Rapports de Recherche

N° 478

OPTIMUM ROBUST DETECTION
OF CHANGES IN THE AR PART
OF A MULTIVARIABLE
ARMA PROCESS

Anne ROUGÉE
Michèle BASSEVILLE
Albert BENVENISTE
Georges MOUSTAKIDES

Janvier 1986
OPTIMUM ROBUST DETECTION OF CHANGES IN
THE AR PART OF A MULTIVARIABLE ARMA PROCESS

Anne Rougé (IRISA/Univ), Michèle Basseville (IRISA/CNRS),
Albert Benveniste (IRISA/INRIA), Georges Moustakides (IRISA/INRIA)

Résumé.— On étudie les propriétés d'un nouveau test statistique, dit instrumental,
récemment proposé [5] pour la détection et le diagnostic de changements dans
la partie AR d'un processus ARMA vectoriel. On exhibe le nombre optimal
d' instruments et la matrice de réduction optimale. On établit le lien avec la
précision de la méthode d'identification par variables instrumentales [15]. On
effectue la comparaison avec les tests locaux de vraisemblance.

Ces tests ont été développés comme solution au problème de la surveillance
des vibrations pour les plateformes offshore.

Abstract.— We investigate the theoretical properties of a new instruments-
based test statistics recently proposed [5] for detection and diagnosis of changes
in the AR part of a multivariable ARMA process. The optimum number of
instruments and reduction matrix are exhibited. The connection with the accuracy
of the IV identification method [15] is established. The comparison with local
likelihood tests is done.

These tests have been developed as a solution to the problem of vibration
monitoring for offshore platforms.
OPTIMUM ROBUST DETECTION OF CHANGES IN THE AR PART OF A MULTIVARIABLE ARMA PROCESS

Anne Rougée, Michèle Basseville*, Albert Benveniste**, Georges Moustakides***

Abstract.- We investigate the theoretical properties of a new instruments-based test statistics recently proposed [5] for detection and diagnosis of changes in the AR part of a multivariable ARMA process. The optimum number of instruments and reduction matrix are exhibited. The connection with the accuracy of the IV identification method [15] is established. The comparison with local likelihood tests is done.

These tests have been developed as a solution to the problem of vibration monitoring for offshore platforms.

This work was supported by IFREMER grant n°84/7392 and by CNRS GRECO SARTA

The authors are with IRISA, Campus de Beaulieu, 35042 Rennes Cédex, France

Tél.: 99.36.20.00

*Also with CNRS
**Also with INRIA
I. INTRODUCTION

The problem of detecting changes in the properties of systems or signals is of particular interest both for failure detection in dynamical systems, industrial processes..., and for segmentation of digital signals in view of recognition. This problem has received an increasing attention these last fifteen years, in many fields of applications as can be seen from the survey papers \[20\] \[13\] \[10\] and the more methodological works \[9\] \[15\] \[4\].

In the present paper, we address the problem of detecting changes in the AR part of a multivariable ARMA process, or equivalently changes in the eigenstructure of a multivariable system. In addition to the detection problem, we also consider diagnosis problem, namely the decision concerning which poles and modes have changed. These types of problems arise for example in the domain of vibration monitoring of structures submitted to natural excitation, such as offshore platforms. In such a case, the change detection and diagnosis problems are still more complex: because the excitation is not measurable and nonstationary, it has to be considered as a nuisance parameter. For a complete description of this application and the underlying motivations, we refer the reader to \[5\]. In order to solve these problems which cannot be solved by likelihood methods \[5\], we recently proposed some new instrumental tests, the numerical properties of which have been investigated in \[3\] for scalar signals and in \[5\] for multivariable systems.

The purpose of this paper is the investigation of the theoretical properties of these tests with special emphasis on the connection between our
instrumental tests and, on one hand, recent results concerning the instrumental variable identification method \cite{19}, and, on the other hand, robustness properties of likelihood ratio tests. Accordingly, the paper is organized as follows. In section II, after problem formulation, we consider the question of optimization of the number of instruments and of the reduction matrix to be used in the test. In section III, we give a special attention to the AR case, for which we show that the optimum power of the test is attained with a finite number of instruments. In section IV, the comparison with the accuracy of the IV identification method is done. In section V, we investigate the connections between local likelihood tests and our instrumental tests, and show that these last ones are min-max optimal in the scalar case. In section VI, we show that, in the scalar case, the use of a finite number of filtered instruments results in attaining the asymptotically optimal power of the test. Finally, some conclusions are presented in section VII.

II. AN INSTRUMENTAL STATISTICS: DEFINITION AND OPTIMIZATION

In this section, we first introduce our instrumental statistics for solving the problem of detection of changes in the AR part of a multivariable ARMA process.

Then we exhibit the optimum number of instruments and reduction matrix to be used in order to optimize the power of the test. We recall that, in this paper, we consider the stationary case. We refer the interested reader to \cite{5} for problem statement, test implementation and numerical results related to the nonstationary case.
1. Problem statement and test definition

We consider a multivariable process, described either by the state-space representation:

\[
\begin{align*}
X_{t+1} &= F X_t + V_{t+1} \\
Y_t &= H X_t
\end{align*}
\]  \hspace{1cm} (1)

where \( X_t \in \mathbb{R}^n \), \( Y_t \in \mathbb{R}^p \), \( \text{cov} (V_{t+1}) = Q \),

or equivalently by the ARMA representation:

\[
Y_t = \sum_{i=1}^{p} A_i Y_{t-i} + \sum_{j=0}^{p-1} B_j E_{t-j}
\]  \hspace{1cm} (2)

where \( (E_t) \) is a standard white noise.

As mentioned in the Introduction, we are interested in detecting and diagnosing small changes in the state transition matrix \( F \) (resp. in the AR parameters \( (A_i) \)), while the state noise covariance matrix \( Q \) (resp. the M.A. part \( (B_j) \)) is unknown (and time-varying in \(|5|\)). We assume that a nominal observable model \( (H_0, F_0) \) is available, and thus a nominal AR model \( \theta^0 \) also, where:

\[
\theta^T = (A_p, \ldots, A_1)
\]  \hspace{1cm} (3)
A new sample $Y_1, \ldots, Y_n$ is observed, and, following a model validation approach, we wish to test whether it is conveniently described by the reference model $\theta^0$ or not. For this purpose, we use a local approach \cite{11, 7, 14}, i.e. we consider the following hypotheses:

$$H_0 : \theta = \theta^0 \text{ no change}$$

$$H_1 : \theta = \theta^0 + \delta \theta$$

where $\delta \theta$ is a possible change direction. We define what we call an instrumental statistics by:

$$U_n(s) = \frac{1}{\sqrt{s}} \sum_{t=1}^{s} Z_t^N w_t^T$$

where

$$Z_{NT}^T = (Y_{t-p}^T, \ldots, Y_{t-p-N+1}^T)$$

is the vector of instruments

$$w_t = Y_t - \sum_{i=1}^{p} A_1 Y_{t-1} \Delta Y_t - \theta^0 t_1 \phi_t$$

$$\phi_t = (Y_{t-p}^T, \ldots, Y_{t-1}^T)$$

Notice that $U_n(s)$ may be generated in another way, using the following formula:

$$U_n(s) = \mathcal{B}_{p+1,N}(s) \begin{pmatrix} \theta^0 \\ -I_r \end{pmatrix} \begin{pmatrix} \theta^0 \\ \end{pmatrix}$$
where

$$
\mathcal{H}_{p,q}(s) = \begin{pmatrix}
    R_0(s) & \cdots & R_{q-1}(s) \\
    \vdots & \ddots & \vdots \\
    R_{p-1}(s) & \cdots & R_{p+q-2}(s)
\end{pmatrix}
$$

