SOURCE DETECTION AND SEPARATION IN POWER PLANT PROCESS MONITORING: APPLICATION OF THE BOOTSTRAP

Rostom Aouada[1,2], *Saïd Aouada*[3], *Guy d'Urso*[2] *and* *Abdelhak M. Zoubir*[3]

[1] Sciences de l'Automatique & du Traitement du Signal, SUPELEC,
3, rue Joliot-Curie, 91192 Gif-sur-Yvette Cedex, France.
rostom_aouada@ieee.org

[2] Direction Etudes & Recherches, Electricité de France (EDF),
6, quais Watier, 78401 Chatou Cedex, France.
Guy.Durso@der.edf.fr

[3] Signal Processing Group (SPG), Darmstadt University of Technology,
Merckstrasse 25, D-64283, Darmstadt, Germany.
{aouada,zoubir}@ieee.org

ABSTRACT

We consider source enumeration and identification in the context of monitoring the cooling circuit of a pressurized-water-reactor (PWR) nuclear plant. We employ a linear instantaneous-mixture to describe the system. We use the Gerschgorin radii of the transformed covariance matrix of the data to detect the number of sources. In particular, we illustrate the advantage of employing the bootstrap in a scenario where no or little *a priori* knowledge is available on the statistical properties of the measured data. A specific denoising procedure is also applied to the data to alleviate the effect of small variations of the noise power over the sensors, and allow a more accurate source separation. The results show the potential of both Gerschgorin-based detection and the bootstrap in practice.

1. PROBLEM FORMULATION

We plan a non-destructive assessment of individual physical variables and their effect on the behavior of the cooling process in a nuclear power plant. As part of this assessment, a number of measurements is conducted around the cooling circuit, without direct access to the coolant (water). Three major steps compose our analysis: 1) a simple but representative model of the measurements is established, 2) the number of latent signals is identified, and 3) the signals of interest are estimated. In the following, we summarize these steps.

We have p narrowband latent signals measured simultaneously by a set of M sensors. The source signals correspond to the physical quantities to be monitored, which are mainly the *temperature*, *pressure* and *flow* of the water. We therefore predict 2 to 3 independent major sources in the system, given that the water temperature and pressure can be strongly correlated. Yet, for our blind analysis, the number of sources p is assumed unknown and is to be estimated and validated. The number of sensors M is much larger than the number of sources p. The modulation and conditioning of the sources are conducted in a narrowband framework.

Although the most precise modeling of the cooling circuit (as most processes in a nuclear plant) introduces convolutive mixtures, we approximate the behavior of our system by an instantaneous-mixture linear model, as a trade-off between precision and computational complexity. The size of the measured data sample is therefore defined to partially corroborate this assumption, i.e., a large M and a large number of collected snapshots, L. Hence, the received signal vector at instant t is modeled as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \qquad t = 1, \ldots, L \tag{1}$$

where $\mathbf{A}$ is the $(M \times p)$-dimensional mixing matrix, $\mathbf{s}(t)$ is the p-dimensional vector of the sources, and $\mathbf{n}(t)$ is the M-dimensional vector of sensor noise. In the sequel, $\mathbf{A}$ and $\mathbf{s}(t)$ satisfy the usual rank and parameter identifiability conditions.

The sources and the noise are assumed to be uncorrelated and their respective distributions are unknown. The data covariance matrix is given by

$$\mathbf{R} = \mathsf{E}\left\{\mathbf{x}(t)\mathbf{x}^H(t)\right\} = \mathbf{A}\mathbf{R}_{\mathbf{s}}\mathbf{A}^H + \sigma^2\mathbf{I} \tag{2}$$

where $(.)^H$ denotes Hermitian transpose, $\mathsf{E}(.)$ is expectation, $\mathbf{R}_{\mathbf{s}} = \mathsf{E}\left\{\mathbf{s}(t)\mathbf{s}^H(t)\right\}$ is the covariance matrix of the sources, and σ^2 is the *ideal* power of the noise.

Depending on the combination of measured physical quantities, the technology of the particular sensors is variable. In addition, they are not necessarily located in a homogeneous pattern around the process, so that the most representative information is collected. Moreover, given the relatively large values of M and L, it appears that a uniform calibration of all the sensors over the total duration of the measurement is very difficult. Although the measurements are pre-processed to level the power of the noise, some residual perturbation may appear from one sensor to another.

