On the use of higher-order statistical tests in the analysis of time series associated with space data

Utilisation de tests basés sur des statistiques d'ordre supérieur dans l'analyse de séries temporelles mesurées dans l'espace

par A. MASSON, B.B. SHISHKOV, F. LEFEUVRE

Laboratoire de Physique et Chimie de l'Environnement, Centre National de la Recherche Scientifique, 45071 Orléans, Cedex 2, France

abstract and key words

Tests of hypotheses based on Higher Order Statistics (HOS) are reviewed in the particular context of the identification of non-linear processes in space plasma. The time series under study are associated with the measurements of electric or/and magnetic field components, or/and counting rates of particles. The basic principles of HOS techniques are reviewed. A general and unified procedure is suggested in order to construct statistical tests: (1) for detecting a non-gaussian or transient signal in a gaussian (or non-gaussian) noise, (2) testing a stochastic time series for non-gaussianity (including non-linearity), (3) studying non-linear wave interactions by using the kth-order coherency function. Asymptotic theory of estimates of the kth-order spectra is implemented in a digital signal processing framework. The effectiveness of the signal detection algorithms is demonstrated through computer simulations. Examples of application on the analysis of satellite data are given.

Higher-order statistics; non-linear phenomena; transient phenomena; space plasmas.

résumé et mots clés

Des tests d'hypothèses basés sur des statistiques d'ordre supérieur sont revus dans le contexte particulier de l'identification de processus non-linéaires dans les plasmas spatiaux. Les séries temporelles étudiées sont associées à la mesure de composantes du champ électrique et/ou magnétique d'ondes ou de turbulences, et/ou de données particules. Les principes de base des statistiques d'ordre supérieur sont brièvement rappelés. Une procédure générale et unifiée est suggérée afin de construire des tests statistiques permettant : (1) de détecter des signaux non-gaussiens ou transitoires au sein d'un bruit gaussien (ou non-gaussien), (2) de tester si une série temporelle est associée ou non à un processus stochastique issu d'un processus non-linéaire, (3) d'étudier des interactions non-linéaires à plusieurs ondes par l'utilisation de la fonction de cohérence d'ordre k. La théorie asymptotique des estimés des spectres d'ordre k est mise en œuvre dans le cas discret. L'efficacité des algorithmes de détection est démontrée par le biais de simulations numériques. Des exemples d'applications à des données satellites sont présentés.

Statistiques d'ordre supérieur ; phénomènes non linéaires ; phénomènes transitoires ; plasmas spatiaux.

1. introduction

Nonlinear processes are known to play a major role in the exchange of energy in space plasmas. Their identification is one of the main objectives of space plasma experimenters. The studies are generally performed from time series associated with the measurements of electromagnetic wave field components and/or of particle distributions. The analysis techniques which are involved are based on principles developed in statistical physics. However, in the absence of review papers directed to the problems they have to solve, experimenters have a tendency either to make use of a small amount of the capabilities of the methods only, or to use hypotheses they don't check.

An illustration of the problem encountered by space experimenters is given in [32]. To point out the presence of a nonlinear interaction in the ionosphere, they are looking for phase relationships between the waveform of an electric field component measured on-board a satellite and the time series associated with density fluctuations observed at the same times on the satellite. As the two signals are clearly non-stationary at the time intervals where an energy exchange seems to take place, the authors use properties of higher-order correlations and spectra for bandlimited deterministic transients, as recalled by Pflug *et al*. in a series of papers devoted to the study of acoustic signals [37, 38]. All references are given on the basic papers from which it is possible to check whether the procedure is applicable or not. But, the checking is not done for two reasons: first, the needed information is scattered in several papers, which makes difficult a full understanding of the principles; second, real data are so different from the synthetic data on which the principles have been tested that it seems worthless to try to test all hypotheses.

The same remarks can be made for others papers using higher-order statistics published by space plasma experimenters [10, 28, 52]. The authors are faced to the same experimental constraints:

– limitation in the measured field components and/or particle distributions, with the consequence that the most relevant data are not always available,

– signals recorded at a moving point in space, with the consequence : (i) that the time intervals on which the analysis is possible is very limited, (ii) that the observations are generally made off the interaction regions, *i.e.* that the original phenomena are partially masked by propagation effects,

– observation of intermediate regimes between strong turbulence (characterised by the apparition of transient events and/or clear frequency peaks on the power spectra) and weak turbulence (which look like stochastic phenomena).

Before applying higher-order statistics, they make clear their hypotheses. But they don't re-examine basic questions of applicability. It is the objective of the present paper to provide elements to fill this gap and to present examples of applications on satellite data.

Let start the re-examination of the applicability or higherorder statistics right from the beginning. When using the power spectral density (PSD) estimation, the process under consideration is treated as a superposition of statistically uncorrelated harmonic components [35]. As a consequence, phase relations between frequency components are suppressed [20]. In other words, second-order statistics (SOS) are phase-blind and only describe linear mechanisms governing the process. However, there are practical situations where we would have to look beyond the power spectrum or autocorrelation domain.

The general motivation behind the use of HOS in signal processing is threefold [34]

1. to extract information due to deviations from normality,

2. to estimate the phase of non-Gaussian parametric signals,

3. to detect and characterise the nonlinear properties of mechanisms that generate time series *via* phase relations of their harmonic components.

Higher-order spectra are quite natural tools to analyse the nonlinearity of a system operating under a random input. However, general relations are not available for arbitrary random data passing arbitrary nonlinear systems. Each type of non-linearity has to be investigated as a special case. Despite the fact that progress has been achieved in developing the theoretical properties of nonlinear models, only a few statistical methods for the detection and characterisation of non-linearity from a finite set of observations are available [35].

In [17], [18], [19], [22], [24] and [49], by using 3^{rd}-order coherency function, several statistical tests have been proposed:

– to detect a non-gaussian or transient signal in a gaussian (or non-gaussian) noise,

– to test a stochastic time series for non-gaussianity (including non-linearity),

– to study non-linear wave interactions.

They refer to fundamental books and papers such as [1], [5], [6], [7], [34], [35], [41], [42], and [47].

The aim of the present paper is to give experimenters a basic theoretical background without having to consider a variety of approaches that are often accompanied by errors and/or incomplete simulations. A general and unified procedure of higher-order spectra analysis technique is suggested to reconstruct the statistical tests listed above. It allows to determine the smallest $(k-1)$ dimensional region in the frequency domain (or respectively the k^{th}-order discrete principal time domain) whose k^{th}-order spectral values uniquely specify the entire k^{th}-order spectrum. Using this result, one is able to construct the k^{th}-order principal test-statistic.. A slight modification of the principal sta-

tistical tests allows passing easily from one algorithm to another. Asymptotic theory of estimates of the k^{th}-order spectra is implemented in a digital signal processing framework. The effectiveness of the signal detection algorithms is demonstrated through numerical simulations and applied on real satellite data. All the notations and abbreviations used in this paper are listed at the end.

2. general background

2.1. moments and cumulants

Before outlining some of the methods that have been proposed, let first introduce some useful background definitions and properties of higher-order statistics [7, 33, 35, 51].
Let $v = (v_1, v_2, \ldots, v_k)$ be real parameters and $x = (x_1, x_2, \ldots, x_k)$ a collection of random variables. The k^{th}-order joint moments are given by [35]

$$\begin{aligned}\text{Mom}[x_1, x_2, \ldots, x_k] &= E\{x_1, x_2, \ldots, x_k\} \\ &= \{-j\}^k \frac{\partial^k \Phi(v_1, v_2, \ldots, v_k)}{\partial v_1 \partial v_2 \ldots \partial v_k}\Big|_{v_1=v_2=\ldots=v_k=0}\end{aligned} \tag{1}$$

where

$$\Phi(v_1, v_2, \ldots, v_k) = E\left\{\exp[j(v_1x_1 + v_2x_2 + \ldots + v_kx_k)]\right\}$$

is their joint characteristic function and $j^2 = -1$.
The joint cumulants (also called semi-invariants of order k), $\text{Cum}[x_1, x_2, \ldots, x_k]$, of the same set of random variables, are defined as the coefficients of $(v_1, v_2, \ldots, v_k)$ in the Taylor series expansion (provided it exists) of the cumulant-generating function

$$\Psi(v_1, v_2, \ldots, v_k) = \ln\, \Phi(v_1, v_2, \ldots, v_k) \tag{2}$$

about zero, *i.e.*

$$\text{Cum}[x_1, x_2, \ldots, x_k] = (-j)^k \frac{\partial^k \ln\Phi(v_1, v_2, \ldots, v_k)}{\partial v_1 \partial v_2 \ldots \partial v_k}\Big|_{v_1=v_2=\ldots=v_k=0} \tag{3}$$

Some important properties of cumulants can be found in the Appendix.

2.2. moments and cumulants of stationary processes

Consider a real stationary random process $\{x_n\} = \{x(n)\}$, $n = 0, \pm1, \pm2, \ldots$ and its moments up to order k exist. The k^{th}-order moment and cumulant of this process, denoted $m_{kx}(\tau_1, \ldots, \tau_{k-1})$ and $c_{kx}(\tau_1, \ldots, \tau_{k-1})$, are defined as the joint k^{th}-order moment and cumulant of the random variables $x(n)$, $x(n + \tau_{k_1}), \ldots, x(n + \tau_{k-1})$, *i.e.*

$$m_{kx}(\tau_1, \ldots, \tau_{k-1}) = E[x(n)x(n + \tau_1) \ldots x(n + \tau_{k-1})] \tag{4}$$

$$c_{kx}(\tau_1, \ldots, \tau_{k-1}) = \text{Cum}[x(n), x(n + \tau_1), \ldots, x(n+\tau_{k-1})] \tag{5}$$

They depend only on the time differences $\tau_1, \ldots, \tau_{k-1}, \tau_i = 0, \pm1, \pm2, \ldots$ for all i. Note that if $\{x(n)\}$ were non-stationary, all m_{kx} and c_{kx} would also be time dependent.
For the gaussian case, il can be shown that all m_{kx} and c_{kx} for *k* odd are identical to zero [26]. Let us stress out that the even-order moments, and specially m_{4x}, are, generally speaking, not equal to zero, while the even-order cumulants are null.
For the non-gaussian case, it can also be shown [26] that, for $k = 3, 4$, m_{kx} and c_{kx} are related such that

$$c_{kx}(\tau_1, \ldots, \tau_{k-1}) = m_{kx}(\tau_1, \ldots, \tau_{k-1}) - m_{kg}(\tau_1, \ldots, \tau_{k-1}) \tag{6}$$

where $\{g(n)\}$ is a Gaussian random process with the same second order statistics as $\{x(n)\}$. Cumulants therefore, not only display the amount of higher-order correlation, but also provide a measure of the distance of the random process from gaussianity. Finally, using the properties mentioned above, it can be easily demonstrated that the 2nd, 3rd and 4th order cumulants or the random process $\{x(n)\}$ in the non-gaussian case, are given by

$$c_{2x}(\tau) = E\left[x(n)x(n + \tau)\right] \tag{7}$$

$$c_{3x}(\tau_1, \tau_2) = E\left[x(n)x(n + \tau_1)x(n + \tau_2)\right] \tag{8}$$

$$\begin{aligned}c_{4x}(\tau_1, \tau_2, \tau_3) = E\left[x(n)x(n + \tau_1)x(n + \tau_2)x(n + \tau_3)\right] \\ - c_{2x}(\tau_1)c_{2x}(\tau_2 - \tau_3) \\ - c_{2x}(\tau_2)c_{2x}(\tau_3 - \tau_1) \\ - c_{2x}(\tau_3)c_{2x}(\tau_1 - \tau_2)\end{aligned} \tag{9}$$

By setting $\tau = \tau_1 = \tau_2 = \tau_3 = 0$ in (7), (8) and (9), we get

$$\begin{aligned}\gamma_{2x} &= E[x^2(n)] = c_{2x}(0) \quad \text{(variance)} \\ \gamma_{3x} &= E[x^3(n)] = c_{3x}(0, 0) \text{ (normalized skewness)} \\ \gamma_{4x} &= E[x^4(n)] - 3[\gamma_{2x}]^2 = c_{4x}(0, 0, 0) \text{ (normalized kurtosis)}\end{aligned} \tag{10}$$

2.3. cumulant spectra

Assuming that $c_{kx}(\tau_1, \ldots, \tau_{k-1})$ is absolutely summable, the k^{th}-order discrete spectra or polyspectra of the process $\{x(n)\}$ are defined as the $(k-1)$– dimensional discrete-time Fourier transform of the k^{th}-order cumulant, *i.e.*

$$C_{kx}(f_1,\ldots,f_{k-1}) = \sum_{\tau_1}\ldots\sum_{\tau_{k-1}} c_{kx}(\tau_1,\ldots,\tau_{k-1}) \exp\{-j2\pi(f_1\tau_1+\ldots+f_{k-1}\tau_{k-1})\} \quad (11)$$

It is a periodic function with period 1,

$$|f_i| \leqslant \frac{1}{2} \text{ for } i = 1,2,\ldots,k-1 \text{ and}$$
$$|f_1+f_2+\ldots+f_{k-1}| \leqslant \frac{1}{2}.$$

The idea of a spectral representation for higher-order moments of a time series appears in [4], and was further developed in [47]. A spectral representation for cumulants (attributed to Kolmogorov) appears in [47]. Higher-order spectra are derived from the first principles in [5, 6]. The quantities $C_{2x}(f)$, $C_{3x}(f_1,f_2)$ and $C_{4x}(f_1,f_2,f_3)$ are called the spectrum, the bispectrum, and the trispectrum, respectively. These quantities will also be denoted as $S_x(f), B_x(f_1,f_2), T_x(f_1,f_2,f_3)$, throughout this paper. According to the alternative direct method [5], they can be presented in the discrete frequency domain as follows

$$\begin{aligned} S_x(f) &= E[X(f)X^*(f)] \text{ (spectrum)} \\ B_x(f_1,f_2) &= E[X(f_1)X(f_2)X^*(f_1+f_2)] \text{ (bispectrum)} \\ T_x(f_1,f_2,f_3) &= E[X(f_1)X(f_2)X(f_3)X^*(f_1+f_2+f_3)] \\ &\quad \text{(trispectrum)} \end{aligned} \quad (12)$$

where

$$X(f) = X\left(\frac{m}{N}\right) = \sum_{n=0}^{N-1} x(n)\exp\left\{-j\frac{2\pi mn}{N}\right\}, \quad m = 0,1,\ldots,N/2 \quad (13)$$

is an N-point discrete Fourier transform (DFT).