(6)

is the empirical Hankel matrix in which:

$$
R_m(s) = \sum_t y_{t+m} y_t^T.
$$

Then, denote by:

$$
\mathcal{H}_{p,q} = \mathcal{E}_0 \left( \frac{1}{n} \mathcal{H}_{p,q}(s) \right)
$$

(6')

the expected value of (6) under the hypothesis $H_0$.

Finally, we introduce the corresponding vectors:

$$
\mathcal{U}_N(s) \triangleq \text{col} \left( U_N^T(s) \right)
$$

(7)

$$
= \sum_{t=1}^{N} z_t^N \otimes W_t
$$

obtained by stacking the $Nr$ columns of $U_N^T(s)$ on top of each other, and:

$$
\mathcal{G} = \text{col} \left( \theta^T \right)
$$

(8)
Under the no change hypothesis $H_0$, $W_t$ in (4) is a MA process, uncorrelated with $Z_t^N$, and $U_N(s)$ is zero-mean. Under the hypothesis $H_1$, we have:

$$
\mathbb{E}_1(U_N(s)) = \mathbb{E}_1\left[ \frac{1}{\sqrt{s}} \sum_{t=1}^{s} Z_t^N (Y_t - \theta^T \phi_t) + \frac{1}{\sqrt{s}} \sum_{t=1}^{s} Z_t^N \phi_t^T (\theta - \theta^0) \right]
$$

$$= \mathcal{H}^T_{p,N} \delta \theta
$$

Correspondingly,

$$
\mathbb{E}_1(U_N(s)) = \left( \mathcal{H}^T_{p,N} \otimes I_r \right) \delta \Theta
$$

Therefore, using $U_N(s)$, we will be able to detect any change $\delta \theta$ belonging to the range of $\mathcal{H}^T_{p,N}$. Let us now investigate this point.

We first introduce the following classical notations:

$$
\Theta^p(H_0,F_0) = \begin{pmatrix}
H_0 \\
H_0F_0 \\
\vdots \\
H_0F_0^{p-1} \\
H_0F_0^p
\end{pmatrix}
$$

(10)

and

$$
\mathcal{C}_N(F_0,G_0) = (G_0,F_0G_0,\ldots,F_0^{N-1}G_0)
$$

(11)

where

$$
G_0 = \mathbb{E}_0(X_t, Y_t^T).
$$

From now on, we assume that the nominal representation $(H_0,F_0)$ (1) is observable and that the following factorization holds:

$$
\mathcal{H}_{p,N} = \Theta^p(H_0,F_0) \cdot \mathcal{C}_N(F_0,G_0)
$$

(12)
where \( \mathbf{Q}_N(F_0, G_0) \) is of full row rank \( n \). In such a case, because of (12), the only changes on \( \theta \) we will not be able to detect with the aid of \( U_N(s) \) are those that satisfy:

\[
\mathbf{Q}_p^T(H_0, F_0) \delta \theta = 0
\]

(13)

But these that last changes do not correspond to any change in the minimal representation (1) of the system. The reason for that is as follows. It is well known that the representations (1) and (2) are connected through the equation:

\[
\mathbf{Q}_p^T (H_0, F_0) \theta^0 = 0
\]

(14)

and any \( \theta^0 \) satisfying this relation leads to a valid ARMA representation of the system. But two different parameters \( \theta^0 \) and \( \theta^0 + \delta \theta \) satisfying both (14) are precisely related through (13).

Consequently (9) means that any change in the minimal representation of the system will result in a change in the mean value of the process \( U_N(s) \) (4).

Furthermore, it may be shown that (even under nonstationarity assumptions [12]) the following local asymptotic normalities hold:

under \( H_0 \), \( U_N(s) \xrightarrow{s \to \infty} \mathcal{N}(0, \Sigma_N) \)

under \( H_1 \), \( U_N(s) \xrightarrow{s \to \infty} \mathcal{N}(\mathbf{Q}_p^T \otimes \mathbf{I}_r \delta \mathbf{Q}, \Sigma_N) \)
where:
\[ \Sigma_N = \sum_{i=1-p}^{p-1} E_0 (Z_N Z_{NT}^T \otimes W_t W_{t-1}^T) \]  
\[(15)\]

is the covariance matrix of \( U_N \).

Finally, before defining our test, we recall a classical result in gaussian hypothesis testing. Let \( U \) be a random variable distributed as \( \mathcal{N}(u, \Sigma) \). For testing \( u = 0 \) against \( u = M \nu \), the log-likelihood ratio is:

\[ T = -\frac{1}{2} (U - M \nu)^T \Sigma^{-1} (U - M \nu) + \frac{1}{2} U^T \Sigma^{-1} U \]

\[ = U^T \Sigma^{-1} M \nu - \frac{1}{2} \nu^T M^T \Sigma^{-1} M \nu \]  
\[(16)\]

If \( M \) is of full column-rank, the maximum likelihood estimate of \( \nu \) is:

\[ \hat{\nu} = (M^T \Sigma^{-1} M)^{-1} M^T \Sigma^{-1} U \]  
\[(17)\]

and including it in (16), we get:

\[ T = \frac{1}{2} U^T \Sigma^{-1} M (M^T \Sigma^{-1} M)^{-1} M^T \Sigma^{-1} U \]  
\[(18)\]

If \( M \) is not of full column-rank, let \( E \) be the matrix containing the basis vectors of a complement of the kernel of \( M \); then:

\[ \nu = E \nu + \nu_0. \]
where \( \psi \in \text{Ker}(M) \), and thus:

\[
\mu = M\nu = \varepsilon \beta
\]

with ME of full column-rank. In this case, the \( \chi^2 \) statistics:

\[
U^T \Sigma^{-1} \text{ME} (E^T M^T \Sigma^{-1} \text{ME})^{-1} E^T M^T \Sigma^{-1} U
\]  

is the GLR test for testing \( \mu = 0 \) against \( \mu = \varepsilon \beta \).

