

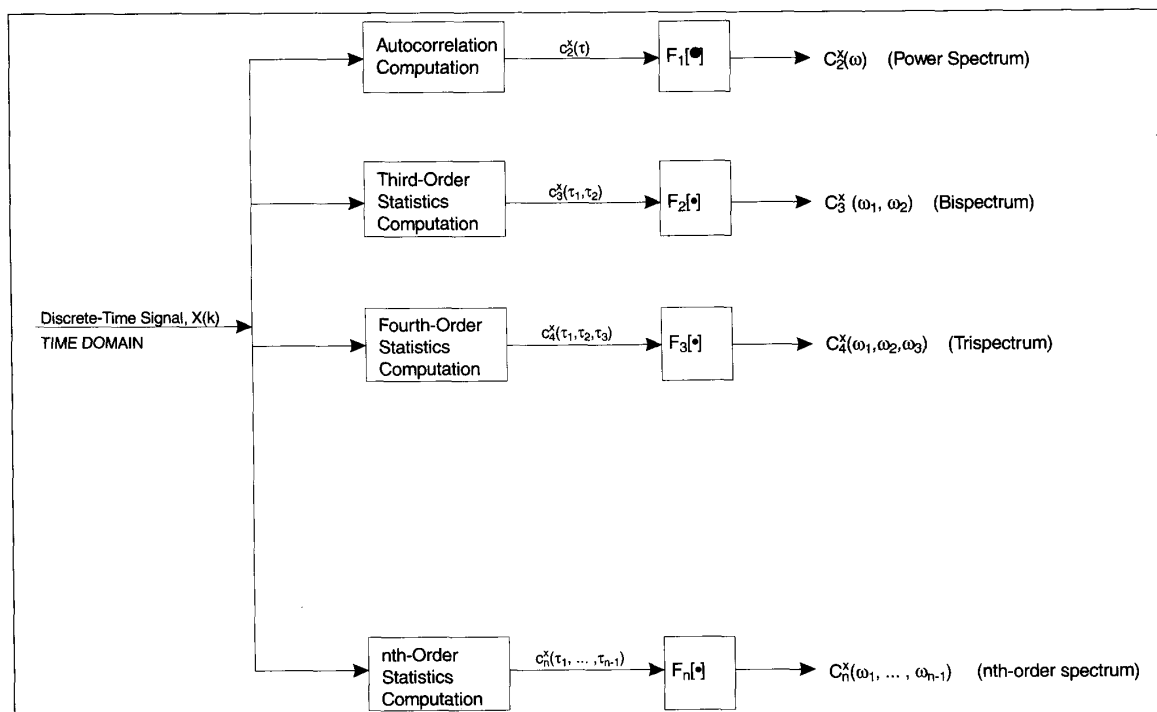
Signal Processing with Higher-Order Spectra

The estimation of the power spectral density or simply the power spectrum of discrete-time deterministic or stochastic signals has been a useful tool in digital signal processing for more than thirty years. Power spectrum estimation techniques have proved essential to the creation of advanced radar, sonar, communication, speech, biomedical, geophysical, and other data processing systems. The available power spectrum estimation techniques may be considered in a number of separate classes, namely, conventional (or "Fourier type") methods, maximum-likelihood method of Capon with its modifications, maximum-entropy, minimum energy, and minimum-cross-entropy methods, as well as methods based on autoregressive (AR), moving-average (MA) and ARMA models; and harmonic decomposition methods such as Prony, Pisarenko, MUSIC, ESPRIT, and Singular Value Decomposition. Developments in this area have also led to signal modeling, and to extensions to multi-dimensional, multi-channel, and array processing problems. Each one of the aforementioned techniques has certain advantages and limitations not only in terms of estimation performance but also in terms of computational complexity and, therefore, depending on the signal environment, one has to choose the most appropriate [Marple, 1987; Kay, 1988; Haykin, 1983].

In power spectrum estimation, the signal under consideration is processed in such a way that the distribution of power among its frequency components is estimated. As such, phase relations between frequency components are suppressed. The information contained in the power spectrum is essentially that which is present in the autocorrelation sequence; this would suffice for the complete statistical description of a Gaussian signal. However, there are practical situations where we have to look beyond the power spectrum (autocorrelation) of a signal to extract information regarding deviations from Gaussianity and presence of phase relations [Nikias and Raghuveer, 1987; Mendel 1991].

Higher order spectra (also known as polyspectra), defined in terms of higher order statistics ("cumulants") of a signal, do contain such information. Particular cases of higher order spectra are the third-order spectrum also called the bispectrum which is, by definition, the Fourier transform of the third-order statistics, and the trispectrum (fourth-order spectrum) which is the Fourier transform of the fourth-order statistics of a stationary signal. The power spectrum is, in fact, a member of the class of higher order spectra, i.e., it is a second-order spectrum [Rosenblatt, 1985]. Figure 1 illustrates the higher-order spectra classification map of a given discrete-time signal. Although higher-order statistics and spectra of a signal can be defined in terms

Chrysostomos L. Nikias and Jerry M. Mendel



1. The higher-order spectra classification map of a discrete signal $X(k)$. $F[\bullet]$ denotes n -dimensional Fourier Transform.

of moments and cumulants, moment and moment spectra can be very useful in the analysis of deterministic signals (transient and periodic) whereas cumulants and cumulant spectra are of great importance in the analysis of stochastic signals [Nikias and Petropulu, 1993].

There are several general motivations behind the use of higher-order spectra in signal processing. These include techniques to: (1) suppress additive colored Gaussian noise of unknown power spectrum; (2) identify non-minimum phase systems or reconstruct nonminimum phase signals; (3) extract information due to deviations from Gaussianity; and (4) detect and characterize nonlinear properties in signals as well as identify nonlinear systems [Nikias and Raghuvver, 1987].

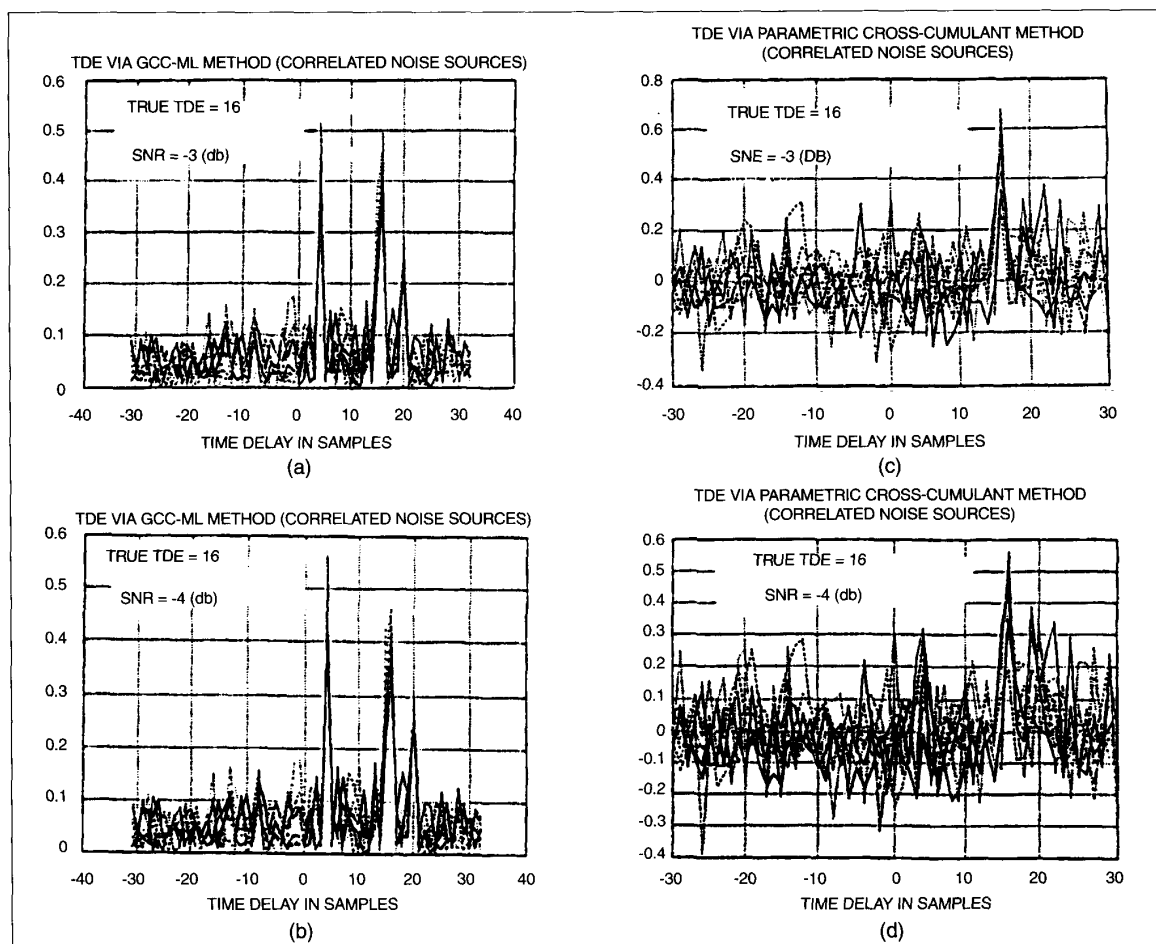
The first motivation is based on the property that for Gaussian signals only, all cumulant spectra of order greater than two are identically zero. If a non-Gaussian signal is received along with additive Gaussian noise, a transform to a higher-order cumulant domain will eliminate the noise. Hence, in these signal processing settings, there will be certain advantages to detecting or/and estimating signal parameters from cumulant spectra of the observed data. In particular, cumulant spectra can become high signal-to-noise ratio (SNR) domains in which one may perform detection, parameter estimation or even entire signal reconstruction. Figure 2 illustrates time-delay parameter estimation results obtained by a cross-correlation (i.e., 2nd-order statistics) method and a technique based on cross third-order cumulants. The signal of interest is assumed to be non-Gaussian whereas the additive noise is Gaussian and spatially correlated. From Fig. 2, it is apparent that the third-order cumulants do sup-

press the effect of Gaussian noise and thus provide better time-delay estimates (TDE), especially in low SNR [Nikias and Pan, 1988].

The second motivation is based on the fact that polyspectra (cumulant and moment) preserve the true phase character of signals. For modeling time series data in signal processing problems, second-order statistics are almost exclusively used because they are usually the result of least-squares optimization criteria. However, the autocorrelation domain suppresses phase information. An accurate phase reconstruction in the autocorrelation (or power-spectrum) domain can only be achieved if the signal is minimum phase. On the other hand, non-minimum phase signal reconstruction or system identification can be achieved in higher-order spectrum domains due to the ability of polyspectra to preserve both magnitude and non-minimum phase information. Figure 3 illustrates two different signals which have identical autocorrelation but different third-order statistics. Consequently, these two signals have identical power spectra and different bispectra.

The third motivation is based on the observation that most "real world" signals are non-Gaussian and thus have non-zero higher-order spectra. As Fig. 1 demonstrates, a non-Gaussian signal can be decomposed into its higher-order spectral functions where each one of them may contain different information about the signals. This can be very useful in signal classification problems where distinct classification features can be extracted from higher-order spectrum domains.

Finally, introduction of higher-order spectra is quite natural when we try to analyze the nonlinearity of a system



2. Time delay estimation in the presence of spatially correlated Gaussian noise. Assume that $X(k)$ and $Y(k)$ are two available sensor measurements satisfying $X(k) = S(k) + W_1(k)$ and $Y(k) = A S(k - D) + W_2(k)$, where $S(k)$ is an unknown non-Gaussian signal, D is the time delay and $W_1(k)$, $W_2(k)$ are spatially correlated Gaussian noises. The problem is to estimate D . A generalized cross-correlation method based on maximum likelihood window (ML) estimates the autocorrelation of $S(k)$ which peaks at D ($D = 16$) and the spatial correlation between $W_1(k)$, $W_2(k)$ as shown in (a) and (b). On the other hand, a cross-cumulant method suppresses the effect of noise and thus makes the estimation of D more reliable, as shown in (c) and (d).