When standard source detection methods are employed in such a scenario, their performance can degrade significantly due to the variation of the eigenvalues of $\mathbf{R}$. The quality of estimation of the signal subspace and its dimension based solely on these eigenvalues shows a threshold effect, with respect to the variation of the Signal-to-Noise Ratio (SNR), and the Worst-Noise-Power Ratio (WNPR). Figure 1 shows the Mean-Square Error (MSE) of the cosine of the angle between the estimated and ideal signal subspace as defined in [1], in terms of SNR and WNPR. It is clear that beyond a certain threshold value of SNR or WNPR, the separation of the signal subspace is not possible from the ordered eigenvalues of $\mathbf{R}$. In what follows, instead of relying on the ordered eigenvalues, we will use ordered subspace projectors (Gerschgorin radii) which are more robust to small perturbations of the noise power over the sensors.

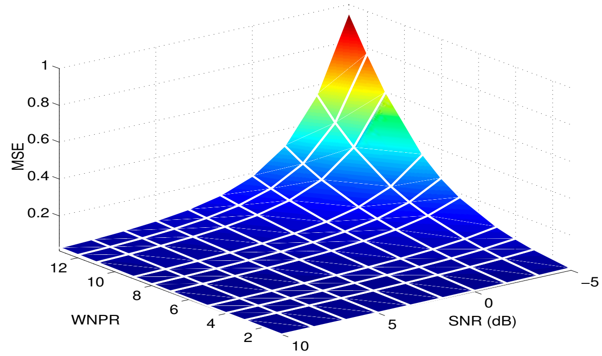


Fig. 1. Simulated MSE of $\cos(\gamma)$, γ being the angle between the estimated and ideal signal subspace, with a Vandermonde-type $\mathbf{A}$, $M = 10$, and $P = 2$.

2. DATA SAMPLE AND PRE-PROCESSING

The data sample of interest is collected and formatted by Electricité de France (EDF), from its Nogent-sur-Seine site. The collected data is of size $M = 142$ by $L = 3 \times 10^5$. The data are sampled above the Nyquist rate, at a period of 1 min. The sensor outputs are correlated and have the same general shape over time. As an example, the output of sensor #113 is illustrated in Fig. 2.

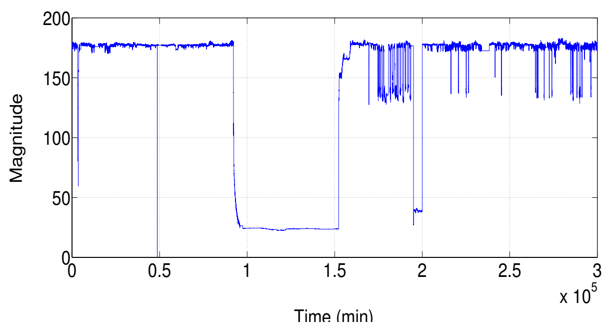


Fig. 2. Output of sensor #113.

Two different operation modes of the power plant can be easily identified. First, between $t = 0$ and $t = 10^5$, a *steady-state* mode, where the measurements are mostly stationary around their nominal value, and second, between $t = 1.5 \times 10^5$ and $t = 3 \times 10^5$, a *network-steered* mode, where the measurements are non-stationary and strongly vary with the load on the network. In this analysis, we are interested only in the steady-state operation, and thus limit the sample size to $L = 10^5$.

Note that other segments are visible. These are also discarded from our analysis. They correspond to irrelevant situations, such as an interrupted measurement due to the opening of the reactor's heart, or constant measured values due to a likely dysfunction of the sensors.

The collected data of interest are further cleaned from short constant segments and outlier points, and then tested for stationarity. Depending on the application, the data can also be whitened and centered.

3. SOURCE NUMBER DETECTION

If the contribution of the M-th (last) array element is suppressed, a unitary transformed covariance matrix $\boldsymbol{\mathcal{R}}$ is obtained from the original covariance matrix $\mathbf{R}$ [2].