Using (9) and a reduction matrix \( D \) such that:

\[
M = H_{p,N}^T D^T \otimes I_r
\]

is of full column-rank, (19) results in our instrumental test:

\[
t_0(s) = U_{N}(s) \Sigma^{-1} M (M^T \Sigma^{-1} M)^{-1} M^T \Sigma^{-1} U_{N}(s)
\]  

with \( M \) given by (20).

As shown in [12], [16], \( H_{p,N} \) and \( \Sigma_N \) may be replaced by their corresponding sample means computed on the new record, even in case of nonstationary excitation. Therefore, the test (21) may be also implemented in this last case of nonstationary MA part.

Numerical experiments emphasizing the efficiency of this test are reported in [3] for scalar signals and [5] for multivariable systems.
2. A criterion for performance evaluation

From now on, we investigate the performance of the test (21) using the following classical performance index: for a fixed level (or false alarm probability) $\alpha$, the threshold is given by:

$$P_0 (t_0 > \lambda) = \alpha$$  \hspace{1cm} (22)

and the performance index is the power of the test:

$$\beta = P_1 (t_0 > \lambda)$$  \hspace{1cm} (23)

which is to be optimized.

Because of the asymptotical normality (15) of $U_N(s)$ under both hypotheses $H_0$ and $H_1$ \cite{12}, the asymptotic distribution of $t_0$ (21) is a $\chi^2$ distribution with $n-r$ degrees of freedom under both $H_0$ and $H_1$. Under $H_1$, this distribution is non central, with non centrality parameter equal to:

$$\gamma = \delta \otimes \Gamma_{N,D} \delta \otimes \Gamma$$  \hspace{1cm} (24)

where

$$\Gamma_{N,D} = (\mathcal{H}_{p,N} \otimes I_r) \Lambda_N^{-1} M (M^T \Lambda_N^{-1} M)^{-1} M^T \Lambda_N^{-1} (\mathcal{H}_{p,N}^T \otimes I_r)$$  \hspace{1cm} (25)

and $M$ is given by (20).
Consequently the test threshold $\lambda$ is chosen in (22) independently of $N$ and $D$, and the asymptotic power of the test, for a fixed level $\alpha$, is:

$$
\beta_{N,D} = \mathbb{P}(\chi_r^2(m, r) > \lambda)
$$

(26)

As the function $\gamma \rightarrow \mathbb{P}(\chi_r^2(m, \gamma) > \lambda)$ is increasing whatever $\lambda$ is, the power $\beta_{N,D}$ is an increasing function of $\gamma$ (24), which is a quadratic form. Therefore we will consider the optimization of this quadratic form defined by $\Gamma_{N,D}$.

3. Influence of the reduction matrix

We first show that $\Gamma_{N,D}$ does not depend upon the reduction matrix $D$ introduced in (20). As $D$ is such that:

$$
E \begin{bmatrix} \Lambda & D \\ 0 & 0 \end{bmatrix} \begin{bmatrix} H_0 & F_0 \end{bmatrix}
$$

is invertible, and because of the factorization (12) of the Hankel matrix, we have:

$$
M^T = E \begin{bmatrix} \mathcal{L}_0^0 & \mathcal{I}_r \end{bmatrix}_N
$$

and

$$
(M^T \Sigma^{-1}_N M)^{-1} = (E^{-T} \otimes \mathcal{I}_r) \left( (\mathcal{L}_0^0 \otimes \mathcal{I}_r) \Sigma^{-1}_N (\mathcal{L}_r^0 \otimes \mathcal{I}_r) \right)^{-1} (E^{-1} \otimes \mathcal{I}_r)
$$

Therefore, we may rewrite (25) as:

$$
\Gamma_{N,D} = (\mathcal{H}_p,N \otimes \mathcal{I}_r) \Sigma^{-1}_N (\mathcal{L}_r^0 \otimes \mathcal{I}_r) \left( (\mathcal{L}_0^0 \otimes \mathcal{I}_r) \Sigma^{-1}_N (\mathcal{L}_r^0 \otimes \mathcal{I}_r) \right)^{-1} (\mathcal{L}_0^0 \otimes \mathcal{I}_r) \\
= (\mathcal{L}_p^0 \otimes \mathcal{I}_r) \Sigma^{-1}_N (\mathcal{L}_r^0 \otimes \mathcal{I}_r) (\mathcal{H}_p,N \otimes \mathcal{I}_r)
$$

(27)
In all these computations, we make extensive use of the properties of the Kronecker product, as summarized in [21] for example.

4. Optimizing the number $N$ of instruments

We now show that the sequence $(\Gamma_n^N)_{N \leq N^p}$ is nondecreasing, with respect to the ordering of positive matrices; for this purpose, we introduce the following partitioned matrices for computing $\Gamma_{N+1} - \Gamma_N$:

$$
\mathbf{U}_{N+1}^{(s)} = \begin{pmatrix}
\mathbf{U}_N^{(s)} \\
\frac{1}{\sqrt{s}} \sum_{t=1}^s \gamma_{t-p-N} \otimes \mathbf{w}_t
\end{pmatrix}
$$

where $\mathbf{U}_N^{(s)}$ is defined in (7):

$$
\mathbf{H}_{p,N+1} = (\mathbf{H}_{p,N} : \mathbf{H}_{N+1})
$$

and

$$
\mathbf{L}_{N+1} = \begin{pmatrix}
\mathbf{L}_N & \mathbf{S}_2 \\
\mathbf{S}_2 & \mathbf{S}_1
\end{pmatrix}
$$

where

$$
\mathbf{S}_1 = \mathcal{E}_0 \left( \sum_{i=1-p}^{p-1} \gamma_{t-p-N} \gamma_{t-p-N-1}^T \otimes \mathbf{w}_t \mathbf{w}_t^T \right)
$$

$$
\mathbf{S}_2 = \mathcal{E}_0 \left( \sum_{i=1-p}^{p-1} \epsilon_{t-p-N} \epsilon_{t-p-N-1}^T \otimes \mathbf{w}_t \mathbf{w}_t^T \right)
$$

(28)
Then, using (27) and the inversion formula for partitioned matrices, we get:

\[
\Gamma_{N+1} - \Gamma_N = \left(\mathcal{H}_P, N \bigotimes I_r\right) L_N^{-1} S_2 \Delta^{-1} S_2^T L_N^{-1} \left(\mathcal{H}_P, N \bigotimes I_r\right)
\]

\[
- \left(\mathcal{H}_N, N \bigotimes I_r\right) S_2^T L_N^{-1} \left(\mathcal{H}_P, N \bigotimes I_r\right)
\]

\[
- \left(\mathcal{H}_P, N \bigotimes I_r\right) L_N^{-1} S_2 \Delta^{-1} \left(\mathcal{H}_N, N \bigotimes I_r\right)
\]

\[
+ \left(\mathcal{H}_N, N \bigotimes I_r\right) \Delta^{-1} \left(\mathcal{H}_N, N \bigotimes I_r\right)
\]

\[
= \left(\left(\mathcal{H}_N, N \bigotimes I_r\right) - \left(\mathcal{H}_P, N \bigotimes I_r\right) L_N^{-1} S_2 \Delta^{-1} \left(\left(\mathcal{H}_N, N \bigotimes I_r\right) - \left(\mathcal{H}_P, N \bigotimes I_r\right) L_N^{-1} S_2 \right) \right)^T
\]

where: \( \Delta = S_2^T L_N^{-1} S_2 \).