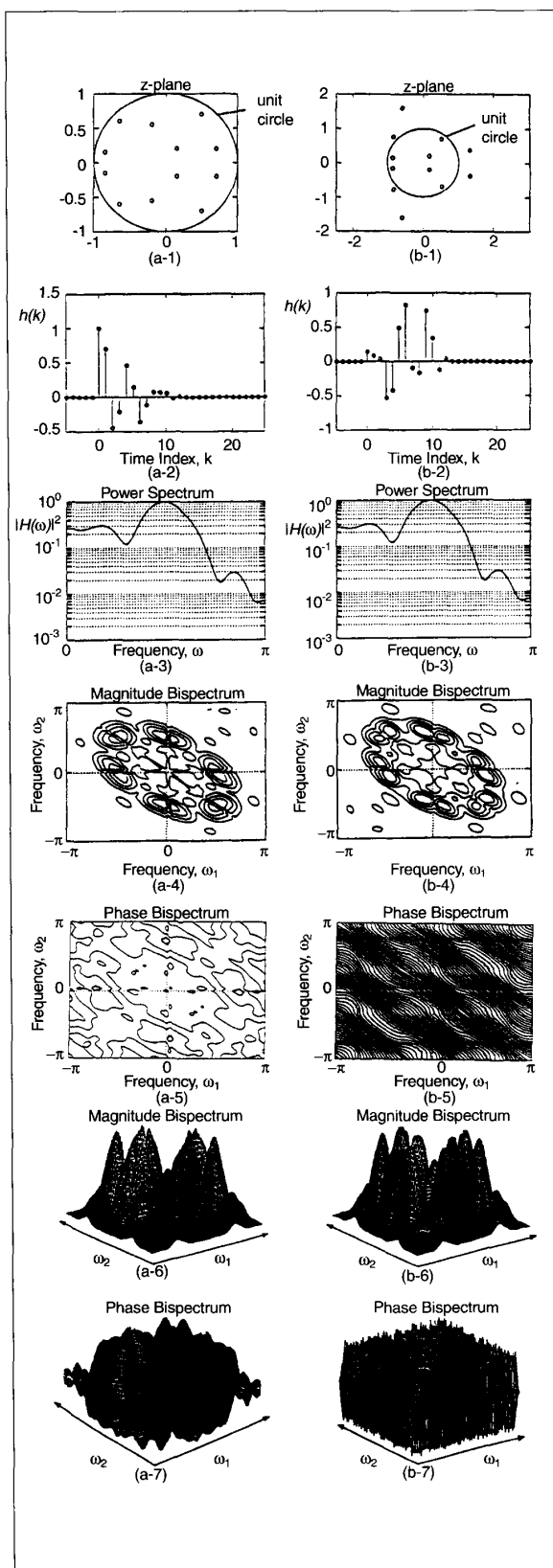
operating under a random input. General relations for arbitrary stationary random data passing through arbitrary linear systems have been studied quite extensively for many years. In principle, most of these relations are based on power spectrum (autocorrelation) matching criteria. On the other hand, general relations are not available for arbitrary stationary random data passing through arbitrary nonlinear systems. Instead, each type of nonlinearity has to be investigated as a special case [Brillinger, 1977]. Polyspectra can play a key role in detecting and characterizing the type of nonlinearity in a system from its output data [Schetzen, 1989]. Several signal processing methods for the detection and characterization of nonlinearities in time series using higher-order spectra have been developed [Rao and Gabr, 1984; Nikias and Petropulu, 1993].

The organization of the article is as follows. First we discuss the strengths and limitations of correlation-based

signal processing methods. Then we introduce the definitions, properties and computation of higher-order statistics and spectra with emphasis on the bispectrum and trispectrum. Following that, we describe parametric and non-parametric expressions of polyspectra of linear processes. The following section addresses polyspectra of nonlinear processes. We conclude the article with a discussion of the applications of higher-order spectra in signal processing.

Correlation-Based Signal Processing: Strengths and Limitations

In this section, we review important results from correlation-based signal processing in order to point out its strengths and limitations. All of our results are for a discrete-time random signal, because we are interested in digital signal processing applications and the data we work with is assumed to be sampled.



Given a real, stationary, zero-mean random signal $\{X(k)\}$, its **autocorrelation function**, $c_2^x(\tau)$, provides a measure of how the sequence is correlated with itself at different time points:

$$c_2^x(\tau) = E \{X(k) X(k + \tau)\} \quad (1)$$

Note that $c_2^x(\tau)$ is a symmetric function about $\tau = 0$, i.e., $c_2^x(-\tau) = c_2^x(\tau)$; hence, $c_2^x(\tau)$ is a zero-phase function, which means that all phase information about $X(k)$ is lost in $c_2^x(\tau)$.

Given two real, stationary, zero-mean random signals $\{X(k)\}$ and $\{Y(k)\}$, their **cross-correlation function**, which is defined below, provides a measure of how the two signals are correlated with each other at different time points:

$$c_{xy}(\tau) = E \{X(k) Y(k + \tau)\} \quad (2)$$

Note that $c_{xy}(\tau)$ is not a symmetric function about $\tau = 0$, but that $c_{yx}(-\tau) = c_{xy}(\tau)$.

The discrete-time Fourier transforms of $c_2^x(\tau)$ and $c_{xy}(\tau)$ are known as the **power spectrum** of signal $\{X(k)\}$ and the **cross-power spectrum** between signals $\{X(k)\}$ and $\{Y(k)\}$, respectively, i.e.,

$$C_2^x(\omega) = F \{c_2^x(\tau)\} \text{ so that } c_2^x(\tau) = F^{-1} \{C_2^x(\omega)\} \quad (3)$$

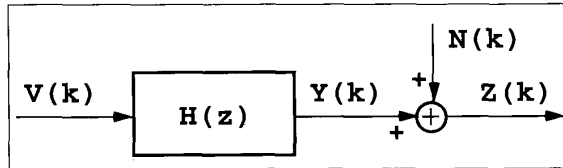
and

$$C_{xy}(\omega) = F \{c_{xy}(\tau)\} \text{ so that } c_{xy}(\tau) = F^{-1} \{C_{xy}(\omega)\} \quad (4)$$

where $F\{\bullet\}$ and $F^{-1}\{\bullet\}$ denote Fourier and inverse Fourier transforms. Note that $C_2^x(\omega)$ is a real function of ω whereas $C_{xy}(\omega)$ is a complex function of ω .

If a discrete random signal is uncorrelated from one time point to the next, i.e., $E\{X(i) X(j)\} = 0$ for all $i \neq j$, then we refer to that signal as (discrete-time) **white noise**. In general, white noise does not have to be stationary, i.e., its variance may vary with time. If it is stationary, so that $E\{X^2(k)\} = \gamma_2^x$, then $C_2^x(\omega) = \gamma_2^x$ for all values of ω ; in this case, we associate white noise with a signal whose power spectrum is "flat" for all frequencies. Because this commonly-used description of

3. Signal representation: (a-1) zeros of minimum phase system, (a-2) minimum phase sequence, (a-3) power spectrum of the minimum phase sequence, (a-4) contour of the magnitude of the bispectrum of the minimum phase sequence; (a-5) contour of the phase of the bispectrum of the minimum phase sequence, (a-6) 3-D plot of the magnitude of the bispectrum of the minimum phase sequence, (a-7) 3-D plot of the phase of the bispectrum of the minimum phase sequence; (b-1) zeros of spectrally-equivalent nonminimum phase system, (b-2) nonminimum phase sequence, (b-3) power spectrum of the nonminimum phase sequence, (b-4) contour of the magnitude of the bispectrum of the nonminimum phase sequence. Signal representation: (b-5) contour of the phase of the bispectrum of the nonminimum phase sequence, (b-6) 3-D plot of the magnitude of the bispectrum of the nonminimum phase sequence, (b-7) 3-D plot of the phase of the bispectrum of the nonminimum phase sequence.



4. Single-channel system.

white noise only involves second-order statistics, we could refer to such white noise as "second-order" white noise. We have not done so in the past, because we have not needed the notion of a white noise that is characterized by its higher-than second-order statistics. We will return to a discussion of such higher-order white noise.

A single channel model is depicted in Fig. 4. In this model, $V(k)$ is stationary white noise with finite variance γ_2^V ($V(k)$ may be Gaussian or non-Gaussian); $H(z)$ [$h(k)$] is causal and stable; $N(k)$ is Gaussian (white or colored); and, $V(k)$ and $N(k)$ are statistically independent. The convolutional equation describing the noise-free output of this system is

$$Y(k) = h(k) * V(k) = \sum_{i=0}^k h(i) V(k-i) \quad (5)$$

which can also be expressed in the complex z -domain or frequency-domain, as

$$Y(z) = H(z)V(z) \text{ or } Y(\omega) = H(\omega)V(\omega) \quad (6)$$

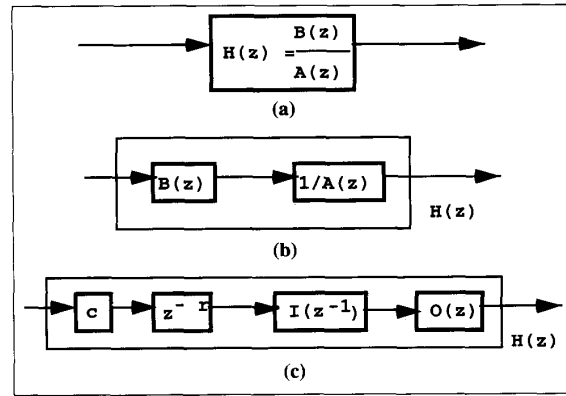
where $z = \exp(j\omega T)$, and we have assumed that sampling time T equals unity.

Three very popular parametric channel models are: (1) **Moving Average (MA)**, in which $H(z) = B(z)$ and $B(z)$ is a rational polynomial in z , one with q zeros, so that $H(z)$ is an all-zero model; (2) **Autoregressive (AR)**, in which $H(z) = 1/A(z)$ and $A(z)$ is a rational polynomial in z , one with p zeros, so that $H(z)$ is an all-pole model; and, (3) **Autoregressive-Moving Average (ARMA)**, in which $H(z) = B(z)/A(z)$ where $A(z)$ and $B(z)$ are as described for the AR and MA models, respectively; hence, an ARMA model has q zeros and p poles. An MA model has a finite impulse response (FIR), whereas AR and ARMA models have infinite impulse responses (IIR).

When all the zeros of $H(z)$ lie inside the unit circle in the complex z -plane, then $H(z)$ is said to be **minimum phase**. When all of the zeros of $H(z)$ lie outside the unit circle, then $H(z)$ is said to be **maximum phase**. When some of $H(z)$'s zeros lie inside or outside of the unit circle, then $H(z)$ is said to be mixed phase or **nonminimum phase**. Many real-world systems (channels) are nonminimum phase.

Because of the linearity of the system in Fig. 4, different representations of the ARMA model, $H(z) = B(z)/A(z)$, are possible, including those depicted in Fig. 5. The representation in 5a considers the ARMA model in one shot; 5b considers the ARMA model as a cascade of an MA model followed by an AR model; and, 5c is based on expressing a nonminimum phase system in terms of its poles and zeros as follows:

$$H(z) = B(z)/A(z) = c z^{-r} I(z^{-1}) O(z) \quad (7)$$



5. Equivalent representations for channel $H(z)$.

where c is a constant, r is an integer,

$$I(z^{-1}) = \frac{\prod_{i=1}^{L_1} (1 - a_i z^{-1})}{\prod_{i=1}^{L_2} (1 - c_i z^{-1})} \quad (8)$$

is the minimum phase component of $H(z)$, with poles $\{c_i\}$ and zeros $\{a_i\}$ inside the unit circle, i.e., $|c_i| < 1$ and $|a_i| < 1$ for all $\{i\}$, and

$$O(z) = \prod_{i=1}^{L_2} (1 - b_i z) \quad (9)$$

is the maximum phase component of $H(z)$, with zeros outside the unit circle at $1/b_i$, where $|b_i| < 1$ for all $\{i\}$. The different representations for $H(z)$ are each useful in their own right.

The **complex cepstrum** is widely known in digital signal processing circles (e.g., Oppenheim and Schaffer, 1989, ch. 12). One starts with the transfer function $H(z)$, takes its logarithm, $\hat{H}(z) = \ln H(z)$, and then takes the inverse z -transform of $\hat{H}(z)$ to obtain the complex cepstrum $\hat{h}(k)$.

