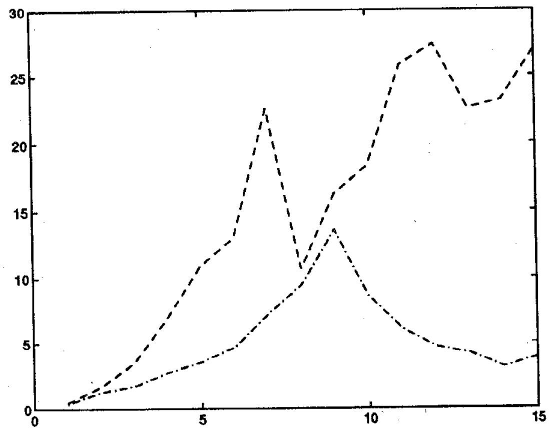

Fig. 3. Mean square error (— — IMMA, — · — exact filter).

IMM estimate (see Fig. 3). Thus, the presented results show clearly that the proposed approach can improve the estimation performances in a significant way over the well-known IMM algorithm.

IV. Conclusion

This note presents an explicit solution to the problem of estimating the comprehensive state of a hybrid dynamic system. From $q_n^j(z)$, it is a direct exercise to find the normalized state density

$$p_n^j(z) = \frac{q_n^j(z)}{\sum_{i=1}^{N} \int_{\mathbb{R}} q_n^i(z)\, dz}.$$

From $p_n^j(z)$, the $\mathcal{Y}_k$-conditional expectations of x_k and Z_k are easily computed. Note that this need only be done at time points at which the estimates are required by an external node. Although the growth in the number of terms grows geometrically, if $q_0^i(z)$ is written as a Gaussian sum, the specific weights, means, and variances of each term in the sum are explicit, and these statistics can be used in a suitable pruning algorithm. The abridgement issue is currently being explored by one of the authors in the context of an application in wireless communication (GMSK).

References

[1] G. A. Ackerson and K. S. Fu, "On state estimation in switching environments," *IEEE Trans. Automat. Contr.*, vol. 15, no. 1, pp. 10–17, 1970.

[2] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Automat. Contr.*, vol. 33, no. 8, pp. 780–783, 1988.

[3] J. R. Cloutier, J. H. Evers, and J. J. Feeley, "Assessment of air-to-air missile guidance and control technology," *IEEE Contr. Syst. Mag.*, vol. 9, pp. 27–34, 1989.

[4] J. R. Elliott, "A general recursive discrete-time filter," *J. App. Probability*, vol. 30, pp. 575–588, 1993.

[5] R. J. Elliott, L. Aggoun, and J. Moore. *Hidden Markov Model, Estimation and Control.* New York: Springer-Verlag, 1995.

[6] M. Mariton, *Jump Linear Systems in Automatic Control.* New York: Marcel Dekker, 1990.

[7] P. S. Maybeck and R. D. Stevens, "Reconfigurable flight control via multiple model adaptive control methods," *IEEE Trans. Aerospace Electronic Syst.*, vol. 27, pp. 470–480, 1991.

[8] D. D. Sworder, R. Vojak, and R. G. Hutchins, "Gain adaptive tracking," *J. Guidance, Contr. Dynamics*, vol. 16, no. 5, pp. 865–873, 1993.

On the Identification of Noisy MA Models

F. Desbouvries, I. Fijalkow, and Ph. Loubaton

***Abstract*—In this paper, we address the identification problem of p-input q-output MA models corrupted by a white noise with an 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 from the autocovariance function of the observation. Some simple algebraic identification procedures are also given.**

I. Introduction

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

$$y_n = [H(z)]v_n + w_n \tag{1}$$

where $H(z) = \Sigma_{k=0}^{M} H_k z^{-k}$ is a $q \times p$ finite impulse response transfer function, v_n is a p-dimensional (unobservable) white noise sequence for which $E(v_n v_n^T) = I$, and w_n is an additive q-dimensional white noise (i.e., $E(w(n)w^T(m)) = 0$ if $n \neq m$). It is assumed that i) $p < q$, i.e., the dimension of y is strictly greater than the dimension of the input v and ii) $H(z)$ is outer, i.e., that $\text{Rank}(H(z)) = p$ for $|z| > 1$. Given the (exact) autocovariance function of y, we address in this paper the identification problem of $H(z)$ (up to a $p \times p$ constant orthogonal matrix) *in the case where the covariance matrix* $\Sigma = E(w_n w_n^T)$ *is unknown.* This question is motivated by the so-called blind equalization problem, arising in the context of digital communications. For more information, see [7], [1], [2], and [8], for example.

Let us denote by R^y the autocovariance function of y defined by $R_n^y = E(y_{n+k} y_k^T)$, and denote accordingly by R the autocovariance sequence of the signal $[H(z)]v_n$. As the noise w is assumed to be white, it is clear that $R_0^y = R_0 + \Sigma$ and that $R_n^y = R_n$ for $n \geq 1$. As Σ is unknown, R_0^y does not contain any information on R_0 except the inequality $R_0 < R_0^y$. As it seems difficult to use, the approach developed here consists of identifying $H(z)$ from the sole knowledge of the truncated sequence $(R_n)_{n\geq 1}$. It can be reformulated as follows.

