DETECTION OF SPECTRALLY EQUIVALENT PARAMETRIC PROCESSES USING HIGHER ORDER STATISTICS

Martial Coulon and Jean-Yves Tourneret

ENSEEIHT / GAPSE
2 Rue Camichel, BP 7122
31071 Toulouse Cedex, France
coulon@len7.enseeiht.fr tournere@len7.enseeiht.fr

Ananthram Swami

Army Research Lab
2800 Powder Mill Road
Adelphi, USA
Ananthram_Swami@emh3.arl.mil

ABSTRACT

The paper addresses the problem of detecting two spectrally equivalent parametric processes (SEPP): the noisy AR process and the ARMA process. Higher-order statistics (HOS) are shown to be effective for detection. Two HOS based detectors are derived and compared. The first detector studies the singularity of an HOS-based Yule-Walker matrix. The second detector filters the data by an AR filter estimated from the data; the residual HOS are then shown to be effective for the SEPP detection problem.

1. INTRODUCTION

The detection and classification of signals contaminated by noise has been intensively studied in the literature [11]. Optimal detectors, based on the Neyman-Pearson criterion, can be derived, when statistical properties regarding signal and/or noise are available. Unfortunately, these detectors can be difficult to implement, because of intractable computations. Moreover, the signal and/or noise statistics can be unknown. In such cases, one must resort to suboptimal detectors. These detectors are based on discriminating features, relative to a parametric or non parametric analysis.

This paper studies the suboptimal detection of spectrally equivalent processes (SEP's). SEP's have been observed in many signal processing applications including communication systems. For instance, it is well known that many modulation signals (PSK, QAM) have the same mean and the same power spectral density (or equivalently second-order statistics). Consequently, suboptimal detectors based on first or second-order statistics yield poor results. Several alternatives based on higher-order moments or cumulants have then been studied for the classification of these modulation signals [1][10]. This paper proposes to model the SEP's by two parametric spectrally equivalent models (SEM's): the 'noisy AR model' (see eq. (1)), and its SE ARMA model (see eq. (3)). It is well-known that these two models have the same AR parameters. Consequently, the AR parameters are not suitable for SEP detection. The main contribution of the paper is to show that the SEM MA parameters are effective for the SEP detection problem. The paper is organized as follows. Section 2 formulates the problem. Section 3 presents a detector based on the Singularity of a HOS-based Yule-Walker Matrix (SM). Section 4 studies a new detector based on HOS of filtered data, where the filter parameters are estimated from the data. Section 5 studies the test function related to both detectors. Simulation results and conclusion are reported in sections 6 and 7 respectively.

2. PROBLEM FORMULATION

Consider a 'noisy AR time series' $y_0(n)$ defined by:

$$y_0(n) = x(n) + b(n) \tag{1}$$

where $x(n)$ is an AR process driven by an iid non-Gaussian sequence $e(n)$:

$$x(n) = -\sum_{j=1}^{p} a_j x(n-j) + e(n) \tag{2}$$

and $b(n)$ is the additive Gaussian noise independent of $e(n)$. It is well-known that $y_0(n)$ has the same mean and the same power spectral density as an ARMA(p,p) process (with the same AR parameters) driven by an iid sequence $g(n)$ defined by:

$$y_1(n) = -\sum_{j=1}^{p} a_j y_1(n-j) + \sum_{j=0}^{p} b_j g(n-j) \tag{3}$$

The spectral equivalence property for $y_0(n)$ and $y_1(n)$ yields:

$$\gamma_{2g} = \gamma_{2b} \frac{a_p}{b_p} \tag{4}$$

Moreover, parameters $\{b_j\}_{j=1,\dots,p}$ can be computed from γ_{2e}, γ_{2b} and $\{a_j\}_{j=1,\dots,p}$ via spectral factorization. The SEP detection problem is the following binary hypothesis test:

$$H_0 : y(n) = y_0(n)$$
$$H_1 : y(n) = y_1(n)$$

In principle, if the pdfs of the various innovation processes are known, and if appropriate priors are available, the Bayesian/Maximum Likelihood approach could be used. However, this approach is intractable in the non-Gaussian context, and, of course, the pdfs have to be known. Hence, we propose two suboptimal detectors based on higher-order cumulants.