When $H(z)$ is decomposed as in Eq. 7 into the product of its minimum-phase and maximum-phase components, $i(k)$ and $o(k)$, respectively, then

$$h(k) = c i(k) * o(k) * \delta(k-r). \quad (10)$$

It is well known that $i(k)$ and $o(k)$ can be computed recursively in terms of $\hat{h}(k)$ using the following formulas [Oppenheim and Schaffer (1989), Pan and Nikias (1988)]:

$$i(k) = \begin{cases} 0, & k < 0 \\ \exp(\hat{h}(0)), & k = 0 \\ -\frac{1}{k} \sum_{j=1}^k A^{(j)} i(k-j) & k > 0 \end{cases} \quad (11)$$

Table 1. Minimum-, maximum- and mixed-phase systems with identical power spectra (or autocorrelations): $0 < a < 1$, $0 < b < 1$ (Nikias and Raghuveer, 1987).

System	Minimum Phase	Maximum Phase	Mixed Phase
$H(z)$	$(1 - az^{-1})(1 - bz^{-1})$	$(1 - az)(1 - bz)$	$(1 - az)(1 - bz^{-1})$
Autocorrelations			
$c_2(0)$	$1 + a^2b^2 + (a + b)^2$		
$c_2(1)$	$-(a + b)(1 + ab)$		
$c_2(2)$	ab		

and

$$o(k) = \begin{cases} 0, & k > 0 \\ 1, & k = 0 \\ \frac{1}{k} \left[\sum_{j=k}^k B^{(-j)} o(k-j) \right] & k < 0 \end{cases} \quad (12)$$

in which the "A" and "B" cepstral coefficients are related to the minimum-delay and maximum-delay zeros, respectively, i.e.,

$$A^{(i)} = \sum_{j=1}^{L_1} a_j^i - \sum_{j=1}^{L_2} c_j^i \quad (13)$$

and

$$B^{(i)} = \sum_{j=1}^{L_2} b_j^i \quad (14)$$

If the cepstral coefficients can be computed, then $i(k)$ and $o(k)$ can be reconstructed from Eqs. 11 and 12, respectively, after which $h(k)$ can be reconstructed (to within a scale factor and a pure delay) from Eq. 10.

It is also well known (e.g., Papoulis, 1991), that

$$c_2^z(\tau) = c_2^y(\tau) + c_2^N(\tau) = \gamma_2^y \sum_{i=0}^{\infty} h(i)h(i+\tau) + c_2^N(\tau) \quad (15)$$

$$C_2^z(z) = \gamma_2^y H(z)H(z^{-1}) + C_2^N(z) \quad (16a)$$

or

$$C_2^z(\omega) = \gamma_2^y |H(\omega)|^2 + C_2^N(\omega), \quad (16b)$$

and,

$$c_{vz}(\tau) = \gamma_2^y h(\tau). \quad (17)$$

From Eqs. 16, we see that all phase information about $H(\omega)$ has been lost in the spectrum (or in the autocorrelation); hence, we say that **correlation or spectra are phase blind**. Observe, from Eq. 17, that if we have access to both the input and output of the system, then we can reconstruct the correct

phase IR, $h(k)$. In many important signal processing applications, we only have access to the output signal of the system; hence, we cannot use Eq. 17 in those applications.

Suppose that $H(z)$ has zeros both inside and outside the unit circle. When these zeros are reflected to their complementary locations [i.e., (some or all of) those inside the unit circle

are reflected outside the unit circle, and (some or all of) those outside the unit circle are reflected inside the unit circle], we see that the power spectrum remains unchanged, e.g., $C_2(z^{-1}) = C_2(z)$. When all of $H(z)$'s zeros that are outside the unit circle are reflected inside the unit circle, so that the resulting transfer function is minimum phase, i.e., $H(z) \rightarrow H_{MP}(z)$, we again find that

$$\gamma_2^y H_{MP}(z)H_{MP}(z^{-1}) + C_2^N(z) = \gamma_2^y H(z)H(z^{-1}) + C_2^N(z).$$

The minimum phase transfer function $H_{MP}(z)$ is said to be "spectrally equivalent minimum phase" (SEMP) equivalent to $H(z)$. Table 1 provides a simple example of three MA systems that have exactly the same autocorrelations and spectra. It illustrates the point that it is impossible to reconstruct the correct phase model just from autocorrelation or power spectrum information.

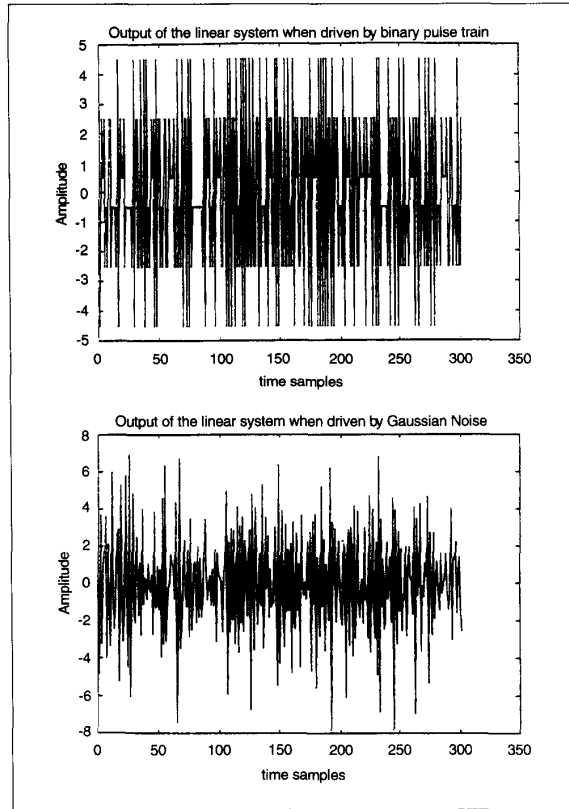
Figure 6 depicts two time series that have exactly the same power spectrum. One is Gaussian and the other is non-Gaussian. The top time series was obtained by exciting the non-minimum phase MA system $H(z) = (1 - \frac{1}{2}z^{-1})(1 - 2z^{-1})$ with a ± 1 binary pulse train (unity variance, and -2 fourth-order cumulant), whereas the bottom time series was obtained by exciting the same system with a Gaussian sequence (unity variance). Observe that these time series look quite different. If we only use correlation information, with its associated SEMP model, then we will be extracting the wrong information from the data when the actual model is nonminimum phase. In the next section we demonstrate that we can extract correct information from the data when we work with higher-order statistics.

Higher-Order Statistics and Spectra

a. Definitions and Properties

In this section, we introduce the definitions, properties and computation of higher-order statistics, i.e., moments and cumulants, and their corresponding higher-order spectra. The description is given for both stochastic and deterministic signals. However, the emphasis of the discussion is placed on the 2nd-, 3rd-, and 4th-order statistics and their respective Fourier transforms: power spectrum, bispectrum, and trispectrum.

If $\{X(k)\}$, $k = 0, \pm 1, \pm 2, \pm 3, \dots$ is a real stationary discrete-time signal and its moments up to order n exist, then



6. Time series with identical power spectra. One is Gaussian; the other is non-Gaussian.

$$m_n^X(\tau_1, \tau_2, \dots, \tau_{n-1}) \triangleq E \{X(k) X(k + \tau_1) \dots X(k + \tau_{n-1})\} \quad (18)$$

represents the n th-order moment function of the stationary signal, which depends only on the time differences $\tau_1, \tau_2, \dots, \tau_{n-1}$, $\tau_i = 0, \pm 1, \dots$ for all i . Clearly, the 2nd-order moment function, $m_2^X(\tau_1)$, is the autocorrelation of $\{X(k)\}$ whereas $m_3^X(\tau_1, \tau_2)$ and $m_4^X(\tau_1, \tau_2, \tau_3)$ are the 3rd- and 4th-order moments, respectively. $E\{\bullet\}$ denotes statistical expectation.

The n th-order cumulant function of a non-Gaussian stationary random signal $X(k)$ can be written as (for $n = 3, 4$ only):

$$c_n^X(\tau_1, \tau_2, \dots, \tau_{n-1}) = m_n^X(\tau_1, \tau_2, \dots, \tau_{n-1}) - m_n^G(\tau_1, \tau_2, \dots, \tau_{n-1}) \quad (19)$$

where $m_n^X(\tau_1, \dots, \tau_{n-1})$ is the n th-order moment function of $X(k)$ and $m_n^G(\tau_1, \tau_2, \dots, \tau_{n-1})$ is the n th-order moment function of an equivalent Gaussian signal that has the same mean value and autocorrelation sequence as $X(k)$. Clearly, if $X(k)$ is Gaussian, $m_n^X(\tau_1, \dots, \tau_{n-1}) = m_n^G(\tau_1, \dots, \tau_{n-1})$ and thus $c_n^X(\tau_1, \tau_2, \dots, \tau_{n-1}) = 0$. Note, however, that although Eq. 19

is only true for orders $n = 3$ and 4 , $c_n^X(\tau_1, \tau_2, \dots, \tau_{n-1}) = 0$ for all n if $X(k)$ is Gaussian. The properties of cumulants are summarized in Table 2.

The following relationships between moment and cumulant sequences of $\{X(k)\}$ exist for orders $n = 1, 2, 3, 4$.

1st-order cumulants:

$$c_1^X = m_1^X = E\{X(k)\} \quad (\text{mean value}) \quad (20)$$

2nd-order cumulants:

$$\begin{aligned} c_2^X(\tau_1) &= m_2^X(\tau_1) - (m_1^X)^2 \quad (\text{covariance sequence}) \\ &= m_2^X(-\tau_1) - (m_1^X)^2 = c_2^X(-\tau_1) \end{aligned} \quad (21)$$

where $m_2^X(-\tau_1)$ is the autocorrelation sequence. Thus, we see that the 2nd-order cumulant sequence is the *covariance* while the 2nd-order moment sequence is the *autocorrelation*.

3rd-order cumulants:

$$\begin{aligned} c_3^X(\tau_1, \tau_2) &= m_3^X(\tau_1, \tau_2) \\ &- m_1^X [m_2^X(\tau_1) + m_2^X(\tau_2) + m_2^X(\tau_1 - \tau_2)] + 2(m_1^X)^3 \end{aligned} \quad (22)$$

where $m_3^X(\tau_1, \tau_2)$ is the 3rd-order moment sequence. This follows if we combine Eqs. 18 and 19.

4th-order cumulants:

Combining Eqs. 18 and 19, we get

$$\begin{aligned} c_4^X(\tau_1, \tau_2, \tau_3) &= m_4^X(\tau_1, \tau_2, \tau_3) \\ &- m_2^X(\tau_1) \cdot m_2^X(\tau_3 - \tau_2) - m_2^X(\tau_2) \cdot m_2^X(\tau_3 - \tau_1) \\ &- m_2^X(\tau_3) \cdot m_2^X(\tau_2 - \tau_1) - m_1^X [m_3^X(\tau_2 - \tau_1, \tau_3 - \tau_1) \\ &+ m_3^X(\tau_2, \tau_3) + m_3^X(\tau_2, \tau_4) + m_3^X(\tau_1, \tau_2)] \\ &+ (m_2^X)^2 [m_1^X(\tau_1) + m_2^X(\tau_2) + m_2^X(\tau_3) + m_2^X(\tau_3 - \tau_1) \\ &+ m_2^X(\tau_3 - \tau_2) + m_2^X(\tau_2 - \tau_1)] - 6(m_1^X)^4. \end{aligned} \quad (23)$$

If the signal $\{X(k)\}$ is zero-mean $m_1^X = 0$, it follows from Eqs. 21 and 22 that the second- and third-order cumulants are identical to the second- and third-order moments, respectively; however, to generate the fourth-order cumulants, we need knowledge of the fourth-order and second-order moments in Eq. 23, i.e.,

$$\begin{aligned} c_4^X(\tau_1, \tau_2, \tau_3) &= m_4^X(\tau_1, \tau_2, \tau_3) - m_2^X(\tau_1) \cdot m_2^X(\tau_3 - \tau_2) \\ &- m_2^X(\tau_2) \cdot m_2^X(\tau_3 - \tau_1) - m_2^X(\tau_3) \cdot m_2^X(\tau_2 - \tau_1). \end{aligned} \quad (24)$$

By putting $\tau_1 = \tau_2 = \tau_3 = 0$ in Eqs. 21–23 and assuming $m_1^X = 0$, we obtain

$$\begin{aligned} \gamma_2^X &= E\{X^2(k)\} = c_2^X(0) \quad (\text{variance}) \\ \gamma_3^X &= E\{X^3(k)\} = c_3^X(0, 0) \quad (\text{skewness}) \\ \gamma_4^X &= E\{X^4(k)\} - 3[\gamma_2^X]^2 = c_4^X(0, 0, 0) \quad (\text{kurtosis}). \end{aligned} \quad (25)$$

Normalized kurtosis is defined as $\gamma_4^X/[\gamma_2^X]^2$. Equation 25 gives the variance, skewness and kurtosis measures in terms of cumulants at zero lags.