The first $M-1$ Gerschgorin radii of the transformed matrix $\boldsymbol{\mathcal{R}}$, denoted ρ_m, $m = 1, \ldots, M-1$, in ascending order of magnitude, are the basis for the separation between the signal and noise subspaces, instead of the ordered eigenvalues of $\mathbf{R}$, since ideally these radii satisfy the following [2]:

$$\rho_1 \geq \rho_2 \geq \ldots \geq \rho_p \geq \rho_{p+1} = \rho_{p+2} = \ldots = \rho_{M-1} = 0 \quad (3)$$

where the first p elements, $\rho_1, \ldots, \rho_p$, correspond to the signal subspace. In practice, due to the finite data length, the sample covariance matrix $\hat{\mathbf{R}}$ is used, leading to radii $\hat{\rho}_m$, $m = 1, \ldots, M-1$. From (3), it can be deduced that estimation of the number of sources p can be achieved by checking simultaneously for zero the Gerschgorin radii $\hat{\rho}_m$, $m = p+1, \ldots, M-1$, corresponding to the noise subspace.

One way to automatically make a decision on the correct number of sources is to apply an appropriate information criterion. Due to its high robustness to small noise-power perturbations over the sensors, the variant of MDL of [3], labeled NU-MDL, can be employed in our scenario. The limitation of NU-MDL however, is that it is specifically derived for the stochastic Gaussian case. If Gaussianity of the data is not verified, combined with small samples and a low SNR, the performance of the criterion is expected to degrade significantly.

As an alternative to NU-MDL, a sequential hypothesis test (SHT) is formulated in [5] to statistically detect the smallest radii corresponding to the candidate noise subspace. To this end, the following two test statistics can be employed with different performance levels

$$T_{1_q} = \sum_{m=q+1}^{M-1} \hat{\rho}_m^2 \quad (4)$$

$$T_{2_q} = \left(\frac{1}{M-1-q} \sum_{m=q+1}^{M-1} \hat{\rho}_m \right) - \left(\prod_{m=q+1}^{M-1} \hat{\rho}_m^{\frac{1}{M-1-q}} \right) \quad (5)$$

$$\text{for} \quad q = 0, \ldots, M-2.$$

Note that no reference is made to the ordered eigenvalues of the covariance matrix $\mathbf{R}$.

For an arbitrary or unknown distribution of the measured data, or for non-asymptotic scenarios, the bootstrap is employed to estimate the null distribution of the test statistics (4) and (5)[6, 5]. The principle of the bootstrap is that the data sample represents an empirical estimate of the true distribution. Thus, resampling from this estimate creates bootstrap data sets which are used to conduct inference. Because of their robustness, it is convenient to apply the above bootstrap-based detectors to our data set, as no *a priori* knowledge on its statistical characteristics is available.

Given the b-th bootstrap resample of the data, denoted $\mathbf{x}_b^\star(t)$, for $t = 1, \ldots, L$ and $b = 1, \ldots, B$, the estimate of the empirical distribution of the test statistics under the null can be obtained by $\hat{T}_q^{\mathsf{H}}(b) = T_q^\star(b) - T_q$, where T_q is the test statistics evaluated from the data $\mathbf{x}(t)$, while $T_q^\star(b)$ is the test statistics evaluated from the resample $\mathbf{x}_b^\star(t)$ [6, 7]. For a given level of significance α, when no sources are present (global null), the probability of correctly deciding that $\hat{p} = 0$ must be maintained at $1 - \alpha$.

It is worth mentioning that numerical results [3, 5] showed a satisfactory performance of Gerschgorin-based detectors with WNPR values beyond 10, whereas conventional eigenvalue-based detectors fail for WNPR even less than 2.

4. BLIND IDENTIFICATION OF THE SOURCES

An efficient control of the cooling circuit requires an individual monitoring of the measured physical variables. In order to be able to act on each component separately, the latter must be observed with sufficient accuracy, considering that the cooling water is not directly accessible. Our analysis therefore involves the application of blind source separation algorithms to the data. We employed three different methods. These are: JADE [8], SOBI [9], and FastICA [10]. JADE is based on the whitening and joint diagonalization of fourth-order cumulant matrices (non-Gaussian data). Based on the same joint diagonalization principle, SOBI requires that the sources have different non-zero spectral components and uses shifted correlations of the data. FastICA maximizes the non-Gaussianity of the marginal densities of the whitened data, using a fixed-point iterative procedure instead of more complex gradient descent methods.