\( \Delta \) is positive definite because the same is true for \( \Gamma^{-1}_{N+1} \) \( \left[12\right] \).

Therefore, for all \( N \geq p \):

\[
\Gamma_{N+1} \succ \Gamma_N
\]

and \( \Gamma_{N+1} = \Gamma_N \) if and only if:

\[
\left(\mathcal{H}_P, N \bigotimes I_r\right) L_N^{-1} S_2 = \mathcal{H}_N, N \bigotimes I_r
\]  \( \left(29\right) \)

As a conclusion, let us summarize the results of this section by the following theorem:
Theorem 1

1) the asymptotic power of the instrumental test defined in (21) does not depend upon the reduction matrix \(D\);

2) this power increases with the number \(N\) of instruments, uniformly with respect to the change \(\delta\).

The most powerful instrumental test thus corresponds to the choice of an infinite number of instruments.

III. THE AR CASE

In this short section, we give a special attention to the case of AR processes, because of the properties of the instrumental test (21) in this situation: actually the condition (29) is satisfied when \((Y_t)\) is an AR process.

In fact, in this case (15) may be written as:

\[
\Sigma_N = E_0(Z_t^N Z_t^N)^T \otimes E_t E_t^T
\]

\[
= T_N \otimes \Lambda
\]

where \(T_N\) is the \(Nr \times Nr\) block-Toeplitz covariance matrix of \((Y_t)\) and

\[
\Lambda = E_0 (E_t E_t^T).
\]
Similarly, we rewrite (28) as:

\[ S_2 = \begin{pmatrix} R_N \\ \vdots \\ R_1 \end{pmatrix} \otimes \Lambda \]

Therefore (29) is equivalent to:

\[ \bigotimes_{p,N} T_N^{-1} \begin{pmatrix} R_N \\ \vdots \\ R_1 \end{pmatrix} = \begin{pmatrix} R_N \\ \vdots \\ R_{N+p-1} \end{pmatrix} \]

or in other words:

\[ \begin{pmatrix} R_k, \ldots, R_{N+k-1} \end{pmatrix} T_N^{-1} \begin{pmatrix} R_N \\ \vdots \\ R_1 \end{pmatrix} = R_{N+k} \quad (31) \]

for all \( k : 0 \leq k \leq p-1 \).

But, starting from:

\[ R_m = A_1 R_{m-1} + \ldots + A_p R_{m-p} \quad (m \geq 1) \]

we can show, by induction on \( k \), the existence of matrices \( A^{(k)} \) such that:

\[ R_{m+k-1} = \sum_{i=1}^{p} A^{(k)}_{i} R_{m-1} \quad (m \geq 2) \]

\[ = \begin{pmatrix} A^{(k)}_1 & \ldots & A^{(k)}_p & 0 & \ldots & 0 \end{pmatrix} \begin{pmatrix} R_{m-1} \\ \vdots \\ R_{m-N} \end{pmatrix} \quad (32) \]
Consequently:

\[
\begin{pmatrix}
R_k \\
R_{N+k-1}
\end{pmatrix} = \begin{pmatrix}
A(k) \\
\vdots \\
A(k) \\
\vdots \\
1 \\
\vdots \\
A(k) \\
\vdots \\
p \\
\vdots \\
0 \\
\vdots \\
0
\end{pmatrix}
\begin{pmatrix}
R_0 \\
R_1 \\
R_2 \\
\vdots \\
R_N \\
\vdots \\
R_{N+1} \\
R_0
\end{pmatrix}
\]

and thus:

\[
\begin{pmatrix}
R_k \\
R_{N+k-1}
\end{pmatrix} \Gamma_N^{-1} = \begin{pmatrix}
A(k) \\
\vdots \\
A(k) \\
\vdots \\
1 \\
\vdots \\
A(k) \\
\vdots \\
p \\
\vdots \\
0 \\
\vdots \\
0
\end{pmatrix}
\]  
\tag{33}

Using (33) and the relation (32) for \( m = N+1 \), we get (31) and thus (29).

Thus, in the AR case, we have:

\[
\Gamma_{N+1} = \Gamma_N = \Gamma_p \quad (\forall \ N \leq p)
\]  
\tag{34}

We have thus shown the following:

**Theorem 2**

If \( \{Y_t\} \) is an AR process, the sequence \( \{\Gamma_N\} \) is constant in \( N \).

This theorem means that the instrumental test (21) is as much powerful with \( p \) instruments as with an infinite number of instruments, in the AR case.

**IV. COMPARISON WITH THE ACCURACY OF THE IV IDENTIFICATION METHOD**

We now go back to the (general) case of ARMA processes. Consider the following particular case:

\[
\text{rank } (\mathcal{H}_{P,N}) = n = pr
\]  
\tag{35}

which is always satisfied in the scalar case \( r = 1 \).
We investigate the connection between the asymptotic power of our instrumental test (21) and the asymptotic accuracy of the IV identification method in the following extended form considered in [18]:

\[
\hat{\theta}_{IV} = \text{Arg min}_{\theta} \| \text{col} \left[ \sum_{t=1}^{s} Z_{t}^{N} (Y_{t} - \theta^T \phi_{t}) \right] \|_2^2
\]

where \(Z_{t}^{N}\) and \(\phi_{t}\) are defined in (4) and \(Q\) is a \(N r^2 \times N r^2\) symmetric positive definite matrix.

Using (8), the IV estimate is defined by:

\[
\hat{\Theta}_{IV} = \text{Arg min}_{\Theta \in R^{pr^{2}}} \| (\sum_{t=1}^{s} Z_{t}^{N} \phi_{t}^T \otimes I_r) \Theta - \sum_{t=1}^{s} (Z_{t}^{N} \otimes I_r) Y_{t} \|_Q^2 \tag{36}
\]

and is computed as the least squares solution of the following system:

\[
Q^{1/2} M_{s} \Theta = Q^{1/2} (\frac{1}{s} \sum_{t=1}^{s} Z_{t}^{N} \otimes Y_{t}) \tag{37}
\]

where \(M_{s} = \frac{1}{s} \sum_{t=1}^{s} Z_{t}^{N} \phi_{t}^T \otimes I_r\)

\[= \frac{1}{s} \mathcal{H}_{p, N(s)} \otimes I_r \]

