based FIR blind identification methods were introduced for the case $\Sigma = \sigma^2 I$ (see e.g. [12,11] in the SIMO case $p = 1$, and [2,14] in the MIMO case $p > 1$). Later on, it was shown that these methods could be generalized to the case where Σ is unknown: the case $p = 1$ was considered in [1], and the case $p > 1$ in [5] (in the polynomial case), and [6] (in the rational case). Lastly, an alternate approach is proposed in [13].

We now specify the content of the paper. The results presented in Section 2 are based on the identification of the unknown coefficient R_0. Section 2.1 outlines the general approach, Section 2.2 reformulates some results of Desbouvries et al. [5] in order to introduce more clearly the benefit of the idea presented in Section 2.3. Finally, we introduce in Section 3 a new identifiability result based on the Wiener–Hopf factorization theory.

In this paper, we essentially focus on identifiability results based on the knowledge of the true autocovariance coefficients $(R_n)_{1\leqslant n\leqslant M}$. However, effective estimation algorithms can be derived from the results of Section 2. The practical use of the material presented in Section 3 is more involved, and is out of the scope of this paper. Finally, we note that the new results presented in this paper can be adapted immediately to the case where $H(z)$ is rational. We have chosen to restrict ourselves to the polynomial case for sake of clarity.

2. Identifying the missing covariance lag R_0

2.1. Outline of the results

In this section, the general idea is as follows. Since $\text{Rank}(S(z)) = p < q$ for all z, the structural equation $H(z)H^T(z^{-1}) = S(z) = \sum_{k=-M}^{M} R_k z^{-k}$ provides an implicit relation among the covariance lags $\{R_k\}_{k=0}^{M}$ which, in turn, enables (under certain sufficient conditions) to recover the central lag R_0 from $\{R_k\}_{k=1}^{M}$. Once R_0 is known, computing $H(z)$ from $H(z)H^T(z^{-1})$ is a classical problem, due to assumptions (C1) and (C2), and is thus omitted.

Our results are based on the following observation. Since $q > p$, there exist $1 \times q$ polynomials $g(z) = \sum_{k=0}^{N} g_k z^{-k}$ (where N is to be determined) satisfying

$$g(z)H(z) = 0 \quad \text{for all } z, \tag{2.2}$$

or, equivalently, satisfying $g(z)H(z)H^T(z^{-1}) = 0$ for all z. Let us assume that we can compute a set of r such polynomials $g_j(z)$, and let us set $G(z) = [g_1^T(z), \dots, g_r^T(z)]^T = \sum_{n=0}^{N} G_n z^{-n}$. Put $S(z) = \sum_{n=-M}^{M} R_n z^{-n} = R_0 + T(z)$. We have $0 = G(z)H(z) = G(z)S(z) = G(z)(R_0 + T(z))$, and thus

$$G(z)R_0 = -G(z)T(z), \tag{2.3}$$

where the right-hand side can be computed from the data. Identifying on both sides the coefficient of z^{-k} for $0 \leqslant k \leqslant N$, and setting $\mathcal{G} = [G_0^T, \dots, G_N^T]^T$, we see that we can compute the matrix $\mathcal{G}R_0$. If $\mathcal{G}$ has full column rank, R_0 can be retrieved from (2.3).

The method thus relies on the possibility of computing from the data (a sufficient number of) polynomials $g(z)$ satisfying (2.2). In [5], a method for computing $(M-1)$-degree such polynomials was proposed. However, this approach is valid only under very strong assumptions on $H(z)$. After a brief review of some results of Desbouvries et al. [5] (see Section 2.2), we shall thus propose an interesting alternative approach to evaluate polynomials satisfying (2.2) (see Section 2.3). Since both methods rely heavily on the notion of polynomial bases of rational subspaces, we first begin with briefly recalling some useful results on that subject.

A brief review of rational subspaces: Let us first recall that the set $\mathcal{F}_q$ of all $q \times 1$ rational transfer functions is a q-dimensional subspace over the field $\mathcal{F}_1$ of all scalar rational transfer functions. Let $\mathcal{S}$ be a p-dimensional ($p < q$) subspace of $\mathcal{F}_q$. The set $\{F_i(z)\}_{i=1}^p$ is a basis of $\mathcal{S}$ if and only if the $q \times p$ rational matrix $F(z) = [F_1(z), \dots, F_p(z)]$ has rank p for almost all z ($F(z)$ is said to have a normal rank equal to p). $\mathcal{S}$ admits polynomial bases. A polynomial basis $\{F_i(z)\}_{i=1}^p$ is said to be minimal if $\sum_{i=1}^p \deg(F_i(z))$ is minimum (see [8] for more details). Minimal polynomial bases are characterized by the well-known criterion (see [8,9]):