All the aforementioned methods incorporate data whitening as a first step. This operation supposes the knowledge of the ideal noise power σ^2 in (1). If the SNR is high enough to considerably reduce the effect of noise, the above source separation methods will not be strongly affected by the noise nonuniformity. If the power perturbations over the sensors are very small, an average σ^2 can be easily estimated through the eigen-decomposition (or singular decomposition) of $\mathbf{R}$. Otherwise, the lack of whitening will require the matrices of the contrast function to be symmetrized. For JADE and SOBI for example, this will invoke the use of some non-orthogonal joint diagonalization technique at a higher computational cost.

From the development of [2, 5], we can use the same subspace separation criterion, i.e., the ordered Gerschgorin radii of (3), to *denoise* the system by estimating the perturbed noise powers. Similarly to the derivation of NU-MDL and the bootstrap-based SHT, this estimation is conducted at a cost of discarding one sensor.

Given the previously estimated number of sources $\hat{p}$, denote by $\mathbf{E}_n$ the $((M-1)\times(M-1-\hat{p}))$-dimensional basis made of the eigenvectors spanning the noise subspace. These eigenvectors are obtained from the reduced-data sample covariance matrix, $\hat{\mathbf{R}}_M$, and correspond to the radii ρ_m, $m = p+1,\ldots,M-1$. Also denote the m-th column of $\hat{\mathbf{R}}_M$ by $\hat{\mathbf{r}}_m$, for $m = 1,\ldots,M-1$. Next, note that the vector $\hat{\mathbf{r}}_m - \sigma_m^2\mathbf{i}_m$ is completely noise-free and characterizes only the sources, where $\mathbf{i}_m$ is given as

$$\mathbf{i}_m(k) = \begin{cases} 1 & \text{for } k = m; \\ 0 & \text{else} \end{cases} \tag{6}$$

and σ_m^2 denotes the perturbed noise power as measured on the m-th sensor.

Exploiting the principle of orthogonality between the signal and noise subspaces, we can easily verify that ideally, we have

$$\hat{\sigma}_m^2 = \arg\min_{\sigma_m^2}\left\{\left(\hat{\mathbf{r}}_m - \sigma_m^2\mathbf{i}_m\right)^H \mathbf{E}_n\mathbf{E}_n^H\left(\hat{\mathbf{r}}_m - \sigma_m^2\mathbf{i}_m\right)\right\} \tag{7}$$

and the solution for $\hat{\sigma}_m^2$ follows straightforward, for $m = 1,\ldots,M-1$.

With the estimated $M-1$ noise powers, arranged in a vector $\hat{\mathbf{q}}_M$, a noise-free reduced-data covariance matrix can be deduced as

$$\tilde{\mathbf{R}}_M = \hat{\mathbf{R}}_M - \hat{\mathbf{Q}}_M \tag{8}$$

where $\hat{\mathbf{Q}}_M = \text{diag}\{\hat{\mathbf{q}}_M\}$.

The noise-free covariance matrix $\tilde{\mathbf{R}}_M$ can then be used directly with the conventional source separation algorithms.

Similarly to [3], the total number of sensors can be used to yield M different estimates of each noise power, and averaging over the M estimates resulting in a higher accuracy.

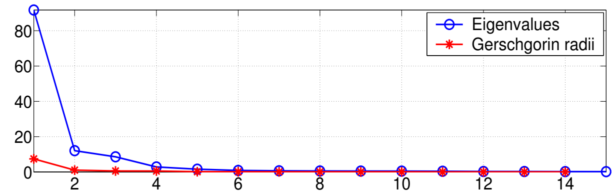


Fig. 3. Sample eigenvalues and Gerschgorin radii.

