NONLINEAR H-$ARMA$ MODELS

David Declercq and Patrick Duvaut

ETIS URA CNRS 2235 6, avenue du PONCEAU 95014, Cergy-Pontoise, France
e-mail : declercq@ensea.fr

ABSTRACT

We present, in this contribution, some aspects of nongaussian H-$ARMA$ models. After recalling that an H-$ARMA$ process is obtained by passing an $ARMA$ process through a Hermite polynomial nonlinearity, we describe the theoretical analysis of their cumulants and cumulant spectra. The main advantage of this kind of model is that the cumulant structure of the output can be deduced directly from the input covariance sequence. We give the analytic forms of these cumulants, together with some comments on their estimation. Then, we present the problems we are facing concerning the identification of the model's parameters, and give a first (and naive) method for their estimation. We give some results obtained on synthetic data and finally conclude with some remarks on this class of processes.

1. BRIEF PRESENTATION OF H-$ARMA$ MODELS

H-$ARMA$ models have been introduced in [1] with a detailed study of the estimation of their polycovariances. The mean of generation of a H-$ARMA$ process is twofold. Having a stationary white gaussian input with zero mean and unity variance, we first obtain a colored gaussian $ARMA$ process with a linear filter and then we apply it to an instantaneous nonlinear filter in order to generate a nongaussian behaviour. The choice of the nonlinearity is obviously crucial to provide an efficient modelisation of various nongaussian behaviours. A polynomial nonlinear filter satisfies this condition and allows to derive theoretical results concerning the cumulants and the distribution of such processes, what is useful to build identification methods. First, we have worked on the simplest polynomial - a square - and the results we have obtained [2] led us to consider Hermite polynomials as an extention. The nonlinearity of H-$ARMA$ models is therefore a linear combination of Hermite polynomials of various degrees to which we will refer as *Hermite filter* (cf. figure 1). The equation of an H-$ARMA(P,p,q)$ model is given by

$$z[n] = \sum_{k=1}^{P} \alpha_k H_k\left(y[n]\right) \tag{1}$$

$$y[n] = \sum_{i=1}^{p} a_i y[n-i] + \sum_{j=1}^{q} b_j x[n-j] + x[n] \tag{2}$$

where $x[n]$ is a stationary white gaussian noise.
Because nongaussian processes can not be identified only with the covariance function or the spectrum, we need for that purpose, analytic expressions of higher order cumulants or cumulant spectra. The use of Hermite polynomials ensures that we have access to tractable - and sometimes simple - expressions for the cumulants of our models at any order. Moreover, these expressions depend directly from the covariance function of the $ARMA$ process, what will help us to build inversion algorithms. In section 2, we present briefly and with no proof the means to obtain such expressions, both from the time and frequency standpoints. Section 2.3 is devoted to a first try to solve the problem of identifying the coefficients of the nonlinearity and the $ARMA$ filter, with a sequential approach. We discuss in the conclusion of the advantages and the drawbacks of our models, followed with the possible applications to real data and some ideas to tackle with the inversion of H-$ARMA$ models.

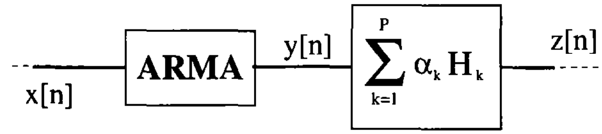

Figure 1: generation of H-$ARMA$ models

2. STATISTICAL PROPERTIES

We need in this section the use of some useful formulas based on n-linear expansions in Hermite polynomials series. These expansions allow to derive expressions for the expectation of products of Hermite polynomials when applied to a colored gaussian variable. The well known bilinear case - the Mehler formula - leads to

$$\mathbb{E}\left[H_m(x)H_n(y)\right] = n!\rho^n\delta(m-n) \tag{3}$$

where x and y are normal variables with covariance ρ.
Kibble has extended the Mehler formula to the n-linear case, and Slepian has proved the same result a few years later with a more general approach [3] [4]. We are therewith able to give the general expression of the cumulants of an H-$ARMA$ process at any order because it is only a linear combination of expectations of Hermite polynomials products. Unfortunately, this expression requires some special notations and more space than we have in this contribution. We will then restrict our study to the second and third order cumulants and we will report the general case in a future paper.

remark

For convenience and to make the formulas described above still valid, the $ARMA$ process $y[n]$ must be standardised - what we may do without loss of generality. In that case, a H-$ARMA$ process is always centered (with zero mean), property that comes from the Hermite polynomials: $\mathbb{E}\left[H_i(y[n])\right] = 0 \quad \forall i \geq 1$.