Proposition 1. *The polynomial basis $\{F_i(z)\}_{i=1}^p$ is minimal if and only if the matrix polynomial $F(z) = [F_1(z), \dots, F_p(z)]$ is irreducible and column reduced.*

All minimal polynomial bases share the same set of degrees $\{M_i = \deg(F_i(z))\}_{i=1}^p$. Usually, the minimal degrees $\{M_i\}_{i=1}^p$ are called the Kronecker indices associated to $\mathcal{S}$. The "orthogonal" $\mathcal{B}$ of $\mathcal{S}$ is the $(q-p)$-dimensional subspace of all $1 \times q$ rational transfer functions $g(z)$ satisfying $g(z)f(z) = 0$ for all $f \in \mathcal{S}$. The Kronecker indices $\{M_j^\perp\}_{j=1}^{q-p}$ of $\mathcal{B}$ are the so-called dual Kronecker indices of $\mathcal{S}$. They satisfy the important equality:

$$\sum_{i=1}^{p} M_i = \sum_{j=1}^{q-p} M_j^\perp. \tag{2.4}$$

We now turn back to our problem. Since $H(z)$ satisfies conditions (C1) and (C2), the rational space generated by its columns is p-dimensional. From now on, this subspace is denoted by $\mathcal{S}$ and its $(q-p)$-dimensional dual subspace by $\mathcal{B}$. Let $\{M_j^\perp\}_{j=1}^{q-p}$, with $0 \leqslant M_1^\perp \leqslant \cdots \leqslant M_{q-p}^\perp$, denote the Kronecker indices of $\mathcal{B}$. Assume that we have a method for computing from the data the set of all polynomials $g(z)$ satisfying (2.2) and of degree lower or equal to N. If $N < M_1^\perp = \min_j M_j^\perp$, then (2.2) holds if and only if $g(z) = 0$ for all z, while if $M_s^\perp \leqslant N < M_{s+1}^\perp$, then there exist exactly s linearly independent polynomials $g(z)$, with degree at most N, satisfying $g(z)H(z) = 0$. In particular, if $N \geqslant M_{q-p}^\perp = \max_j M_j^\perp$, then there exists a polynomial basis of $\mathcal{B}$ consisting of polynomials of degree at most N.

2.2. *Making use of the block-Hankel matrix associated to the sequence $\{R_n\}_{1 \leqslant n \leqslant M}$*

In order to prepare the reader to the content of Section 2.3, we reformulate (and hopefully simplify) a method proposed in [5] for computing polynomials of $\mathcal{B}$, of degree $(M-1)$, from the data $(R_n)_{1 \leqslant n \leqslant M}$. Let (C1) and (C2) hold. For the approach of Desbouvries et al. [5] to be valid, $H(z)$ also has to satisfy the following extra assumptions (note that (C4) implies (C2)):

- (C3) The columns $H_i(z)$ of $H(z)$ all share the same degree M;
- (C4) $H(z)$ is irreducible, i.e., $\mathrm{Rank}(H(z)) = p$ for all $z \neq 0$, including $z = \infty$;
- (C5) $H(z)$ is column reduced, and thus $\mathrm{Rank}(H_M) = p$, because of (C3).

Let $\mathcal{H}$ be the $qM \times qM$ block Hankel matrix given by

$$\mathcal{H} = \begin{bmatrix} R_1 & \cdots & R_M \\ \vdots & \iddots & \\ R_M & & 0 \end{bmatrix}. \tag{2.5}$$

$\mathcal{H}$ can be factored as

$$\mathcal{H} = \begin{bmatrix} H_1 & \cdots & H_M \\ \vdots & \iddots & \\ H_M & & 0 \end{bmatrix} \begin{bmatrix} H_0^T & & 0 \\ \vdots & \ddots & \\ H_{M-1}^T & \cdots & H_0^T \end{bmatrix} = \mathcal{O}\mathcal{C}^T. \tag{2.6}$$

Since H_M and H_0 have full column rank, $\mathcal{O}$ and $\mathcal{C}$ have full column rank Mp. Therefore, the rank of $\mathcal{H}$ is also equal to Mp. Let J be the q-block exchange matrix: $J = J_{M \times M} \otimes I_q$, where $\otimes$ denotes the