Let $H(z) = \Sigma_{k=0}^{M} H_k z^{-k}$ be a $q \times p$ outer polynomial matrix (of the variable z^{-1}), and let $(R_k)_{k=0,M}$ be the autocovariance function associated to the "spectral density" $H(z)H^T(z^{-1})$. How does one identify $H(z)$ (up to a $p \times p$ orthogonal matrix) from the knowledge of the sequence $(R_k)_{k\geq 1}$?

To our best knowledge, this problem has been addressed in only two previous contributions related to the specific case $p = 1$. In [3], it is shown that the unknown $q \times 1$ transfer function $H(z)$ is not necessarily identifiable. In case of identifiability, an identification procedure based on the stochastic realization theory is proposed. However, it is based on a difficult nonconvex optimization problem for which no satisfying solution has been proposed. In [2], $H(z)$ is assumed nonzero for each z, i.e., the components of $H(z)$ have no common zero. This additional condition (which is quite relevant in our context) makes the problem considerably simpler. In particular, under a simple sufficient identifiability condition, a "subspace-like"

Manuscript received September 19, 1995; revised March 22, 1996.

F. Desbouvries is with INT/SIM, 91011 Evry Cedex, France (e-mail: desbou@int-evry.fr).

I. Fijalkow is with ENSEA/ETIS, 95014 Cergy-Pontoise Cedex, France.

Ph. Loubaton is with the Université de Marne la Vallée Unité de Formation SPI, 93166 Noisy le Grand Cedex, France.

Publisher Item Identifier S 0018-9286(96)08444-9.

identification algorithm, based on a singular value decomposition, has been derived.

This paper is organized as follows. In Section II, we first briefly review the results presented in [3] and [2]. Then in Section III, we address the case $p > 1$ in which new difficulties arise. With regard to possible applications to blind equalization, our aim is to derive sufficient identifiability conditions leading to simple algebraic identification procedures.

II. A Review of the Case $p = 1$

In this section, we briefly review the results of [3] and [2].

In [3], necessary and sufficient identifiability conditions of a rational transfer function $H(z)$ are derived. From these conditions, it is easy to exhibit situations in which $H(z)$ is not identifiable from the sequence $(R_n)_{n\geq 1}$ in the case where $q = 2$. When $q \geq 3$, the situation becomes simpler, and it is established in [3] and [2] that if

C1) $\dim(\text{span}\{H(z)/z \in \mathbb{C}\}) > 2$

then $H(z)$ is identifiable from the sequence $(R_n)_{n\geq 1}$. In other words, as soon as the values taken by the transfer function $H(z)$ do not lie in a constant two-dimensional subspace (in which case, one would be back to the difficult case $q = 2$), $H(z)$ is identifiable.

Now, under the additional irreducibility assumption

C2) $H(z) \neq 0$ for each $z \neq 0$ (including $z = \infty$)

a simple identification procedure has been presented in [2] based on the $qM \times qM$ matrix $\mathcal{R}$ defined by

$$\mathcal{R} \stackrel{\text{def}}{=} \begin{bmatrix} R_M & & 0 \\ \vdots & \ddots & \\ R_1 & \cdots & R_M \end{bmatrix} + \begin{bmatrix} R_M^T & \cdots & R_1^T \\ & \ddots & \vdots \\ 0 & & R_M^T \end{bmatrix}. \quad (2)$$

Since $R_n = \Sigma_{i=n}^{M} H_i H_{i-n}^T$ for $n \geq 1$, it is easy to show that

$$\mathcal{R} = T_{M-1}(H) \mathcal{J}_M T_{M-1}^T(H) \quad (3)$$

in which $T_N(H)$ is the $(N+1)q \times (M+N+1)$ generalized Sylvester matrix (or $(N+1)q \times (M+N+1)p$ in the subsequent general case $p \geq 1$), and $\mathcal{J}_K$ the signature matrix, defined respectively by

$$T_N(H) = \begin{bmatrix} H_0 & \cdots & H_M & & 0 \\ & \ddots & & \ddots & \\ 0 & & H_0 & \cdots & H_M \end{bmatrix}$$
$$\mathcal{J}_K = \begin{bmatrix} 0_K & I_K \\ I_K & 0_K \end{bmatrix}. \quad (4)$$

Using standard results related to Sylvester matrices (see formula 3.12 below), Condition C2) implies that $T_{M-1}(H)$ is full column-rank. Therefore, the left kernel of $\mathcal{R}$, denoted by $\text{Ker}^l(\mathcal{R})$ (i.e., the set of all qM-dimensional row vectors g for which $g\mathcal{R} = 0$), coincides with that of the matrix $T_{M-1}(H)$ itself, which can thus be retrieved from the sequence $(R_n)_{n=1\cdots M}$. The key point is that under Conditions C2) and C1), $H(z)$ is uniquely defined up to a constant scalar factor by the space $\text{Ker}^l(T_{M-1}(H))$.