2.1. cumulants

Knowing the autocovariance sequence $\Gamma[n]$ of the $ARMA$ process $y[n]$, we can give the expression of the covariance $\gamma_2[n]$ of the H-$ARMA$ nongaussian process $z[n]$ (figure 1) with the help of the Mehler formula

$$\gamma_2[n] = \mathbb{E}\left[z[t]z[t+n]\right] = \sum_{k=1}^{P} \alpha_k^2\, k!\, \Gamma^k[n] \tag{4}$$

Using the Kibble-Slepian formula at the third order, we can follow the same approach to derive the expression of the bicovariance of $z[n]$

$$\begin{aligned}
\gamma_3[m,n] &= \mathbb{E}\left[z[t]z[t+m]z[t+n]\right] \\
&= \sum_{k_1,k_2,k_3=1}^{P} \alpha_{k_1}\alpha_{k_2}\alpha_{k_3}\, \mathbb{E}\left[H_{k_1}\left(y[t]\right) H_{k_2}\left(y[t+m]\right) H_{k_3}\left(y[t+n]\right)\right] \\
&= \sum_{k_1,k_2,k_3=1}^{P} \frac{\alpha_{k_1}\alpha_{k_2}\alpha_{k_3}\, k_1!k_2!k_3!}{K_1!K_2!K_3!} \Gamma[m]^{K_1}\Gamma[n]^{K_2}\Gamma[n-m]^{K_3}
\end{aligned} \tag{5}$$

where $K_1 = \frac{k_1+k_2-k_3}{2}$, $K_2 = \frac{k_1+k_3-k_2}{2}$, $K_3 = \frac{k_2+k_3-k_1}{2}$ and we keep in the sum only the terms for which $k_1+k_2+k_3$ is even. This tedious result becomes very simple when the *Hermite filter* restricts to only one polynomial of degree k. In that case, we have

$$\begin{aligned}
\gamma_3[m,n] &= \frac{(k!)^3}{\left(\frac{k}{2}\right)^3}\left(\Gamma[m]\Gamma[n]\Gamma[n-m]\right)^{\frac{k}{2}} && \text{if k is even} \\
&= 0 && \text{if k is odd}
\end{aligned} \tag{6}$$

We have considered, as an example, a process H-$ARMA(2,3,3)$ with Hermite filter $H2-2H1$, with 3 zeros and 3 poles for the $ARMA$ part:

$$\begin{aligned}
zeros &= (0.9, -0.9, 0.5) \\
poles &= \left(0.7, 0.9e^{2j\pi 0.075}, 0.9e^{-2j\pi 0.075}\right)
\end{aligned}$$

We have drawn on figure 2 the shape of the theoretical bicovariance of this model.

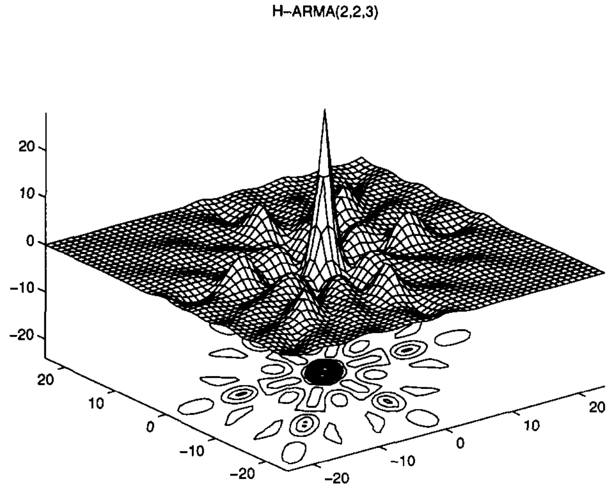

Figure 2: bicovariance shape of a H-$ARMA(2,3,3)$

The discrete bispectrum can also be obtained. After a few calculation, we finally have the following expression

$$\begin{aligned}
\mathring{\gamma}_3\,[s_1,s_2] = &\sum_{k_1,k_2,k_3=1}^{P} \frac{\alpha_{k_1}\alpha_{k_2}\alpha_{k_3} k_1!k_2!k_3!}{N\ \ K_1!K_2!K_3!} * \\
&\sum_{l=0}^{N-1} \mathring{\Gamma}\,[l]^{*K_1}\, \mathring{\Gamma}\,[s_1+l]^{*K_2}\, \mathring{\Gamma}\,[s_2-l]^{*K_3}
\end{aligned} \tag{8}$$

where K_1, K_2 and K_3 are defined in the preceeding section.
We have drawn on figure 3 the theoretical bispectrum of the same model as in figure 2