Kronecker product, and $(J_{M\times M}(i,j)=\delta_{i+j-(M+1)})_{i,j=1}^{M}$. It is easy to check that a qM-dimensional row vector $g=(g_0,\ldots,g_{M-1})$ (in which each g_k is q-dimensional) satisfies

$$gJ\mathscr{H}=0 \quad \text{and} \quad \mathscr{H}g^{\mathrm{T}}=0$$

if and only if $gT_{M-1}(H)=0$. So $\mathrm{Ker}^{\mathrm{l}}(T_{M-1}(H))=\mathrm{Ker}^{\mathrm{l}}(J\mathscr{H})\cap\mathrm{Ker}^{\mathrm{l}}(\mathscr{H}^{\mathrm{T}})$ ($\mathrm{Ker}^{\mathrm{l}}$ stands for the left kernel). Let $g(z)=\sum_{k=0}^{M-1}g_kz^{-k}$ be the $(M-1)$-degree polynomial associated to g. We have $gT_{M-1}(H)=0$ if and only if $g(z)H(z)=0$ for all z. Therefore, there is a one to one correspondence between the space $\mathrm{Ker}^{\mathrm{l}}(J\mathscr{H})\cap\mathrm{Ker}^{\mathrm{l}}(\mathscr{H}^{\mathrm{T}})$ and the set of all polynomials, of degree less than or equal to $M-1$, belonging to $\mathscr{B}$.

We first have to check if this subset of $\mathscr{B}$ is not trivially reduced to $\{0\}$. Since $H(z)$ is irreducible and column reduced, its columns form a minimal polynomial basis of $\mathscr{S}$. Therefore, the Kronecker indices of $\mathscr{S}$ all coincide with M. Hence, relation (2.4) becomes $pM=\sum_{j=1}^{q-p}M_j^{\perp}$. If $(M-1)<M_1^{\perp}$, we have $pM>(q-p)(M-1)$ or, equivalently, $q<2p+p/(M-1)$. Consequently, if $q\geqslant 2p+p/(M-1)$, there exists at least one non-zero $(M-1)$-degree polynomial $g(z)$ satisfying: $g(z)H(z)=0$ for all z. This condition, which is assumed to hold from now on in this section, implies in particular that $q>2p$.

It remains to investigate under which conditions there exist $1\times q$ polynomials $g_j(z)$, of degree $M-1$, such that the matrix $\mathscr{G}$ associated to $G(z)$ has full column rank (see the beginning of Section 2). Intuitively, for this condition to hold, the number s of linearly independent rows (over the field $\mathscr{F}_1$) of $G(z)$ should be as large as possible. It is thus not surprising that in case $s=q-p=\dim(\mathscr{B})$ (the maximum possible value), R_0 can be identified from (2.3), as we now see.

Let us thus assume that $(M-1)\geqslant M_{q-p}^{\perp}=\max_{j=1,q-p}M_j^{\perp}$ (which implies that $q\geqslant 2p+p/(M-1)$). Then there exist $(q-p)$ linearly independent $1\times q$ polynomials $\{g_j(z)\}_{j=1}^{q-p}$, of degree $M-1$, satisfying $g_j(z)H(z)=0$ (see the end of Section 2.1). Let $G(z)=[g_1^{\mathrm{T}}(z),\ldots,g_{q-p}^{\mathrm{T}}(z)]^{\mathrm{T}}=\sum_{k=0}^{M-1}G_kz^{-k}$, and let $\mathscr{G}=[G_0^{\mathrm{T}},\ldots,G_{M-1}^{\mathrm{T}}]^{\mathrm{T}}$. The condition: ($\mathscr{G}$ has full column rank), is equivalent to the condition: ($G(z)x_0=0$ for all $z\Rightarrow x_0=0$). Let us thus consider a constant vector x_0 satisfying $G(z)x_0=0$ for all z. Since the $(q-p)$ rows of $G(z)$ are linearly independent, $G(z)$ is a basis of $\mathscr{B}$. Therefore, the condition $G(z)x_0=0$ for all z holds if and only if the constant polynomial $x(z)=x_0$ belongs to the dual space of $\mathscr{B}$, i.e. to $\mathscr{S}$. But the Kronecker indices of $\mathscr{S}$ are all equal to M, and $M\geqslant 1$. Therefore, $\mathscr{S}$ does not contain non-zero constant vectors, and $x_0=0$.

We now summarize the discussion so far in the following theorem:

Theorem 1. *Let $H(z)=\sum_{k=0}^{M}H_kz^{-k}$ be a $q\times p$ polynomial, and let $\{R_n\}$ be the autocovariance function associated to the "spectral density" $S(z)=H(z)H^{\mathrm{T}}(z^{-1})$. Let $\mathscr{S}$ be the rational subspace generated by the columns of $H(z)$, with Kronecker indices $\{M_i\}_{i=1}^{p}$ and dual Kronecker indices $\{M_j^{\perp}\}_{j=1}^{q-p}$. Assume that* (C3)–(C5) *hold, and that $(M-1)\geqslant M_{q-p}^{\perp}=\max_{j=1,q-p}M_j^{\perp}$ (which, in particular, implies that $q\geqslant 2p+p/(M-1)$). Then $H(z)$ is identifiable (up to a $p\times p$ orthogonal matrix) from the truncated sequence $\{R_n\}_{1\leqslant n\leqslant M}$.*