Theorem 2.1: Assuming that Conditions C2) and C1) hold, let Π be the orthogonal projection matrix on the space $\text{Ker}^l(\mathcal{R})$. Then, a degree M polynomial $q \times 1$ transfer function $F(z)$ satisfies

$$\Pi T_{M-1}(F) = 0 \quad (5)$$

if and only if $F(z) = H(z)$ up to a constant scalar factor.

In other words, the transfer function can be identified up to a scalar factor by solving the linear system (5). The determination of the remaining scalar factor is straightforward.

This result is a consequence of well-known results related to the so-called minimal polynomial bases of a rational subspace (see [4] and [5], for example). As Theorem 2.1 is extended to the case $p > 1$ in the next section, the proof is omitted.

III. The Case $p > 1$

This section is devoted to the case $p > 1$. To derive simple identification procedures, we assume from now on that the following conditions hold.

C3) The columns $(H_i(z))_{i=1,p}$ of $H(z)$ have the same degree M.

C4) $Rank(H(z)) = p$ for each $z \neq 0$, i.e., $H(z)$ is irreducible.

C5) $H(z)$ is column-reduced (i.e., $\text{Rank}(H_M) = p$ due to $C3$)).

We propose two approaches here. The first is based on a specific state-space realization of $(R_n)_{n=1,M}$. The second extends in some sense the subspace approach of [2].

Note that as in the case $p = 1$, it seems difficult to derive simple sufficient identifiability conditions on $H(z)$ when $q \leq 2p$. In this case, it is possible to build examples of unidentifiable transfer functions $H(z)$ as in the case $p = 1$. In particular, all the sufficient conditions presented in the next sections hold only if $q > 2p$.

A. The Realization Approach

First, one should note that the three conditions on $H(z)$ above imply that the MacMillan degree of $H(z)$ coincides with pM. As $H(z)$ is outer, it represents a minimal degree causal spectral factor of the "spectral density" $H(z)H^T(z^{-1})$, and it is a standard result that the MacMillan degree of the transfer function $\Sigma_{n=1}^{M} R_n z^{-n}$ also coincides with Mp. Hence, the block Hankel matrix $\mathcal{H}$ defined by

$$\mathcal{H} = \begin{bmatrix} R_1 & \cdots & R_M \\ \vdots & & \\ R_M & & 0 \end{bmatrix} \quad (6)$$

has rank Mp. On the other hand, $\mathcal{H}$ can be factored as

$$\mathcal{H} = \begin{bmatrix} H_1 & \cdots & H_M \\ \vdots & & \\ H_M & & 0 \end{bmatrix} \begin{bmatrix} H_0^T & & 0 \\ \vdots & \ddots & \\ H_{M-1}^T & \cdots & H_0^T \end{bmatrix} = \mathcal{O}\mathcal{C}^T. \quad (7)$$

As H_M and H_0 are full column-rank, it is clear that $\mathcal{O}$ and $\mathcal{C}$ have rank Mp. The factorization $\mathcal{O}\mathcal{C}^T$ is therefore minimal, and it is easily seen that $\mathcal{O}$ and $\mathcal{C}^T$ represent the observability and the controllability matrices associated to the (A, B, C) realization given by $(A(i,j) = \delta_{i-(j+p)})_{i,j=1}^{Mp}$ (A is the p-block shift matrix), $C = [H_1 \cdots H_M]$, and $B = [H_0 \cdots H_{M-1}]^T$ (i.e., $R_n = CA^{n-1}B$ for each $n \geq 1$). Let J be the q-block exchange matrix $J = J_{M\times M} \otimes I_q$, where $\otimes$ denotes the Kronecker product, and $(J_{M\times M}(i,j) = \delta_{i+j-(M+1)})_{i,j=1}^{M}$. $(\mathcal{C}, J\mathcal{O})$ coincides with the Sylvester matrix $T_{M-1}(H)$ [defined by (4)] associated with $H(z)$, and the equality

$$[I_{(M-1)q}, 0]\mathcal{C}[0_p, \cdots, 0_p, I_p]^T = [0, I_{(M-1)q}]J\mathcal{O}[I_p, 0_p, \cdots, 0_p]^T \quad (8)$$

holds. Conversely, let any other minimal state-space representation (A, B', C') of $(R_n)_{n\geq 1}$ (with related observability and controllability matrices $\mathcal{O}'$ and $\mathcal{C}'^T$), with the same transition matrix A, and which satisfies

$$[I_{(M-1)q}, 0]\mathcal{C}'[0_p, \cdots, 0_p, I_p]^T = [0, I_{(M-1)q}]J\mathcal{O}'[I_p, 0_p, \cdots, 0_p]^T. \quad (9)$$