2.4. the k^{th}-order coherency function

A normalized cumulant spectrum, or k^{th}-order coherency function, is a function that combines two completely different entities, namely, the cumulant spectrum of order k, $C_{kx}(f_1,\ldots,f_{k-1})$ and the power spectrum $C_{2x}(f)$ of a process. The k^{th}-order coherency function is defined as follows

$$P_{kx}(f_1,\ldots,f_{k-1}) = \frac{C_{kx}(f_1,\ldots,f_{k-1})}{[C_{2x}(f_1)\ldots C_{2x}(f_{k-1})C_{2x}(f_1+\ldots+f_{k-1})]^{1/2}} \quad (14)$$

The second-order $(k=2)$, the third-order $(k=3)$ and the fourth-order $(k=4)$ coherency function are called coherency, bicoherency and tricoherency respectively and will be also denoted as $s_x(f), b_x(f_1,f_2), t_x(f_1,f_2,f_3)$. The k^{th}-order coherency function is very useful for the detection and the characterization of non-linearities in time series via phase relations of their harmonic components.

The magnitude of the k^{th}-order coherency, $|P_{kx}(f_1,\ldots,f_{k-1})|$, is called the coherency index.

Many symmetries have been pointed out [34, 35, 36, 41] in the argument of $c_{kx}(\tau_1,\ldots,\tau_{k-1})$ and $C_{kx}(f_1,\ldots,f_{k-1})$ which make their calculations manageable. It is very important to know the k^{th}-order discrete-time principal domain PD_k (or respectively the smallest $(k-1)$-dimensional region in the frequency domain whose k^{th}-order spectral values uniquely specify the entire k^{th}-order spectrum) [8].

Finally a logical question to ask is "Why do we need fourth-order cumulants, *i.e.*, aren't third-order cumulants sufficient?" If a random process is symmetrically distributed, then its third-order cumulant equals zero; hence, for such a process, we must use fourth-order cumulants.

3. statistical tests based on HOS: detection scheme and HOS estimators

Throughout most of this paper, we will consider the basic problem of detecting the presence or the absence of a signal $\{s(n); n=0,1,\ldots,N-1\}$ in a set of measurements $\{x(n); n=0,1,\ldots,N-1\}$, corrupted by a noise $\{\nu(n);\ n=0,1,\ldots,N-1\}$. In this section, $\{x(n)\}$ is assumed to be stationary and $\{s(n)\}$ to be non-gaussian. The noise is supposed to be an additive white Gaussian noise (AWGN), and independent of the signal. Although both of these assumptions are often violated in applications, they are invoked here to focus our attention on the effects of the signal model in the detector design. The effect of a non-Gaussian noise will be discussed in section 4 and 6.

3.1. detection scheme

Since $\{\nu(n)\}$ is additive and independent of the signal, the detection scheme can be described mathematically in terms of an hypothesis test between the following pair of statistical hypotheses

$$H_0 : x(n) = \nu(n), n = 0,1,\ldots,N-1$$
versus (15)
$$H_1 : x(n) = \nu(n) + s(n), n = 0,1,\ldots,N-1$$

Assuming all the moments of the signal and the noise exist up to the order k, the gaussianity of $\{\nu(n)\}$ implies (*c.f.* equation (6)) that for $k \geqslant 3$ the cumulants $c_{k\nu}(\tau_1, \ldots, \tau_{k-1}) \equiv 0$, or respectively the higher-order spectra $C_{k\nu}(f_1, \ldots, f_{k-1}) \equiv 0$, over their corresponding time and frequency domains. Thus, assuming a non-Gaussian signal process $\{s(n)\}$ to have a nonzero k^{th}-order spectrum, the discrete-time hypotheses (15) can be turned into the following test

$$\begin{aligned} &H_0 : C_{kx}(f_1, \ldots, f_{k-1}) \equiv 0 \quad \text{over } PD_k \\ &\textit{versus} \\ &H_1 : C_{kx}(f_1, \ldots, f_{k-1}) \neq 0 \quad \text{over } PD_k \end{aligned} \tag{16}$$

where PD_k is the k^{th}-order discrete-time principal domain [5, 8] (*e.g.* $PD_2 = 1/2, PD_3 = 1/16$ if the process $\{x(n)\}$ is not aliased and $PD_4 = 1/72$).

As one can notice, this test depends only on the k^{th}-order spectra of the observation process $\{x(n)\}$. So, before going any further (see section 4.2.), we need to describe briefly the computation and the asymptotic properties of a consistent estimator of C_{kx} with reduced estimation variance.

3.2. computation and asymptotic properties of a consistent polyspectra estimator

In the $(k-1)$-dimensional frequency domain, a consistent estimator of $C_{kx}(f_1, \ldots, f_{k-1})$, based upon N samples of $\{x(n); n = 0, 1, \ldots, N-1\}$ and denoted $\widehat{C}_{kx}^{(N)}(f_1, \ldots, f_{k-1})$, can be formed as follows [5, 6]. As a starting point, one has to check the mean value of the signal. Indeed, if the mean value is significantly different from zeo, it might impact the dynamics of the signal. A way to avoid this problem is to centre the signal such that: $x(t) = x(t) - \frac{1}{T}\int_T x(t)dt$, with $T = NT_s$ and T_s the sampling period set to unity. Then, The $N = JL$ samples are divided into J non-overlapping segments of L samples each (L is assumed even). An L-point DFT is then performed over each segment

$$X^{(j)}(m) = \sum_{n=0}^{L-1} x^{(j)}(n) \exp\left\{-i\frac{2\pi mn}{L}\right\} \quad \text{for } m = 0, 1, \ldots, \text{ and } j = 1, 2, \ldots, J \tag{17}$$

at the point $(f_1, \ldots, f_{k-1}) = \left(\frac{m_1}{L}, \ldots, \frac{m_{k-1}}{L}\right)$, is given by

$$\widehat{C}_{kx}^{(j)}(f_1, \ldots, f_{k-1}) = \frac{1}{L} X^{(j)}(f_1) \ldots X^{(j)}(f_{k-1}) X^{(j)^*}(f_1 + \ldots + f_{k-1} \tag{18}$$

The frequency resolution (or bandwidth) of this periodogram is $1/L$ in each direction. In order to reduce the variance of this estimator, the k^{th}-order periodograms from each of the J non-overlapping segments are averaged to form

$$\widehat{C}_{kx}^{(av)}(f_1, \ldots, f_{k-1}) = \frac{1}{J}\sum_{j=1}^{J} \widehat{C}_{kx}^{(j)}(f_1, \ldots, f_{k-1}) \tag{19}$$

In order to further reduce the variance, an additional smoothing over a hypercube containing M resolution cells on a side (a total of M^{k-1} discrete-frequency values) is performed in the frequency domain, assuming the k^{th}-order spectrum is relatively smooth in that neighborhood. At this stage the exact smoothing window shape is not critical, so we will use a uniform window for simplicity.

This frequency averaging finally yields to

$$\widehat{C}_{kx}^{(N)}(f_1, \ldots, f_{k-1}) = \frac{1}{M^{k-1}} \sum_{q_1=-M/2}^{M/2-1} \cdots \sum_{q_{k-1}=-M/2}^{M/2-1} \widehat{C}_{kx}^{(av)} \left(f_1 + \frac{q_1}{L}, \ldots, f_{k-1} + \frac{q_{k-1}}{L}\right) \tag{20}$$

The resulting number of smoothed estimates $\widehat{C}_{kx}^{(N)}(f_1, \ldots, f_{k-1})$ over PD_k is approximately equal to $(PD_k)/\Delta_{kN}^{k-1}$.

The bandwidth of $\widehat{C}_{kx}^{(N)}$ in each dimension of the combined time-and frequency-domain smoothing is

$$\Delta_{kN} = M\frac{1}{L} = \frac{JM}{N} \tag{21}$$

If we define Δ_{kN} to be the smoothing bandwidth of a general estimate such as (20), then the asymptotic behavior of the resulting estimates is described by the following theorem

THEOREM 1 [6]: Given that the cumulants of $\{x(n)\}$ exist and satisfy a weak summability condition and that for $N \to \infty, \Delta_{kN}$ satisfies

$$\Delta_{kN} \to 0, \Delta_{kN}^{k-1} N \to \infty \tag{22}$$

then asymptotically as $N \to \infty$, the estimate is complex Gaussian distributed with

$$\widehat{C}_{kx}^{(N)}(f_1, \ldots, f_{k-1}) \sim N_c \left\{C_{kx}(f_1, \ldots, f_{k-1}), \frac{1}{\Delta_{kN}^{k-1} N} C_{2x}(f_1) \ldots C_{2x}(f_{k-1}) C_{2x}(f_1 + \ldots + f_{k-1})\right\} \tag{23}$$

The key result associated with this approach is that for sufficiently large N it provides approximately unbiased estimates

$$E\{\widehat{C}_{kx}^{(N)}(f_1, \ldots, f_{k-1})\} \cong C_{kx}(f_1, \ldots, f_{k-1}) \tag{24}$$

with asymptotic variances

$$\begin{aligned}&\mathrm{Var}\left\{\mathrm{Re}\left[\widehat{C}_{kx}^{(N)}(f_1,\ldots,f_{k-1})\right]\right\}\\ &\cong\ \mathrm{Var}\left\{\mathrm{Im}\left[\widehat{C}_{kx}^{(N)}(f_1,\ldots,f_{k-1})\right]\right\}\\ &\cong\ \frac{1}{\Delta_{kN}^{k-1}N}C_{2x}(f_1)\ldots C_{2x}(f_{k-1})C_{2x}(f_1+\ldots+f_{k-1})\end{aligned} \quad (25)$$

In addition, if the sampling grid width in the principal domain of the $(k-1)$-dimensional frequency domain is larger than the bandwidth Δ_{kN}, then the estimates at the different frequency points are asymptotically independent. Explicitly, $\widehat{C}_{kx}^{(N)}(f_{j_1},\ldots,f_{j_{k-1}})$ is independent of $\widehat{C}_{kx}^{(N)}(f_1,\ldots,f_{1_{k-1}})$ if for $j_i\neq 1_i, |f_{j_i}-f_{1_i}|\geqslant\Delta_{kN}, i=1,\ldots,k-1$. Henceforth we assume that Δ_{kN} satisfies condition (22), and that Δ_{kN} is chosen as the frequency spacing between points in the $(k-1)$-dimensional frequency domain at which $\widehat{C}_{kx}^{(N)}(f_{j_1},\ldots,f_{j_{k-1}})$ is evaluated, *i.e.* $f_j=f_{1_j}=1_j\Delta_{kN}, 1_j=0,1,\ldots$.

Henceforth we assume that Δ_{kN} satisfies condition (22), and that Δ_{kN} is chosen as the frequency spacing between points in the $(k-1)$-dimensional frequency domain at which $\widehat{C}_{kx}^{(N)}(f_1,\ldots,f_{1_{k-1}})$ is evaluated, *i.e.* $f_j=f_{1_j}=1\Delta_{kN}, 1_j=0,1\ldots$

It has to be reminded that the above procedure describes the multi-periodogram method to estimate polyspectra based on moments. For several reasons (see section 2), the polyspectra based on cumulants are preferred. Provided the signal has been centred, the two estimates are equivalent at order 3. But this is not the case for order 4 (trispectra).

3.3 computation of the trispectra based on cumulants

Assuming a real zero mean stationary random process $\{x(n); n=0,1,\ldots,N-1\}$, the frequency-domain counterpart of equation (9) is given by

$$\begin{aligned}C_{4x}(f_1,f_2,f_3)=\ &C_{4x}^m(f_1,f_2,f_3)\\ &-C_{2x}(f_1)C_{2x}(f_2)\delta(f_1+f_2)\\ &-C_{2x}(f_1)C_{2x}(f_3)\delta(f_1+f_3)\\ &-C_{2x}(f_2)C_{2x}(f_3)\delta(f_2+f_3)\end{aligned} \quad (26)$$

where C_{4x}^m stands for the trispectrum based on moments and C_{4x} for the trispectrum based on cumulants.

As one can notice, this relation reveals the presence of 3 planes of singularities, namely:

$$\begin{cases}f_1+f_2=0\\ f_1+f_3=0\\ f_2+f_3=0\end{cases} \quad (27)$$

Outside these planes, both trispectra are identical. Consequently, the estimation of the trispectrum based on cumulants not only requires to center the signal beforehand but also to subtract the last 3 terms of equation (26). Indeed, it is easy to show on numerical simulations that these last terms may be much greater than the first, i.e. may masked the trispectrum values based on moments C_{4x}^m.