3. SM-BASED DETECTOR

This section recalls a SEP detector proposed in [3], based on the Higher-Order Yule-Walker Equations (HOYWE) for the SEP detection. The HOYWE for an $AR(p)$ process are defined by [8]:

$$\sum_{j=0}^{p} a_j C_k^x(m-j,0,\ldots,0)=0\ ,\ \forall m>0 \tag{5}$$

where $C_k^x(\rho_1,\rho_2,\ldots,\rho_{k-1})$ denotes the kth-order cumulant of the AR process at lag $\underline{\rho}=(\rho_1,\rho_2,\ldots,\rho_{k-1})$. The property $C_k^{y_0}(\underline{\rho})=C_k^x(\underline{\rho})$ holds $\forall k>2, \forall \underline{\rho}\in \mathbb{Z}^{k-1}$, since the additive noise is Gaussian and independent of $x(n)$. Denote $\mathbf{\Delta}_p(\underline{\xi})=\det(\mathbf{R}_p(\underline{\xi}))$ where $\mathbf{R}_p(\underline{\xi})$ is the Toeplitz matrix whose first row and first column are $(\xi_{p+1},\ldots,\xi_1)$ and $(\xi_{p+1},\ldots,\xi_{2p+1})^T$ respectively, with $\underline{\xi}=(\xi_1,\ldots,\xi_{2p+1})\in\mathbb{R}^{2p+1}$. Denote

$$\mathbf{C}_k^0=(C_k^{y_0}(1-p,0,\ldots,0),\ldots,C_k^{y_0}(1+p,0,\ldots,0))^T \tag{6}$$

The concatenation of eq.'s (5) for $m\in\{1,\ldots,p+1\}$ yields

$$\mathbf{R}_p(\mathbf{C}_k^0).(1,a_1,\ldots,a_p)^T=0 \tag{7}$$

hence

$$\Delta_0\triangleq\mathbf{\Delta}_p(\mathbf{C}_k^0)=0 \tag{8}$$

On the other hand, for an $ARMA(p,p)$ process, eq.'s (5) hold for $m>p$, but not for $m\in\{1,\ldots,p\}$. A large number of simulations have shown that $\Delta_1\triangleq\mathbf{\Delta}_p(\mathbf{C}_k^1)\neq 0$, where $\mathbf{C}_k^1$ is defined as in (6) with cumulants of the $ARMA$ process. Therefore, we assume that it is true in the rest of the paper. The SE noisy AR and $ARMA$ process detection can then be expressed as a simple binary hypothesis testing problem:

$$\begin{array}{lll} H_0: & \text{(Noisy AR process)} & \Delta=\Delta_0=0\\ H_1: & \text{(ARMA process)} & \Delta=\Delta_1\neq 0\end{array} \tag{9}$$

Define $\widehat{\mathbf{C}}_k$ as the sample cumulant vector obtained by replacing the true cumulants in (6) by their usual estimates, and denote $\widehat{\Delta}\triangleq\mathbf{\Delta}_p(\widehat{\mathbf{C}}_k)$. The noisy AR and $ARMA$ process cumulant vector estimates are asymptotically unbiased Gaussian vectors with: $\lim\limits_{N\to+\infty} NE[(\widehat{\mathbf{C}}_k-\mathbf{C}_k^i)(\widehat{\mathbf{C}}_k-\mathbf{C}_k^i)^T\mid H_i]=\mathbf{\Sigma}_k^i$ [8]. According to ([2], p. 211), the determinant estimate $\widehat{\Delta}$ is asymptotically an unbiased Gaussian variable with:

$$\lim_{N\to+\infty} NE[(\widehat{\mathbf{\Delta}}-\mathbf{\Delta}_i)^2/H_i]=D_k^{iT}\mathbf{\Sigma}_k^i D_k^i\triangleq\sigma_i^2 \tag{10}$$