Then, it is easily seen that $\mathcal{O}'$ and $\mathcal{C}'$ are given by

$$\mathcal{O}' = \begin{bmatrix} H_1' & \cdots & H_M' \\ \vdots & & \\ H_M' & & 0 \end{bmatrix} \quad \text{and} \quad \mathcal{C}'^T = \begin{bmatrix} H_0'^T & & 0 \\ \vdots & \ddots & \\ H_{M-1}'^T & \cdots & H_0'^T \end{bmatrix}$$

for some matrices $H_0', \cdots, H_M'$. In particular, the correlation coefficients $(R_n')_{n\geq 1}$ associated with the "spectral density" $H'(z)H'^T(z^{-1})$ coincide with $(R_n)_{n\geq 1}$. From this, we deduce a sufficient identifiability condition of $H(z)$ as well as a simple identification algorithm.

Theorem 3.1: Suppose that the transfer function $H(z)$ satisfies the condition

C6) $T_{M-2}(H)$ is a full column-rank matrix.

Then, $H(z)$ is identifiable.

Proof: Our proof provides both the proof of the Theorem and the identification procedure. Let (A, B_0, C_0) be a minimal realization of $(R_n)_{n\geq 1}$ with transition matrix, the block shift matrix A (such a realization can be easily calculated from $(R_n)_{n=1,M}$). To show that $H(z)$ is identifiable, it is sufficient to establish that there exists a unique (up to an orthogonal matrix) realization (A, B', C') satisfying (9). As (A, B', C') is a minimal realization, there exists a unique invertible $Mp \times Mp$ matrix P for which $PA = AP, C' = C_0P, B' = P^{-1}B_0$. In terms of observability and controllability matrices, this can be written as $\mathcal{O}' = \mathcal{O}_0P, \mathcal{C}' = \mathcal{C}_0P^{-T}$. As P commutes with A, it is easy to show that P is a lower triangular block-Toeplitz matrix and that P^{-T} is an upper triangular block-Toeplitz matrix. Let $[P_1^T, P_2^T, \cdots, P_M^T]^T$ be the first block-column of P and $[Q_M^T, \cdots, Q_1^T]^T$ the last block-column of P^{-T}. Then, (9) holds if and only if

$$\begin{aligned}(0, I_{(M-1)q})J\mathcal{O}_0[P_1^T, P_2^T, \cdots, P_M^T]^T \\ = (I_{(M-1)q}, 0)\mathcal{C}_0[Q_M^T, \cdots, Q_1^T]^T.\end{aligned}$$

Let T_0 be the matrix $T_0 = [(I_{(M-1)q}, 0)\mathcal{C}_0, (0, I_{(M-1)q})J\mathcal{O}_0]$. Then, the above relation can be written as

$$T_0[-Q_M^T, \cdots, -Q_1^T, P_1^T, \cdots, P_M^T]^T = 0. \tag{10}$$

Let us now study the dimension of the kernel of T_0. Since (A, B_0, C_0) and the canonical realization (A, B, C) introduced above are similar, there exists an invertible matrix P_0 such that $\mathcal{C}_0 = \mathcal{C}P_0^{-T}$ and $\mathcal{O}_0 = \mathcal{O}P_0$. Therefore, T_0 satisfies

$$T_0 = [(I_{(M-1)q}, 0)\mathcal{C}, (0, I_{(M-1)q})J\mathcal{O}]\begin{bmatrix} P_0^{-T} & 0 \\ 0 & P_0 \end{bmatrix}.$$

However

$$(I_{(M-1)q}, 0)\mathcal{C} = \begin{bmatrix} H_0 & \cdots & H_{M-2} & H_{M-1} \\ & \ddots & \vdots & \vdots \\ 0 & & H_0 & H_1 \end{bmatrix}$$

and

$$(0, I_{(M-1)q})J\mathcal{O} = \begin{bmatrix} H_{M-1} & H_M & & 0 \\ \vdots & \vdots & \ddots & \\ H_1 & H_2 & \cdots & H_M \end{bmatrix}.$$

Therefore, the rank of T_0 coincides with the rank of $T_{M-2}(H)$. As $T_{M-2}(H)$ is supposed to be full column-rank, the kernel of T_0 is p-dimensional. Hence, the matrix $[-Q_M^T, \cdots, -Q_1^T, P_1^T, \cdots, P_M^T]^T$ is uniquely defined up to an invertible $p \times p$ matrix by (10). This remaining matrix can be straightforwardly identified up to an orthogonal matrix using $R_M = H_0H_M^T$. ∎

Note that the rank of the Sylvester matrix $T_N(H)$ can be expressed in terms of N and of certain Kronecker indexes associated to the rational subspace generated by the columns of $H(z)$ (see (12) below). So, the full rank condition of $T_{M-2}(H)$ can be made more explicit.

B. The Subspace Method for $p > 1$

In this section, we shall show how the results of [2] can be generalized to the case $p > 1$. For this purpose, we need to review some well-known properties related to minimal polynomial bases of a rational subspace.