	L ($\times 1000$)								
	6	9	12	15	18	21	24	27	31
MDL	15	15	15	15	15	15	15	15	15
NU-MDL	4	4	4	4	7	7	7	7	7
T_1	2	2	2	2	3	3	3	3	3
T_2	2	2	2	3	3	3	3	3	3

Table 1. Estimated number of sources

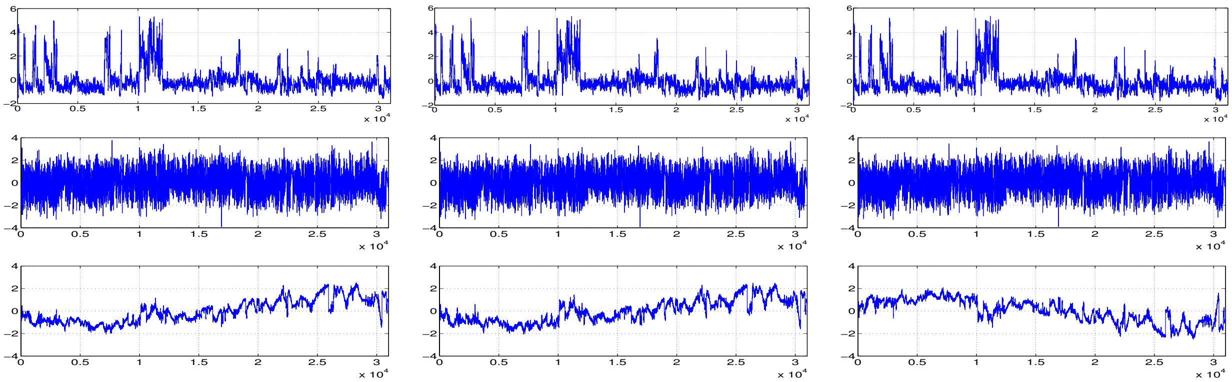
5. EXPERIMENTAL RESULTS

As a first step, the number of sources is estimated using MDL, NU-MDL and the bootstrap SHT, both with T_1 and T_2 of (4) and (5), respectively. In our example, the number of sensors used for this estimation is limited to 15 (#103 through #118) to reduce the computational load (recall that the complexity associated with the eigen-decomposition of $\hat{\mathbf{R}}$ is of the order $O(M^3)$). The test significance level is $\alpha = 5\%$, whereas the number of bootstrap resamples is set to $B = 3000$.

Fig. 3 shows the ordered eigenvalues and Gerschgorin radii. Note that we have a strongly dominating eigenvalue, which translates to a large Gerschgorin radius, although the transformation reduces the magnitude dispersion. Hence, in our relatively high SNR case, we expect the detectors to yield a small number of sources. It should be kept in mind however, that the dominating source can possibly mask the smaller sources. At the same time, because of the poor clustering of the smallest eigenvalues, the information criteria can typically over-model the system [11].

Application of the detectors (25 independent trials) results in the estimates shown in Table 1. The accuracy of the detection improves with an increasing sample size L. The Gerschgorin-based criteria outperform the conventional MDL. Despite the high SNR, the main reason for the failure of MDL is the non-Gaussianity of the embedded sources, which are thus strongly mismodeled by the informative part of MDL. The non-Gaussianity of the sources affects the NU-MDL in the same negative way. The bootstrap detectors on the other hand are more suitable for unknown distributions of the data. In addition, determination of the signal and noise subspace dimensions is carried-out through the ordered Gerschgorin radii rather than the ordered eigenvalues, thus reducing the possibility of masking smaller sources. The result is more sensible, irrespective of the sample size L. Despite the increased computational complexity, this result validates the relevance of the bootstrap in practice, especially for off-line analysis.

Upon determination of the order of the model, blind source separation is applied to the total data sample. All three methods, i.e., JADE, SOBI and FastICA, provide almost identical results (up to a rotation transformation). This result supports the linear model of (1). The separation methods are applied to the data both directly and after estimation of the perturbed noise powers, following (7). Given the relatively high SNR, the results in both cases are very close, therefore only the results after denoising are presented for conciseness. It is worth mentioning however, that if the SNR were low, prior estimation of the noise powers would significantly improve the results. Although the quality of estimation of the signal and noise subspace is dependent on the quality of estimation of the



(a) (b) (c)

Fig. 4. Separated sources using: (a) SOBI, (b) JADE, and (c) FastICA.

Gerschgorin radii ρ_m, $m = 1, \ldots, M-1$, the denoising step of (8) can only improve the overall result, by significantly reducing the loading due to noise on the diagonal of $\mathbf{R}$.

The spectral content of the signals is variable with time and allows the use of shifted correlation matrices for SOBI. This illustrates that for the size of the data, the information provided by its second-order moment is sufficient for the synthesis of the embedded sources. The non-Guassianity of the data is also well exploited as illustrated by the performance of JADE and FastICA.