Table 2. Cumulant Properties

Cumulants can be treated as an operator, just as we treat expectation as an operator. This is due to the following six cumulant properties.

[CP1] Cumulants of scaled quantities (where the scale factors are non-random) equal the product of all the scale factors times the cumulant of the unscaled quantities, i.e., if $\lambda_i, i = 1, 2, \dots, n$, are constants, and $x_i, i = 1, 2, \dots, n$, are random variables, then (note that for third- or fourth-order cumulants, $n = 3$ or 4)

$$\text{cum}(\lambda_1 x_1, \dots, \lambda_n x_n) = \left(\prod_{i=1}^n \lambda_i \right) \text{cum}(x_1, \dots, x_n)$$

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

$$\text{cum}(x_1, \dots, x_n) = \text{cum}(x_{i_1}, \dots, x_{i_n})$$

where (i_1, \dots, i_n) is a permutation of $(1, \dots, n)$; this means we can interchange the arguments of the cumulant anyway we wish without changing the value of the cumulant, e.g., $c_3^x(\tau_1, \tau_2, \tau_3) = c_3^x(\tau_3, \tau_1, \tau_2) = c_3^x(\tau_2, \tau_3, \tau_1)$, etc.

[CP3] Cumulants are additive in their arguments, i.e., cumulants of sums equal sums of cumulants; hence the name "cumulant." so, for example, even when x_0 and y_0 are **not** statistically independent, it is true that

$$\text{cum}(x_0 + y_0, z_1, \dots, z_n) = \text{cum}(x_0, z_1, \dots, z_n) + \text{cum}(y_0, z_1, \dots, z_n)$$

[CP4] Cumulants are blind to additive constants, i.e., if α is a constant, then

$$\text{cum}(\alpha + z_1, z_2, \dots, z_n) = \text{cum}(z_1, z_2, \dots, z_n)$$

[CP5] Cumulants of a sum of statistically independent quantities equal the sum of the cumulants of the individual quantities, i.e., if the random variables $\{x_i\}$ are independent of the random variables $\{y_i\}$, $i = 1, 2, \dots, k$, then

$$\text{cum}(x_1 + y_1, \dots, x_n + y_n) = \text{cum}(x_1, \dots, x_n) + \text{cum}(y_1, \dots, y_n)$$

Note that if x_i and y_i were not independent, then by [CP3] there would be $2n$ terms on the right-hand side of $\text{cum}(x_1 + y_1, \dots, x_n + y_n)$. Statistical independence reduces this number to two terms.

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

$$\text{cum}(x_1, \dots, x_n) = 0$$

Proofs of all these properties can be found in Section B of the Appendix in Mendel (1991).

Example. Suppose $Z(k) = aY(k) + V(k)$ where $Y(k)$ and $V(k)$ are statistically independent, and a is a constant; then, from [CP5] and [CP1], $c_n^z(\tau_1, \dots, \tau_n) = a^n c_n^y(\tau_1, \dots, \tau_n) + c_n^v(\tau_1, \dots, \tau_n)$. If, in addition, $V(t)$ is a statistically independent process, then, from [CP6], $c_n^v(\tau_1, \dots, \tau_n) = 0$, so that $c_n^z(\tau_1, \dots, \tau_n) = a^n c_n^y(\tau_1, \dots, \tau_n)$.

Note that $c_n^v(\tau_1, \dots, \tau_n) = 0$ as long as $V(k)$ is Gaussian. This is because all higher-order cumulants of any kind of Gaussian process are zero, be they colored or white processes.

A 1-D slice of the n -th order cumulant is obtained by freezing $(n - 2)$ of its $n - 1$ indexes. Many types of 1-D slices are possible, including radial, vertical, horizontal, diagonal, and offset-diagonal. A diagonal slice is obtained by setting $\tau_i = \tau, i = 1, 2, \dots, n - 1$. All these 1-D slices are very useful in applications of cumulants in signal processing.

A logical question to ask is "Why do we need fourth-order cumulants, i.e., aren't third-order cumulants good enough?" 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. For example, Laplace, uniform, Gaussian, and Bernoulli-Gaussian distributions are symmetric, whereas exponential, Rayleigh and K-distributions are nonsymmetric. Additionally, some processes have extremely small third-order cumulants and much larger fourth-order cumulants; hence, for such processes we would also use the latter. Finally, in some specific applications

(e.g., retrieval of harmonics and cubic phase coupling) third-order cumulants equal zero whereas fourth-order cumulants are nonzero.

Higher-order spectra are defined in terms of either cumulants (e.g., cumulant spectra) or moments (e.g., moment spectra). As explained later, in the case of stochastic signals there are certain advantages to using cumulants; while for deterministic signals, it is better to use moments. Simply stated, higher-order spectra are multi-dimensional Fourier transforms of higher-order statistics. Thus, the power spectrum, bispectrum and trispectrum are defined in terms of cumulants as follows.

Power Spectrum:

$$C_2^x(\omega) = \sum_{\tau=-\infty}^{+\infty} c_2^x(\tau) \exp[-j(\omega\tau)], \quad (26)$$

$|\omega| \leq \pi$ where $c_2^x(\tau)$ is the covariance sequence of $\{X(k)\}$. Eq. 26 is also known as the Wiener-Khinchine theorem. From Eqs. 21 and 26, we have

$$\begin{aligned} c_2^x(\tau) &= c_2^x(-\tau) \\ C_2^x(\omega) &= C_2^x(-\omega) \\ C_2^x(\omega) &\geq 0 \quad (\text{real, nonnegative function}) \end{aligned} \quad (27)$$

Although the power spectrum was previously discussed, its definition and properties are repeated here, so they can be easily compared with the bispectrum and trispectrum.

Bispectrum: $n = 3$

$$C_3^x(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{+\infty} \sum_{\tau_2=-\infty}^{+\infty} c_3^x(\tau_1, \tau_2) \exp\{-j(\omega_1\tau_1 + \omega_2\tau_2)\} \quad (28)$$

$$|\omega_1| \leq \pi, |\omega_2| \leq \pi, |\omega_1 + \omega_2| \leq \pi$$

where $c_3^x(\tau_1, \tau_2)$ is the third-order cumulant sequence of $\{X(k)\}$ described by Eq. 22.

Important symmetry conditions follow from the properties of moments and Eq. 22:

$$\begin{aligned} c_3^x(\tau_1, \tau_2) &= c_3^x(\tau_2, \tau_1) = c_3^x(-\tau_2, \tau_1 - \tau_2) \\ &= c_3^x(\tau_2 - \tau_1, -\tau_1) = c_3^x(\tau_1 - \tau_2, -\tau_2) \\ &= c_3^x(\tau_1, \tau_2 - \tau_1) \end{aligned} \quad (29)$$

As a consequence, knowing the third-order cumulants in any of the six sectors, I through VI, shown in Fig. 7a, would enable us to find the entire third-order cumulant sequence. These sectors include their boundaries so that, for example, *sector I* is an infinite wedge bounded by the lines $\tau_2 = 0$ and $\tau_1 = \tau_2, \tau_1 \geq 0$.

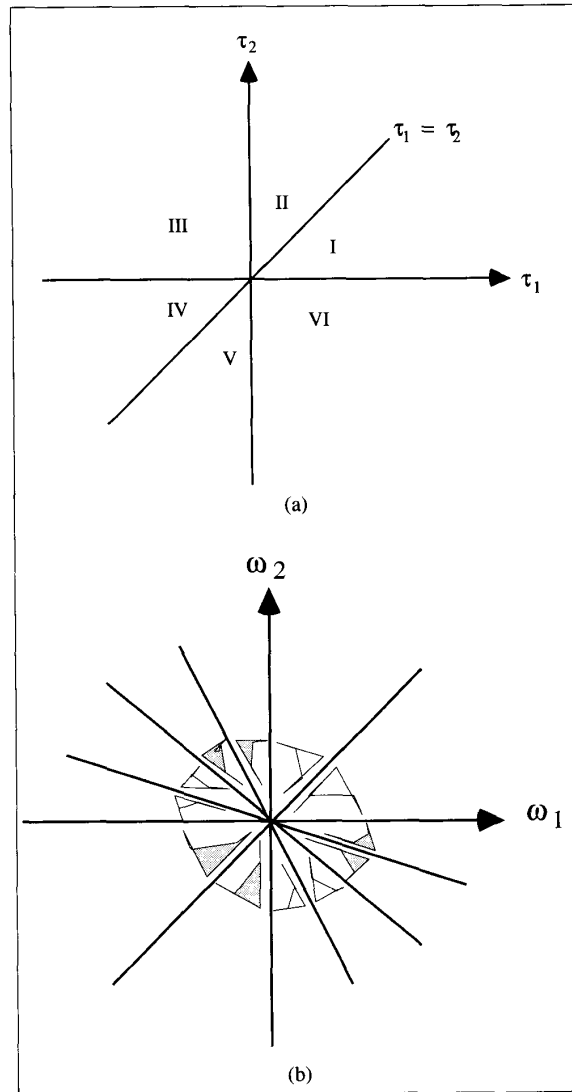
The definition of the bispectrum in Eq. 28 and the properties of third-order cumulants in Eq. 29 give

$$\begin{aligned} C_3^x(\omega_1, \omega_2) &= C_3^x(\omega_2, \omega_1) \\ &= C_3^{x*}(-\omega_2, -\omega_1) = C_3^x(-\omega_1 - \omega_2, \omega_2) \\ &= C_3^x(\omega_1, -\omega_1 - \omega_2) = C_3^x(-\omega_1 - \omega_2, \omega_1) \\ &= C_3^x(\omega_2, -\omega_1 - \omega_2). \end{aligned} \quad (30)$$

Thus knowledge of the bispectrum in the triangular region $\omega_2 \geq 0, \omega_1 \geq \omega_2, \omega_1 + \omega_2 \leq \pi$ shown in Fig. 7b is enough for a complete description of the bispectrum. For real processes, the bispectrum has 12 symmetry regions.

Trispectrum: $n = 4$

$$\begin{aligned} C_4^x(\omega_1, \omega_2, \omega_3) &= \\ \sum_{\tau_1=-\infty}^{+\infty} \sum_{\tau_2=-\infty}^{+\infty} \sum_{\tau_3=-\infty}^{+\infty} c_4^x(\tau_1, \tau_2, \tau_3) \exp\{-j(\omega_1\tau_1 + \omega_2\tau_2 + \omega_3\tau_3)\} \\ |\omega_1| \leq \pi, |\omega_2| \leq \pi, |\omega_3| \leq \pi, |\omega_1 + \omega_2 + \omega_3| \leq \pi \end{aligned} \quad (31)$$



7. (a) Symmetry regions of third-order cumulants. (b) Symmetry regions of the bispectrum (Nikias and Petropulu, 1993).

where $c_4^x(\tau_1, \tau_2, \tau_3)$ is the fourth-order cumulant sequence given by Eq. 23. From the Eq. 31's definition of fourth-order cumulants, a lot of symmetry properties can be derived for the trispectrum, similar to those of the bispectrum. Pflug et. al. [1992] point out that the trispectrum of a real signal has 96 symmetry regions.