1) A 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 $\mathcal{F}_q$. $\mathcal{S}$ admits polynomial bases. A polynomial basis $(F_1(z), \cdots, F_p(z))$ is said to be minimal if $\Sigma_{i=1}^p \deg(F_i(z))$ is minimum (see [5] for more details). Let $M_i = \deg(F_i(z))$. All minimal bases share the same degrees $(M_i)_{i=1,p}$ and are characterized by the well-known criterion (see [4] and [5]).

Proposition 3.1: The polynomial basis $(F_1(z), \cdots, F_p(z))$ is minimal if and only if the matrix polynomial $F(z) = (F_1(z), \cdots, F_p(z))$ is irreducible and column reduced. In this case, a polynomial $f(z)$ belongs to $\mathcal{S}$ if and only if there exists a p-dimensional polynomial vector $r(z)$ for which $f(z) = F(z)r(z)$. Finally, if $M_i = M$ for each i, two minimal bases $F(z)$ and $F'(z)$ coincide up to a constant invertible $p \times p$ matrix.

Usually, the minimal degrees $(M_i)_{i=1,p}$ are called the Kronecker indexes associated with $\mathcal{S}$. 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 each $f \in \mathcal{S}$. $\mathcal{B}$ admits Kronecker indexes denoted by $(M_j^\perp)_{j=1,q-p}$ which satisfy the important equality

$$\sum_{i=1}^{p} M_i = \sum_{j=1}^{q-p} M_j^\perp. \tag{11}$$

Finally, we recall that if $F(z) = (F_1(z), \cdots, F_p(z))$ is a minimal polynomial basis of $\mathcal{S}$, then the rank of the Sylvester matrix $T_N(F)$ is given by [6]

$$\text{Rank}(T_N(F)) = q(N+1) - \sum_{j, M_j^\perp \leq N} (N + 1 - M_j^\perp). \tag{12}$$

2) A Direct Generalization of the Subspace Method to the Case $p > 1$: Let $\mathcal{S}$ denote the p-dimensional rational subspace generated by the columns of $H(z)$ and $\mathcal{B}$ its $(q-p)$-dimensional "orthogonal." Since $H(z)$ is irreducible and column reduced, it represents a minimal polynomial basis of $\mathcal{S}$. Keeping this in mind, Theorem 2.1 can be generalized as follows.

Theorem 3.2: Let $(M_i^\perp)_{i=1,q-p}$ be the Kronecker indexes of $\mathcal{B}$. Suppose that

C7) $M_i^\perp \leq M - 1$ for each $i = 1, q-p$.

Let Π be the orthogonal projection matrix on $\text{Ker}^l(\mathcal{R})$, and let $F(z) = (F_1(z), \cdots, F_p(z))$ be a full rank $q \times p$ polynomial transfer function[1] such that $\deg(F_i(z)) = M$ for $i = 1, p$. Then

$$\Pi T_{M-1}(F) = 0 \tag{13}$$

if and only if $F(z) = H(z)R$ for some invertible constant matrix R.

[1]The matrix polynomial $F(z)$ is said to be full rank if its columns, considered as elements of the rational space $\mathcal{F}_q$, are linearly independent. This, of course, is equivalent to $\text{Rank}(F(z)) = p$ for almost all z.

$$S'(z) = \begin{bmatrix} A(z) & B(z) \\ C(z) & D(z) \end{bmatrix} = \begin{bmatrix} V(R_0 + S_+(z) + S_+(z^{-1}))V^T & V(R_0 + S_+(z) + S_+(z^{-1}))V'^T \\ V'(R_0 + S_+(z) + S_+(z^{-1}))V^T & V'R_0V'^T + V'(S_+(z) + S_+(z^{-1}))V'^T \end{bmatrix}$$

Proof: The matrix $\mathcal{R}$ defined by (2) satisfies (3). On the other hand, using (12), it is easily seen that Conditions C7) and C3)–C5) imply that $T_{M-1}(H)$ is full column-rank. Hence, as in the case $p = 1, \text{Ker}^l(\mathcal{R}) = \text{Ker}^l(T_{M-1}(H))$. Let $g = (g_0, \cdots, g_{M-1})$ be a row vector of $\mathbb{R}^{qM}$, and let $g(z) = \Sigma_{k=0}^{M-1} g_k z^{-k}$ denote the corresponding transfer function. They satisfy

$$gT_{M-1}(H) = 0 \Leftrightarrow g(z)H(z) = 0 \quad (14)$$

and so the left kernel of $T_{M-1}(H)$ yields the polynomials $g(z) \in \mathcal{B}$, the degrees of which are less than or equal to $M-1$. Now, due to C7), $\mathcal{B}$ admits a polynomial basis, the elements of which have a degree less than $M - 1$. It is easy to show that a $q \times p$ polynomial matrix $F(z)$ satisfies $\Pi T_{M-1}(F) = 0$ if and only if the columns of $F(z)$ belong to $\mathcal{S}$. If $F(z)$ is assumed to be a full rank matrix polynomial, and if $\deg(F_i(z)) = M$ for $i = 1,p$, then $F(z)$ is a minimal polynomial basis of $\mathcal{S}$. As $H(z)$ is also a minimal polynomial basis of $\mathcal{S}$, by Proposition 3.1 there exists a constant invertible $p \times p$ matrix R such that $F(z) = H(z)R$. ∎