In practice, if the conditions of the theorem are satisfied, this result is equivalent to the following property: let the Mq-dimensional rows of $U=(U_0,\ldots,U_{M-1})$ form an orthonormal basis of $\mathrm{Ker}^{\mathrm{l}}(J\mathscr{H})\cap\mathrm{Ker}^{\mathrm{l}}(\mathscr{H}^{\mathrm{T}})$; then, the matrix $\mathscr{U}=(U_0^{\mathrm{T}},\ldots,U_{M-1}^{\mathrm{T}})^{\mathrm{T}}$ has full column rank, so that R_0 can be identified from the product $\mathscr{U}R_0$. This leads to a practical estimation algorithm based on the empirical autocovariance coefficients of the observation.

2.3. Making use of the derivative of $S(z)$

The approach presented in Section 2.2 enables to compute polynomials of $\mathscr{B}$; however, the degree of these polynomials is bounded by $M-1$. This feature is the main limitation of the approach, because, on the other hand, the condition $M-1\geqslant M_{q-p}^{\perp}$ can be restrictive. Consider for example the case $p=2$, $q=5$ and $M=5$. Among the 14 triples $M_1^{\perp},M_2^{\perp},M_3^{\perp}$ satisfying $0\leqslant M_1^{\perp}\leqslant M_2^{\perp}\leqslant M_3^{\perp}$ and $\sum_{j=1}^{3}M_j^{\perp}=10$, only two satisfy the above condition. If $M-1<M_{q-p}^{\perp}$, one can compute $s<(q-p)$ linearly independent polynomials of $\mathscr{B}$ of degree $M-1$. However, we no longer have any reasonable condition guaranteeing that the corresponding matrices $\mathscr{G}$ have full column rank.

In this section, we propose an alternative approach which overcomes this drawback. It is based on the use of the derivative $S'(z)$ of $S(z)$ (w.r.t. the variable z^{-1}). This function obviously does not depend on R_0, and is therefore known. From $S(z)=H(z)H^{\mathrm{T}}(z^{-1})$, we have

$$S'(z) = [\, H'(z)\ H(z)\,] \begin{bmatrix} I_p & 0 \\ 0 & z^2 I_p \end{bmatrix} \begin{bmatrix} H^{\mathrm{T}}(z^{-1}) \\ -H'^{\mathrm{T}}(z^{-1}) \end{bmatrix}, \tag{2.7}$$

where $H'(z)$ stands for the derivative of $H(z)$ w.r.t. z^{-1}. In all this section, we shall assume that (C2) holds, that $q > 2p$ and that (C6): (Normal Rank $([H'(z), H(z)]) = 2p$). We shall make an extensive use of the $(q-2p)$-dimensional dual space $\mathscr{B}'$ of $\mathscr{S}'$. The Kronecker indices of $\mathscr{S}'$ are denoted by $\{M_i'\}_{i=1}^{2p}$, with $0 \leqslant M_1' \leqslant \cdots \leqslant M_{2p}'$, and those of $\mathscr{B}'$ by $\{M_j'^{\perp}\}_{j=1}^{q-2p}$, with $0 \leqslant M_1'^{\perp} \leqslant \cdots \leqslant M_{q-2p}'^{\perp}$.

Let us first outline the approach. Since $S'(z)$ is known, one can extract for each $N \geqslant M_1'^{\perp}$ the set of all degree N $1\times q$ polynomials $g(z) = \sum_{k=0}^{N} g_k z^{-k}$ satisfying $g(z)S'(z) = 0$ for all z; this is because $g(z)S'(z) = 0$ if and only if the row vector $g = (g_0, \ldots, g_N)$ belongs to the left kernel of the generalized Sylvester matrix associated to $S'(z)$. Now, (C6) ensures that the rational spaces spanned by the columns of $S'(z)$ and of $[H'(z), H(z)]$, respectively, are equal, and thus that $g(z)S'(z) = 0$ if and only if $g(z)[H'(z), H(z)] = 0$. This, in turn, implies that $g(z)H(z) = 0$ for all z. Therefore, using the derivative of $S(z)$ enables to extract polynomials of $\mathscr{B}$ of degree N for all $N \geqslant M_1'^{\perp}$. However, these elements do not span the whole set $\mathscr{B}$ because $\dim(\mathscr{B}') < \dim(\mathscr{B}) = q - p$. Nevertheless, one can use them to build a matrix polynomial $G(z)$ for which the associated matrix $\mathscr{G}$ has full column rank, as we now see.