One way to solve this problem has been suggested by [15] and explained in more details in [27]. As we know where are located the singularities, one may substitute the values around the zones of singularities by principle of analytical continuation using the values outside the zones of singularities.

4. detection of non-gaussian signals by third-order spectral analysis

In this section, we focus on the following problem: detecting a non-Gaussian stationary random signal in the presence of an additive Gaussian (or non-Gaussian) noise statistically independent from the signal, by means of 3rd-order spectral analysis. In order to do so, an application of some of the results of section 3 is first presented. A statistical gaussianity test is then described for the gaussian noise case in subsection 4.2 and for the non-gaussian noise case in 4.3. Finally, the probability of detection of the test as a function of the signal-to-noise ratio (SNR), the skewness of the signal and the smoothing parameter, is presented at the end of this section, by means of numerical simulations.

4.1. computation and asymptotic properties of consistent 3rd-order spectra estimators

The subsection 3.2 was dedicated to the computation and the asymptotic properties of k^{th}-order spectra estimators. Here is just presented an application of the main results of this subsection at the 3rd-order.

As a reminder, a discrete stationary random process $\{x(n); n=0,1,\ldots,N-1\}$ is considered. The periodogram method is applied: the $N=JL$ samples are divided into J non-overlapping segments of L samples each. The computation of a consistent estimate of the bispectrum with reduced variance, $\widehat{B}_x^{(N)}$, is then carried out using the following estimates (*c.f.* (18), (19) and (20))

$$\widehat{B}_x^{(j)}(f_1,f_2)=\frac{1}{L}X^{(j)}(f_1)X^{(j)^*}(f_2)X^{(j)*}(f_1+f_2)$$

$$\widehat{B}_x^{(av)}(f_1,f_2)=\frac{1}{L}\sum_{j=1}^{J}\widehat{B}_x^{(j)}(f_1,f_2)$$

$$\widehat{B}_x^{(N)}(f_1,f_2)=\frac{1}{M^2}\sum_{q_1=M/2}^{M/2-1}\sum_{q_2=M/2}^{M/2-1}\widehat{B}_x^{(av)}\left(f_1+\frac{q_1}{L},f_2+\frac{q_2}{L}\right) \quad (28)$$

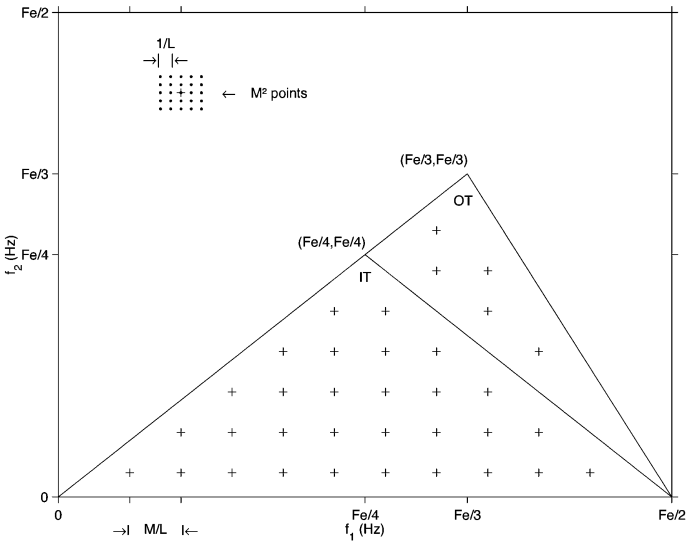


Figure 1. – Illustration of the smoothing bispectral procedure using $M=5$. Note that all the squares containing M^2 points covering both the IT and OT domains have to be neglected. All squares with less than M^2 points are also ignored.

Fig.1 illustrates the procedure to smooth the bispectral estimates $\widehat{B}_x^{(av)}$ in the discrete principal frequency domain PD_3. Note (*c.f.* subsection 6.2) that if the process $\{x(n)\}$ is aliased then PD_3 is equal to the principal domain $IT+OT$ of Fig. 1 (originally defined in [6]), $PD_3=(IT+OT)=1/12$. If not, the subset IT of the principal domain should be used only. In this case: $PD_3=(IT)=1/16$.

Using theorem 1, one can easily deduce the asymptotic behavior of $\widehat{B}_x^{(N)}$

$$\widehat{B}_x^{(N)}(f_1,f_2)\sim N_c\left\{B_x(f_1,f_2),\frac{1}{\Delta_{3N}^2N}S_x(f_1)S_x(f_1+f_2)\right\} \quad (29)$$

Finally, the bicoherency function is defined by

$$\hat{b}_x^{(N)}(f_1,f_2)=\frac{\widehat{B}_x^{(N)}(f_1,f_2)}{\left[\frac{1}{\Delta_{3N}^2N}S_x(f_1)S_x(f_2)S_x(f_1+f_2)\right]^{1/2}} \quad (30)$$

where $\Delta_{3N}=M/L;\Delta_{3N}>\frac{1}{\sqrt{N}}$[16]

From theorem 1, it follows that the $\hat{b}_x^{(N)}(f_1,f_2)$'s are asymptotically complex-normal with unit variance and independent one from another.

4.2. detecting a non-gaussian signal in an additive independent gaussian noise

The detection scheme defined in (16) for an additive independent gaussian noise becomes when applied at the 3rd-order

$$H_0:\forall(f_{i_1},f_{i_2})\in PD_3, B_x(f_{i_1},f_{i_2})\equiv 0$$
versus (31)
$$H_1:\exists(f_{i_1},f_{i_2})\in PD_3, B_x(f_{i_1},f_{i_2})\neq 0$$

Having decided on the null and alternative hypotheses, the next step is to determine a statistic, labeled $\widetilde{T}_g^2$, which will show up any departure from the null hypothesis.

Let us suppose that the bicoherency function is estimated at K different pairs of frequencies $\{(f_{i_1},f_{i_2}),i=1,2,...,K\}$, over PD_3. Note that K is approximately equal to PD_3/Δ_{3N}^2. Each estimated bicoherence $\hat{b}_x^{(N)}(f_{i_1},f_{i_2})$ is asymptotically complex normal, with unit variance, independent one from another (*cf.* section 4.1. and theorem 1). Therefore, the statistic $T_g^2=2\sum_{i=1}^{K}\left|\hat{b}_x^{(N)}(f_{i_1},f_{i_2})\right|^2$ suggested in [49, 17], is asymptotically a sum of $2K$ independent squared gaussian variables (K variables related to the real part of the $\hat{b}_x^{(N)}(f_{i_1},f_{i_2})$ and K variables related to the imaginary part of $\hat{b}_x^{(N)}(f_{i_1},f_{i_2})$). So, T_g^2 has asymptotically a central chi-square distribution with $2K$ degrees of freedom (χ_{2K}^2).

As one can notice in equation (30), the estimation of the bicoherency function at a given frequency pair (f_{i_1},f_{i_2}), requires the knowledge of the true spectrum $S_x(f)$ of the signal. As the true spectrum is usually unknown for real world data, one may substitute in (30) the true spectrum $S_x(f)$ by a consistent spectrum estimator $\hat{S}_x(f)$. Asymptotic properties of SOS were discussed for example in [44, 45]. As shown in [17], if the spectrum is estimated by averaging M adjacent periodogram ordinates, then the statistic T_g^2 is also asymptotically χ_{2K}^2

$$\widetilde{T}_g^2=2\sum_{i=1}^{K}\left|\hat{bic}_x^{(N)}(f_{i_1},f_{i_2})\right|^2\sim\chi_{2k}^2 \quad (32)$$

where $\hat{bic}_x^{(N)}(f_{i_1},f_{i_2})=$

$$\frac{\widehat{B}_x^{(N)}(f_{i_1},f_{i_2})}{\left[\frac{1}{\Delta_{3N}^2N}\hat{S}_x(f_{i_1})\hat{S}_x(f_{i_2})\hat{S}_x(f_{i_1}+f_{i_2})\right]^{1/2}}$$

Under the null hypothesis H_0, $B_x(f_{i_1}, f_{i_2}) \equiv 0$ over PD_3. Thus, each bicoherence estimate $\widehat{B}_x^{(N)}(f_{i_1}, f_{i_2})$ is (approximately) a complex standard normal variable (zero mean and unit variance). Therefore, $\widetilde{T}_g^2$ under H_0 is equal to the sum of $2K$ independent squared (approximately) standard normal variables. In other words, $\widetilde{T}_g^2$ is asymptotically $\chi^2_{2K}(0)$ distributed

$$\{\widetilde{T}_g^2 | H_0\} \sim \chi^2_{2K}(0) \tag{33}$$

Using this result, the gaussianity test may be applied as follows. Suppose that K is not large enough to approximate the χ^2_{2K} distribution of $\widetilde{T}_g^2$ by a normal distribution. Then, for a given level of significance α (a typical value of α is 0.05), the critical value T_α^2 of the test statistic $\widetilde{T}_g^2$ may be computed using: $\alpha = \int_{-\infty}^{T_\alpha^2} f(z)dz = \text{Prob}\left(\widetilde{T}_g^2 > T_\alpha^2\right)$, where $f(z)$ is the probability density of $\chi^2_{2K}(0)$.

If K is large enough (typically $K \geqslant 15$), as it is well known [20], the chi-square distribution $\chi^2_{2K}(0)$ can be approximated by a normal distribution with mean $2K$ and variance $4K$, which implies

$$\{\widetilde{T}_g^2 | H_0\} \sim N(2K, 4K), \text{ for large } K \tag{34}$$

Consequently, the variable $z = \dfrac{\widetilde{T}_g^2 - 2K}{\sqrt{4K}}$ is a standard normal random variable.

Practically, for large K, the gaussianity test may be applied as follows. First of all, one has to choose a level of significance α. Corresponding to this level of significance, the critical value T_α^2 of the test statistic $\{\widetilde{T}_g^2 | H_0\}$ is calculated using the following relationship

$$\alpha = 1 - \Phi\left(\frac{T_\alpha^2 - 2K}{\sqrt{4K}}\right), \text{ for large } K \tag{35}$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-u^2/2)du$, is the cumulative distribution function (CDF) of a standard normal variable.

If $\widetilde{T}_g^2 > T_\alpha^2$, then the null hypothesis is rejected at the level of confidence $(1-\alpha)$. In this case, the gaussian assumption must be rejected at the level of confidence $(1-\alpha)$. If, not the process may be non-gaussian, but the data are consistent with a zero bispectrum.

The following section is a variation of this test when the noise is non-gaussian.

4.3. detecting a non-gaussian signal in an additive independent non-gaussian noise

In the case of an additive and independent non-gaussian noise, the null and alternative hypotheses defined in (29) may be turned into the following ones

$$\begin{aligned} &H_0 : \forall (f_{i_1}, f_{i_2}) \in PD_3, B_x(f_{i_1}, f_{i_2}) = B_\nu(f_{i_1}, f_{i_2}) \\ &\textit{versus} \\ &H_1 : \forall (f_{i_1}, f_{i_2}) \in PD_3, B_x(f_{i_1}, f_{i_2}) \\ &\qquad = B_s(f_{i_1}, f_{i_2}) + B_\nu(f_{i_1}, f_{i_2}) \end{aligned} \tag{36}$$

with B_s and B_v the bispectrum of the signal and the noise respectively.

Using the same kind of argumentation as in the previous subsection, it is not difficult to show that the distribution of

$$T_{ng}^2 = 2\Delta_{3N}^2 N \sum_{i=1}^{K} \frac{\left|\widehat{B}_x^{(N)}(f_{i_1}, f_{i_2}) - B_\nu(f_{i_1}, f_{i_2})\right|^2}{\hat{S}_x(f_{i_1})\hat{S}_x(f_{i_2})\hat{S}_x(f_{i_1}+f_{i_2})} \tag{37}$$

is approximately χ^2_{2K}. In order to allow a practical implementation of the test, a further assumption is needed, namely: a noise-only sample of the data is available. In this case, we can replace $B_\nu(f_{i_1}, f_{i_2})$ by $\widehat{B}_\nu^{(N)}(f_{i_1}, f_{i_2})$, estimated from the noise-only part of the data. We can then define the statistic $\widetilde{T}_{ng}^2$ by

$$\widetilde{T}_{ng}^2 = 2\Delta_{3N}^2 N \sum_{i=1}^{K} \frac{\left|\widehat{B}_x^{(N)}(f_{i_1}, f_{i_2}) - \widehat{B}_\nu^{(N)}(f_{i_1}, f_{i_2})\right|^2}{\hat{S}_x(f_{i_1})\hat{S}_x(f_{i_2})\hat{S}_x(f_{i_1}+f_{i_2})} \sim \chi^2_{2K} \tag{38}$$

Like in subsection 4.2, it can be shown that under $H_0, \widetilde{T}_{ng}^2$ is also asymptotically $\chi^2_{2K}(0)$ distributed. Consequently, for a given level of significance α, the corresponding critical value $\widetilde{T}_\alpha^2$ of the test statistic $\{\widetilde{T}_{ng}^2 | H_0\}$ can be computed as in the gaussian noise case (because T_α^2 depends only on the asymptotic distribution of the statistic).