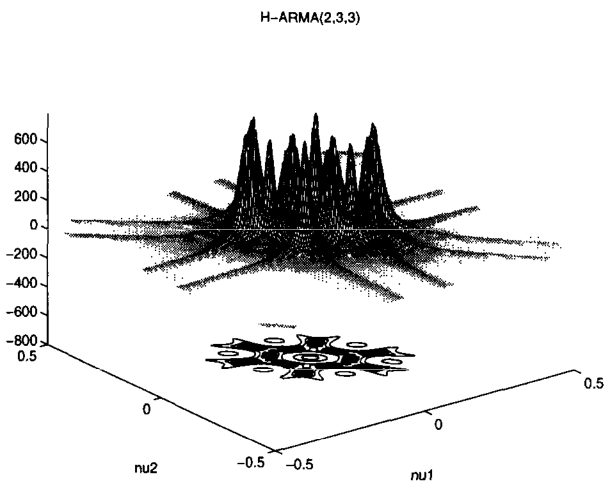

Figure 3: bispectrum of a H-$ARMA(2,3,3)$

2.2. cumulant spectra

The cumulant spectra of H-$ARMA$ models exists at any order as the Fourier transform of the corresponding cumulant if $z[n]$ has an absolutely continuous spectrum - which is true because it is a instantaneous nonlinear transformation of a gaussian $ARMA$ process. The expression of the spectrum of similar processes is given in ([5], pp.82-88) but with no results for the bispectrum.
The discrete spectrum of an H-$ARMA$ model with respect to s, the discrete frequency, is given by

$$\mathring{\gamma}_2\,[s] = DFT(\gamma_2[n]) = \sum_{k=1}^{P} \alpha_k^2\, k!\, \left(\mathring{\Gamma}\,[s]\right)^{*k} \tag{7}$$

where $\mathring{\Gamma}\,[s] = \sum_{n=1}^{N} \Gamma[n]e^{-2j\pi\frac{sn}{N}}$ is the discrete spectrum of the $ARMA$ process $y[n]$ and $\mathring{\Gamma}\,[s]^{*k}$ is the nth-order discrete convolution of $\mathring{\Gamma}\,[s]$ (we have to keep in mind that s is a discrete frequency).

2.3. empirical estimators

A complete study of empirical estimators of the cumulants can be found in [6] [7], in an asymptotic standpoint only: the estimators are consistant and asymptotically normal. The special structure of $H\text{-}ARMA$ models give us the ability to obtain the *exact* variance of estimation - that means for finite sample sizes - for the covariance or the bicovariance. We have described the method for obtaining such a variance in [1], following the ideas pointed out in [2], supported by extensive Monte Carlo simulations. The main result is that, for a particular process, we can give the minimum number of estimation points to reach a given variance of estimation (both for the covariance and the bicovariance).

3. MODEL IDENTIFICATION

3.1. description of the method

The estimation of the coefficients of a nonlinear model is of a great importance and is never easily solved. In most cases, the dimension of a model is approached with an Akaike like criterion, and the coefficients itselves with a maximum likelihood approach. Moreover, some model structures allow to consider moment based methods (see [8] for a wide range of examples). In $H\text{-}ARMA$ models, the only prior is the gaussian density of the input, and we have choosen - in a first attempt of identification - to solve the inversion problem separately: the nonlinear part, first, and then the $ARMA$ coefficients. We are aware that this is obviously a sub-optimal approach, but the identification of these coefficients is so hard that we must proceed step by step. We will present in the sequel of this section, the two steps identification knowing the dimension of the model, that means that we take as a prior the maximal degree of the Hermite filter and the order of the $ARMA$ filter.