Let again $H_i(z)$ (resp. $H_i'(z)$) denote the ith column of $H(z)$ (resp. of $H'(z)$). From (2.4), $\sum_{j=1}^{q-2p} M_j'^{\perp} = \sum_{i=1}^{2p} M_i' < \sum_{i=1}^{p} (\deg(H_i(z)) + \deg(H_i'(z)))$ (the inequality is strict because $[H'(z)\ H(z)]$ is not column-reduced). Therefore, $M_{q-2p}'^{\perp} < \sum_{i=1}^{p} (\deg(H_i(z)) + \deg(H_i'(z)))$. If N is chosen greater than $\sum_{i=1}^{p} (\deg(H_i(z)) + \deg(H_i'(z)))$, then $N \geqslant M_{q-2p}'^{\perp}$, and thus there exist $(q-2p)$ linearly independent $1 \times q$ polynomials $\{g_j(z)\}_{j=1}^{q-2p}$ of degree N satisfying $g_j(z)H(z) = 0$. Let $G(z) = [g_1^{\mathrm{T}}(z), \ldots, g_{q-2p}^{\mathrm{T}}(z)]^{\mathrm{T}} = \sum_{k=0}^{N} G_k z^{-k}$, and let $\mathscr{G} = [G_0^{\mathrm{T}}, \ldots, G_N^{\mathrm{T}}]^{\mathrm{T}}$. For the rest of the discussion we proceed as in Section 2.2. Let x_0 satisfy $\mathscr{G}x_0 = 0$ or, equivalently, $G(z)x_0 = 0$ for all z. Since $G(z)$ is a basis of $\mathscr{B}'$, this holds if and only if the constant polynomial vector $x(z) = x_0$ belongs to the dual space of $\mathscr{B}'$, i.e., to $\mathscr{S}'$. Let us assume that $M_1' \geqslant 1$. Then x_0 must be reduced to 0, $\mathscr{G}$ has full column rank, and R_0 can be identified from (2.3).

This result shows that R_0 can be identified from the left Kernel of $T_N(S')$ (if N is chosen large enough) provided $M_1' \geqslant 1$. This condition implies in particular that $\deg(H_i(z)) \geqslant 2$ for each i. We note however that this new approach needs less restrictive assumptions on $H(z)$ than the exploitation of the block Hankel matrix $\mathscr{H}$. In particular, we no longer need that all channels $H_i(z)$ have equal length (see condition (C3)), and the sufficient identifiability condition $M - 1 \geqslant M_{q-p}^{\perp}$ of Theorem 1 is relaxed. Moreover, the reader may check that the approach remains valid if $H(z)$ is rational.

We now summarize the discussion in the following theorem:

Theorem 2. *Let $H(z) = \sum_{k=0}^{M} H_k z^{-k}$ be a $q \times p$ minimum phase polynomial, and let $\{R_n\}$ be the autocovariance function associated to the "spectral density" $S(z) = H(z)H^{\mathrm{T}}(z^{-1})$. Assume that $q > 2p$ and that* (C6) *holds. Let $\{M_i'\}_{i=1}^{2p}$, with $0 \leqslant M_1' \leqslant \cdots \leqslant M_{2p}'$, be the Kronecker indices of the rational space $\mathscr{S}'$ generated by the columns of $[H'(z), H(z)]$. Let us assume that $\mathscr{S}'$ does not contain non-zero constant vectors, i.e., that its smallest Kronecker index M_1' is non-zero. Then $H(z)$ is identifiable* (*up to a $p \times p$ orthogonal matrix*) *from the truncated sequence $\{R_n\}_{1 \leqslant n \leqslant M}$.*

Finally, let us briefly outline the principle of a possible identification algorithm. The equation $G(z)S'(z) = 0$ is equivalent to $[G_0, \ldots, G_N]T_N(S') = 0$. The generalized Sylvester matrix $T_N(S')$ associated with $S'(z)$ plays the same role as the matrix $[J\mathscr{H}\ \ \mathscr{H}^{\mathrm{T}}]$ in the preceding algorithm. Hence, let now $U = (U_0, \ldots, U_N)$ be a matrix, the rows of which form an orthonormal basis of $\mathrm{Ker}^{\mathrm{l}}(T_N(S'))$. Then, provided $M_1' \geqslant 1$, the matrix $\mathscr{U} = [U_0^{\mathrm{T}}, \ldots, U_N^{\mathrm{T}}]^{\mathrm{T}}$ has full column rank, so that R_0 can be identified from the product $\mathscr{U}R_0$. Note that the dimension of $\mathrm{Ker}^{\mathrm{l}}(T_N(S'))$ is generally much greater than the dimension $q - 2p$ of $\mathscr{B}'$. This is because the linear independence of the rows of U does not imply the linear independence over the field $\mathscr{F}_1$ of their associated $1 \times q$ polynomials. In other words, the rows of $U(z) = \sum_{k=0}^{N} U_k z^{-k}$ span $\mathscr{B}'$, but are not linearly independent over $\mathscr{F}_1$. Similar remarks also apply to the algorithm of Section 2.2.