The bispectrum and trispectrum are generally complex functions, i.e., they have magnitude and phase:

$$\begin{aligned} C_n^x(\omega_1, \omega_2, \dots, \omega_{n-1}) &= \\ \left| C_n^x(\omega_1, \omega_2, \dots, \omega_{n-1}) \right| \cdot \exp j\psi_n^x(\omega_1, \omega_2, \dots, \omega_{n-1}) \\ n &= 3, 4, \dots \end{aligned} \quad (32)$$

Cumulant spectra are more useful in the processing of random signals than are moment spectra, because: (a) cumu-

Table 3. Higher-Order Moment Spectra of Deterministic Signals

We give the definitions, as in Nikias and Petropulu [1993, Ch. 3], of the bispectrum and trispectrum of deterministic signals (energy and periodic).

◆ ENERGY SIGNALS

Let $\{x(k)\}$ $k = 0, \pm 1, \pm 2, \dots$ be a finite energy deterministic signal. Then

$$\text{Fourier transform: } X(\omega) = \sum_{k=-\infty}^{\infty} x(k)e^{-j\omega k}$$

$$\text{Energy spectrum: } M_2^x(\omega_1, \omega_2) = X(\omega)X^*(\omega)$$

$$\text{Bispectrum: } M_3^x(\omega_1, \omega_2) = X(\omega_1)X(\omega_2)X^*(\omega_1 + \omega_2)$$

$$\text{Trispectrum: } M_4^x(\omega_1, \omega_2, \omega_3) = X(\omega_1)X(\omega_2)X(\omega_3)X^*(\omega_1 + \omega_2 + \omega_3)$$

◆ PERIODIC SIGNALS

Let $\hat{x}(k)$ be a periodic signal with period N , i.e., $\hat{x}(k) = \hat{x}(k + N)$. Then

$$\text{Fourier series: } X(\lambda) = \sum_{k=0}^{N-1} \hat{x}(k)e^{-j\frac{2\pi}{N}\lambda k}; \quad \lambda = 0, 1, \dots, N-1$$

$$\text{Power spectrum: } R_2^{\hat{x}}(\lambda) = \frac{1}{N} X(\lambda)X^*(\lambda)$$

$$\text{Bispectrum: } R_3^{\hat{x}}(\lambda_1, \lambda_2) = \frac{1}{N} X(\lambda_1)X(\lambda_2)X^*(\lambda_1 + \lambda_2)$$

$$\text{Trispectrum: } R_4^{\hat{x}}(\lambda_1, \lambda_2, \lambda_3) = \frac{1}{N} X(\lambda_1)X(\lambda_2)X(\lambda_3)X^*(\lambda_1 + \lambda_2 + \lambda_3)$$

lant spectra of order $n > 2$ are zero if the signal is Gaussian and thus non-zero cumulant spectra provide a measure of extent of non-Gaussianity; (b) cumulants provide a suitable measure of extent of statistical dependence in time series; (c) just as the covariance function of white noise is an impulse function and its spectrum is flat, the cumulants of (higher-order) white noise are multidimensional impulse functions and the polyspectra of this noise is multidimensionally flat; and (d) the cumulant of two statistically independent random processes equals the sum of the cumulants of the individual random processes, whereas the same is not true for higher-order moments. This last property lets us work with the cumulant very easily as an operator.

A normalized higher-order spectrum or n th-order coherency index is a function that combines the cumulant spectrum of order n with the power spectrum ($n = 2$) of a signal. The 3rd- and 4th-order coherency indexes are respectively defined by

$$P_3^x(\omega_1, \omega_2) = \frac{C_3^x(\omega_1, \omega_2)}{\sqrt{C_2^x(\omega_1)C_2^x(\omega_2)C_2^x(\omega_1 + \omega_2)}} \quad (33a) \quad \text{(biocoherency)}$$

$$P_4^x(\omega_1, \omega_2, \omega_3) = \frac{C_4^x(\omega_1, \omega_2, \omega_3)}{\sqrt{C_2^x(\omega_1)C_2^x(\omega_2)C_2^x(\omega_3)C_2^x(\omega_1 + \omega_2 + \omega_3)}} \quad (33b) \quad \text{(tricoherency)}$$

These functions are very useful in the detection and characterization of nonlinearities in time series and in *discriminating linear processes from nonlinear ones*. In fact, a signal is said to be a linear non-Gaussian process of order n if the magnitude of the n th-order coherency, $|P_n^x(\omega_1, \dots, \omega_{n-1})|$, is constant over all frequencies; otherwise, the signal is said to be a non-linear process [Hinich, 1982; Raghuveer and Nikias, 1985].

The higher-order spectrum of a linear non-Gaussian signal, $C_n^x(\omega_1, \omega_2, \dots, \omega_{n-1})$, can always be written in the form

$$C_n^x(\omega_1, \dots, \omega_{n-1}) = \gamma_n^x H(\omega_1) \dots H(\omega_{n-1}) H^*(\omega_1 + \dots + \omega_{n-1}) \quad (34)$$

where γ_n^x is a scalar constant and $H(\omega)$ is the frequency transfer function of a linear time-invariant (LTI) system. If a non-Gaussian signal, $X(k)$, is generated by exciting a LTI system, $H(\omega)$, with non-Gaussian white noise, $V(k)$, as shown later, then Eq. 34 is true for all orders. Special cases include

$$C_2^x(\omega) = \gamma_2^y |H(\omega)|^2 \quad (\text{Power Spectrum}) \quad (35a)$$

$$C_3^x(\omega_1, \omega_2) = \gamma_3^y H(\omega_1) \cdot H(\omega_2) \cdot H^*(\omega_1 + \omega_2) \quad (\text{Bispectrum}) \quad (35b)$$

$$C_4^x(\omega_1, \omega_2, \omega_3) = \gamma_4^y H(\omega_1) \cdot H(\omega_2) \cdot H(\omega_3) \cdot H^*(\omega_1 + \omega_2 + \omega_3) \quad (\text{Trispectrum}) \quad (35c)$$

However, there are non-Gaussian signals that can be linear in the bispectrum domain, i.e.,

$$C_3^x(\omega_1, \omega_2) = T(\omega_1) T(\omega_2) T^*(\omega_1 + \omega_2)$$

but non-linear in the trispectrum or higher-order spectrum domains.

Although the power spectrum of either a linear or non-linear process can always be written in the form of Eq. 35a, higher-order spectra can take the form of Eq. 34 only if the process is linear (has flat magnitude coherency). Combining Eq. 35a with Eq. 35b and Eq. 35c, we see that the power spectrum of a linear process can be reconstructed from its bispectrum or trispectrum up to a constant term (provided $H(0) \neq 0$), i.e.,

$$C_3^x(\omega, 0) = C_2^x(\omega) H(0) \cdot \left(\frac{\gamma_3^y}{\gamma_2^y} \right) \quad (36a)$$

$$C_4^x(\omega_1, \omega_2, 0) = C_3^x(\omega_1, \omega_2) H(0) \cdot \left(\frac{\gamma_4^y}{\gamma_3^y} \right) \quad (36b)$$

as well as the bispectrum from its trispectrum using

$$C_4^x(\omega_1, \omega_2, 0) = C_3^x(\omega_1, \omega_2) H(0) \cdot \left(\frac{\gamma_4^y}{\gamma_3^y} \right) \quad (37)$$

Equations 36a and 36b can be very useful in obtaining better estimates of the power spectrum of the signal from its bispectrum/trispectrum when the additive noise is Gaussian.

If $W(k)$ is a stationary non-Gaussian process with $E\{W(k)\} = 0$ and n th-order cumulant sequence

$$C_n^w(\tau_1, \dots, \tau_{n-1}) = \gamma_n^w \cdot \delta(\tau_1, \dots, \tau_{n-1}) \quad (38a)$$

where $\delta(\tau_1, \dots, \tau_{n-1})$ is the delta function, then $W(k)$ is said to be n th-order white. By taking $(n-1)$ -dimensional Fourier transform of Eq. 38a, we obtain

$$C_n^w(\omega_1, \dots, \omega_{n-1}) = \gamma_n^w \quad (38b)$$

which is a flat spectrum for all frequencies. Important special cases of higher-order spectra of a white process include

$$\begin{aligned} C_2^w(\omega) &= \gamma_2^w \quad (\text{Power Spectrum}) \\ C_3^w(\omega_1, \omega_2) &= \gamma_3^w \quad (\text{Bispectrum}) \\ C_4^w(\omega_1, \omega_2, \omega_3) &= \gamma_4^w \quad (\text{Trispectrum}) \end{aligned} \quad (39)$$

where γ_2^w , γ_3^w , and γ_4^w are the variance, skewness and kurtosis of $\{W(k)\}$, respectively.

The definitions of higher-order moment spectra of deterministic signals are summarized in Table 3.

b. Higher-Order Spectra Computation from Data

The problem met in practice is one of estimating the higher-order spectra of a process when a finite set of measurements is given. Two of the most popular conventional approaches are the *indirect* and *direct* methods which may be seen as direct approximations of the definition of higher-order spectra. While these approximations are straightforward, sometimes the required computations may be expensive despite the use of fast Fourier Transform (FFT) algorithms. We describe both conventional methods here as bispectrum estimators. Their extension to trispectrum estimation is described in [Nikias and Petropulu, 1993].

Indirect Method

Let $\{X(1), X(2), \dots, X(N)\}$ be the given data set. Then we have the following:

1. Segment the data into K records of M samples each, i.e., $N = KM$.
2. Subtract the average value of each record.
3. Assuming that $\{x^{(i)}(k), k = 0, 1, \dots, M-1\}$ is the data set per segment $i = 1, 2, \dots, K$, obtain an estimate of the third-moment sequence

$$r^{(i)}(m, n) = \frac{1}{M} \sum_{\ell=s_1}^{s_2} x^{(i)}(\ell) x^{(i)}(\ell+m) x^{(i)}(\ell+n)$$

where

$$i = 1, 2, \dots, K$$

$$s_1 = \max(0, -m, -n)$$

$$s_2 = \min(M-1, M-1-m, M-1-n).$$

4. Average $r^{(i)}(m, n)$ over all segments

$$\hat{c}_3^x(m, n) = \frac{1}{K} \sum_{i=1}^K r^{(i)}(m, n)$$

5. Generate the bispectrum estimate

$$\hat{C}_3^x(\omega_1, \omega_2) = \sum_{m=-L}^L \sum_{n=-L}^L \hat{c}_3^x(m, n) W(m, n) \exp[-j(\omega_1 m + \omega_2 n)]$$

where $L < M-1$ and $W(m, n)$ is a two-dimensional window function. Let us note that the computations of the bispectrum estimate may be substantially reduced if the symmetry properties of third-order cumulants (Eq. 29) are taken into account for the calculations of $r^{(i)}(m, n)$ and if the symmetry properties of the bispectrum shown in Eq. 30 are also incorporated in the computations.

As in the case of conventional power spectrum estimation, to get better estimates, suitable windows should be used. The window functions should satisfy the following constraints:

- a) $W(m, n) = W(n, m) = W(-m, n-m) = W(m-n, n)$ (symmetry properties of third order cumulants);
b) $W(m, n) = 0$ outside the region of support of $\hat{C}_3^x(m, n)$;
c) $W(0, 0) = 1$ (normalizing condition);
d) $W(\omega_1, \omega_2) \geq 0$, for all (ω_1, ω_2) .

A class of functions which satisfies the constraints for $W(m, n)$ is:

$$W(m, n) = d(m)d(n)d(n-m) \quad (40)$$

where

$$\begin{aligned} d(m) &= d(-m) \\ d(m) &= 0, m > L \\ d(0) &= 1 \end{aligned}$$

$$D(\omega) \geq 0, \text{ for all } \omega$$

Equation 40 allows a reconstruction of two-dimensional window functions for bispectrum estimation using standard one-dimensional lag windows. Two good choices of 1-D windows [Nikias and Raghuveer, 1987] are:

- a) Optimum window (minimum bispectrum bias supremum):

$$d_0(m) = \begin{cases} \frac{1}{\pi} \left| \sin \frac{\pi m}{L} \right| + \left(1 - \frac{|m|}{L}\right) \cos \frac{\pi m}{L}, & |m| \leq L \\ 0, & |m| > L. \end{cases} \quad (41a)$$

- b) Parzen window:

$$d_p(m) = \begin{cases} 1 - 6\left(\frac{|m|}{L}\right)^2 + 6\left(\frac{|m|}{L}\right)^3, & |m| \leq \frac{L}{2} \\ 2\left(1 - \frac{|m|}{L}\right)^3, & \frac{L}{2} \leq |m| \leq L \\ 0, & |m| > L. \end{cases} \quad (41b)$$

More windows and their properties for higher-order spectrum estimation can be found in [Nikias and Petropulu, 1993].