If $\widetilde{T}_{ng}^2 \geqslant T_\alpha^2$, then the gaussian assumption of the signal must be rejected at the level of confidence $(1-\alpha)$. If, not the signal may be non-gaussian, but the data are consistent with the bispectrum noise level.

The next section is dedicated to the effectiveness of both tests as a function of the averaged SNR, the averaged skewness of the signal and the smoothing parameter M.

4.4. probability of detection of the gaussianity test

As one can notice, the statistic $\widetilde{T}^2_{ng}$, defined in equation (38), is equal to the statistic $\widetilde{T}^2_g$ (equation (32)) when the noise is gaussian. Consequently, by using the statistic $\widetilde{T}^2_{ng}$ we cover at once the non-gaussian noise case and the gaussian noise case. Hereinafter, the statistic $\widetilde{T}^2_{ng}$ will be used only.

In order to characterize the effectiveness of a statistical test, let calculate the probability of detection (P_D) and the probability of false alarm (P_F) of the test. In our case, they can be defined as follows

$$\begin{aligned} P_F &= \mathrm{Prob}\{\widetilde{T}^2_{ng} > T^2_\alpha | H_0\} \text{ (probability of false alarm)} \\ P_D &= \mathrm{Prob}\{\widetilde{T}^2_{ng} > T^2_\alpha | H_1\} \text{ (probability of detection)} \end{aligned} \tag{39}$$

with T^2_α the confidence limit defined in the subsection 4.2.

Under the alternative hypothesis H_1, the random variable $\dfrac{\widehat{B}^{(N)}_x(f_{i_1}, f_{i_2}) - \widehat{B}^{(N)}_\nu(f_{i_1}, f_{i_2})}{\left[\dfrac{1}{\Delta^2_{3N} N} \hat{S}_x(f_{i_1})\hat{S}_x(f_{i_2})\hat{S}_x(f_{i_1}+f_{i_2})\right]^{1/2}}$ is (approximately) normally distributed with mean

$B_s(f_{i_1}, f_{i_2}) / \left\{\dfrac{1}{\Delta^2_{3N} N} S_x(f_{i_1})S_x(f_{i_2})S_x(f_{i_1}+f_{i_2})\right\}^{1/2}$ and unit variance. If we now define the parameter λ by

$$\lambda = 2\Delta^2_{3N} N \sum_{i=1}^{K} \frac{|B_s(f_{i_1}, f_{i_2})|^2}{S_x(f_{i_1})S_x(f_{i_2})S_x(f_{i_1}+f_{i_2})} \tag{40}$$

it can be shown that the statistic $\widetilde{T}^2_{ng}$ is asymptotically distributed as follows

$$\{\widetilde{T}^2_{ng} | H_1\} \sim \chi^2_{2K}(0) + N(\lambda, 4\lambda) \tag{41}$$

and that these two added distributions are related to two independent random variables. Note that it is equivalent to say that $\{\widetilde{T}^2_{ng} | H_1\}$ is asymptotically distributed as a chi-square with $2K$ degrees of freedom and noncentrality parameter λ.

Therefore, for large K, by using the equation (34)

$$\{\widetilde{T}^2_{ng} | H_1\} \sim N(2K + \lambda, 4K + 4\lambda), \text{ for large } K \tag{42}$$

Consequently, under H_1, the random variable $y = \dfrac{\widetilde{T}^2_{ng} - 2K - \lambda}{2\sqrt{K+\lambda}}$ is a standard normal random variable. Thus, the probability P_D that $\{\widetilde{T}^2_{ng} | H_1\}$ will exceed the threshold T^2_α is equal to the probability that the standard normal random variable y will exceed the threshold $\left(\dfrac{T^2_\alpha - 2K - \lambda}{2\sqrt{K+\lambda}}\right)$

$$P_D = 1 - \Phi\left(\frac{T^2_\alpha - 2K - \lambda}{2\sqrt{K+\lambda}}\right) \tag{43}$$

with Φ the function defined in (35).

Let now define the following two quantities

$$\Gamma_s = 2\Delta^2_{3N} N \sum_{i=1}^{K} \frac{\left|\widehat{B}_s(f_{i_1}, f_{i_2})\right|^2}{\hat{S}_s(f_{i_1})\hat{S}_s(f_{i_2})\hat{S}_s(f_{i_1}+f_{i_2})}$$ (averaged skewness of the signal)

$$\rho = \sigma^2_s / \sigma^2_v \text{ (averaged SNR)}$$

In order to show the behaviour of P_D versus Γ_s, ρ and the smoothing parameter M, the following numerical simulations have been performed.

The non-gaussian signal has a log-normal distribution whose probability density function $f(y)$ is of the form

$$f(y) = \frac{1}{\sigma y \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right\} \tag{44}$$

Choosing different values of σ, while keeping μ constant, allows to get different values of Γ_s.

In the simulations presented here, the noise is gaussian. The reason why is that comparable results are obtained for a gaussian or a non-gaussian noise. The other parameters of the simulations are: $N = 10\,000$, $J = 10$, $L = 1\,000$ (which implies $K = 600$) and the level of significance α is set to 10^{-3}.

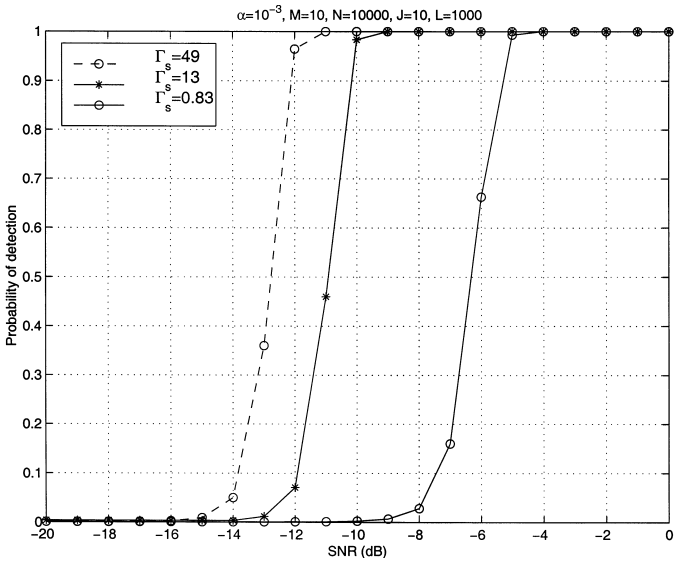


Figure 2. – Probability of detection of the gaussianity test as a function of a function of SNR, for a simulated non-gaussian signal distributed as a log-normal corrupted by a gaussian noise, in the case $M = 10$.

In Figure 2 are presented 3 curves of the probability of detection P_D versus the averaged SNR$_\rho$, for 3 different values of Γ_s. As one can see, the higher Γ_s, the better the detection. The gaussianity test, for this set of parameters, remains valid up to $-12\,dB$.

In the case of Fig. 2, $M = 10$. If we now increase M up to 16 (Fig. 3), then one can observe a $1dB$ improvement in the detection, with comparable levels of Γ_s.

As a conclusion, the bispectral detection scheme is (asymptotically) optimal for testing, whether the received signal has a nonzero bispectrum or not. However, it is only suboptimal for signal detection since the signal might carry additional detection information in its higher (greater than 3) spectra.

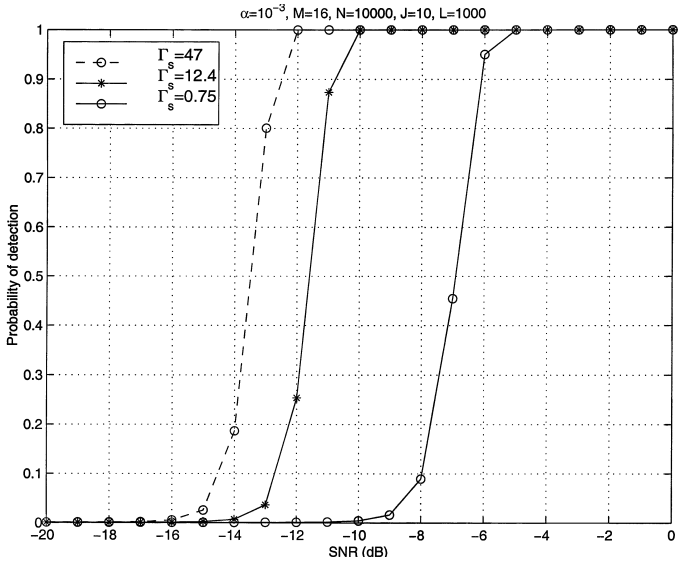


Figure 3. – Probability of detection of the gaussianity test as a function of a function of SNR, for a simulated non-gaussian signal distributed as a log-normal corrupted by a gaussian noise, in the case $M = 16$.

5. testing for linearity of stationary time series

5.1. squared bicoherence property of a linear process

Let us assume that the time serie $\{x(n); n = 0, 1...N-1\}$ is generated by a finite parameter linear stationary random process (*e.g.* an autoregressive _AR_ process). In this case, $\{x(n)\}$ is linear and its n^{th} element is of the form

$$x(n) = \sum_{\tau=0}^{\infty} h(\tau)\varepsilon(n-\tau) \qquad (45)$$

where the $\varepsilon(n)$ are independent identically distributed random innovations with $E[\varepsilon(n)] = 0$. The stationary process $\{\varepsilon(n)\}$ may be seen as the input of a time invariant linear filter whose impulse response is $\{h(n); n = 0, 1, \ldots\}$. If we now further assume that $\sum_{n=0}^{\infty} h^2(n) \langle \infty$, then the covariance function of the stationary output process $\{x(n)\}$ is finite. Therefore, if the input is Gaussian, then the output is Gaussian and its covariance function completely determines the joint distributions of the process. But suppose that the input is not gaussian and that $\gamma_{3\varepsilon} = E[\varepsilon^3(n)] \neq 0$, then the third order cumulant of the output $\{x(n)\}$ is non null for many values of τ_1 and τ_2

$$c_{3x}(\tau_1, \tau_2) = E\{x(n)x(n+\tau_1)x(n+\tau_2)\} \neq 0\,,$$

respectively for many frequency pairs (f_1, f_2) inside PD_3 (PD_3 is defined in subsection 4.1)

$$B_x(f_1, f_2) = \gamma_{3\varepsilon} H(f_1)H(f_2)H(f_1+f_2) \neq 0 \qquad (46)$$

where $H(f) = \sum_{n=0}^{\infty} h(n)\exp(-j2\pi fn)$ may be seen as the filter transfer function.

Note that it is also true if $\{x(n)\}$ is generated *via* a nonlinear filtering operation satisfying a Volterra functional expansion [7].

As $\{x(n)\}$ is a linear stationary random process, its spectrum is given by $S_x(f) = |H(f)|^2\sigma_\varepsilon^2$, where $\sigma_\varepsilon^2 = E[\varepsilon^2(n)]$. Thus from (46) and (14) it follows that

$$b_x(f_1, f_2) = \frac{\gamma_{3\varepsilon} H(f_1)H(f_2)H^*(f_1+f_2)}{(\sigma_\varepsilon^2)^{3/2}|H(f_1)||H(f_2)||H^*(f_1+f_2))|} \qquad (47)$$

Consequently, its squared bicoherence, also called skewness function, is constant over PD_3

$$|b_x(f_1, f_2)|^2 = \frac{\gamma_{3\varepsilon}^2}{\sigma_\varepsilon^6} \quad \text{over } PD_3 \qquad (48)$$

In order to illustrate this point, we performed the following numerical simulations. The process $\{x(n)\}$ is generated using

$$x(n) = \varepsilon(n) + 0.9\,\varepsilon(n-1) + 0.385\,\varepsilon(n-2) - 0.771\varepsilon(n-3)$$

where $\{\varepsilon(n)\}$ is distributed as a log-normal (*c.f.* equation (44)) with $\sigma = 1.6$.

In this case, $\{x(n)\}$ is a linear non gaussian process. In Fig. 4, the skewness function of $\{x(n)\}$ is displayed using $N = 7680$, $J = 30, L = 256$ and $M = 1$. As one can observe the skewness function is nearly constant with mean 0.21 and standard deviation 0.05.

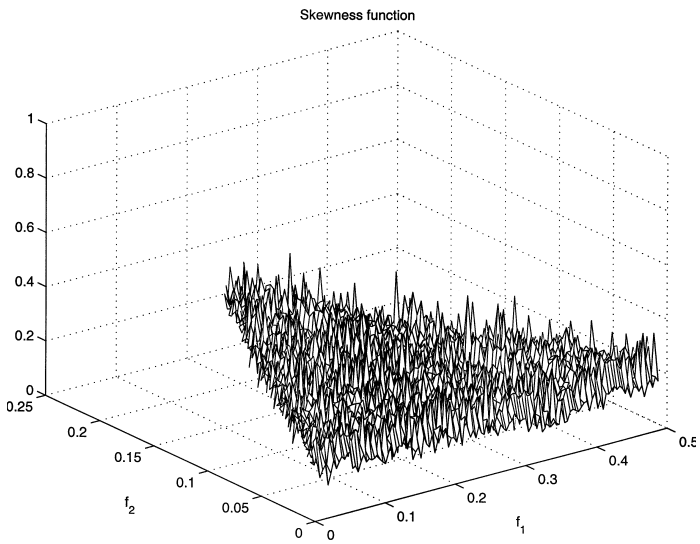


Figure 4. – Skewness function of a simulated linear non-gaussian (d.d.p. log-normal, with sigma=1,6) signal. The mean and the standard deviation of the skewness function values are equal to 0.21 and 0.05 respectively.