3. A Wiener–Hopf factorization based approach

In this final section, we still make use of the factorization (2.7) but in a different way. Under simple sufficient conditions on $H(z)$, we shall see that the factor

$[H'(z)\ H(z)]$ in (2.7) is essentially unique (because of its particular structure), which solves the identifiability problem of $H(z)$ from $\{R_k\}_{k=1}^{M}$.

Let us first set

$$W_-(z) = [H'(z), H(z)], \qquad \Lambda(z) = \begin{bmatrix} I_p & 0 \\ 0 & z^2 I_p \end{bmatrix}$$

and

$$W_+(z^{-1}) = \begin{bmatrix} H^{\mathrm{T}}(z^{-1}) \\ -H'^{\mathrm{T}}(z^{-1}) \end{bmatrix}, \tag{3.8}$$

so that (2.7) reads $S'(z) = W_-(z)\Lambda(z)W_+(z^{-1})$. Throughout this section, we shall assume that $q > 2p$, and that

(C7): $\mathrm{Rank}(W_-(z)) = 2p$
for all z such that $|z| \geqslant 1$ (including $z = \infty$)

(which implies (C2)). Let Γ denote the unit circle, $\Gamma_+ = \{z \in \mathbb{C},\ \text{s.t.}\ |z| < 1\}$ its interior, and $\Gamma_- = \{z \in \mathbb{C},\ \text{s.t.}\ |z| > 1\} \cup \infty$ the complement of $\Gamma_+ \cup \Gamma$ in the Riemann sphere $\overline{\mathbb{C}} = \mathbb{C} \cup \infty$. (C7) implies that $W_-(z)$ is analytic and has full rank $2p$ in $\Gamma_- \cup \Gamma$, while $W_+(z^{-1})$ is analytic and has full rank $2p$ in $\Gamma_+ \cup \Gamma$. Thus (2.7) is indeed on particular so-called non-canonical left Wiener–Hopf factorization of $S'(z)$ with respect to the contour Γ (for sake of brevity, we shall simply use the term: WH factorization) [4]. As is seen from the following theorem, the three factors $W_-(z)$, $\Lambda(z)$, and $W_+(z^{-1})$ are uniquely defined up to certain non-trivial indeterminacies:

Theorem 3. *Let* (C7) *hold. Then the decomposition* $S'(z) = W_-(z)\Lambda(z)W_+(z^{-1})$ *in* (2.7) *is a WH factorization of* $S'(z)$. *Let* $S'(z) = \tilde{W}_-(z)\tilde{\Lambda}(z)\tilde{W}_+(z^{-1})$ *be another WH factorization of* $S'(z)$, *i.e.,* $\tilde{W}_-(z)$ *is analytic and has full rank* $2p$ *in* $\Gamma_- \cup \Gamma$, $\tilde{W}_+(z^{-1})$ *is analytic and has full rank* $2p$ *in* $\Gamma_+ \cup \Gamma$, *and* $\tilde{\Lambda}(z)$ *is a diagonal matrix with diagonal entries* $(z^{k_1}, \ldots, z^{k_{2p}})$, *where the indices* $k_1 \leqslant \cdots \leqslant k_{2p}$ *belong to* $\mathbb{Z}$. *Then*

$$\tilde{\Lambda}(z) = \Lambda(z), \qquad \tilde{W}_-(z) = W_-(z)C_-^{-1}(z)$$

and

$$\tilde{W}_+(z^{-1}) = \Lambda^{-1}(z)C_-(z)\Lambda(z)W_+(z^{-1}), \tag{3.9}$$

where $C_-(z)$ *belongs to the multiplicative group* $\mathscr{C}_-$ *of* $2p \times 2p$ *block upper triangular matrices*

$$C_-(z) = \begin{bmatrix} C_{1,1} & C_{1,2}(z) \\ 0 & C_{2,2} \end{bmatrix} \tag{3.10}$$

satisfying $C_{1,1}$ *and* $C_{2,2}$ *are constant regular matrices, and* $C_{1,2}(z) = \sum_{i=0}^{2} C_{1,2}^{(i)} z^{-i}$ *is a polynomial matrix in* z^{-1} *of degree at most* 2. *Conversely, if* $C_-(z)$ *is any such matrix, and if* $\tilde{W}_-(z)$, $\tilde{W}_+(z^{-1})$ *and* $\tilde{\Lambda}(z)$ *are given by* (3.9), *then the factorization* $S'(z) = \tilde{W}_-(z)\tilde{\Lambda}(z)\tilde{W}_+(z^{-1})$ *is a WH factorization as defined above.*

Proof. Theorem 3 is a particular application of the general theorems [4, Theorems 1.1 and 1.2] to the present situation of interest. Note that definitions and results about existence and unicity are given in [4, Chapter 1] for the case of square factors, but the generalization to the present situation ($q > 2p$) is immediate. □

We shall now show that taking into account the particular structure of $W_-(z)$ (i.e., its left block is the derivative of its right block) enables to raise the indeterminacies evoked in Theorem 3.