In (10), D_k^i is a vector whose j^{th} element is $D_k^i(j)=(\partial\mathbf{\Delta}_p(\underline{\xi})/\partial\xi_j)(\mathbf{C}_k^i)$. It can be proved that $D_k^i(j)=\sum_{m=1}^{|\xi_j|}(Cof(\mathbf{R}_p(\mathbf{C}_k^i))_m$, where $|\xi_j|=p+1-|p+1-j|$ denotes the number of ξ_j in $\mathbf{R}_p(\underline{\xi})$ and $(Cof(\mathbf{R}_p(.)))_m$ the $\mathbf{R}_p(.)$ matrix cofactor computed at the m^{th} occurrence of ξ_j. The statistical properties of $\widehat{\Delta}$ can then be asymptotically derived under both hypotheses:

$$\begin{array}{lll} H_0: & \text{Noisy AR process} & \sqrt{N}\widehat{\Delta}\sim N(0,\sigma_0^2)\\ H_1: & \text{ARMA process} & \sqrt{N}\left(\widehat{\Delta}-\Delta_1\right)\sim N(0,\sigma_1^2)\end{array} \tag{11}$$

4. MA DETECTOR

Denote by $z_i(n)$ the output of the FIR filter with Z-transform $A(z)=\sum_{k=0}^{p}a_k z^{-k}$ driven by $y_i(n)$. The SEP detection problem can be rewritten as:

$$\begin{array}{c} H_0: z(n)=z_0(n)=e(n)+\sum_{l=0}^{p}a_l b(n-l)\\ H_1: z(n)=z_1(n)=\sum_{l=0}^{p}a_l y(n-l)=\sum_{l=0}^{p}b_l g(n-l)\end{array} \tag{12}$$

Eq. (12) shows that 1) $z_0(n)$ is the sum of a Gaussian MA(p) sequence and an iid non-Gaussian sequence $e(n)$. Consequently, the kth-order cumulants ($k>2$) of $z_0(n)$ are zero except at lag $\rho=0$; 2) $z_1(n)$ is a pure non-Gaussian MA(p) sequence, whose kth-order cumulants are non-zero for a specific set of lags. Define $C_k^{z_i}(\rho)$ as in (5). Giannakis [5] established that the MA parameters of a non Gaussian ARMA(α,β) model can be identified uniquely from cumulants which belong to the following set:

$$I_k=\{\rho\,|\max(\alpha,\beta)-\alpha\leq\rho_2\leq\rho_1\leq\beta+2\alpha, \rho_l=0, l=3,..,k-1\}$$

Using basic properties of kth-order cumulants, the following binary hypothesis testing problem can be considered:

$$\begin{array}{c} H_0: g=g_0=0\\ H_1: g=g_1\neq 0\end{array} \tag{13}$$

where g_i is the theoretical HOS vector whose elements are $C_k^{z_i}(\rho)$, $\rho\in I_k, \rho\neq 0$. Denote by $\widehat{g}$ the vector obtained by replacing the true cumulants in g by their usual estimates computed from N samples. The asymptotic statistical behavior of the HOS vector estimate $\widehat{g}$ is [4]:

$$\begin{array}{c} \sqrt{N}\widehat{g}\sim N(0,\Sigma_0)\\ \sqrt{N}(\widehat{g}-g_1)\sim N(0,\Sigma_1)\end{array} \tag{14}$$

where Σ_0 and Σ_1 are two matrices independent of N. The asymptotic statistics of $\widehat{g}$ can be used to derive kth-order cumulant based likelihood ratio detectors. In practical applications, AR parameter vector $a=[1, a_1,\ldots a_p]^T$ is unknown. Consequently, it has to be estimated. Any higher-order cumulant based method can be used since it is blind to the additive noise. In this case, g is formed by cumulants of the output of the FIR filter with Z-transform $\widehat{A}(z)=\sum_{k=0}^{p}\widehat{a}_k z^{-k}$ driven by $y_i(n)$.