If we now use exactly the same signal $\{\varepsilon(n)\}$ and the same set of parameters (N, J, L, M) as in Fig. 4 but turn $\{x(n)\}$ non-linear using the following relation:

$$x(n) = \varepsilon(n) * \varepsilon(n-3).$$

then the process $\{x(n)\}$ is both non-gaussian and non-linear. The resulting skewness function, shown in Fig. 5, is far less constant. Indeed, the mean value (0.24) of the skewness function is now of the order of its standard deviation (0.22).

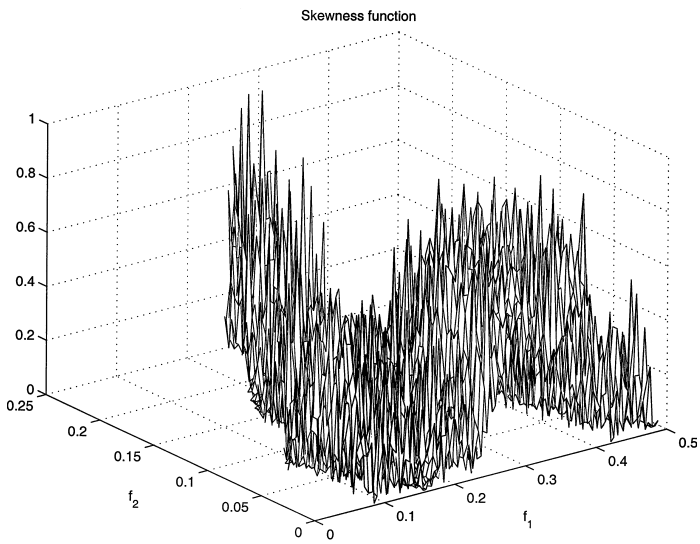


Figure 5. – Skewness function of a simulated non-linear non-gaussian (d.d.p. log-normal, with sigma=1,6) signal. The mean and the standard deviation of the skewness function values are equal to 0,24 and 0,22 respectively.

5.2. linearity test

To the author's best knowledge, two linearity tests, based on the constancy of the skewness function when the process is linear, have been developed so far, one in [49], the other in [17]. The following test is the one developed in [17], using $c = 1/2$ (c is defined in [17] section 3).
As we have seen in subsection 5.1, before testing if a signal is linear or not, one has to be confident in the fact that the signal is non-gaussian (*e.g.* by using the gaussianity test of section 4). Indeed, if the signal is gaussian then its squared bicoherence is zero and consequently there is no need to test if the squared bicoherence is a non-zero constant... So let us assume that the set of measurements $\{x(n)\}$ is a non-gaussian stationary random process. Using the additional assumptions of the previous subsection, the skewness function constancy of a linear process can be tested using the following set of hypotheses

$$H_0 : |b_x(f_1, f_2)|^2 = \frac{\gamma_{3\varepsilon}^2}{\sigma_\varepsilon^6} \quad \text{over } PD_3$$

versus (49)

$$H_1 : |b_x(f_1, f_2)|^2 = \frac{\gamma_{3\varepsilon}^2}{\sigma_\varepsilon^6} \quad \text{over } PD_3$$

In practice, even if $\{x(n)\}$ is linear (null hypothesis), its estimated skewness function will not be exactly flat. Nevertheless, its flatness may be characterized by averaging the K skewness function values estimated over the frequency domain PD_3. This average is equal to $\hat{\lambda}_0/2$, where $\hat{\lambda}_0$ is given by

$$\hat{\lambda}_0 = (2/K) \sum_{i=1}^{K} |\hat{bic}_x^{(N)}(f_{i_1}, f_{i_2})|^2 \tag{50}$$

where $\hat{bic}_x^{(N)}(f_{i_1}, f_{i_2})$ is defined by the equation (32).
Using the results of subsection 4.2., we know that each estimate $\hat{bic}_x^{(N)}(f_{i_1}, f_{i_2})$ is normally distributed. Therefore, under the null hypothesis, the random variable $2\left|\hat{b}_x^{(N)}(f_{i_1}, f_{i_2})\right|^2$ is asympto-tically distributed as a chi-square with 2 degrees of freedom and non-centrality parameter $\hat{\lambda}_0$

$$2\left|\hat{b}_x^{(N)}(f_{i_1}, f_{i_2})\right|^2 \sim \chi_2^2(\hat{\lambda}_0) \tag{51}$$

The theoretical interquartile range of a $\chi_2^2(\hat{\lambda}_0)$ density, denoted $f(\chi^2)$, is $\xi_3 - \xi_1$, where ξ_1 and ξ_3 are respectively the first and third quartile of $f(\chi^2)$. ξ_1 and ξ_3 can be computed using

$$\text{Prob}(\chi_2^2(\hat{\lambda}_0) < \xi_1) = \text{Prob}(\chi_2^2(\hat{\lambda}_0) > \xi_3) = 1/4 \tag{52}$$

If we now define the following statistic:

$$L = (2/K) \sum_{i=1}^{K} \left| \hat{b}_x^{(N)}(f_{i_1}, f_{i_2}) \right|^2 \qquad (53)$$

Thus, the interquartile range of the statistic L, dénoted $R = \xi_3 - \xi_1$, can be calculated using :

$$\text{Prob}\left(\chi^2_{2K}(\hat{\lambda}_0) < \xi_1\right) = \text{Prob}(\chi^2_{2K}(\hat{\lambda}_0) > \xi_3\Big) = \frac{1}{4} \qquad (54)$$

The linearity test can then be turned into the following question: is the interquartile range of the statistic L significantly larger or smaller than the theoretical interquartile range of $\chi^2_2(\hat{\lambda}_0)$? If yes, then the null hypothesis of linearity must be rejected. In mathematical terms this can be described by turning the hypothesis test (49) into the following one

$$\begin{array}{ll} H_0 : & R = \xi_3 - \xi_1 \\ \textit{versus} & \\ H_1 : & R \neq \xi_3 - \xi_1 \end{array} \qquad (55)$$

This kind of test is sometimes called a two-tailed test. A practical implementation of this test will now be presented for large K.

When K is large, it is shown in [16] that

$$R \sim N(\xi_3 - \xi_1, \sigma_0^2) \quad \text{for large } K \qquad (56)$$

where

$$\sigma_0^2 = (16K)^{-1} \left[3f^{-2}(\xi_1) - 2f^{-1}(\xi_1)f^{-1}(\xi_3) + 3f^{-2}(\xi_3)\right]$$

Consequently, for large $K, z = \dfrac{R - (\xi_3 - \xi_1)}{\sigma_0}$ is a standard normal variable. For a given level of significance α, the critical value $z_{\alpha/2}$ can then be calculated using the following relation:

$$\frac{\alpha}{2} = 1 - \Phi(z_{\alpha/2}) \qquad (57)$$

with Φ the function defined in (35).

Let denote by R_0 the observed value of R and by $z_0 = \dfrac{R_0 - (\xi_3 - \xi_1)}{\sigma_0}$ the observed value of z.

If $|z_0| > z_{\alpha/2}$, the null assumption of linearity is rejected at the level of confidence $(1 - \alpha)$. If not, the process may be non-linear but the interquartile range of the statistic L is not significantly different from the theoretical interquartile range of a $\chi^2_2(\hat{\lambda}_0)$.

6. detecting a transient signal by bispectral analysis

This section describes a method for detecting an unknown deterministic transient signal in broadband noise. The overall idea of this method is the following. A sliding time window scans the data at regular time step. At each time step, a test is preformed. This test is derived from a fundamental property of the principal domain of the bispectrum. This property is presented in the second part of this section. The test itself is described in the subsection 6.3 and its probability of detection in 6.4. The following part is dedicated to the assumptions of the test.

6.1. assumptions of the test

First of all, it is assumed that the observed process $\{x(n)\}$ is first passed through a low-pass filter whose cutoff frequency is f_0. Then, $\{x(n)\}$ is sampled at the rate $F_s = 2f_0$. In other words, the observed process is not aliased.

Secondly, we suppose that the deterministic transient signal $\{s(n)\}$ is entirely present in the sliding time window, of duration T, when the null hypothesis of noise alone is not true. In other words, we neglect the problem of determining the time of arrival of the transient signal. All that is known about $\{s(n)\}$ is that its frequency band lies in the interval $[0, f_0]$. Consequently, matched filtering detection [14] is not appropriate for this problem since the signal waveform is unknown. Moreover, the mean of the transient is assumed to be nearly zero, which is the case in most applications. The "variance" of the transient signal is denoted by σ_s^2.

Thirdly, the noise, $\{\nu(n)\}$, is assumed to be a six$^{\text{th}}$ order stationary zero mean random process with bounded moment functions of all order. The stationarity assumption of the noise and a generalised finite memory condition are required for the asymptotic sampling distribution of the spectral and the bispectral estimators to be gaussian (for more details see [6]). Note that the noise is not assumed to be gaussian. This assumption is not needed.

Moreover, we suppose that the noise can be observed for a relatively long period when there is no signal. Consequently, standard spectral estimates will yield to consistent and asymptotically unbiased estimates of the noise spectrum.

Finally, it is assumed that the noise can be linearly pre-whitened with great precision, so the problem can be simplified by assuming the noise to be white. We denote its standard deviation by σ_ν, a known parameter.

The test described in 6.3 is based on a property of the bispectrum principal domain of a discrete time band-limited stationary process. So, let first present this property.

6.2. bispectrum principal domain of a discrete time band-limited stationary process

A usual mistake is made upon the principal domain of the bispectrum. Indeed, outside the statistics community, many researchers and engineers have been using bispectral techniques but generally by estimating the bispectrum over the IT domain (*c.f.* Fig 1). As it will be now presented, this IT domain is only a subset of the bispectrum principal domain of a discrete time band-limited stationary process.

Suppose a real stationary zero mean continuous-time random process $\{x(t)\}$. Let also assume that all expected values, sums and integrals used in what follows exist.

As we have seen in section 2, the third-order cumulant of $\{x(t)\}$ is defined by $c_{3x}(\tau_1,\tau_2) = E[x(t)x(t+\tau_1)x(t+\tau_2)]$. c_{3x} is independent of t because $\{x(t)\}$ is stationary. The bispectrum $B_x(f_1,f_2)$ of $\{x(t)\}$ is defined as the Fourier transform of the third-order cumulant

$$B_x(f_1,f_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} c_{3x}(\tau_1,\tau_2)e^{-i2\pi(f_1\tau_1+f_2\tau_2)}d\tau_1 d\tau_2 \tag{58}$$

The relationship between $B_x(f_1,f_2)$ and $E[X(f_1)X(f_2)X(f_3)]$ mentioned in equation (12) may be detailed as follows when $\{x(t)\}$ is stationary

$$\begin{aligned} &E[X(f_1)X(f_2)X(f_3)] \\ &= E\left[\iiint x(t)x(r)x(s)e^{-i2\pi(f_1t+f_2r+f_3s)}dtdrds\right] \\ &= \iint E[x(t)x(t+\tau_1)x(t+\tau_2)])e^{-i2\pi(f_1\tau_1+f_1\tau_2)}d\tau_1 d\tau_2 \\ &\cdot \int e^{-i2\pi(f_1+f_2+f_3)t}dt \\ &= B_x(f_1,f_2)\delta(f_1+f_2+f_3) \end{aligned} \tag{59}$$

So, $B_x(f_1,f_2) = E[X(f_1)X(f_2)X(f_3)]$ only if $f_3 = -f_1 - f_2$. Since $E[X(f_1)X(f_2)X(f_3)]$ is symmetric about the lines $f_1 = f_2, f_1 = f_3$ and $f_2 = f_3$ (and the conjugate symmetry lines), a non-redundant domain (or principal domain) of the continuous time bispectrum lies in the following C cone of the (f_1,f_2) plane

$$C = \{f_1,f_2 : 0 \leqslant f_1 \text{ and } f_2 \leqslant f_1\} \tag{60}$$

Now, suppose that $\{x(n)\}$ is sampled at the rate $F_s = 2f_0$. The sampling interval is $T_s = 1/F_s$. The bispectrum of the sampled process $\{x(n)\}$ is defined analogous with equation (59) by

$$B_x(f_1,f_2) = \sum_{j=-\infty}^{+\infty}\sum_{k=-\infty}^{+\infty} c_{3x}(jT_s,kT_s)e^{-i2\pi(f_1jT_s+f_2T_s)}e^{-i2\pi(f_1+f_2+f_3)T_s} \tag{61}$$

where now $f_1+f_2+f_3$ is not only constrained to be 0 but can be equal to n/T_s for any signed integer n. Consequently, the sampling implies that the cone C is first cut by the symmetry line $2f_1+f_2 = F_s$. Thus, the bispectrum principal domain of a stationary process sampled at the frequency F_s is given by (*c.f.* Fig 1)

$$IT + OT = \{f_1,f_2 : f_2 \leqslant f_1 \text{ and } 0 \leqslant f_1+f_2 \leqslant (F_s - f_1)\} \tag{62}$$

Moreover, because the process $\{x(n)\}$ is not aliased, its Fourier transform $X(f)$ is such that $X(f) = 0$ if $f > f_0$. Consequently, $B_x(f_1,f_2) = 0$, if $f_1+f_2 > 0$ or if $f_1 > 0$. The principal domain $IT + OT$ can then be restricted to the subset IT (*c.f.* Fig 1) defined by

$$IT = \{f_1,f_2 : 0 \leqslant f_1 \leqslant f_0, f_2 \leqslant f_1 \text{ and } 0 \leqslant f_1+f_2 \leqslant f_0\} \tag{63}$$

When the process is not stationary, the principal domain can be considered as that of an aliased stationary process. In this case, the bispectrum will usually be nonzero over the whole principal domain $IT + OT$. An extreme case is the one of the Dirac function. Its bispectrum has a constant nonzero value over the whole $IT + OT$ domain. Now, as shown by Lacoume *et al.* (1997), the bicoherence of the $\text{sinc}(x)$ function allows to point out the importance of the signal bandwidth. For a bandwidth greater than $F_s/3$ nonzero values are seen out of the IT domain : *i.e.* in the OT domain. They are due to overlapping. For a bandwidth lower than $F_s/3$ non-zero values are seen in the IT domain only. These observations suggest to look for non-zero values in the OT domain to identify wide – band non-stationary signals. From what we know, one may consider the transient processes observed in space as wide – band non-stationary phenomena.