Direct Method

Let $\{X(1), X(2), \dots, X(N)\}$ be the available set of observations for bispectrum estimation. Let us assume that f_s is the sampling frequency and $\Delta_0 = f_s/N_0$ is the required spacing between frequency samples in the bispectrum domain along horizontal or vertical directions; thus, N_0 is the total number of frequency samples. Then we have the following:

- a) Segment the data into K segments of M samples each, i.e., $N = KM$, and subtract the average value of each segment. If necessary, add zeros at the end of each segment to obtain a convenient length M for the FFT.
b) Assuming that $\{x^{(i)}(k), k = 0, 1, 2, \dots, M-1\}$ are the data of segment $\{i\}$, generate the DFT coefficients

$$Y^{(i)}(\lambda) = \frac{1}{M} \sum_{k=0}^{M-1} x^{(i)}(k) \exp(-j2\pi k\lambda/M), \lambda = 0, 1, \dots, \frac{M}{2}$$

$$i = 1, 2, \dots, K.$$

- c) In general, $M = M_1 \times N_0$, where M_1 is a positive integer (assumed odd number), i.e., $M_1 = 2L_1 + 1$. Since M is even and M_1 is odd, we compromise on the value of N_0 (closest integer). Form the bispectrum estimate based on the DFT coefficients

$$\hat{b}_i(\lambda_1, \lambda_2) = \frac{1}{\Delta_0^2} \sum_{k_1=-L_1}^{L_1} \sum_{k_2=-L_1}^{L_1} Y^{(i)}(\lambda_1+k_1) Y^{(i)}(\lambda_2+k_2) Y^{(i)*}(\lambda_1+\lambda_2+k_1+k_2)$$

over the triangular region $0 \leq \lambda_2 \leq \lambda_1, \lambda_1 + \lambda_2 = f_s/2$. For the special case where no averaging is performed in the bispectrum domain, $M_1 = 1, L_1 = 0$, and therefore:

$$\hat{b}_i(\lambda_1, \lambda_2) = \frac{1}{\Delta_0^2} Y^{(i)}(\lambda_1) Y^{(i)}(\lambda_2) Y^{(i)*}(\lambda_1 + \lambda_2)$$

- d) The bispectrum estimate of the given data is the average over the K pieces

$$\hat{C}_3^x(\omega_1, \omega_2) = \frac{1}{K} \sum_{i=1}^K \hat{b}_i(\omega_1, \omega_2)$$

where

$$\omega_1 = \left(\frac{2\pi f_s}{N_0}\right)(\lambda_1) \text{ and } \omega_2 = \left(\frac{2\pi f_s}{N_0}\right)(\lambda_2).$$

c. Properties of Conventional Estimators and Asymptotic Behavior

The statistical properties of the indirect and direct conventional methods for higher-order spectrum estimation have been studied by Rosenblatt and Van Ness [1965], Van Ness [1968], Brillinger and Rosenblatt [1967], Rao and Gabr [1984]; their implications in signal processing have been studied by Nikias and Petropulu [1993].

Let us consider the case of the bispectrum. Assume that $C_2^x(\omega)$ and $C_3^x(\omega_1, \omega_2)$ are respectively the true power spectrum and bispectrum of a stationary zero-mean signal. Let $\hat{C}_3^x(\omega_1, \omega_2)$ be a consistent bispectrum estimate computed by indirect or direct conventional methods using a single realization of the signal of length N . The key result associated with these methods is that for sufficiently large record size M and total length N , both direct and indirect approaches provide approximately unbiased estimates; viz.:

$$E[\hat{C}_3^x(\omega_1, \omega_2)] \equiv C_3^x(\omega_1, \omega_2) \quad (42)$$

with asymptotic variances

$$\begin{aligned} \text{var}[\text{Re}[\hat{C}_3^x(\omega_1, \omega_2)]] &\equiv \text{var}[\text{Im}[\hat{C}_3^x(\omega_1, \omega_2)]] \\ &\equiv \frac{1}{2} \sigma_2^3(\omega_1, \omega_2), \end{aligned} \quad (43a)$$

where

$$\sigma_3^2(\omega_1, \omega_2) = \begin{cases} \frac{VL^2}{MK} C_2^x(\omega_1) C_2^x(\omega_2) C_2^x(\omega_1 + \omega_2) & \text{(indirect)} \\ \frac{N_0^2}{MK} C_2^x(\omega_1) C_2^x(\omega_2) C_2^x(\omega_1 + \omega_2) & \text{(direct)} \end{cases} \quad (43b)$$

for $0 < \omega_2 < \omega_1$, where K is the number of records, M the number of samples per record and V is the total energy of the bispectrum window, which is unity for a rectangular window; L is defined in step (5) of the indirect method and $N_0 = (M/M_1)$ is defined in the description of the direct method. From Eqs. 43a and 43b, it is apparent that if a rectangular window is used with the indirect method and $L = N_0$, the two conventional methods give approximately the same estimates.

Brillinger and Rosenblatt [1967] and Rosenblatt [1985] showed that for large M and N , the error bicoherence

$$\frac{\hat{C}_3^x(\omega_1, \omega_2) - C_3^x(\omega_1, \omega_2)}{\sigma_3(\omega_1, \omega_2)} \sim N_c(0, 1) \quad (44)$$

is approximately complex Gaussian variant with mean zero and variance equal to unity. Another equally important large sample result that follows from the asymptotic results developed by Brillinger and Rosenblatt [1967] is that these statistics can be treated as independent random variables over the grid in the principal domain if the grid width is larger than or equal to the bispectrum bandwidth; i.e., $\hat{C}_3^x(\omega_j, \omega_k)$ and $\hat{C}_3^x(\omega_r, \omega_s)$ are independent for $j \neq r$ or $k \neq s$ if $|\omega_{j+1} - \omega_j| \geq 2\pi\Delta_0$ or $|\omega_{r+1} - \omega_r| \geq 2\pi\Delta_0$, where

$$\Delta_0 \triangleq \begin{cases} \frac{1}{L} & \text{(indirect)} \\ \frac{1}{N_0} & \text{(direct)}. \end{cases} \quad (45)$$

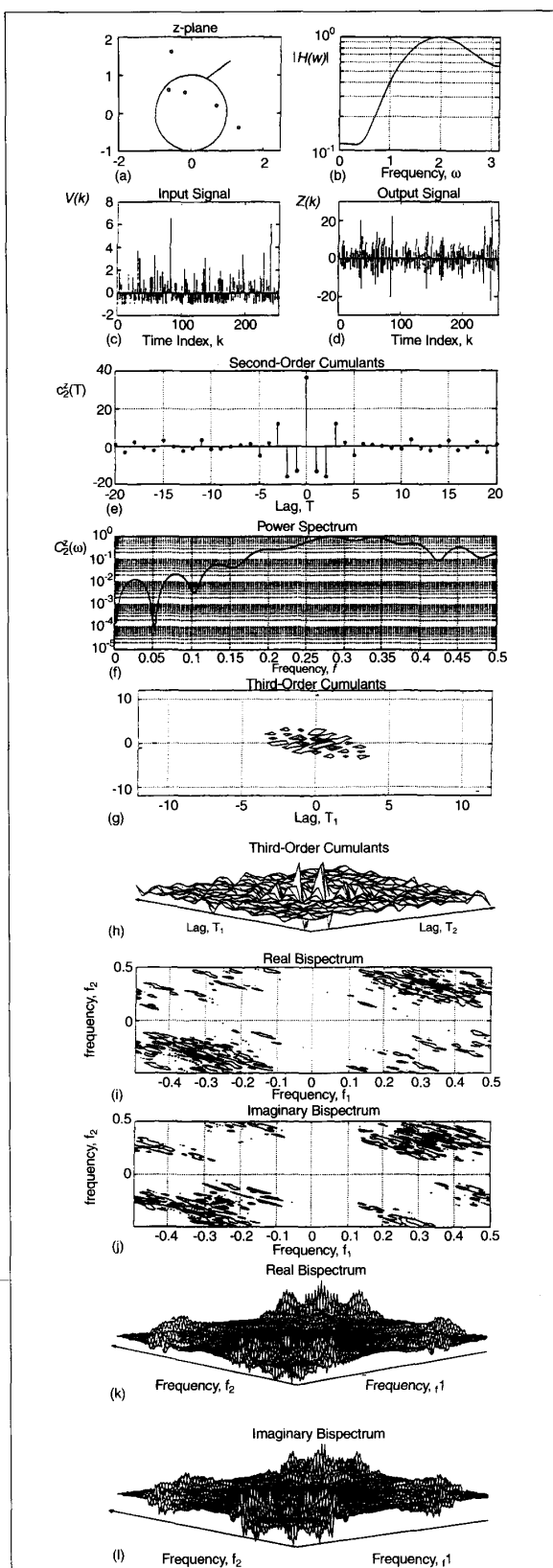
The asymptotic independence and Gaussianity imply that the magnitude squared bicoherence statistic [Hinich, 1982]

$$ch_s(\omega_1, \omega_2) = \frac{|\hat{C}_3^x(\omega_1, \omega_2) - C_3^x(\omega_1, \omega_2)|^2}{\sigma_3^2(\omega_1, \omega_2)} \quad (46)$$

is approximately a noncentral chi-square statistic with 2 degrees of freedom. Hinich [1985] reported that this approximation holds for samples as small as $N = 256$ if $\Delta_0 \equiv \sqrt{N}$. If the process has zero bispectrum $C_3^x(\omega_1, \omega_2) = 0$, then

$ch_s(\omega_1, \omega_2)$ is central chi-square variant with 2 degrees of freedom.

8. (a) Zeros of a nonminimum phase system, (b) magnitude response of the nonminimum phase system, (c) input random sequence (zero-mean one-sided exponential distributed with variance 1), (d) output sequence, (e) 2nd-order cumulants of the output, (f) power spectrum of the output, (g) Contour of the 3rd-order cumulants of the output, (h) 3-D plot of the 3rd-order cumulants of the output, (i) contour of the real part of the bispectrum of the output, (j) contour of the imaginary part of the bispectrum of the output, (k) 3-D plot of the real part of the bispectrum of the output, (l) 3-D plot of the imaginary part of the bispectrum of the output.



From Eqs. 43a and 43b, it is apparent that the bispectrum variance associated with conventional estimators can be reduced by: (i) increasing the number of records, (ii) reducing the size of the region of support of the window in the cumulant domain (L) or increasing the size of the frequency smoothing window (M_1), and (iii) increasing the record size M . However, increasing the number of records (K) is demanding on the computer time and may introduce potential nonstationarities. On the other hand, frequency domain averaging over large rectangles of size (M_1) or use of cumulant domain windows with small values of L_0 reduce frequency resolution and may increase bias. In the case of "short length" data, K could be increased by using overlapping records [Nikias and Raghuvver, 1987].

The conventional methods have the advantage of ease of implementation (FFT algorithms can be used) and provide good estimates for sufficiently long data. However, because of the "uncertainty principle" of the Fourier transform, the ability of the conventional methods to resolve harmonic components in polyspectra domains is limited. Raghuvver and Nikias [1985] point out that this could pose a problem in detecting quadratic phase coupling at closely-spaced frequency pairs.

One of the key advantages of conventional bispectrum (or higher-order spectrum) estimates is their asymptotic Gaussian properties illustrated by Eqs. 44–46. These asymptotic results serve as the bridge between Likelihood Ratio Test (LRT) detectors and Maximum Likelihood (ML) theory on the one hand, and higher-order spectra (HOS) on the other hand. Detection, parameter estimation and classification schemes can be developed for important signal processing applications using the asymptotic statistical properties of HOS estimates and LRT or ML theory. See, for example, Hinich [1985], Forster and Nikias [1991], and Giannakis and Tsatsanis [1992].