Lemma 3.1. *Assume that the matrix* $[H_1, H_2]$ *has full column rank. Let* $\tilde{W}_-(z)$ *and* $\tilde{W}_+(z^{-1})$ *be the left and right factors, respectively, of some WH factorization of* $S'(z)$. *Then* $H(z)$ *is identifiable from these factors up to a constant* $p \times p$ *orthogonal matrix.*

Proof. Set $\tilde{W}_-(z) = [P_1(z), P_2(z)]$, and $P_i(z) = \sum_{k=0}^{M} P_{i,k} z^{-k}$ for $i = 1, 2$. Then (3.9) reads

$$[H'(z), H(z)] = [P_1(z), P_2(z)] \begin{bmatrix} C_{1,1} & C_{1,2}(z) \\ 0 & C_{2,2} \end{bmatrix}. \tag{3.11}$$

This equation leads in particular to $(P_2(z)C_{2,2} + P_1(z)C_{1,2}(z))' = P_1(z)C_{1,1}$. Developing and equating on both sides the coefficients of z^{-k} for $0 \leqslant k \leqslant M$ leads immediately to the matrix equation

$$\begin{bmatrix} 0 & P_{1,0} & P_{1,1} & P_{1,0} \\ 2P_{1,0} & 2P_{1,1} & 2P_{1,2} & P_{1,1} \\ 3P_{1,1} & 3P_{1,2} & 3P_{1,3} & P_{1,2} \\ 4P_{1,2} & 4P_{1,3} & 4P_{1,4} & P_{1,3} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} -C_{1,2}^{(2)} C_{2,2}^{-1} \\ -C_{1,2}^{(1)} C_{2,2}^{-1} \\ -C_{1,2}^{(0)} C_{2,2}^{-1} \\ C_{1,1} C_{2,2}^{-1} \end{bmatrix} = \begin{bmatrix} P_{2,1} \\ 2P_{2,2} \\ 3P_{2,3} \\ 4P_{2,4} \\ \vdots \end{bmatrix}. \tag{3.12}$$

Since $P_1(z) = H'(z)C_{1,1}^{-1}$, it is easy to check that the first matrix of the left-hand side of (3.12) is equal to

$$\begin{bmatrix} 0 & H_1 & 2H_2 & H_1 \\ 2H_1 & 4H_2 & 6H_3 & 2H_2 \\ \vdots & & & \vdots \end{bmatrix} \begin{bmatrix} C_{1,1}^{-1} & & 0 \\ & \ddots & \\ 0 & & C_{1,1}^{-1} \end{bmatrix},$$

and thus has full column rank if $[H_1, H_2]$ has full column rank $2p$. If this holds, then Eq. (3.12) provides the matrices $C_{1,2}^{(0)}C_{2,2}^{-1}$, $C_{1,2}^{(1)}C_{2,2}^{-1}$, $C_{1,2}^{(2)}C_{2,2}^{-1}$ and $C_{1,1}C_{2,2}^{-1}$. Injecting into (3.11), we see that one can recover the matrix $\bar{H}(z) = H(z)C_{2,2}^{-1}$ from any left Wiener–Hopf factor of $S'(z)$.

It remains to identify the unknown matrix $C_{2,2}$. To that end, let $\tilde{W}_+(z^{-1}) = [Q_1(z^{-1}), Q_2(z^{-1})]^{\mathrm{T}}$. From $W_+(z^{-1}) = \Lambda^{-1}(z)C_-^{-1}(z)\Lambda(z)\tilde{W}_+(z^{-1})$, we see that $-H'^{\mathrm{T}}(z^{-1}) = C_{2,2}^{-1}Q_2^{\mathrm{T}}(z^{-1})$. In particular, $Q_2(0) = -H_1C_{2,2}^{\mathrm{T}} = -\bar{H}_1 \times (C_{2,2}C_{2,2}^{\mathrm{T}})$. Now, $\bar{H}_1$ has full column rank and thus admits a left inverse $\bar{H}_1^{-\mathrm{L}}$. We thus see that the matrix $A = C_{2,2}C_{2,2}^{\mathrm{T}} = -\bar{H}_1^{-\mathrm{L}}Q_2(0)$ can be computed from the data. Let $A^{1/2}$ be a square root of A. Then $A^{1/2} = C_{2,2}Q$ for some orthogonal matrix Q, and finally $\bar{H}(z)A^{1/2} = H(z)Q$. □

This identifiability result can be used to derive a concrete estimation algorithm of $H(z)$ provided we have at hand a simple constructive Wiener–Hopf factorization algorithm. This non-obvious problem is out of the scope of the present paper, and is under investigation.