6.3. detecting a transient signal by bispectral analysis

The null and the alternative hypotheses of this test are

$$\begin{aligned} &H_0 : x(n) = \nu(n) && \text{(noise alone)} \\ &\textit{versus} \\ &H_1 : x(n) = s(n) + \nu(n) && \text{(signal plus noise)} \end{aligned} \tag{64}$$

with $n = 0,1,\ldots,N-1$.

Since $\{\nu(n)\}$ is assumed stationary (*c.f.* subsection 6.1), the observed process $\{x(n)\}$ is stationary under H_0. Additionally, the sampled process $\{x(n)\}$ is not aliased. Consequently, as shown in 6.2, its bispectrum $B_x(f_{i_1}, f_{i_2})$ is null for any frequency couple $(f_{i_1}, f_{i_2}) \in OT$, even if the noise is nongaussian. Under H_1, the deterministic transient signal is present which makes the observed process nonstationary. Since the stationary noise is additive, $B_x(f_{i_1}, f_{i_2}) = B_s(f_{i_1}, f_{i_2})$ under H_1, for any $(f_{i_1}, f_{i_2}) \in OT$. So, the hypothesis test (64) becomes

$$\begin{aligned} &H_0 : \forall (f_{i_1}, f_{i_2}) \in OT,\ B_x(f_{i_1}, f_{i_2}) = 0\,, \\ &\textit{versus} \\ &H_1 : \forall (f_{i_1}, f_{i_2}) \in OT,\ B_x(f_{i_1}, f_{i_2}) = B_s(f_{i_1}, f_{i_2}) \end{aligned} \tag{65}$$

In order to test this set of hypotheses, the statistic defined in equation (32) will be used. Since the noise is assumed to be white, this statistic becomes under H_0

$$\widetilde{T}_c^2 = 2\Delta_{3N}^2 N \sum_{i=1}^{K} \frac{|\widehat{B}_x^{(N)}(f_{i_1}, f_{i_2})|^2}{\sigma_\nu^6} \tag{66}$$

Note that the estimation of $\widehat{B}_x^{(N)}(f_{i_1}, f_{i_2})$ is performed as in subsection 4.1., but as the signal is deterministic, only one record is used: $J = 1$ which implies $N = L$.

If H_0 is true, then $\widetilde{T}_c^2$ is approximately distributed as a central chi-square with $2K$ degrees of freedom $\chi_{2K}^2(0)$, where K is the number of frequency pairs in OT. According to our studies, $K = PD_3(OT)/\Delta_{3N}^2$, where $PD_3(OT) = 1/48$, $PD_3(IT) = 1/16$, $PD_3(IT + OT) = 1/12$.

For a selected level of significance α, the critical value T_α^2 of the test statistic $\widetilde{T}_\alpha^2$ can be calculated as in subsection 4.2.

If $\widetilde{T}_c^2 > T_\alpha^2$, then the presence of a deterministic transient signal is accepted at the level of confidence $(1 - \alpha)$.

In the case of a stationary gaussian noise, $\forall (f_{i_1}, f_{i_2}) \in (IT + OT), B_v(f_{i_1}, f_{i_2}) = 0$. Thus, the test statistic (66) can also be used in the IT triangle to detect a transient signal. Note that the OT test does not require the assumption of Gaussian noise, whereas the IT test is invalid if the additive noise has a significantly nonzero skewness function.

Finally, if the skewness function of the signal has not a significant value, the kurtosis function of the signal $T_s(f_1, f_2, f_3)$ has to be investigated. For detector schemes using fourth-order statistics, see [9, 13].

6.4. probability of detection of the test

If the signal is present (H_1 is true), the test's power function $\mathrm{Prob}(T_c^2 \rangle T_\alpha^2 | H_1)$ depends upon the magnitude of $B_x^{(N)}(f_{i_1}, f_{i_2})$ for each pair of frequencies in the OT triangle. Under H_1, the statistic $\widetilde{T}_c^2$ has a non-central $\chi_{2K}^2(\lambda)$ distribution with non-centrality parameter λ

$$\lambda = 2\left(\frac{JM}{N}\right)^2 N \sum_{i=1}^{K} \frac{|B_s(f_{i_1}, f_{i_2})|^2}{\sigma_\nu^6} \tag{67}$$

Keeping in mind that the bispectrum of the noise is zero in OT, the following approximation holds

$$\frac{|\widehat{B}_x^{(N)}(f_{i_1}, f_{i_2})|^2}{\sigma_\nu^6} \sim \frac{|B_s(f_{i_1}, f_{i_2})|^2}{\sigma_\nu^6} \sim \left(\frac{\sigma_s^2}{\sigma_\nu^2}\right)^3 \sim \rho^3 \tag{68}$$

and

$$\rho = \frac{\sigma_s^2}{\sigma_\nu^2} \tag{69}$$

is the signal to noise ratio. So the skewness function of the signal $B_s(f_1, f_2)$ plays again an important role on the signal detectability.

In Figure 7, is displayed a plot of the probability of detection as a function of SNR$= \sigma_s^2/\sigma_\nu^2$, for the following numerical simulation. The transient signal is a chirp signal whose frequency band lies in the interval $[0, 500\ \mathrm{Hz}]$. More precisely, the chirp signal is generated using a linear swept-frequency cosine generator. The noise is gaussian since the algorithm performs equally well when the noise is gaussian or not the other parameters of the simulation are: $J = 1, N = L = 2500, M = 50$ and the level of significance $\alpha = 10^{-3}$. As one can notice, the power of test P_D allows a detection of the transient signal up to $-8dB$.

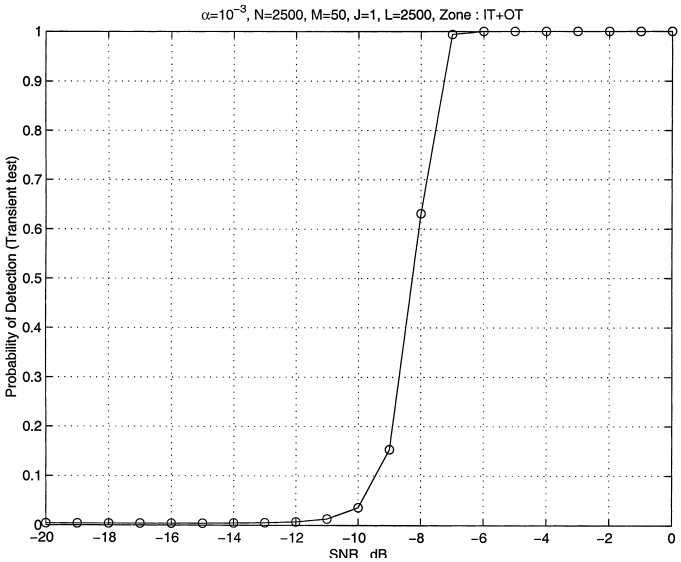


Figure 6. – Probability of detection of the transient test as a function of SNR, for a simulated chirp transient signal corrupted bu a gaussian noise.

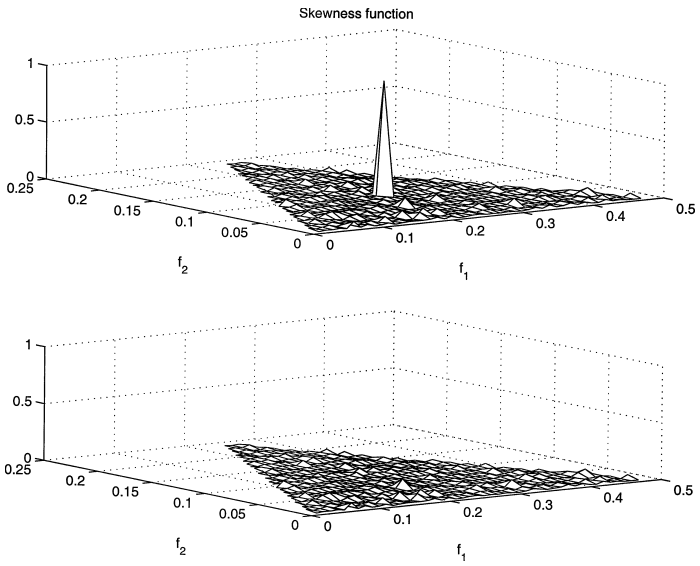


Figure 7. – Top panel: skewness function of a simulated signal composed of a gaussian noise together with 3 waves with frequencies (0.25, 0.10, 0.35) quadratically phased coupled; bottom panel: skewness function of a signal composed of a gaussian noise together with 3 uncorrelated Fourier components.

7. applications of the HOS technique to nonlinear wave-wave interactions

A data set of any fluctuating physical quantity (*e.g.* plasma density fluctuations) may be regarded, to a first approximation, as a superposition of statistically uncorrelated waves. Consequently, it can be described by its power spectrum which shows the frequency distribution of the power of the fluctuations.

However, linear spectral analysis techniques are of limited interest when various spectral components interact one with another due to some nonlinear or parametric processes. In the case of a non-linear wave-wave interaction, the interaction between two (three) waves gives rise to phase coupled new waves. More precisely, three (four) waves with frequencies f_k and phases $\theta_k, k = 1, 2, 3 (k = 1, 2, 3, 4)$, are said to be quadratically (cubically) phase coupled if $f_3 = f_1 \pm f_2, \theta_3 = \theta_1 \pm \theta_2$ ($f_4 = f_1 + f_2 \pm f_3, \theta_4 = \theta_1 + \theta_2 \pm \theta_3$). As it will be now presented, the detection of such phase coherence may be carried out using the higher-order spectra technique [23, 24, 26, 39].

Depending on the bispectrum (trispectrum) estimation method employed, the conventional techniques for the detection and the quantification of quadratic (cubic) phase coupling are divided into three categories [35]

1. the direct class of techniques,
2. the indirect class of techniques,
3. the complex demodulates class of techniques.

We use the first one. This direct class of techniques is based on the use of consistent HOS estimators of the bicoherency $\hat{b}_x^{(N)}(f_1, f_2)$ and the tricoherency $\hat{t}_x^{(N)}(f_1, f_2, f_3)$, defined by

$$\hat{b}_x^{(N)}(f_1, f_2) = \frac{\widehat{B}_x^{(N)}(f_1, f_2)}{\left[\frac{1}{\Delta_{3N}^2 N}\hat{S}_x(f_1)\hat{S}_x(f_2)\hat{S}_x(f_1+f_2)\right]^{1/2}} \quad (70)$$

$$\hat{t}_x^{(N)}(f_1, f_2, f_3) = \frac{\widehat{T}_x^{(N)}(f_1, f_2, f_3)}{\left[\frac{1}{\Delta_{4N}^3 N}\hat{S}_x(f_1)\hat{S}_x(f_2)\hat{S}_x(f_3)\hat{S}_x(f_1+f_2+f_3)\right]^{1/2}} \quad (71)$$

A detailed presentation of the bicoherency estimation has been given in section 4. Note that the smoothing operation (*c.f.* subsection 4.1.) of the HOS estimators is inadequate when studying non-linear wave-wave interactions (which implies $M = 1$). For the tricoherency estimator, see [26].

The definition of the bispectrum (trispectrum), given in equation (12), explains why HOS can quantify the phase degree coupling of three (four) waves. Indeed, if the waves at frequencies f_1, f_2 and $(f_1, f_2, f_3$ and $f_1 + f_2 + f_3)$ are excited independently, the phase of each wave will be randomly distributed over $[-\pi, \pi]$. Thus, when a statistical averaging is carried out (due to the expectation operator in (12)), the amplitude of the bicoherence (tricoherence) for this set of frequencies will tend to zero, due to the random phase mixing effect. Now, if the three (four) spectral components are nonlinearly phase coupled, then the sum of the phases of the waves will be no longer random (although the phase of each wave is randomly changing for each realization). Consequently, the statistical averaging will not lead to a null value of the amplitude of the bicoherence (tricoherence).

In order to illustrate this point, the following numerical simulations have been performed

$$x_1(n) = \cos(2\pi f_1 n + \theta_1) + \cos(2\pi f_2 n + \theta_2) + \cos(2\pi f_3 n + \theta_3) + \nu(n) \quad (72)$$

and

$$x_2(n) = \cos(2\pi f_1 n + \theta_1) + \cos(2\pi f_2 n + \theta_2) + \cos(2\pi f_3 n + (\theta_1 + \theta_2)) + \nu(n) \quad (73)$$

where $f_1 = 0.25$, $f_2 = 0.13$, $f_3 = 0.38$; $n = 1, 2, ..., N$ with $N = 10\,100$.