- Hermite filter coefficients
  A first - although naive - idea to estimate the coefficients α_k of the nonlinear filter is to build a penalty function and minimize it over a field of possible values of α_k (that's the basis of mean-square approaches). Because our specific problem is ill-conditionned, we have regularised it by taking a fractional distance between two probability density functions. First, we estimate the histogram $\hat{f}(i)$ of the observed data $z[n]$ on a discrete set of classes and numerically [1] calculate the posterior density of $z[n]$, knowing α_k. We minimize thereafter the distance (9) with a dichotomised grid search (various step lengths).

$$d_\beta(\hat{f}, f) = \left(\sum_i \left(\hat{f}(i) - f(i|\alpha_k) \right)^\beta \right)^{1/\beta} \qquad \beta < 1 \tag{9}$$

  After a few tries, we have empirically choosen the value $\beta = 0.1$.
  We have applied this approach to synthetic data in the next section, but the computer time to complete this kind of minimisation is extremely large and we should have to speed up the convergence rate with classical methods (gradient search ...). Unfortunately, the shape of (9) is extremely chaotic and such methods don't provide reliable estimates (in our case). Some ideas to regularise the problem and to find other ways to achieve convergence in a reasonable computer time are under investigation, they will be reported in a future paper.

  remark

  We have choosen this type of distance (9) because $\mathcal{L}_1$ or $\mathcal{L}_2$ distances were providing very bad results and classical methods based on the likelihood (Kullback-Leiber distances, EM, MCMC) are unefficient here because it does not provide a unique maximum but rather a continuous crest of maxima. Some comments and explanations can be found in [9].

- $ARMA$ coefficients
  We present in this paper a method of identifying the $ARMA$ coefficients of the $H\text{-}ARMA$ models based on the cumulants. The developpement of other methods (maximum likelihood, composite criterion, MCMC approach ...) is still in progress and a comparison study on real data will be presented in [9].
  Having identified the Hermite filter with the first step (estimated the values of $\hat{\alpha}_k$) and giving estimates of the covariance and the bicovariance on a principle axis ($m = 0$) of the observed data, we can deduce the covariance function $\Gamma[n]$ of the $ARMA$ process with the following polynomials (see eq. (4)-(5))

$$\begin{cases} P_1\left(\Gamma[n]\right) = \displaystyle\sum_{k=1}^{P} \alpha_k^2\, k!\, \Gamma^k[n] - \hat{\gamma}_2[n] \\ P_2\left(\Gamma[n]\right) = \displaystyle\sum_{k_1,k_2,k_3=1}^{P} \frac{\alpha_{k_1}\alpha_{k_2}\alpha_{k_3} k_1! k_2! k_3!}{K_1! K_2! K_3!} \Gamma[n]^{k_3} - \hat{\gamma}_3[0,n] \end{cases}$$

  An estimated value $\hat{\Gamma}[n] \quad \forall n > 0$ is equal to the common zero of P_1 and P_2. In the general case, P_1 or P_2 have P zeros, where P is the maximum degree of the Hermite nonlinearity - then we must choose only one solution which is the estimation of $\Gamma[n]$. A first way to solve the problem of the choice of a zero at any n is therefore to consider the common zero (if it exists of P_1 and P_2). A more global answer to this problem - also usefull when P_1 and P_2 have more than one common zero - is to decide at a time n which zero is the right one by testing the positive definite property of the sequence $\Gamma[i] \quad \forall i \leq n$.
  Finaly, we make use of a Levinson type algorithm, based on the covariance sequence, to estimate the order and the coefficients of the $ARMA$ filter. Because of its sequential steps, this method is not satisfactory: it depends on the accuracy of the estimates $\hat{\alpha}_k$. The advantage of this approach is that the estimation of the autocovariance function of $y[n]$ depends directly from the estimations of the covariance and the bicovariance of $z[n]$ on which we have proved theoretical results of convergence (see section 2.3).

3.2. results on synthetic data

We have considered 2000 samples of an $H\text{-}ARMA(3,2,0)$ process with Hermite filter $H_3 - 1.4H_2 + 3.2H_1$ and AR poles $(0.9e^{2j\pi 0.075}, 0.9e^{-2j\pi 0.075})$. The results obtained with the methods described in the preceeding section are reported in table 1. For the identification of the Hermite filter, we have started the grid search with the boundary regions: $-10 \leq \alpha_1 \leq 10$, $-5 \leq \alpha_2 \leq 5$, $-5 \leq \alpha_3 \leq 5$. The coefficients of the odd degree Hermite

[1] because we don't have an analytic expression for the p.d.f. of such a nongaussian process

polynomials are determined up to a sign because the distribution of the input is symmetrical (gaussian density) and a coefficient of odd order and its opposite yield the same ouput behaviour. Taking those values of α_k, we obtain the two polynomials:

$$\begin{cases} P_1(\Gamma[n]) = 6.46\,\Gamma^3[n] + 4.05\,\Gamma^2[n] + 10.24\,\Gamma[n] - \hat{\gamma}_2[n] \\ P_2(\Gamma[n]) = -167.24\,\Gamma^3[n] - 164.3\,\Gamma^2[n] - 115.14\,\Gamma[n] - \hat{\gamma}_3[0,n] \end{cases}$$

We have estimated 20 points of the ouput covariance and bicovariance and both polynomials yield to the same covariance function of $y[n]$. We have ploted on figure 4 the estimated covariance function with the theoretical one in dashed line. The results are good for this very simple example, but studies on classical modelisation criteria such as fitting real data or prediction must be developped in future works.