In other words, under C7), $H(z)$ can be identified up to a constant matrix by solving (13). The remaining constant matrix R can be extracted by using the relation $R_M = H_0 H_M^T$. However, in contrast with the case $p = 1$ (where, by (11), C1) and C7) are equivalent), C7) is difficult to interpret and seems, moreover, quite restrictive. Consider, for instance, the case $p = 2, q = 5$, and $M = 5$. Among the 14 triples $M_1^\perp, M_2^\perp, M_3^\perp$ satisfying $0 \leq M_1^\perp \leq M_2^\perp \leq M_3^\perp$ and $\Sigma_{j=1}^3 M_j^\perp = 10$ only two satisfy C7). This underscores the need to find alternate techniques.

3) The Improved Subspace Method: In this section, we show how to improve the subspace method. In particular, we show that Condition C7) is actually not necessary to get a simple identification procedure. Assume that $q > 2p$. Using equality (11), it is clear that at least one of the indexes $(M_i^\perp)_{i=1,q-p}$ is less than or equal to $M - 1$. Let r denote the greatest index for which $M_r^\perp \leq M - 1$, and $g_1(z), \cdots, g_r(z)$ r linearly independent polynomials of $\mathcal{B}$ of respective degrees $(M_i^\perp)_{i=1,r}$. Suppose for the moment that the left kernel of $T_{M-1}(H)$ can be extracted from the sequence $(R_n)_{n\geq 1}$. Then, from (14), the space $\text{Ker}^l(T_{M-1}(H))$ potentially allows the reconstruction of a polynomial basis $G(z)$ of the r-dimensional rational subspace generated by the $1 \times q$ polynomials $g_1(z), \cdots, g_r(z)$. Of course, $G(z)$ satisfies $G(z)H(z) = 0$. We shall show how to use this information to identify the unknown coefficient R_0. Since $H(z)$ is outer, the subsequent retrieval of $H(z)$ is then immediate. In the following, $S(z)$ will denote the spectral density $H(z)H^T(z^{-1})$.

Let us first indicate how to retrieve the space $\text{Ker}^l(T_{M-1}(H))$. Note that if Condition C7) does not hold, $T_{M-1}(H)$ is not necessarily full column-rank. Therefore, the equality $\text{Ker}^l(T_{M-1}(H)) = \text{Ker}^l(\mathcal{R})$ does not necessarily hold. However, the following result holds.

Lemma 3.1: Let J be the q-block exchange matrix defined in Section III-A and $\mathcal{H}$ the block Hankel matrix (6). Then

$$\text{Ker}^l(T_{M-1}(H)) = J\ \text{Ker}^l(\mathcal{H}) \cap \text{Ker}^l(\mathcal{H})^\perp. \quad (15)$$

Proof: From the factorization (6), $T_{M-1}(H) = [\mathcal{C}, J\mathcal{O}]$. Therefore, a row vector g satisfies $gT_{M-1}(H) = 0$ if and only if $g \in \text{Ker}^l(\mathcal{C})$ and $g \in J\ \text{Ker}^l(\mathcal{O})$. Since $\mathcal{O}$ and $\mathcal{C}$ are full rank, $\text{Ker}^l(\mathcal{H}) = \text{Ker}^l(\mathcal{O})$ and $\text{Ker}(\mathcal{H}) = \text{Ker}^l(\mathcal{C})$. The remainder is obvious. ∎

Hence, one can compute $\text{Ker}^l(T_{M-1}(H))$ from $\{R_i\}_{i=1}^M$ *without* requiring any longer that $T_{M-1}(H)$ be full rank. Therefore, it is possible to extract a polynomial basis $G(z)$ of the rational space generated by the $1 \times q$ polynomials $g(z) \in \mathcal{B}$ satisfying $\deg(g(z)) \leq M-1$. Of course, $G(z)$ is a $r \times q$ polynomial matrix. We now indicate how to use $G(z)$ to identify $H(z)$.

Theorem 3.3: Let $G(z) = \Sigma_{i=0}^{M-1} G_i z^{-i}$ and $\mathcal{G} \stackrel{\text{def}}{=} [G_0^T \cdots G_{M-1}^T]^T$. If $\mathcal{G}$ is full column-rank q, then $H(z)$ is identifiable.