5. STUDY OF THE TEST FUNCTION FOR BOTH DETECTORS

Eq 's (11) and (14) show that the SEP detection problem can be viewed as two HOS based detection problems, involving Gaussian test statistics. The main difference between the two detectors is that the SM detector is one-dimensional, contrary to the MA detector.

5.1. Known Parameters

Parameters $\{a_j\}_{j=1,...,p}, \{b_j\}_{j=1,...,p}, \gamma_{2e}, \gamma_{2b}$ and γ_{2g} can be determined, when the spectrum of the two SEP's is known (parameters $\{a_j\}_{j=1,...,p}, \gamma_{2e}$ and γ_{2b} are estimated by fitting a noisy AR model to the data and parameters $\{b_j\}_{j=1,...,p}, \gamma_{2g}$ are computed using eq. (4) and spectral factorization). The likelihood ratio detector reduces to compare a quadratic form of normal variables to a suitable threshold, depending on the probability of false alarm (PFA) and the model parameters. The test statistic distribution can be expressed as mixtures of central or non-central χ^2 distributions [7]. It can also be expanded in MacLaurin or Legendre series ([7], pp. 168-173) or computed numerically using appropriate algorithms such as the Imhof algorithm. This paper focuses on practical applications, for which the noisy AR and ARMA parameters are unknown.

5.2. Unknown Parameters

The SEP detection problem defined in (9) and (13) is a composite hypothesis test, when the model parameters are unknown. This part studies a composite hypothesis test very similar to the Hinich linearity test ([9], pp. 46-48). Assume that M independent realizations of $\widehat{\varphi}$ (denoted $\left(\widehat{\varphi}_j\right)_{j=1,...,M}$) are available, where $\widehat{\varphi}$ denotes $\widehat{\Delta}$ or $\widehat{g}$. These M measurements can be obtained from one single signal by segmentation. This segmentation procedure consists of considering a N-sample signal as M segments of K samples (with $N = MK$). Define $\overline{\varphi}$ and $\widehat{S}$ as the sample mean and covariance matrix of the sequence $\left(\widehat{\varphi}_j\right)_{j=1,...,M}$:

$$\begin{aligned} \overline{\varphi} &= \tfrac{1}{M}\textstyle\sum_{j=1}^{M}\widehat{\varphi}_j \\ \widehat{S} &= \textstyle\sum_{j=1}^{M}\left(\widehat{\varphi}_j - \overline{\varphi}\right)\left(\widehat{\varphi}_j - \overline{\varphi}\right)^T \end{aligned} \tag{15}$$

Using the asymptotic normality of vector $(\widehat{\varphi}_1, ..., \widehat{\varphi}_M)^T$, the generalized likelihood ratio detector for the parametric SEP detection problems (11) and (14) is defined by [9]:

$$H_0 \text{ rejected if } T^2 = M\overline{\varphi}^t\widehat{S}^{-1}\overline{\varphi} > \lambda_0 \tag{16}$$

λ_0 is a threshold which can be determined from the distribution of T^2 under the null hypothesis and the PFA. Giri [6] showed that the statistic $\frac{M-q}{q}T^2$ has an F-distribution with $(q, M-q)$ degrees of freedom, under the null hypothesis. Note that for the SM detector, T^2 is the square of a Student statistic with M degrees of freedom. Statistic T^2 has a non-central Hotelling distribution, under hypothesis H_1 [9]. Unfortunately, this distribution is difficult to study. Consequently, the probability of detection (PD) is computed via Monte-Carlo simulations.