	Hermite filter		AR filter		
	true	est.	true	est.	
α_1	3.2	*-3.201*	2	2	order
α_2	-1.4	*-1.424*	0.9	*0.899*	a
α_3	1	*-1.038*	0.075	*0.071*	ν

Table 1: estimations of an H-$ARMA$ model

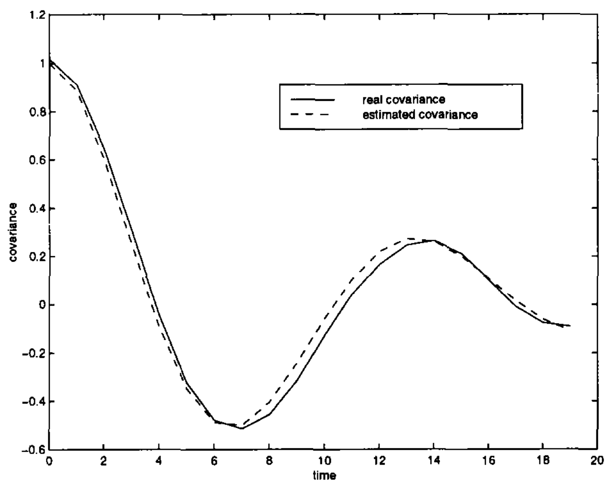

Figure 4: estimation of the covariance of $y[n]$

4. CONCLUSION AND DISCUSSION

The H-$ARMA$ models are a new attempt in modelling nongaussian data by nonlinear transformation of a colored gaussian input. The special structure of these models - instantaneous nonlinear filter, Hermite polynomials ... - ensures a lot of good properties, and especially the ability to derive analytic expressions for their cumulants and for the variance of their estimation. If that kind of nonlinear filter does not provide a very wide range of distribution behaviours, it allows on the other side a simple and intuitive control of the covariance aspects. For that reason, we advice the use of H-$ARMA$ more specifically when the observed data comes from a nonlinear system that is known to have a polynomial behaviour. We can mention for instance the beamforming diagram of antenna where the output measures comes obviously from a quadratic system.

The identification of H-$ARMA$ models is far from being solved and the inversion method described in this paper is clearly subobtimal and computationally intensive. We base the estimation of the $ARMA$ covariance function on the estimated nonlinear coefficients $\hat{\alpha}_k$, and we have no real result on the accuracy of their estimation: a global approach in which we would estimate jointly the nonlinear and the linear coefficients would be better.

5. REFERENCES

[1] D. Declercq and P. Duvaut. Propriétés des processus *h-arma*. In *GRETSI-97*, volume 1, pages 151–154, Grenoble, 1997.

[2] P. Duvaut and D. Declercq. Real and complex qarma processes. *J. Frank. Inst.*, (333-B)(3):413–424, 1996.

[3] W. F. Kibble. An extension of a theorem of mehler's on hermite polynomials. *Proc. Cambridge Phil. Soc.*, (41):12–15, 1945.

[4] D. Slepian. On the symmetrized kronecker power of a matrix and extentions of mehler's formula for hermite polynomials. *SIAM J. Math. Anal.*, (3):606–616, 1972.

[5] E.J. Hannan. *Multiple Time Series.* Wiley series in Prob and Math Stat, 1970.

[6] C.L. Nikias and A.P. Petropulu. *Higher-order Spectra Analysis : A nonlinear signal processing framework.* Prentice Hall, New York, 94.

[7] Rosenblatt and al. *Cumulant spectral estimates: bias and variances.* Coll. Math. Soc. Jan. Bol.: limit theorems in Prob. and Stat. Vol.57, Holland, 89.

[8] D. Guegan. *Séries chronologiques non linéaires à temps discret.* Stat. Math. Prob., Paris, 1994.

[9] D. Declercq and P. Duvaut. Inversion of *h-arma* models. In *submitted to EUSIPCO-98*, Rhodes, 1998.