Proof: The proof is constructive and thus provides an identification procedure. $0 = G(z)H(z)H^T(z^{-1}) = G(z)S(z) = G(z)(R_0 + S_+(z) + S_+^T(z^{-1}))$, with $S_+(z) = \Sigma_{i=1}^M R_i z^{-i}$. So, $G(z)R_0 = -G(z)(S_+(z) + S_+^T(z^{-1}))$. Equaling on both sides the coefficients of $z^0, z^{-1} \cdots, z^{-M+1}$, we get $\mathcal{G}R_0 =$ some known matrix T. R_0 is unique if $\mathcal{G}$ is full column-rank (which implies in particular that $Mr \geq q$), and thus the completion problem $\{R_i\}_{i=1}^M \to R_0$ is solved uniquely. ∎

Let us finally consider the case where $q \leq Mr$ and $\text{Rank}(\mathcal{G}) = q_1$ with $q_1 < q$. Performing an SVD on $\mathcal{G}$, one can write $\mathcal{G} = \mathcal{G}'V$, where $\mathcal{G}' = [G_0'^T \cdots G_{M-1}'^T]^T$ is a full column-rank $Mr \times q_1$ matrix and V is $q_1 \times q$ with $VV^T = I_{q_1}$. This holds if and only if $G(z) = G'(z)V$, where $G'(z) = \Sigma_{i=0}^{M-1} G_i' z^{-i}$ and thus if and only if the rows of $G(z)$ lie for each z in the constant subspace spanned by the rows of V. As mentioned above, $\mathcal{G}R_0 = T$ where T is some known matrix. This now reads $\mathcal{G}'(VR_0) = T$, and so for the moment only VR_0 is identifiable since $\mathcal{G}'$ is full column-rank. We now show that under some additional conditions, R_0 can still be determined by rank properties.

Let us complete V with $q - q_1$ additional rows V' so that $\mathcal{V} \stackrel{\text{def}}{=} [V^T V'^T]^T$ is an orthogonal matrix. Let $S'(z) = \mathcal{V}S(z)\mathcal{V}^T$. Then we have the equation shown at the top of the page, and all the terms in the above partitioning have been identified, except $V'R_0V'^T$ in the $(q - q_1) \times (q - q_1)$ right down corner.

Theorem 3.4: If the normal rank of $A(z)$ is equal to p, then $H(z)$ is identifiable.

Proof: Again, the proof is constructive. Let z_0 be a point for which $\text{Rank}(A(z_0)) = p$. As $\text{Rank}(S'(z_0)) = p$, it is clear that $D(z_0) = C(z_0)A(z_0)^\# B(z_0)$, where $^\#$ denotes any generalized inverse of a matrix. Using the expression of $D(z)$, this last identity determines $V'R_0V'^T$ uniquely. ∎

Let us finally study the condition: normal rank of $A(z) = p$. First, it implies that $q_1 \geq p + r$. In effect, $G(z)H(z) = 0$ for all z implies that $G'(z)A(z) = 0$ for all z. So, $\text{Span}(A(z)) \subset \text{Ker}(G'(z))$ for all z. On the other hand, as the rows of $G(z)$ are linearly independent over the field $\mathcal{F}_1$ of all rational functions, $\text{Rank}(G(z)) = r$ for almost all $z \in \mathbb{C}$. Since $G(z) = G'(z)V$, we also have $\text{Rank}(G'(z)) = r$, i.e., $\dim(\text{Ker}(G'(z)) = q_1 - r$ for almost all z. Hence, $\text{rank}(A(z)) \leq q_1 - r$ for almost all z, i.e., $p \leq q_1 - r$.

Proposition 3.2: The normal rank of $A(z)$ is equal to p if and only if $\text{Span}(H(z)) \cap \text{Ker}(V) = \{0\}$ for almost all z.

Proof: This follows directly from the identity $A(z) = VH(z)H^T(z^{-1})V^T$. ∎

Concluding Remark: One should note that if Conditions C3)–C5) are supposed to hold, p and M can be easily identified from the truncated sequence $(R_k)_{k\geq 1}$. In fact, M is the smallest index for which $R_k = 0$ for each $k \geq M+1$. On the other hand, $R_M = H_0 H_M^T$. Therefore, $p = \mathrm{Rank} R_M$. However, one should note that it seems to be difficult to check that C3)–C5) hold.

References

[1] K. Abed Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," *IEEE Trans. Information Theory*, to appear.

[2] K. Abed Meraim, "Second order blind identification/equalization: Algorithms and statistical performance," Ph.D. dissertation, Ecole Nationale Supérieure des Télécommunications, Paris, July 1995.

[3] I. Fijalkow and P. Loubaton, "Identification of rank one rational spectral densities from noisy observations: A stochastic realization approach," *Syst. Contr. Lett.* vol. 24, pp. 201–205, 1995.

[4] G. D. Forney, "Minimal bases of rational vector spaces, with applications to multivariable linear systems," *SIAM J. Contr.*, vol. 13, no. 3, pp. 493–520, May 1975.

[5] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[6] S. Kung, T. Kailath, and M. Morf, "A generalized resultant matrix for polynomial matrices," in *Proc. IEEE Conf. Decision Contr.*, 1976.