6. SIMULATION RESULTS

Many simulations have been performed to validate the previous theoretical results. Here, simulation results are reported for the composite hypothesis detector (i.e. unknown parameters). The SEP detector performance is evaluated in terms of Receiver Operating Characteristics (ROC's) for different number of samples and different signal to noise ratios (SNR's). These curves represent the probability of detection (PD) as a function of the PFA. First, an $AR(1)$ process with parameters $[1; -0.5]$ driven by a zero mean exponentially distributed i.i.d. input (with variance $\gamma_e^2 = 1$) is considered. This case is very simple since closed form expressions for the $ARMA$ parameters (input variance γ_g^2 and parameters b_j) are available as functions of noisy AR process parameters. The spectrally equivalent $ARMA$ process is driven by a zero mean exponentially distributed input. Numerical results have been computed using Monte-Carlo runs. Fig.'s 1 and 2 show the SM and MA detector ROC's for different SNR's (and a fixed number of samples $M = 5$ and $K = 2000$). The test improves when the SNR decreases. Indeed, when SNR is low, the noisy AR process is close to a Gaussian process, contrary to the SE $ARMA$ process. Thus, the two SE processes can be easily distinguished for low SNR's. Fig.'s 3 and 4 present the SM and MA detector ROC's for different numbers of samples N and a fixed $SNR = 8dB$. Obviously, the higher N, the better the detector performances. These simulations have been carried out for a given $AR(1)$ process. However, other model orders and AR parameters have been studied and give similar performance. For instance, fig's 5 and 6 show the detector ROC's for a noisy $AR(4)$ process with parameters $[1, -0.5, -0.25, -0.125, -0.06]$, for different numer of samples.

7. CONCLUSION

This paper studied two suboptimal detectors based on higher-order cumulants for the detection of spectrally equivalent processes (SEP's). The SEP's were modeled by two spectrally equivalent parametric processes: the noisy AR process and the ARMA process. The first detector was based on the singularity of a HOS-based Yule-Walker matrix. The second filtered the data by an AR filter estimated from the data, and studied the nullity of the residual higher-order cumulants. The distribution of the test statistics was studied. Theoretical results were supported by simulation studies.

8. REFERENCES

[1] D. Boiteau and C. Le Martret, "Classification of linear modulations by mean of a fourth-order cumulant," Proc. of Eusipco, Trieste, 1996.

[2] P. J. Brockwell and R.A. Davis, Time Series: Theory and Methods, Springer Verlag, 2nd Edition, 1990.

[3] M. Coulon and J.Y. Tourneret, "Detection and classification of spectrally equivalent processes: a parametric approach," Proc. of HOS, Banff, Canada, pp 410-4, 1997.

[4] J. A. R. Fonollosa,"Sample Cumulants of Stationary Processes: Asymptotic Results", IEEE Trans. on Signal Processing, vol. 43, April 1995.

[5] G. B. Giannakis, "On the Identifiability of Non-Gaussian ARMA Models using Cumulants," IEEE Trans. on Aut. Cont., vol. 35, N°1, pp 18-26, January 1990.

[6] N. Giri, "On the complex analogues T^2 and R^2 tests," Ann. Math. Statist., vol. 36, pp. 664-670, 1965.

[7] N. L. Johnson and S. Kotz, Continuous Univariate Distributions-2, first edition, John Wiley, 1970.

[8] B. Porat, Digital Processing of Random Signals: Theory and Methods, Prentice Hall, NJ, 1994.

[9] M. B. Priestley, Non-linear and Non-Stationary Time Series Analysis, Academic Press, 1988.

[10] S. S. Soliman, “Signal classification using statistical moments,” IEEE Trans. on Comm., vol. 40, May 1992.

[11] H. L. Van Trees, Detection Estimation and Modulation Theory, Part I. New York: Wiley, 1968.