Because of (35), the matrix:

\[M = \lim_{s \to \infty} M_{s} = \mathcal{H}_{p, N}^T \otimes I_r\]

is of full column rank \(pr^2\). Consequently, for \(s\) large enough, we have:

\[\hat{\Theta}_{IV} = (M_{s}^T Q M_{s})^{-1} M_{s}^T Q \left( \frac{1}{s} \sum_{t=1}^{s} Z_{t}^{N} \otimes Y_{t} \right)\]
and thus:

\[ \sqrt{s} (\hat{\Theta}_ IV - \Theta^0) = (M_s^T Q M_s)^{-1} M_s^T Q (\sum_{t=1}^{8} z_t^N) \otimes (Y_t - \Theta^0 \Phi_t) \]

where \( \Theta^0 \) is the true parameter vector, i.e., the solution of (14) which is unique under the condition (35) because of the factorization (12). Therefore:

\[ \sqrt{s} (\hat{\Theta}_ IV - \Theta^0) = (M_s^T Q M_s)^{-1} M_s^T Q U_N(s) \]  

(38)

Because \( \Theta^0 \) is the true AR parameter, \( U_N(s) \) is asymptotically Gaussian distributed with zero mean and covariance matrix \( \Lambda_N \) (15). Thus we re-obtain the central limit theorem of Stoica et al. [19] :

Theorem 3

Under the condition (35):

\[ \sqrt{s} (\hat{\theta}_ IV - \Theta^0) \xrightarrow{s \to \infty} \mathcal{N}(0, P_{IV}) \]

where

\[ P_{IV} = (M_s^T Q M_s)^{-1} M_s^T Q \sum_{t} Q M (M_s^T Q M)^{-1} \]

(39)

and 

\[ M = \sigma^T_{p,N} \otimes I_r \]

We now compare \( P_{IV} \) (39) and \( \Gamma_N \) (27). Because of (35), \( \Gamma_N \) is invertible and:

\[ \Gamma_N^{-1} = (M_s^T \Sigma_N^{-1} M)^{-1} \]

\[ \Delta = P_N \]
Then, we have:

\[ P_{IV} - P_N = R \Sigma_N R^T \]  \hspace{1cm} (40)

where:

\[ R = (M^T Q M)^{-1} M^T Q - (M^T \Sigma_N^{-1} M)^{-1} M^T \Sigma_N^{-1} M^T \Sigma_N^{-1} . \]

Since \( \Sigma_N \) is positive definite, with (40) we have proved the following theorem:

**Theorem 4**

For any matrix \( Q \), we have:

\[ P_{IV} \succ P_N = \Gamma_N^{-1} \]

and the equality is attained for \( Q = \Sigma_N^{-1} \).

This means that, under condition (35), the inverse of the matrix \( \Gamma_N \), which characterizes the asymptotic power of the instrumental test (21), is equal to the asymptotic covariance matrix of the estimation error of the optimal IV method (36) corresponding to \( Q = \Sigma_N^{-1} \).

Because of theorem 1, the asymptotic power of the test, and thus the asymptotic accuracy of the estimation, increase with the number of instruments, as shown in [19].

Finally, the Cramer-Rao's inequality applied to \( P_N \) for any \( N \gg p \) leads to:

\[ \Gamma_\infty \preceq (\mathcal{F}^{-1})^{-1}_{11} = F_{11} - F_{12} F_{22}^{-1} F_{21} \]  \hspace{1cm} (41)

where

\[ \mathcal{F} = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \]

is the Fisher information matrix of the ARMA process \( (Y_t) \). See Appendix 1.
V. COMPARISON WITH THE LOCAL LIKELIHOOD TEST

In this section, we investigate the connection between local likelihood tests and instrumental tests. We first consider the test based upon an asymptotic local expansion of the likelihood ratio test, and then apply it to the present problem of detecting changes in the AR part of the process \( (Y_t) \) with elimination of the nuisance parameters due to changes in the MA part.

1. Local likelihood test - Min-max approach

Let \( (Y_t) \) be an \( r \)-dimensional ARMA process:

\[
A(q^{-1}) Y_t = B(q^{-1}) E_t \quad (42)
\]

with

\[
A(q^{-1}) = \sum_{i=0}^{p} A_i q^{-i} \quad A_0 = -I_r
\]

\[
B(q^{-1}) = \sum_{i=0}^{q-1} B_i q^{-i} \quad B_1 = +I_r
\]

\[\text{Cov} (E_t) = \Lambda\]

Consider the log-likelihood:

\[L_{\theta}(\psi) \triangleq \log \mathcal{L}(\hat{Y}_1, \ldots, Y_n | \psi)\]

where \( \psi \) is the \( 2pr^2 \) vector of parameters:

\[\psi = \begin{pmatrix} \theta \\ \beta \end{pmatrix}\]
with \( \mathbf{A} \) defined in (8) and (3), and \( \beta \) defined similarly by:

\[
\beta = \text{col} (B_{p-1}, \ldots, B_1)
\]

It is well known that:

\[
L_s(\psi) = -\frac{1}{2} \sum_{t=1}^{S} E_t(\psi)^T \Lambda^{-1} E_t(\psi) - \frac{1}{2} \log|\Lambda| - \frac{r}{2} \log(2\pi)
\]

where \( E_t(\psi) = B^{-1}(q^{-1}) \cdot A(q^{-1}) Y_t \) is the innovation.

Let \( \Delta_s^s \psi \) be the gradient of \( L_s(\psi) \), which is computed in Appendix 1, and \( R_s(\psi^1, \psi^2) \) be the log-likelihood ratio. It is shown in [11] [17] [7] that this ratio has the following second order asymptotic expansion:

\[
R_s(\psi, \psi + \delta \psi) = \frac{1}{\sqrt{s}} \Delta_{\psi}^s (\psi)^T \delta \psi - \frac{1}{2s} \delta \psi^T \cdot F^s(\psi) \cdot \delta \psi + \frac{1}{s} \alpha(Y_1, \ldots, Y_s, \frac{\delta \psi}{\sqrt{s}})
\]

where

- \( \delta \psi \) is an arbitrary change direction,
- \( F^s(\psi) = E_{\psi} (\Delta_{\psi}^s (\psi) \cdot \Delta_{\psi}^s (\psi)^T) \) is the Fisher information matrix
- the residual term \( \alpha \) goes to zero in probability when \( s \) tends to infinity.