The signals x_1 and x_2 are composed of $J = 101$ records. Each record contains $L = 100$ data points. The phases θ_1, θ_2 and θ_3 are generated using a set of random numbers uniformly distributed over $[-\pi, \pi]$. Small level of stationary white gaussian noise $\{\nu(n)\}$ was added to each record.

In the case of the signal x_1, the spectral component at frequency f_3 is an independent spectral component because θ_3 is an independent random phase variable. In the case of the signal x_2, the 3 Fourier components are quadratically phase coupled.

$\{x_1(n)\}$ and $\{x_2(n)\}$ have obviously identical power spectra consisting of impulses at f_1, f_2 and $f_3 = f_1 + f_2$. Therefore power spectral analysis is phase-blind.

The skewness function of $\{x_1(n)\}$ is identically zero over PD_3 due to the random phase mixing effect.

The skewness function of $\{x_2(n)\}$ shows an impulse in PD_3 located at (f_1, f_2) (*c.f.* Fig 7). Thus, bispectral analysis does quantify quadratic phase coupling. This technique is also valid to quantify cubic phase coupling [26] and in general to study nonlinear wave-wave interactions.

8. applications to satellite data

As examples of applications, we will consider here time series on which the results of a bispectral analysis have already been published although the hypotheses were not tested.

8.1. AMPTE-UKS

The first time series we consider is the modulus of the magnetic field, measured by the magnetometer of the AMPTE-UKS satellite, upstream the Earth's bow shock, on day 10/30/1984 between 11:05 and 11:24 UT (see fig. 8). The data are regularly sampled at rate 8 Hz., which implies 8800 samples. This series has been the subject of numerous publications inside the geophysical community ([43, 10, 31] and references therein). The main reason is the presence of SLAMS (Short Large Amplitude Magnetic Structure) of typical duration 10 s (see fig. 8). Indeed, it has been shown that these structures are generated through non-linear quadratic wave-wave interactions. On the contrary, no equivalent presence has been found downstream the bow shock. The second series is composed of 8800 observations measured the same day by AMPTE, between 10:00 and 10:20 UT (see fig. 8) downstream the bow shock. In order to show that the test for gaussianity and linearity explained in section 4 and 5, are able to differentiate between the upstream and the downstream regions, we applied both tests on these two time series.

We first choose to set the smoothing parameter $M = \sqrt{L}$, where L is the number of samples in one of the J intervals. We then vary L from 78 to 220 and apply both tests each time. We limited our analysis to $L = 220$ in order to keep $J \geqslant 40$. The results are given in fig. 9 and fig. 10. In fig. 9, both panels represent the histogram of the 143 probability of false alarms (Pfa) of the gaussianity test. The Pfa is the probability that we will be wrong in accepting the alternative hypothesis. The top panel is related to the first time series (upstream of the bow shock), the bottom panel to the second series (downstream of the bow shock). As one can see in the top panel, in 86 % of the 143 L values, the gaussianity test Pfa is less than 0.05 for the upstream data. The fact that we don't reach 100 % of the 143 L values with Pfa < 0.05, may be due to the fact that by cutting the time series into J pieces we might split sometimes some SLAMS in two. Nevertheless, we can be confident in rejecting the gaussianity hypothesis of the upstream data. In contrast, the results of the bottom panel clearly show that the downstream time series is consistent with a zero skewness function. Nevertheless, it may be non-gaussian if proved at higher order.

As we are confident in rejecting the null assumption of gaussianity for the upstream time series, let apply the linearity test. The result is given in fig. 10. It represents the histogram of the level of confidence of the linearity test, related to the 86 % of the L values above mentioned. As one can see, in 78 % of the L values, the level of confidence is greater than 95 %. Therefore, the null hypothesis of linearity is rejected.

These results confirm the presence of non-gaussianity and nonlinearity in this time series measured in the upstream part of the Earth's bow shock. An obvious advantage of applying these tests is that it can provide a way of scanning the data in order to extract interesting data sets.

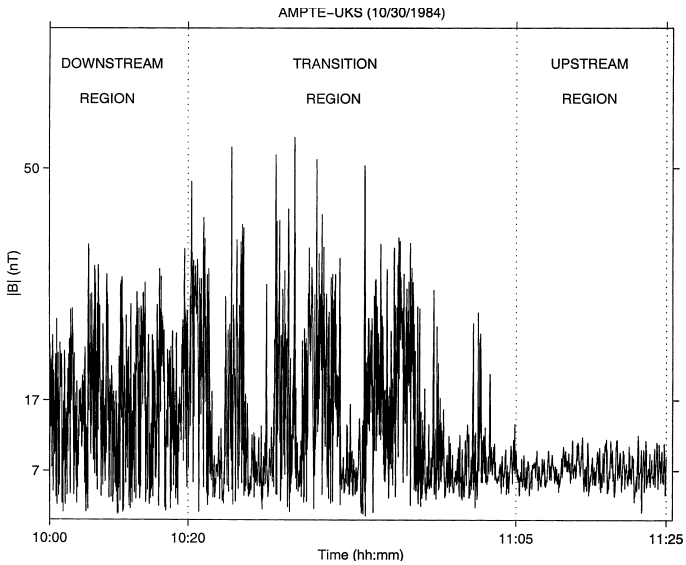


Figure 8. – Modulus of the magnetic field measured on board the AMPTE-UKS satellite as it was crossing the Earth's bow shock.

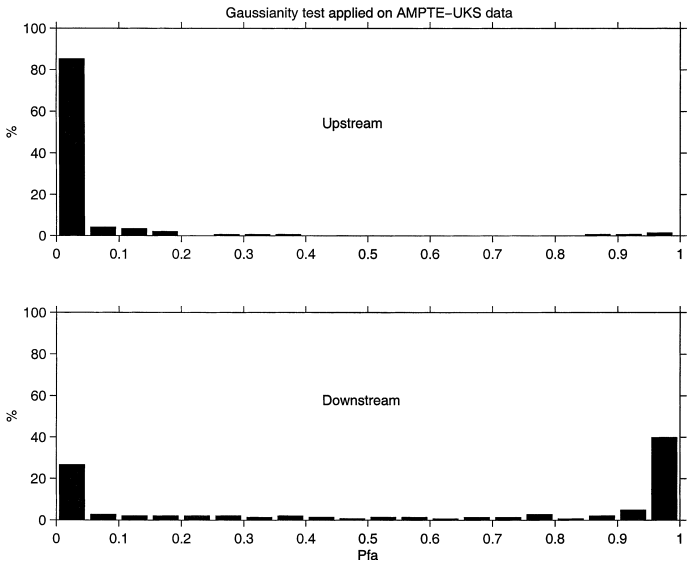


Figure 9. – Top panel: histogram of 143 probability of false alarms of the gaussianity test calculated by varying L from 78 to 220 for the upstream time series; bottom panel: same kind of histogram for the downstream series.

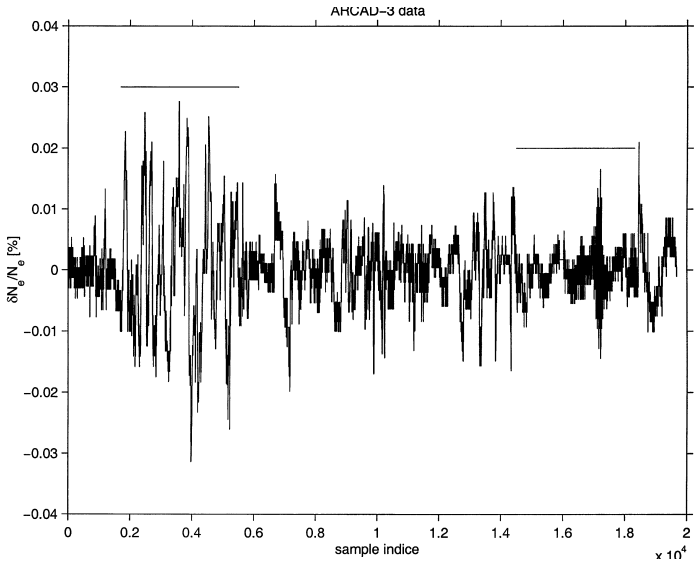


Figure 11. – Electron density fluctuations measured on board ARCAD-3. The two lines over the signal have been plotted to locate the deterministic (leftmost) and the noise-only sample of the data (rightmost).

8.2. ARCAD-3

The third series is the electron density fluctuations measured on board the satellite ARCAD-3, on March 16, 1982, during a crossing of a diffuse aurora region. The altitude was around 650 km and the invariant latitude was 70°. The sampling rate was 5 kHz, which implies 19 672 samples.

This time series has been analyzed together with the simultaneously measured electric field [32]. This analysis revealed the presence of non-linear wave-particle interactions using bispectral tools. One assumption in the use of these tools was that the part of the signal composed of 3 800 samples, located at the beginning of the series, was a deterministic transient signal (see fig. 11). Let now apply the test to detect transient signals, explained in section 6.

We first choose as the noise-only sample of the data, a part of the signal composed of 3 800 samples and located at the end of the series (see fig. 11). We then apply a gaussianity test as in the subsection 8.1. In 100 % of the L values, the Pfa of the gaussianity test is greater than 0.55. Consequently, the gaussianity assumption of the noise only part of the data can not be rejected. Therefore, the statistic of the test is calculated over the entire domain $IT + OT$. The result of the test applied on the presumed transient part is that the presence of a transient deterministic signal is accepted at a level of confidence 0.99. Consequently, the assumption made in [32] is justified by this test.

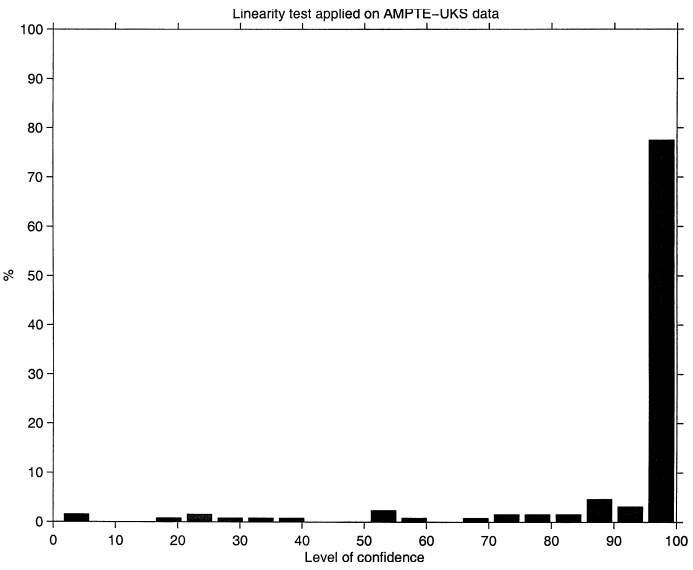


Figure 10. – Histogram of the level of confidence of the linearity test (*cf.* section 5) for the upstream time series.

9. conclusion

A general and unified procedure of higher-order spectral analysis technique is suggested to construct several statistical tests for: (1) detecting a non-gaussian or transient signal in a gaussian or a non-gaussian noise, (2) testing a stochastic time series for non-gaussianity (including non-linearity), (3) studying non-linear wave interactions by using k^{th}-order coherency functions. All the algorithms that have been presented were derived by using asymptotic theory of estimates of k^{th}-order spectra, in

a digital signal processing framework. The effectiveness of the presented algorithms is demonstrated through numerical simulations. Non-gaussianity and non-linearity tests have been applied to several sets of satellite data already analysed by space plasma experimenters. The results show:

1) that the physical hypotheses made by the experimenters, but not tested, were valid,

2) that, applying systematically these tests one has a way to make an automatic classification of the data.

As a conclusion the experimenters are recommended to use statistical tests to classify their data, then, the required hypotheses being validated, to apply the HOS analysis techniques as they usually do.

Notations and abbreviations

a, **a** and ***A*** stand for scalar, vector and matrix in that order. Similarly A^*, A^T, A^H represent the complex conjugate, transpose and transpose complex conjugate of ***A*** respectively. $E[.]$ denotes the expectation operator.

HOS – higher-order spectra

PSD – power spectral density

SOS – second-order statistics

HOS – higher-order statistics

DFT – discrete Fourier transform- kth-order discrete time principal domain

AWGN – additive white Gaussian noise

FFT – fast Fourier transform

GMLR – generalized maximum likelihood ratio

SNR – signal to noise ratio

AR – autoregressive

ARMA – autoregressive moving average

CDF – cumulative distribution function

Appendix

Following are some important properties of cumulants, which are used in theoretical developments [35]:
[CP1] If $a_i, i = 1, \ldots, k$ are constants, and $x_i, i = 1, \ldots, k$, are random variables, then

$$\mathrm{Cum}[a_1, x_1, a_2, x_2 \ldots, a_k x_k] = \left(\prod_{i=1}^{k} a_i \right) \mathrm{Cum}[x_1, x_2, \ldots, x_k].$$

[CP2] Cumulants are symmetric in their arguments.
[CP3] Cumulants are additive in their arguments, *i.e.*

$$\mathrm{Cum}[x_0 + y_0, z_1, \ldots, z_k] = \mathrm{Cum}[x_0 + y_0, z_1 \ldots, z_k] + \mathrm{Cum}[y_0, z_1, \ldots, z_k]$$

[CP4] If α is constant, then

$$\mathrm{Cum}[\alpha + y_0, z_1, \ldots, z_k] = \mathrm{Cum}[z_1, \ldots, z_k]$$

[CP5] If the random variables $\{x_i\}$ are independent of the random variables
$\{y_1\},\ i = 1, 2, \ldots, k$, then

$$\mathrm{Cum}[x_1 + y_1, \ldots, x_k + y_k] = \mathrm{Cum}[x_1, \ldots, x_k] + \mathrm{Cum}[y_1, \ldots, y_k]$$

[CP6] If a subset of the *k* random variables $\{x_i\}$ is independent of the rest, then

$$\mathrm{Cum}[x_1, \ldots, x_k] = 0\,.$$

BIBLIOGRAPHY

[1] T. W. Anderson, An Introduction to Multivariate Statistical Analysis New York: Wiley 1958.