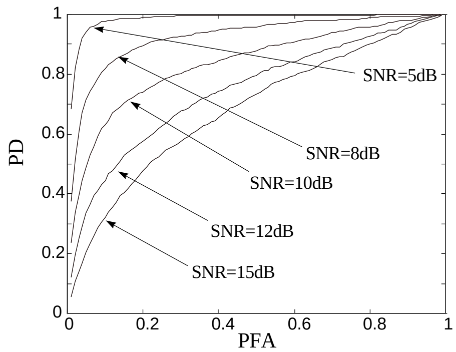

Fig. 1: SM detector ROC's for different SNR's - $N = 10000$

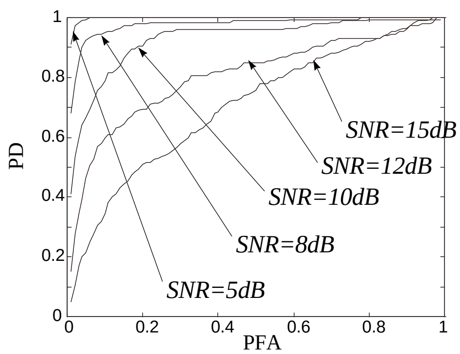

Fig. 2: SM detector ROC's for different SNR's - $N = 10000$

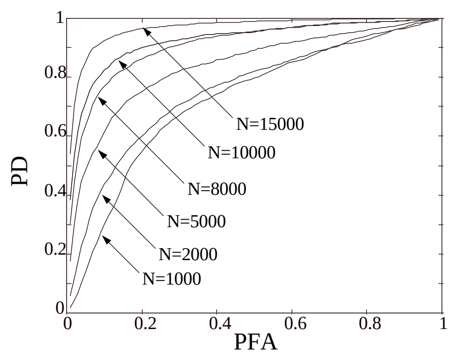

Fig. 3: SM detector ROC's for different number of samples - $SNR = 8dB$

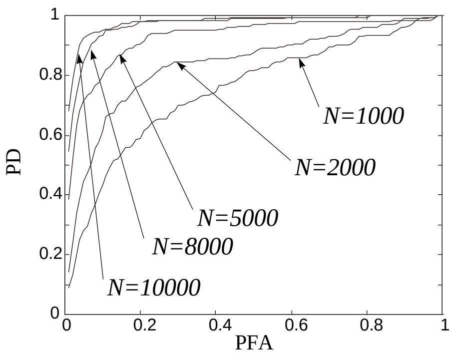

Fig. 4: MA detector ROC's for different number of samples - $SNR = 8dB$

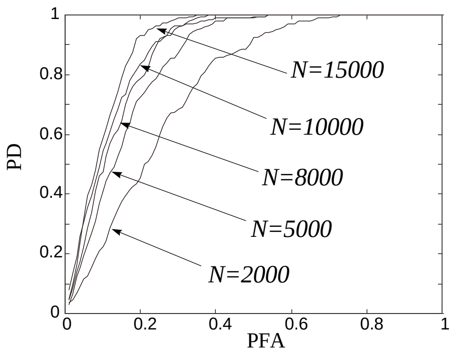

Fig. 5: SM detector ROC's for an $AR(4)$ for different number of samples - $SNR = 8dB$

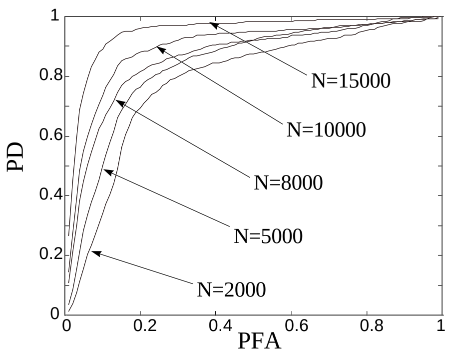

Fig. 6: MA detector ROC's for an $AR(4)$ for different number of samples - $SNR = 8dB$