Linear Processes

a. Cumulants and Polyspectra

A major generalization to Eqs. 15 and 16 was established by Bartlett (1955) and Brillinger and Rosenblatt (1967). In this case the system in Fig. 4 is assumed to be causal and exponentially stable, and $\{V(k)\}$ is assumed to be independent, identically distributed (i.i.d.) and non-Gaussian, i.e.,

$$c_n^V(\tau_1, \dots, \tau_{n-1}) = \begin{cases} \gamma_n^V, & \tau_1 = \tau_2 = \dots = \tau_{n-1} = 0 \\ 0, & \text{otherwise} \end{cases} \quad (47)$$

where γ_n^V denotes the n th-order cumulant of $V(k)$. Additive noise $N(k)$ is assumed to be Gaussian, but need not be white. Their generalizations to Eqs. 15 and 16 are

$$c_n^Z(\tau_1, \dots, \tau_{n-1}) = \gamma_n^V \sum_{k=0}^{\infty} h(k)h(k+\tau_1) \dots h(k+\tau_{n-1}) \quad (48)$$

and

$$C_n^Z(\omega_1, \dots, \omega_{n-1}) = \gamma_n^V H(\omega_1)H(\omega_2) \dots H(\omega_{n-1})H\left(-\sum_{i=1}^{n-1} \omega_i\right) \quad (49a)$$

which can also be expressed in terms of multidimensional z -transforms, as

$$C_n^Z(z_1, \dots, z_{n-1}) = \gamma_n^V H(z_1)H(z_2) \dots H(z_{n-1})H\left(-\prod_{i=1}^{n-1} z_i^{-1}\right). \quad (49b)$$

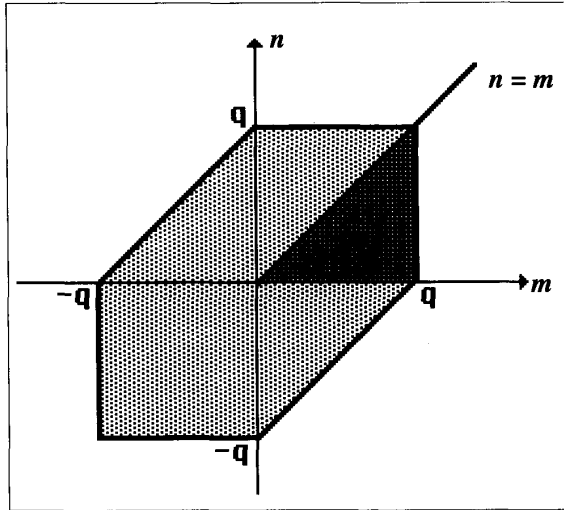
Observe that when $n = 2$, Eqs. 48 and 49 reduce to Eq. 15 [subject to the addition of $c_2^N(\tau)$] and Eq. 16 [subject to the addition of $C_2^N(z)$]. Equations 49a and 49b have been the starting points for many nonparametric and parametric polyspectral techniques that have been developed during the past few years [e.g., Nikias and Raghuvver (1987) and Mendel (1988, 1991)]. One very important use for Eq. 48 is to compute cumulants for models of systems. The procedure for doing this is: (1) determine the model's IR, $h(k)$ $k = 0, 1, 2, \dots, K$; (2) fix the τ_j values, and evaluate Eq. 48; and, (3) repeat step (2) for all τ_j values in the domain of support of interest (be sure to use the symmetry properties of cumulants to reduce computations). This is how the results shown in Fig. 8g and h were obtained. An important use of Eq. 49a is to compute polyspectra for models of systems. The procedure for doing this is: (1) determine the model's IR, $h(k)$ $k = 0, 1, 2, \dots, K$; (2) compute the DFT of $h(k)$, $H(\omega)$; (3) fix the ω_j values, and evaluate Eq. 49a; and, (4) repeat Step (3) for all ω_j values in the domain of support of interest (be sure to use the symmetry properties of polyspectra to reduce computations). Of course, another way to compute the polyspectra is to first compute the cumulants, as just described, and then compute their multi-dimensional DFT.

The generalization of Eqs. 48 and 49 to the case of colored non-Gaussian input is more easily visualized in the polyspectral domain [Bartlett (1955) only considers the $n = 2, 3, 4$ cases; Brillinger and Rosenblatt (1967) provide results for all n], and is:

$$C_n^Z(\omega_1, \dots, \omega_{n-1}) = C_n^V(\omega_1, \dots, \omega_{n-1}) H(\omega_1) H(\omega_2) \dots H(\omega_{n-1}) H\left(-\sum_{i=1}^{n-1} \omega_i\right) \quad (50)$$

Derivations of Eqs. 48–50 can also be found in Section C of the Appendix in Mendel (1991). The generalization of 48 to the colored noise case is also given in Mendel (1991); it is a $(n-1)$ -fold convolution between $c_n^V(\tau_1^{2x}, \dots, \tau_{n-1}^{2x})$ and $c_n^h(\tau_1^{2x}, \dots, \tau_{n-1}^{2x}) = \sum h(k)h(k+\tau_1) \dots h(k+\tau_{n-1})$.

Example 1. (Mendel, 1991): Suppose that $h(k)$ is the impulse response (IR) of a causal MA system. Such a system has a finite IR, and is described by the following model:



9. The domain of support for $c_3^y(m, n)$ for an $MA(q)$ system. The dark shaded c_3^y region is the principal region defined in Eq. 54.

$$Y(k) = \sum_{i=0}^q b(i)V(k-i) \quad (51)$$

The MA parameters are $b(0), b(1), \dots, b(q)$, where q is the order of the MA model, and $b(0)$ is usually assumed equal to unity [the scaling is absorbed into the statistics of $V(k)$]. It is well known that for this model $h(i) = b(i), i = 0, 1, \dots, q$; hence, when $\{V(k)\}$ is i.i.d., we find from Eq. 48 that

$$c_3^y(\tau_1, \tau_2) = \gamma_3^y \sum_{l=0}^q b(l)b(l+\tau_1)b(l+\tau_2) \quad (52)$$

and

$$c_4^y(\tau_1, \tau_2, \tau_3) = \gamma_4^y \sum_{l=0}^q b(l)b(l+\tau_1)b(l+\tau_2)b(l+\tau_3). \quad (53)$$

An interesting question is “for which values of τ_1 and τ_2 is $c_3^y(\tau_1, \tau_2)$ nonzero?” Of course, a comparable question can

be asked for $c_4^y(\tau_1, \tau_2, \tau_3)$. The answer to this question is depicted in Fig. 9. The domain of support for $c_3^y(\tau_1 = m, \tau_2 = n)$ is the six-sided shaded region. This is due to the FIR nature of the MA system. The dark shaded triangular region in the first quadrant is the principal region. In the stationary case, we only have to determine third-order cumulant values in this region, R , where

$$R = \{m, n: 0 \leq n \leq m \leq q\} \quad (54)$$

Observe, from Fig. 9, that the third-order cumulant equals zero for either one or both of its arguments equal to $q+1$. This suggests that it should be possible to determine the order of the MA model, q , by testing, in a statistical sense, the smallness of a third-order cumulant such as $c_3^y(q+1, 0)$. See Giannakis and Mendel (1990) to determine how to do this.

Table 1 provided a simple example of three MA systems that have exactly the same autocorrelations and spectra; hence, they cannot be resolved using second-order statistics. Table 4 continues this example. It gives the third-order cumulants for each $MA(2)$ system for $R = \{m, n: 0 \leq n \leq m \leq 2\}$. Observe that the third-order cumulants for the three systems are sufficiently different so that we suspect that it should be possible to resolve them using these cumulants. We shall prove that this is indeed the case below.

b. Polycepstra

Beginning with the bispectrum

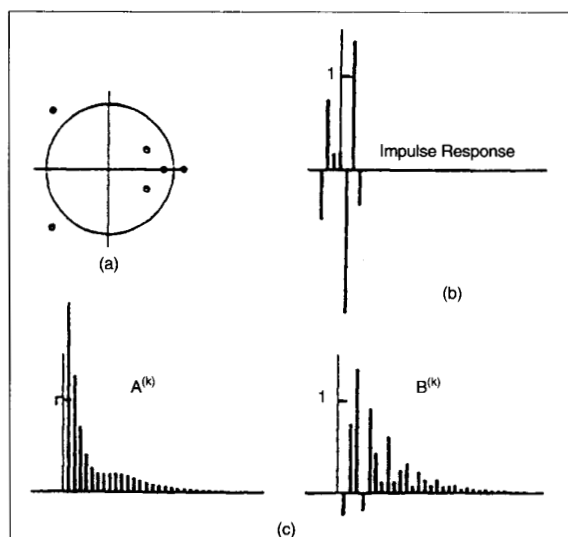
$$C_3^y(z_1, z_2) = \gamma_3^y H(z_1)H(z_2)H(z_1^{-1}, z_2^{-1}), \quad (55)$$

Nikias and Pan (1987) and Pan and Nikias (1988) take the logarithm of $C_3^y(z_1, z_2)$, $\hat{C}_3^y(z_1, z_2) = \ln C_3^y(z_1, z_2)$, and take the inverse 2-D z-transform of $\hat{C}_3^y(z_1, z_2)$ to obtain the **complex bicepstrum** $b_y(m, n)$. Note that these steps parallel the steps which previously led to the complex cepstrum of $h(k)$, $\hat{h}(k)$. Note, also, that comparable results are also obtained by them for the **complex tricepstrum**, $t_y(m, n, l)$. In general, $b_y(m, n)$, $t_y(m, n, l)$, and the even higher-order complex cepstra are known as “polycepstra.”

When $H(z)$ is represented as in Eqs. 7–9 (see, also Fig. 5), then Pan and Nikias have shown that: (1) $b_y(m, n)$ has nonzero values only at: $m = n = 0$, integer values along the m and n

Table 4. Minimum-, maximum- and mixed-phase systems with identical power spectra (or autocorrelations), but with different third-order statistics: $0 < a < 1, 0 < b < 1$ (Nikias and Raghuvver, 1987).

Third-Order Cumulants			
$c_3^y(0, 0)$	$1 - (a+b)^3 + a^3b^3$	$1 - (a+b)^3 + a^3b^3$	$(1+ab)^3 - a^3 - b^3$
$c_3^y(1, 0)$	$-(a+b) + ab(a+b)^2$	$(a+b)^2 - (a+b)a^2b^2$	$a^2(1+ab) - (1+ab)^2b$
$c_3^y(1, 1)$	$(a+b)^2 - (a+b)a^2b^2$	$-(a+b) + ab(a+b)^2$	$-a(1+ab)^2 + (1+ab)b^2$
$c_3^y(2, 0)$	ab	a^2b^2	$-a^2b$
$c_3^y(2, 1)$	$-(a+b)ab$	$-(a+b)ab$	$ab(1+ab)$
$c_3^y(2, 2)$	a^2b^2	ab	$-ab^2$



10. Mixed phase MA system with pronounced resonances: (a) zero location, (b) impulse response, and (c) cepstral coefficients [defined in Eqs. 13 and 14, respectively].

$$b_y(m, n) = \begin{cases} \ln \gamma_3^y c_l & m = 0, n = 0 \\ -1/n A^{(n)} & m = 0, n > 0 \\ -1/m A^{(m)} & n = 0, m > 0 \\ 1/m B^{(-m)} & n = 0, m < 0 \\ 1/n B^{(-n)} & m = 0, n < 0 \\ -1/n B^n & m = n > 0 \\ 1/n A^{(-n)} & m = n < 0 \\ 0 & \text{otherwise} \end{cases} \quad (56)$$

where the "A" and "B" cepstral coefficients are defined in Eqs. 13 and 14, respectively; and, (2) $t_y(m, n, l)$ has nonzero values only at: $m = n = l = 0$, integer values along the m , n and l axes, and at the intersection of these values along the 45-degree line $m = n = l$; for the specific values of $t_y(m, n, l)$, which are comparable to Eq. 56, see Appendix C of Pan and Nikias (1988).