[7] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Proc. Asilomar Conf.*, 1991, pp. 856–860.

[8] A. van der Veen, S. Talwar, and A. Paulraj, "Blind identification of FIR channels carrying multiple finite alphabet signals," in *Proc. ICASSP*, Detroit, MI, May 1995, pp. 1213–1216.

A Single Sample Path-Based Performance Sensitivity Formula for Markov Chains

Xi-Ren Cao, Xue-Ming Yuan, and Li Qiu

***Abstract*—Using a sample path approach, we derive a new formula for performance sensitivities of discrete-time Markov chains. A distinguished feature of this formula is that the quantities involved can be estimated by analyzing a single sample path of a Markov chain. Thus, the formula provides a new direction for sensitivity analysis and can be viewed as an extension of the perturbation realization theory to problems where infinitesimal perturbation analysis does not work well.**

I. Introduction

In this paper, we derive a new formula for performance sensitivities of homogenous discrete-time Markov chains (we shall simply use the term Markov chain hereafter). We use a sample path approach, i.e., we analyze a sample path to determine the performance change due to a change in the transition probability matrix. The formula shows that the derivative of a performance measure equals the weighted sum of the expected value of a quantity, called a *realization factor*, that measures the average performance change when the Markov chain changes from one state to another. The realization factors can be determined by solving a skew-symmetric Lyapunov equation. Furthermore, these realization factors can be estimated by analyzing a single sample path obtained from simulation or a record of a real system.

Manuscript received August 22, 1995; revised March 5, 1996. This work was supported in part by Hong Kong UPGC under Grant HKUST 599/94E and by the Hongkong Telecom Institute of Information Technology.

X.-R. Cao and L. Qiu are with the The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: eecao@ee.ust.hk).

X.-M. Yuan is with the The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong and is on leave from the Chinese Academy of Sciences.

Publisher Item Identifier S 0018-9286(96)08445-0.

The formula provides a new perspective to the sensitivity analysis of Markov chains. It shows that a single sample path contains all the information needed for determining the performance sensitivities in a Markov chain. The approach can be viewed as an extension of the realization theory [1] in infinitesimal perturbation analysis (IPA) [5], [6] to the case where a small change may induce a large change in the sample path. This work was motivated by a recent work of Dai and Ho [4] and can be considered a theoretical justification of the algorithm in [3].

The formula can be easily generalized to continuous-time Markov chains. Since a Markov chain is the main model for many stochastic systems such as queueing systems, the formula developed here may have an impact in other fields, especially in the field of single sample path-based sensitivity analysis and performance optimization.

The paper is organized as follows. Section II introduces the basic concepts derived from a sample path point of view. We consider the simplest, but fundamental, case where one transition probability increases and another transition probability decreases by the same amount. We show how the sensitivity formula was derived by using intuition. Section III provides a rigorous proof for the general results where the transition matrix changes arbitrarily within the constraint of a stochastic matrix. The fundamental case discussed in Section II becomes a special case. Section IV discusses the implication of the results.

II. The Basic Concepts

In this section, by intuition we derive the performance sensitivity formula for the most fundamental case. The basic concepts are introduced. The rigorous proof of the general formula and other results will be provided in the next section.

Consider an irreducible and aperiodic Markov chain $X = \{X_n; n \geq 0\}$ on a finite state space $\mathcal{E} = \{1, 2, \cdots, M\}$ with transition probability matrix $P = [p_{ij}]_{i=1}^M|_{j=1}^M$. Let $f: \mathcal{E} \rightarrow \mathcal{R}$, where $\mathcal{R} = (-\infty, \infty)$ represents the space of real numbers, and $\pi = (\pi_1, \pi_2, \cdots, \pi_M)$ is the steady-state probability vector of X. f is called a *performance function*. The *performance measure* is defined as its expected value with respect to π

$$\eta = E(f) = \sum_{i=1}^{M} \pi_i f(i) = \pi f \tag{1}$$

where $f = (f(1), f(2), \cdots, f(M))^T$ is a column vector, and E denotes the expectation with respect to the steady-state measure π.

Assume that the transition matrix P is perturbed to $P' = [p'_{ij}]_{i=1}^M|_{j=1}^M$, where

$$p'_{ij} = \begin{cases} p_{ls} - \delta, & i = l, j = s \\ p_{lt} + \delta, & i = l, j = t \\ p_{ij}, & \text{otherwise} \end{cases} \tag{2}$$

for some arbitrarily fixed $l, s, t \in \mathcal{E}$, and $\delta > 0$ is any small real number. Let $Q = [q_{ij}]$ be a matrix with $q_{ls} = -1, q_{lt} = 1$, and $q_{ij} = 0$ for all $i \neq l$ or $j \neq s, t$. Then we have $P' = P + \delta Q$. The Markov chain with transition matrix P' is denoted as $X' = \{X'_n; n \geq$