Assume that the nominal parameter \( \psi^0 \) is known. From [7], we know that the following convergences in distribution hold:

under \( H_0 : \psi = \psi^0, \frac{1}{\sqrt{s}} \Delta_{\psi}^s (\psi^0) \xrightarrow{\mathcal{L}_{s \to \infty}} \mathcal{N}(0, \mathcal{F}) \)

under \( H_1 : \psi = \psi^0 + \frac{\delta \psi}{\sqrt{s}}, \frac{1}{\sqrt{s}} \Delta_{\psi}^s (\psi^0) \xrightarrow{\mathcal{L}_{s \to \infty}} \mathcal{N}(\mathcal{F}, \delta \psi, \mathcal{F}) \)

where \( \mathcal{F} = F^1(\psi^0) \) is the Fisher's matrix under \( H_0 \).
Furthermore, the $\chi^2$ statistics:

$$\chi_s \triangleq \frac{1}{\sqrt{s}} \Delta^s (\psi^0)^T \cdot \mathbf{F}^{-1} \cdot \frac{1}{\sqrt{s}} \Delta^s (\psi^0)$$  \hspace{1cm} (45)$$

compared to a threshold is the uniformly most powerful (UMP) test for testing:

$$H_0 : \lambda = 0$$

against

$$H_1 : \lambda \neq 0$$

where $\lambda$ is defined by:

$$\lambda = \delta \psi^T \mathbf{F} \delta \psi \hspace{1cm} (46)$$

This assumes that the components of $\psi$ may change independently of each other, and thus the number of degrees of freedom is equal to the size of $\psi$. If the components of $\psi$ are constrained by:

$$\delta \psi = \mathbf{J} \delta \nu \hspace{1cm} (47)$$

where $\nu$ is another parameterization and $\mathbf{J}$ is of full column rank, the UMP test for $\lambda$ in (46) is defined by the following statistics:

$$\Delta^T \mathbf{F}^{-1} \mathbf{J}^T (\mathbf{J}^T \mathbf{F}^{-1} \mathbf{J})^{-1} \mathbf{J} \mathbf{F}^{-1} \Delta$$  \hspace{1cm} (48)$$

and the number of degrees of freedom is equal to the size of $\nu$. 
Now we are only interested in the changes in $\Theta$, and we look for a test which is robust with respect to the changes in $\beta$. Since the power of the $\chi^2$ test is an increasing function of $\lambda$ (46), we will try to find the least favorable changes $\delta\beta^*$ in $\beta$ minimizing the righthand side of (46) for fixed $\delta \Theta$. For this purpose, we use the following result:

$$
\min_{x,y} (x^T y) \mathcal{F}(x, y) = x^T (F_{11} - F_{12} F_{22}^{-1} F_{21}) x
$$

(49)

where minimum is attained for $y = -F_{22}^{-1} F_{21} x$, and $\mathcal{F}$ is as in (44).

Consequently:

$$
\min_{\delta \beta} \delta \psi^T \cdot \mathcal{F} \cdot \delta \psi = \delta \Theta^T \cdot \mathcal{F} \cdot \delta \Theta
$$

$$
= \delta \psi_*^T \cdot \mathcal{F} \cdot \delta \psi_*
$$

where

$$
\delta \psi_* = \begin{pmatrix} \delta \Theta \\ \delta \beta_* \end{pmatrix} = \begin{pmatrix} I \\ -F_{22}^{-1} F_{21} \end{pmatrix} \delta \Theta
$$

(50)

Consider now the test in (48) with:

$$
\mathcal{F} = \begin{pmatrix} I \\ -F_{22}^{-1} F_{21} \end{pmatrix}
$$
This gives:

\[ \chi_B^* = \frac{1}{\sqrt{s}} \Delta_s^{sT} \cdot \mathcal{F}^{-1} \cdot \frac{1}{\sqrt{s}} \Delta_s^s \]  

(51)

where \( \Delta_s^s \) is defined by:

\[ \Delta_s^s = (I - F_{12} F_{22}^{-1}) \Delta_s^\psi (\psi^0) \]  

(52)

Notice that \( \Delta_s^s \) is the gradient (in \( \psi^0 \)) of the log-likelihood \( L_s(\psi) \) in the direction \( \delta \psi \), because:

\[ \Delta_s^{sT} \cdot \delta \Theta = \Delta_s^\psi (\psi^0)^T \cdot \delta \psi \]  

It is easy to see that, under \( \overline{H}_0 \), \( \chi^* \) is a centered \( \chi^2 \) with a number of degrees of freedom equal to the size of \( \Theta \), and, under \( \overline{H}_1 \), is a non-centered \( \chi^2 \) with same number of degrees of freedom and non-centrality parameter equal to:

\[ \delta \Theta^T \cdot \mathcal{F} \cdot \delta \Theta \]  

(53)

which is independent of \( \delta \Theta \). (53) comes from (44), (52) and the equality:

\[ (I - F_{12} F_{22}^{-1}) \mathcal{F} \delta \psi = (\mathcal{F}, \ 0) \delta \psi \]

\[ = \mathcal{F}, \delta \Theta \]  

Thus we have the following minimax result:
Theorem 5

Let \( P_\beta (T|\delta \beta) \) the power of a test \( T \) for testing \( H_0 \) against \( H_1 \) with any possible change \( \delta \beta \). Then:

\[
\lim_{s \to \infty} P_\beta (X^*_s | \delta \beta) = \lim_{s \to \infty} P_\beta (X^*_s | \delta \beta^*) \geq \lim_{s \to \infty} P_\beta (T_s | \delta \beta^*)
\]

for any other test statistics \( T_s \).

Theorem 5 means that the test \( X^*_s \) (51) is min-max optimal and robust with respect to uncertainties on \( \beta \) (MA part of \( Y_t \)).

2. Min-max optimality of the instrumental test in the scalar case

We now show that, in the scalar case, the test \( X^*_s \) (51) is equivalent to the instrumental test (21) corresponding to the following choice of instruments in (7):

\[
Z_t^* = G_* Z_t^\infty
\]

where \( G_* \) has \( p \) rows, and does not correspond to the use of filtered instruments.

For this purpose, we compute \( \Delta^s_\psi \) defined in (52).

It can be shown \( |14| \) \( |16| \) that, in the scalar case, the gradient of the log-likelihood is:

\[
\Delta^s_\psi (\psi^0) = \frac{1}{\sigma^2} \sum_{t=1}^{s} \frac{1}{B(q^{-1})} \left( \begin{array}{c} \phi_t \\ \epsilon_t \end{array} \right) \cdot \epsilon_t
\]
where $\varepsilon_t \overset{\Delta}{=} \begin{pmatrix} E_{t-p+1} \\ \vdots \\ E_{t-1} \end{pmatrix}$ and $\phi_t$ is defined in (4). (see Appendix 1).

On the other hand, using the definition of $\mathcal{F}$, we have:

$$F_{22} = \text{cov}_0 \left( \frac{1}{B(q^{-1})} \varepsilon_t \right)$$

and

$$F_{12} = \text{cross.cov}_0 \left( \frac{1}{B(q^{-1})} \phi_t, \frac{1}{B(q^{-1})} \varepsilon_t \right).$$

Consequently, we have the following basic geometrical interpretation:

$$F_{12} F_{22}^{-1} \frac{1}{B(q^{-1})} \varepsilon_t = \mathbb{E}_0 (S_t / W) \quad (55)$$

where $S_t = \frac{1}{B(q^{-1})} \phi_t$ and $W = \text{Span} \left( \frac{1}{B(q^{-1})} E_{t-1}, \ldots, \frac{1}{B(q^{-1})} E_{t-p+1} \right)$.