[2] H. Akaike, "Note on higher-order spectra", *Ann. Inst. Statist. Math.*, vol. 18, pp. 123-126,1966.

[3] P. O. Ambard, J. L. Lacoume, and J. M. Brossier, "Transient detection, higher-order time-frequency distributions and the entropy", in Proc. *IEEE Work on Higher-Order Spectral Analysis* (South Lake Tahoe, CA, June 7-9, 1993), pp. 265-269.

[4] A. Blanc-Lapierre and R. Fortet, Théorie des Fonctions Aléatoires. Paris: Masson, 1953.

[5] D. R. Brillinger, "An introduction to polyspectra", *Ann. Math. Statist.*, vol. 36, pp. 1351-1374, 1965.

[6] D. R. Brillinger and M. Rosenblatt, "Asymptotic theory of estimates of k^{th} order spectra", *in Spectral Analysis of Time Series*, pp. 135-188, B. Harris, Ed. New York: Wiley, 1967.

[7] D. R. Brillinger, Time Series, Data Analysis and Theory, expanded. New York: Holt, Rine-hart and Winston, 1981.

[8] V. Chandran and S. Elgar, "A general procedure for the derivation of principal domains of higher-order spectra", *IEEE Trans. Signal Process.*, vol. 42, n° 1, pp. 229-233, Jan. 1994.

[9] R. F. Dwyer, "Detection of non-Gaussian signals by frequency domain kurtosis estimation", in *Proc. 1983 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 607-610, 1983.

[10] T. Dudok de Wit and V. V. Krasnosel'skikh, "Wavelet bicoherence analysis of strong plasma turbulence at the Earth's quasi-parallel bowshock", *Phys. Plasma*, 2, 262-273, 1995.

[11] S. Elgar, "Relationships involving third moments and bispectra of a harmonic process", *IEEE , vol. ASSP-35*, pp. 1725-1726, 1987.

[12] M. M. Gabr and T. Subba Rao, "On the estimation of bispectral density function in the case of randomly missing observations", *IEEE Trans. Signal Processing*, vol. 42, n° 1, pp. 211-216, Jan. 1994.

[13] L. M. Garth and H. Vincent Poor, "Detection of non-Gaussian signals: a paradigm for modern statistical signal processing", *Proc. of the IEEE*, vol. 82, n° 7, pp. 1060-1095, July 1994.

[14] G. B. Gianakis and M. K. Tsatsanis, "Signal detection and classification using matched filtering and higher-order statistics", *IEEE Trans. Acoust. Speech, and Signal Processing*, vol. 38, n° 7, pp. 1284-1296, July 1990.

[15] G. B. Gianakis and M. K. Tsatsanis, "A unifying maximum-likelihood view of cumulant and polyspectral measures for non-gaussian signal classification and estimation", "Signal detection and classification using matched filtering and higher-order statistics", *IEEE Trans. Inform. Theory*, vol. 38, n° 7, pp. 1284-1296, 1990.

[16] M. J. Hinich and C. S. Clay, "The application of the discrete Fourier transform in the estimation of power spectra, coherence, and bispectra of geophysical data", *Rev. Geophys.*, vol. 6, pp. 347-363, 1968.

[17] M. J. Hinich, "Testing for Gaussianity and linearity of a stationary time series", *J. Time Series Anal.*, vol. 3, pp. 169-176, 1982.

[18] M. J. Hinich and G. R. Wilson, "Detection of non-Gaussian signals in non-Gaussian noise using the bispectrum", *IEEE Trans. Acoust., Speech, and Signal Process*., vol. 38, n° 7, pp. 1126-1131, July 1990.

[19] M. J. Hinich, "Detecting a transient signal by bispectral analysis", *IEEE Trans. Acoust, Speech, and Signal Process.*, vol. 38, n° 7, pp. 1277-1283, July 1990.

[20] P. J. Huber, B. Kleiner, T. Gasser, and G. Dumermuth, "Statistical methods for investigating phase relations in stationary stochastic processes", *IEEE Trans. Audio Electro-acoustics*, AV-19, pp. 78-86, 1971.

[21] M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, vol. 1, 3rd ed. New York: Hafner, 1958.

[22] D. Kletter and H. Messer, "Suboptimal detection of non-Gaussian signals by third-order spectral analysis", *IEEE, Trans. Acoust., Speech, and Signal Process.*, vol. 38, n° 6, pp. 901-909, June 1990.

[23] Y. C. Kim and E. J. Powers, "Digital bispectral analysis of self-excited fluctuation spectra", *Phys. Fluids*, 21(8), pp. 1452-1453, August, 1978.

[24] Y. C. Kim and E. J. Powers, "Digital bispectral analysis and its applications to nonlinear wave interactions", *IEEE Trans. Plasma Science*, vol. PS-7, n° 2, pp. 120-131, June 1979.

[25] K. I. Kim and E. J. Powers, "A digital method of modeling quadratically nonlinear systems with a general random input", *IEEE Trans. Acoust., Speech, and Signal Process*., vol. 36, n° 11, pp. 1758-1769, November 1988.

[26] V. Kravtchenko-Berejnoi, F. Lefeuvre, V. Krasnossel'skikh, D. Lagoutte, "On the use of tricoherent analysis to detect non-linear wave-wave interactions", *Signal Processing*, 42, pp. 291-309, 1995.

[27] J.-L. Lacoume, P.-O. Amblard, and P. Comon, *Statistiques d'ordre supérieur pour le traitement du signal*. Paris: Masson, 1997.

[28] D. Lagoutte, F. Lefeuvre, and J. Hanasz, "Application of bicoherence analysis in the study of wave interactions in space plasma", *J. Geophys. Res*., 94, 435, 1989.

[29] D. Lagoutte and P. Latremoliere, SWAN, Software for Waveform Analysis, Version 2.0, LPCE-CNRS, May 1997.

[30] K. S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes", *Ann. Statist.*, vol. 10, n° 4, pp. 1195-1208, 1982.

[31] G. Mann, H. Lühr, W. Baumjohann, "Statistical analysis of short large-amplitude magnetic field structures in the vicinity of the quasi-parallel bow shock", *J. Geophys. Res.*, Vol. 99, n° A7, pp. 13,315, 1994.

[32] A. Masson, F. Lefeuvre, Z. Y. Zhao, D. Lagoutte, J. L. Rauch, "Observation of nonlinear interactions in large scale density enhancements of the high-latitude ionosphere", *J. Geophys. Res*. Vol. 104, N° A10, pp. 22,499, 1999.

[33] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications", *Proc. of the IEEE*, vol. 79, n° 3, pp. 278-305, March 1991.

[34] C. L. Nikias and M. S. Raghuveer, "Bispectrum estimation: a digital signal processing framework", *Proc. of the IEEE*, vol. 75, n° 7, pp. 869-891, July 1987.

[35] C. L. Nikias and A. P. Petropulu, Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[36] L. A. Pflug, G. E. Ioup, J. W. Ioup, R. L. Field, "Properties of higher-order correlations and spectra for bandlimited deterministic transients", *J. Acoust. Soc. Am. 91(2)*, pp. 975-988, February 1992.

[37] L. A. Pflug, G. E. Ioup, J. W. Ioup, K. H. Barues, R. L. Field, G.H. Rayborn, "Detection of oscillatory and impulsive transients using higher-order correlations and spectra", *J. Acoust. Soc. Am. 91(5)*, pp. 2763-2776, May 1992.

[38] L. A. Pflug, G. E. Ioup, J. W. Ioup, R. L. Field, "Prediction of Signal-to-noise ratio gain for passive higher-order correlation detection of energy transients", *J. Acoust. Soc. Am. 98(1)*, pp. 249-260, July 1995.

[39] O. Randriamboarison, V. Kravtchenko-Berejnoi, D. Lagoutte, F. Lefeuvre, "Analysis of models of nonlinear interactions by higher-order statistics", *in Proc. Spatio-Temporal Analysis for Resolving Plasma Turbulence*, pp. 285-288, Aussois, France, 1993.

[40] M. Rosenblatt and J. W. Van Ness, "Estimation of the bispectrum", *Ann. Math. Statist.*, 36, pp. 1120-1136, 1965.

[41] M. Rosenblatt, Stationary Sequences and Random Fields, Boston: Birkhauser, 1985.

[42] G. G. Roussas, Contiguity of Probability Measures. Cambridge University Press, 1972.

[43] S. J. Schwartz, D. Burgess, W. P. Wilkinson, R. L. Kessel, M. Dunlop, and H. Lühr, "Observations of Short Large-Amplitude Magnetic Structures at a Quasi-Parallel Shock", *J. Geophys. Res.*, Vol. 97, n° A4, pp. 4209, 1992.

[44] B. B. Shishkov, "Locally asymptotic normality for the parametric models of time series in the spectral domain", *Compt. Rend. Acad. Bulg. Sc.*, vol. 43, pp. 17-20, 1990.

[45] B. B. Shishkov, "Testing hypotheses on distribution's parameters of time series and applications to signal detection". *Radio Eng. Electron. Phys.*, vol. 39, pp. 411-423, 1993.

[46] B. B. Shishkov, H. Tsuji, J. Xin, Y. Hase, "Spatial and temporal processing of cyclostationary signals in array antennas based on linear prediction model", Submission to Telecommunications Advancement Organization (TAO) of Japan, December, 1997.

[47] A. N. Shiryaev, "Some problems in the spectral theory of higher-order moments", *Theor. Prob. Appl.*, vol. 5, pp. 265-284, 1960.

[48] A. D. Spaulding, "Locally optimum and suboptimum detector performance in a non-Gaussian interference environment", *IEEE Trans. Commun.*, vol. COM-33, n° 6, pp. 509-517, 1985.

[49] T. Subba Rao and M. M. Gabr, "A test for linearity of stationary time series", *J. Time Series Anal.*, vol. 1, n° 2, pp. 145-158, 1980.

[50] T. Subba Rao and M. M. Gabr, "An introduction to bispectral analysis and bilinear time series model", in Lecture Notes in Statistics, D. Brillinger *et al.*, Eds. New York: Springer-Verlag, 1984.

[51] C. W. Therrien, Discrete Random Signals and Statistical Signals Processing. Prentice Hall, Englewood Cliffs, NJ, 1992.

[52] V. Y. Trakhtengerts and M. Hayakawa, "A wave interaction in whistler frequency range in space plasma", *J. Geophys. Res.*, 98, 19,205, 1993.

[53] J. W. Van Ness, "Asymptotic normality of bispectral estimates", *Ann. Math. Statist.*, vol. 37, pp. 1257-1275, 1966.

[54] Workshop on Turbulence in Space Plasmas. Non-Linear Processes and Data Filtering Techniques, Garchy (France), 5-7 May 1998.

Manuscrit reçu le 5 mars 1999.

LES AUTEURS

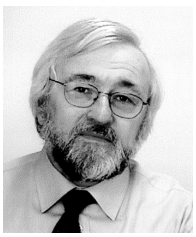
Arnaud MASSON

Arnaud Masson got his Master Thesis in Remote Sensing with honours in 1996 at Paris VI University, France. He then started his PhD under the supervision of Dr. F. Lefeuvre at LPCE/CNRS in Orléans. His main interests are the study of the plasma turbulence in the Earth's magnetosphere, higher-order statistics, wavelet analysis and chaos theory. He was a visiting student at the Swedish Institute of Space Physics (Uppsala) from 1999 till mid 2000, in the Solar Terrestrial Physics group. He obtained a young scientist award from the International Union of Radio Science in 1999 for a digital de-noising technique applied to space data.

Blagovest SHISHKOV

Blagovest Shishkov was born in Topolovgrad, Bulgaria in 1937. He received his PhD in Physics (1974) from the Bulgarian Institute of Electronics and his D.Sc. Tech. degree (1991) from the Technical University of Sofia where he is now Professor in Signal Processing. He has published more than one hundred papers and three books in Signal Transmission and especially in asymptotic methods in parameter estimation, signal detection and identification, pattern recognition and data quantization. He is an official member of the URSI Commission C (Signals and Systems), member of the IEEE, EURASIP and IEICE (Japan).

François LEFEUVRE



François Lefeuvre received his first thesis (3eme cycle) in 1970 from the Paris University and his second thesis (these d'état) in 1977 from the Orleans University. He got a permanent position at CNRS (Centre National de la Recherche Scientifique) in 1972 and became director of research in 1986. He is now Director of the Laboratoire de Physique et Chimie de l'Environnement, in Orléans. His main interest is in the development of analysis techniques for waves and turbulence in space plasmas. He was involved as a CO-I or as a PI of several wave experiments embarked on satellites (GEOS, AUREOL-3, INTERBALL – 2, CLUSTER).