The obtained sources are mostly non-Gaussian, in conformity with the initial assumptions. The first source shows small fluctuations and some impulsive bursts. It is somehow heavytailed. In the steady-state operation mode, it is likely to correspond to the variation of the temperature/pressure of the cooling water. The second source is nearly Gaussian. It is likely to correspond to an internally-generated perturbation, which is rather independent of the variation of the physical quantities of interest. The third source is multi-modal and shows progressive variations which likely correspond to the flow of the cooling water.

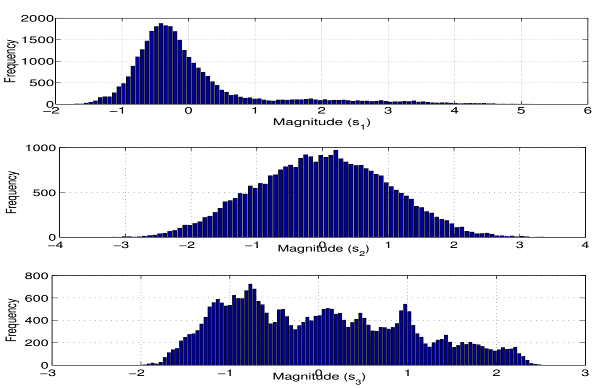


Fig. 5. Histogram of the magnitudes of the estimated sources.

6. CONCLUSION

A linear instantaneous-mixture has been successfully employed to model the cooling process in one of EDF's power plants. The bootstrap, through Gerschgorin-based detection criteria, is applied to the measured data and yields conclusive results. Denoising is applied to alleviate the effect of small variations of the noise power from one sensor to another and allow an accurate source identification. The results show the power of the bootstrap-based detector and its advantage over other methods in a real case where the distribution of the data samples in unknown.

7. REFERENCES

[1] H. Wang and M. Kaveh, "On the Performance of Signal Subspace Processing - Part1: Narrow-Band Systems", *IEEE Tran. Acoustics, Speech, Signal Processing*, Vol. ASSP-34, No. 5, pp. 1201-1209, Oct. 1986.

[2] H.T. Wu, J.F. Yang and F.K. Chen, "Source Number Estimation Using Transformed Gerschgorin Radii", *IEEE Tran. Signal Processing*, Vol. 43, No. 6, pp. 1325-1333, Jun. 1995.

[3] S. Aouada, A.M. Zoubir and C.M.S. See, "Source Detection in the Presence of Nonuniform Noise", *Proc. 29th IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP 2004*, Vol. 2, pp. 165-168, Montreal, Canada, May 2004.

[4] M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria", *IEEE Tran. Acoustics, Speech, Signal Processing*, Vol. ASSP-33, No. 2, pp. 387-392, Feb. 1985.

[5] S. Aouada, D. Traskov, N. d'Heureuse and A.M. Zoubir, "Application of the Bootstrap to Source Detection in Nonuniform Noise", *Proc. 30th IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP 2005*, Vol. 4, pp. 997-1000, Philadelphia, PA, Mar. 2005.

[6] A.M. Zoubir and D.R. Iskander, *Bootstrap Techniques for Signal Processing*, Cambridge, U.K.: Cambridge University Press, 2004.

[7] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. London, U.K.: Chapman & Hall, 1993.

[8] J.F. Cardoso and A. Souloumiac, "Blind Beamforming for Non-Gaussian Signals", *IEE Proceedings-F*, Vol. 140, No. 6, pp. 362-370, Dec. 1993.

[9] A. Belouchrani, K. Abed-Meraim, J.F. Cardoso and E. Moulines, "A Blind Source Separation Technique Using Second-Order Statistics", *IEEE Tran. Signal Processing*, Vol. 45, No. 2, pp. 434-444, Feb. 1997.

[10] A. Hyvärinen and E. Oja, "A Fast Fixed-Point Algorithm for Independent Component Analysis", *Neural Computing*, Vol. 9, No. 7, pp. 1483-1492, 1997.

[11] A.P. Liavas and P.A. Regalia, "On the Behavior of Information Theoretic Criteria for Model Order Selection", *IEEE Tran. Signal Processing*, Vol. 49, No. 8, pp. 1689-1695, Aug. 2001.