Therefore, because of (52), we get:

$$\Delta^s = \frac{1}{\sigma^2} \sum_{t=1}^{S} S_t E_t \quad (56)$$

where $S_t = S_t - \mathbb{E}_0 (S_t / W)$.

We now consider the space $[E]_{t-1}^{-1} = \text{Span} (E_{t-1}, E_{t-2})$, which contains $W$ because $B(q^{-1})$ is stable, and its subspace $V$ such that:

$$[E]_{t-1}^{-1} = V \oplus W$$
It can be easily seen that each component \( \mathcal{S}_t \) of \( \mathcal{S}^t \) in (56) belongs to \( V \). We now look for a convenient basis of \( V \) for computing (56). It turns out that the following lemma holds [16]:

**Lemma**

Let \( F(q^{-1}) \mathcal{E}_{t-1} \) be a generic element of \( \mathcal{E}^{t-1}_{-\infty} \), where \( F \) is a stable filter. Then:

\[
F(q^{-1}) \mathcal{E}_{t-1} \in V \iff F(q^{-1}) = q^{r+1} B(q) G(q^{-1})
\]

where \( G(q^{-1}) \) is also a stable filter.

Consequently, we can write:

\[
V = \operatorname{Span} (B(q) \mathcal{E}_{t-p}, B(q) \mathcal{E}_{t-p+1}, \ldots)
\]

or equivalently:

\[
V = \operatorname{Span} (B(q) \mathcal{Y}_{t-p}, B(q) \mathcal{Y}_{t-p-1}, \ldots)
\]

and, for each component \( \mathcal{S}_t \), we have:

\[
\frac{1}{\sigma^2} \mathcal{S}_t = B(q) G_i(q^{-1}) \mathcal{Y}_{t-p}
\]

(57)

where

\[
G_i(q^{-1}) = \sum_{j=0}^{\infty} G_{ij} q^{-j}
\]
We now compute the components of \( \Delta^S \) in (56):

\[
\Delta^S \left[ i \right] = \sum_{t=1}^{s} \sum_{j=0}^{p-1} B_j G_i(q^{-1}) Y_{t-p+j} E_t
\]

\[
= \sum_{j=0}^{p-1} B_j \sum_{t=1}^{s} G_i(q^{-1}) Y_{t-p+j} E_t
\]

\[
= \sum_{j=0}^{p-1} B_j \sum_{t=1}^{s} G_i(q^{-1}) Y_{t-p} E_{t-j} \quad \text{(neglecting the boundaries)}
\]

\[
= \sum_{t=1}^{s} G_i(q^{-1}) Y_{t-p} \sum_{j=0}^{p-1} B_j E_{t-j}
\]

Therefore, we obtain:

\[
\Delta^S = \sum_{t=1}^{s} Z_t^* \cdot B(q^{-1}) E_t \tag{58}
\]

where

\[
Z_t^* \triangleright \begin{pmatrix} G_i(q^{-1}) Y_{t-p} \\ \vdots \\ G_p(q^{-1}) Y_{t-p} \end{pmatrix}
\]

\[
= G_* Z_t^\infty
\]

with \( G_* \) of size \( p \times \infty \).

Thus:

\[
\frac{1}{\sqrt{s}} \Delta^S = G_* U^\infty \tag{59}
\]

and this holds under \( H_0 \) and, because of the local approach, almost surely under \( H_1 \). (see Appendix 2).
Finally, using the same type of computations as in section II.4., it is possible to show that the local likelihood test and the instrumental test with an infinite number of instruments have the same asymptotic power, i.e. that the equality is attained in (41),

Hence, in the scalar case, Theorem 5 can be enforced into the following:

Theorem 6

(i) the local likelihood test is an instrumental test, for a particular choice of instruments which is not a filtering operation (59);

(ii) for any change $\delta \beta$, we have:

$$
\lim_{s \to \infty} P_s(X^*_s \mid \delta \beta) = \lim_{s \to \infty} P_s(X^*_s \mid \delta \beta^*) \geq \lim_{s \to \infty} P_s(T_s \mid \delta \beta^*)
$$

$$
\| \quad \| 
$$

$$
\lim_{N \to \infty} \lim_{s \to \infty} P_s(t_0(s) \mid \delta \beta) = \lim_{N \to \infty} \lim_{s \to \infty} P_s(t_0(s) \mid \delta \beta^*)
$$

for any other test statistics $T_s$.

The last equality holds because the asymptotic power of $t_0(s)$ does not depend upon $\delta \beta$. 
VI. USING FILTERED INSTRUMENTS IN THE SCALAR CASE

In this last section, we consider in the scalar case, the instrumental tests corresponding to the choice of a finite number of filtered instruments:

\[ U_N^G(s) = \frac{1}{\sqrt{s}} \sum_{t=1}^{5} G(q^{-1}) Z_t^N W_t \]  \hspace{1cm} (60)

where \( G(q^{-1}) \) is a stable invertible filter.

According to (27), we consider, with obvious notations, the following criterion:

\[ I_N^G = \mathcal{H}_p^G \cdot (I_N^G)^{-1} \cdot \mathcal{H}_p^G \]

and we first show that, for any stable invertible filter \( G(q^{-1}) \), the equality:

\[ I^G = I \]  \hspace{1cm} (61)

holds. For this purpose, we introduce the band-Toeplitz matrix \( \mathcal{G} \) defined by:

\[ \mathcal{G} = \begin{pmatrix} G_1 & G_2 & \cdots & \cdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & G_1 & G_2 \end{pmatrix} \]

and such that:

\[ U_N^G(s) = \mathcal{G} U_s^\infty . \]
Thus \( \Gamma_N^G = \Gamma(q_0^G) \), and \( \Gamma_N^G \approx \Gamma_\infty^G \).

Using the same argument for \( G(q^{-1})^{-1} \), which is possible because \( G(q^{-1}) \) is stable and invertible, we show similarly that \( \Gamma_\infty^G \approx \Gamma_\infty^G \).

Therefore (61) holds, this means that the tests corresponding to an infinite number of filtered or non-filtered instruments have the same asymptotic power.