Let us summarize the discussion of this section into the following theorem.

Theorem 4. *Let $H(z) = \sum_{k=0}^{M} H_k z^{-k}$ be a $q \times p$ polynomial, and let $\{R_n\}$ be the autocovariance function associated to the "spectral density" $S(z) = H(z)H^{\mathrm{T}}(z^{-1})$. Assume that $q > 2p$ and that* (C7) *holds. Let us further assume that the matrix $[H_1, H_2]$ has full column rank $2p$. Then $H(z)$ is identifiable* (*up to a $p \times p$ orthogonal matrix*) *from the truncated sequence $\{R_n\}_{1 \leqslant n \leqslant M}$.*

4. Conclusion

In this paper, we proposed two approaches, based on second-order statistics, for identifying a noisy MA model up to a constant orthogonal matrix. The problem has been formulated as the identification of the unknown filter $H(z)$ from the truncated autocovariance sequence $\{R_n\}_{n \neq 0}$ associated to the spectral density $S(z) = H(z)H^{\mathrm{T}}(z^{-1})$, and can thus be viewed as a stochastic realization problem. The first approach consists in identifying R_0 directly from certain $1 \times q$ polynomials $g(z)$ satisfying $g(z)H(z) = 0$. Computing these filters can be done either from the block Hankel matrix associated to the sequence $\{R_n\}_{n \geqslant 1}$, or from the derivative $S'(z)$ of $S(z)$ w.r.t. the variable z^{-1}, which requires less-restrictive assumptions on $H(z)$. Finally, we have introduced an alternative approach based on the Wiener–Hopf factorization theory. In this context, an identifiability result has been proved. The practical use of the Wiener–Hopf factorization approach is currently under investigation.

References

[1] K. Abed-Meraim, Y. Hua, Ph. Loubaton, E. Moulines, Subspace method for blind identification of multichannel FIR systems in noise field with unknown spatial covariance, IEEE Signal Process. Lett. 4 (5) (1997) 135–137.

[2] K. Abed-Meraim, Ph. Loubaton, E. Moulines, A subspace algorithm for certain blind identification problems, IEEE Trans. Inform. Theory 43 (2) (1997) 499–511.

[3] W.S. Burdic, Underwater Acoustic System Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1984.

[4] K. Clancey, I. Gohberg, Factorization of matrix functions and singular integral operators, Operator Theory: Advances and Applications, Vol. 3, Birkhäuser, Basel, 1981.

[5] F. Desbouvries, I. Fijalkow, P. Loubaton, On the identification of noisy MA models, IEEE Trans. Automat. Control 41 (12) (1996) 1810–1814.

[6] F. Desbouvries, P. Loubaton, On the identification of certain rational transfer functions from truncated autocovariance sequences, IEEE Trans. Automat. Control 44 (5) (1999) 1078–1081.

[7] I. Fijalkow, P. Loubaton, Identification of rank one rational spectral densities from noisy observations: a stochastic realization approach, System Control Lett. 24 (1995) 201–205.

[8] G.D. Forney, Minimal bases of rational vector spaces, with applications to multivariable linear systems, SIAM J. Control 13 (3) (1975) 493–520.

[9] T. Kailath, Linear Systems, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[10] J.P. Le Cadre, Parametric methods for spatial signal processing in the presence of unknown colored noise fields, IEEE Trans. Acoust. Speech Signal Process. 37 (7) (1989) 965–983.

[11] E. Moulines, P. Duhamel, J.F. Cardoso, S. Mayrargue, Subspace methods for the blind identification of multichannel FIR filters, IEEE Trans. Signal Process. 43 (2) (1995) 516–525.

[12] L. Tong, G. Xu, T. Kailath, A new approach to blind identification and equalization of multipath channels, Proceedings of the Asilomar Conference, 1991, pp. 856–860.

[13] P. Vandaele, M. Moonen, A stochastic subspace algorithm for blind channel identification in noise fields with unknown spatial color, Proc. ICASSP, Vol. 5, 1999, pp. 2619–2622.

[14] A.J. van der Veen, S. Talwar, A. Paulraj, A subspace approach to blind space-time signal processing for wireless communication systems, IEEE Trans. Signal Process. 45 (1) (1997) 173–190.