We now show that the optimal power \( \Gamma_\infty^G \) may be attained with a finite number of filtered instruments for a particular filter. For that, we investigate the equality condition:

\[
\Gamma_{N+1}^G = \Gamma_N^G \tag{62}
\]

and show that it holds for any \( N > P \) and for the following filter:

\[
G(q^{-1}) = \frac{1}{B^2(q^{-1})} \tag{63}
\]

Using the same notations as in section II.4., (62) is equivalent to:

\[
\text{He}_{p,N} G (\Sigma_N^G)^{-1} S_2^G = H_{N+1}^G
\]
It may be shown \([16]\), that, in the present scalar case, we have:

\[
\mathcal{H}^G_{N+1} = \mathcal{H}^G_{p,N} \mathcal{T}
\]

where

\[
\mathcal{T} = \begin{pmatrix}
0 \\
\vdots \\
0 \\
A_p \\
A_1
\end{pmatrix}
\]

and thus (62) is equivalent to:

\[
\sum^G_N \mathcal{T} = S^G_2
\]  

(64)

On the other hand, it may be easily shown that, for the special filter (63), \(U_N^G(s)\) has the same covariance as:

\[
\frac{1}{\sqrt{s}} \sum_{t=1}^{s} \frac{1}{B(q^{-1})} Z_t^N E_t
\]

and thus:

\[
\sum^G_N = \begin{pmatrix}
\overline{R}_0 & \cdots & \overline{R}_{N-1} \\
\vdots & \ddots & \vdots \\
\overline{R}_{-N+1} & \cdots & \overline{R}_0
\end{pmatrix}
\]

where

\[
\overline{R}_m = \mathcal{E}_0 \left( \frac{1}{B(q^{-1})} Y_{t+m} \frac{1}{B(q^{-1})} Y_t \right)
\]
Then, using the following relationship:

\[ R_m - A_1 R_{m-1} - \ldots - A_p R_{m-p} = \mathcal{E}(E_t \cdot m \quad \frac{1}{B(q^{-1})} Y_t) \]

we prove that (64) holds for the filter (63), which thus satisfies:

\[ \Gamma_p^G = \Gamma_\infty \]

because \( \Gamma_\infty^G \) does not depend upon \( G \) (61).

Therefore, the conclusion of this section is:

**Theorem 7**

The instrumental test corresponding to the choice of \( p \) instruments filtered by \( \frac{1}{B^2(q^{-1})} \), attains the optimal asymptotic power.

It is important to notice that this test is not a local likelihood test, because \( G_0 \) in (59) is not a Toeplitz matrix, and thus does not correspond to the use of filtered instruments (the filters \( G(q^{-1}) \) in (57) do depend upon the index \( i \)).
VII. CONCLUSION

We have investigated the asymptotic power of new instrumental tests which we recently proposed for detecting and diagnosing changes in the AR part of a multivariable ARMA process, in view of solving the problem of vibration monitoring for offshore platforms. The optimization of the number of instruments and the possible filtering operation have been analyzed. The connections with the accuracy of the IV identification method [19] and with the robustness properties [7] of local likelihood tests, have been established.

Finally, the optimum asymptotic power of the tests defined here, can be used as a new criterion for investigating the problem of optimal sensor location for detection. This study will be reported elsewhere.
APPENDIX I

COMPUTATION OF THE FISHER INFORMATION MATRIX

For the parameterization (42) of the process \( (Y_t) \), we compute here the gradient of the log-likelihood and the Fisher information matrix.

From (43), we get:

\[
\Delta_s^\psi (\psi) = - \sum_{t=1}^{s} \left( \frac{\partial E_t}{\partial \psi} \right)^T \Lambda^{-1} E_t (\psi)
\]

where:

\[
E_t (\psi) = B(q^{-1})^{-1} (Y_t - (\phi_t^T \otimes I_r)) \quad (6)
\]

with \( \phi_t \) defined in (4).

Therefore:

\[
\frac{\partial E_t}{\partial \psi} = - B(q^{-1})^{-1} (\phi_t^T \otimes I_r)
\]
On the other hand, we have:

\[
B(q^{-1}) E_t = E_t + (B_{p-1} \ldots B_1) \begin{pmatrix} E_{t-p+1} & \vdots & \vdots & E_{t+1} \\
\end{pmatrix}
\]

\[
= E_t + (\epsilon_t^T \otimes I_r) \beta
\]

where \( \epsilon_t \) is defined in (52).

Thus:

\[
0 = \frac{\partial}{\partial \beta} (B(q^{-1}) E_t)
\]

\[
= B(q^{-1}) \frac{\partial E_t}{\partial \beta} + (\epsilon_t^T \otimes I_r)
\]

and:

\[
\Delta^s \psi (\psi) = \sum_{t=1}^s \begin{pmatrix} B(q^{-1})^{-1} (\phi_t^T \otimes I_r) \\
B(q^{-1})^{-1} (\epsilon_t^T \otimes I_r) \\
\end{pmatrix}^T \Lambda^{-1} E_t
\]

\[
= \sum_{t=1}^s \begin{pmatrix} \phi_t \otimes B(q^{-1})^{-T} \Lambda^{-1} \\
\epsilon_t \otimes B(q^{-1})^{-T} \Lambda^{-1} \\
\end{pmatrix}^T E_t
\]

\[
= \sum_{t=1}^s \begin{pmatrix} (\phi_t) \otimes B(q^{-1})^{-T} \Lambda^{-1} \\
(\epsilon_t) \otimes B(q^{-1})^{-T} \Lambda^{-1} \\
\end{pmatrix}^T E_t
\]
Therefore, the Fisher information matrix is:

\[
\mathcal{F} \triangleq \mathbb{E}_0 \left( \Delta^T_\psi (\psi^0) \Delta^T_\psi (\psi^0)^T \right)
\]

\[
= \text{cov}_0 \left[ \begin{pmatrix} \phi_t \\ \epsilon_t \end{pmatrix} \otimes B(q^{-1})^{-T} \Lambda^{-1/2} \right].
\]
APPENDIX 2

A CONTIGUITY ARGUMENT

We detail here the reasons for which some results true under $H_0$ are also true under $H_1$ defined by:

$$
\mathcal{H}_s = \mathcal{H}^0 + \frac{\delta \mathcal{H}}{\sqrt{s}}
$$

Because of the local feature ($\sqrt{s}$), the process $(Y_t)$ has asymptotically the same second order statistics under $H_0$ and $H_1$, and, in the gaussian case, the classes of laws $P_{\mathcal{H}^0}$ and $\{P_{\mathcal{H}_s}\}$ are contiguous (see [17] - chapter 1). Therefore, for any sequence of events $(A_s)$ such that, for any $s > 0$, $A_s$ is $\sigma(Y_t, t \leq s)$ - measurable, we have the following relation:

$$
P_{\mathcal{H}^0} (A_s) \xrightarrow{s \to \infty} 0 \quad \iff \quad P_{\mathcal{H}_s} (A_s) \xrightarrow{s \to \infty} 0
$$

Especially if $T_s$ is a $\sigma(Y_t, t \leq s)$ - measurable random variable, then:

$$
T_s \xrightarrow{s \to \infty} 0 \quad a.s. \text{ under } P_{\mathcal{H}^0} \quad \iff \quad T_s \xrightarrow{s \to \infty} 0 \quad a.s. \text{ under } P_{\mathcal{H}_s}
$$

This result is used in section V.2.
REFERENCES


14] I.-V. Nikiforov, "Sequential detection of changes in stochastic systems", in [4].