Example 2. (Pan and Nikias, 1988): A noncausal MA(3,3) model with pronounced resonances is depicted in Fig. 10. Its transfer function is

$$H(z) = (1 - 0.869z)(1 + 1.1z + 0.617z^2) \\ (1 - 0.85z^{-1})(1 - 1.2z^{-1} + 0.45z^{-2}) \quad (57)$$

Depicted, also in Fig. 10, are the "A" and "B" cepstral coefficients from which it is straightforward, using Eq. 56, to compute the bicepstral coefficients $b_y(m, n)$.

Table 5. Properties of Cepstral Coefficients

Here we give the properties of cepstrum coefficients as discussed by Nikias and Petropulu [1993, Ch. 5].

Let $h(k)$ be a deterministic energy signal with Fourier transform

$$H(\omega) = |H(\omega)| \exp(j\phi(\omega))$$

and complex cepstrum

$$c_h(m) = F^{-1}[\ln H(\omega)] = \begin{cases} \frac{A^{(m)}}{m}, & m > 0 \\ -\frac{B^{(-m)}}{m}, & m < 0 \end{cases}$$

where $F^{-1}[\bullet]$ denotes inverse Fourier transform.

The cepstrum of $|H(\omega)|$ is a function of $\{A^{(m)} + B^{(m)}\}$ i.e.,

$$p_h^{(m)} = F^{-1}[\ln |H(\omega)|] = \begin{cases} \frac{1}{m} [A^{(m)} + B^{(m)}], & m > 0 \\ \frac{1}{m} [A^{(-m)} + B^{(-m)}], & m < 0 \end{cases}$$

whereas the cepstrum of $e^{j\phi(\omega)}$ is a function of $\{A^{(m)} - B^{(m)}\}$, i.e.,

$$f_h(m) = F^{-1}[\ln \exp(j\phi(\omega))] = \begin{cases} -\frac{1}{m} [A^{(-m)} - B^{(-m)}], & m < 0 \\ -\frac{1}{m} [A^{(m)} - B^{(m)}], & m > 0 \end{cases}$$

The cepstral coefficients $A^{(m)}$, $B^{(m)}$ can be computed directly from higher-order cumulants, without the use of phase unwrapping algorithms by solving a linear system of equations as shown by Pan and Nikias [1988].

The "A" and "B" cepstral coefficients have some very interesting and important properties, which are summarized in Table 5.

c. Identification of Nonminimum Phase Systems

Equation 17 demonstrates that, if we have access to both a system's input and output, then we can reconstruct its IR, $h(k)$, using the cross-correlation function, $r_{VZ}(\tau)$. In some signal processing applications (e.g., equalization, deconvolution in reflection seismology) we only have access to a system's output. Is it possible to determine $h(k)$ (or a scaled version of $h(k)$) just from output measurements? If signals are non-Gaussian, then Lii and Rosenblatt (1982) proved that it is indeed possible to reconstruct the correct phase IR just from output data using higher-order statistics. Their seminal paper has spawned a multitude of new methods for identifying nonminimum phase systems from just noisy output measurements. A very comprehensive survey of these methods is given in Mendel (1991).

Referring to Fig. 4, the problem is: given time-limited noisy measurements $Z(k)$, $k = 1, 2, \dots, N$, estimate $H(z)$'s parameters, when $H(z)$ is represented as in Fig. 5, i.e.,

$$H(z) = B(z)/A(z) = \left[\sum_{j=0}^q b(j)z^{-j} / \sum_{i=0}^p a(i)z^{-i} \right] \quad (58)$$

The parameters to be identified are $b(1), \dots, b(q), a(1), \dots, a(p)$; $a(0) = b(0) = 1$ for scaling purposes, and orders p and q are assumed known. For a discussion of the more realistic case when orders p and q must also be determined, see Giannakis and Mendel (1990).

This output measurement identification problem occurs in many fields, including communications and reflection seismology. In the former, $V(k)$ is a "message," $h(k)$ is a "channel," and $Z(k)$ is a "distorted message" (distorted due to intersymbol interference). An accurate model of the channel is needed in order to restore the message at the receiver. This model is used in many equalization schemes. In reflection seismology $V(k)$ is the earth's "reflectivity sequence" (i.e., the earth's "message") $h(k)$ is the "seismic source wavelet," and $Y(k)$ is the "seismogram." An accurate model of the source wavelet is needed in order to estimate the earth's reflectivity sequence via deconvolution.

When the numerator parameters in Eq. 58 all equal zero except for $b(0)$, we have an all-pole model, in which case we are in the realm of AR parameter estimation. When all of the denominator parameters equal zero except for $a(0)$, we are in the realm of MA parameter estimation. These two special cases have been widely studied not only for their own interest, but also because some methods for estimating ARMA parameters proceed in two steps, by first estimating the AR parameters of the ARMA model and then estimating the MA parameters of the ARMA model, making use of the just-estimated AR parameters.

As a reminder, we use higher-order statistics to solve these identification problems because second-order statistics are phase blind [hence, they can only give rise to minimum phase

or maximum phase models, i.e., to SEMP models], and higher-order statistics are blind to additive Gaussian noise.

Because so many new and interesting methods have been developed for identifying the coefficients of an AR, MA or ARMA model, and these methods are carefully elaborated upon in Mendel (1991), we leave their details for the reader to explore. Here we demonstrate the rather remarkable result that it is possible to determine the parameters of an MA model in closed form. No result of this nature was available before the introduction of higher-order statistics into signal processing.

Example 3. This is a continuation of Example 1. Earlier, we showed that if we have access to both the input and output of a system, then we can reconstruct the correct phase IR, $h(k)$, using Eq. 17. Is there a comparable result, with a simple closed-form formula, for reconstructing the correct phase IR, $h(k)$, using just output measurements? Giannakis (1987) was the first to show that the IR of a q th-order MA system can be calculated just from the system's output cumulants using the following simple closed-form formula (stated here in terms of third-order cumulants; generalizations to arbitrary order cumulants can be found in Swami and Mendel [1988, 1990 (Eq.13)]).

$$h(k) = c_3^z(q, k) / c_3^z(q, 0) \quad k = 0, 1, \dots, q \quad (59)$$

A proof of this interesting result is given in Table 6.

Equation 59 is often referred to as the " **$C(q, k)$ formula**." Lohmann, et al (1983) and Lohmann and Wirmitzer (1984) provide a non-rigorous derivation of the $C(q, k)$ formula for 1-D and 2-D continuous-time processes. Note that Eq. 59 only uses the 1-D slice of the output cumulant along the right-hand side of the darkly shaded right triangle in Fig. 8. Note, also, that it requires exact knowledge of MA order q . It is interesting and important from a theoretical point of view, but is impractical from an actual computation point of view. This is because, in practice, the output cumulant must be estimated, and Eq. 59 does not provide any filtering to reduce the effects of cumulant estimation errors. Fortunately,

Table 6. Derivation of the $C(q, k)$ formula

Here, as in Giannakis (1987), we derive the $C(q, k)$ formula for third-order cumulants. We begin with (Eq. 48) for $k = 3$, i.e.,

$$c_3^z(\tau_1, \tau_2) = \gamma^3 \sum_{k=0}^{\infty} h(k)h(k+\tau_1)h(k+\tau_2) \quad (1)$$

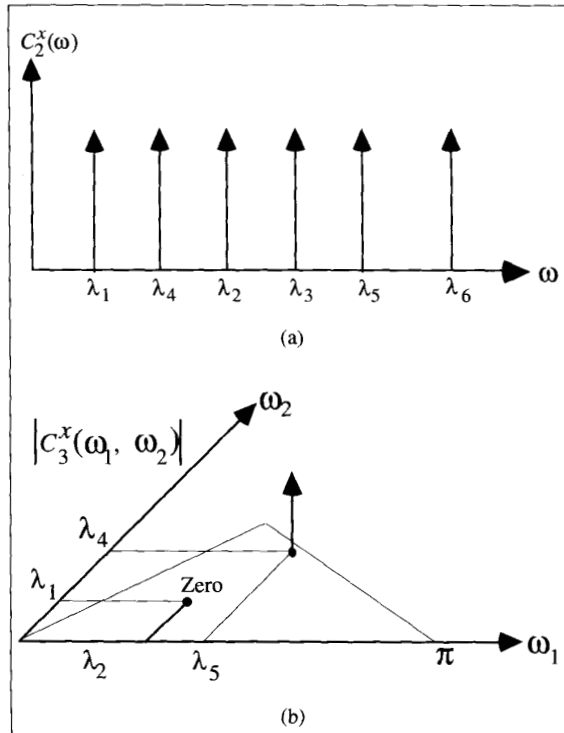
in which $h(0) = 1$, for normalization purposes. Set $\tau_1 = q$ and $\tau_2 = k$ in (1), and use the fact that for an MA(q) system $h(j) = 0$ $j > q$, to see that

$$c_3^z(q, k) = \gamma^3 h(q)h(k) \quad (2)$$

Next, set $\tau_1 = q$ and $\tau_2 = 0$ in (1), to see that

$$c_3^z(q, 0) = \gamma^3 h(q) \quad (3)$$

Dividing (2) by (3) we obtain the $C(q, k)$ formula (Eq. 59).



11. Quadratic phase coupling. (a) Power spectrum of the process described by Eq. 60. (b) Its magnitude bispectrum.

other cumulant-based methods have been developed that do provide such filtering.

Nonlinear Processes

Higher-order moment and cumulant spectra or polyspectra provide a means of detecting and quantifying nonlinearities in stochastic signals. These stochastic signals usually arise when a nonlinear system operates under a random input. General relations for arbitrary stationary random data passing through arbitrary linear systems have been studied quite extensively for many years. In principle, most of these relations are based on autocorrelation, power spectrum, or cross-correlation matching criteria. On the other hand, general relations are not available for arbitrary stationary random data passing through arbitrary nonlinear systems. Instead, each type of nonlinearity has been investigated as a special case. Polyspectra can play a key role in detecting and characterizing the type of nonlinearity in a system from its output sequence. In addition, cross-polyspectra may be used for nonlinear system identification from observations of input and output data.

There are situations in practice in which the interaction between two harmonic components causes contribution to the power at their sum and/or difference frequencies. For example, suppose the signal

$$X(k) = A_1 \cos(\lambda_1 k + \theta_1) + A_2 \cos(\lambda_2 k + \theta_2)$$

is passed through the following simple quadratic nonlinear system:

$$Z(k) = X(k) + \varepsilon X^2(k) \quad (60)$$

where ε is a non-zero constant. The signal $Z(k)$ contains cosinusoidal terms in (λ_1, θ_1) , (λ_2, θ_2) , $(2\lambda_1, 2\theta_1)$, $(2\lambda_2, 2\theta_2)$, $(\lambda_1 + \lambda_2, \theta_1 + \theta_2)$ and $(\lambda_1 - \lambda_2, \theta_1 - \theta_2)$. Such a phenomenon, which gives rise to some or all of these phase relations, that are exactly the same as the frequency relations, is called quadratic phase coupling [Kim and Powers, 1978; Raghuveer and Nikias, 1985]. In certain applications it is necessary to determine if peaks at harmonically related positions in the power spectrum are, in fact, phase coupled. Since the power spectrum suppresses all phase relations, it cannot provide the answer. The third-order cumulants (the bispectrum), however, are capable of detecting and characterizing quadratic phase coupling.

Consider the process

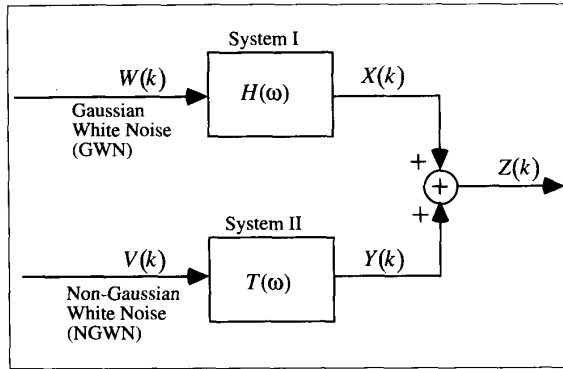
$$X(k) = \sum_{i=1}^6 \cos(\lambda_i k + \phi_i) \quad (61)$$

where $\lambda_2 > \lambda_1 > 0$, $\lambda_5 > \lambda_4 > 0$, $\lambda_3 = \lambda_1 + \lambda_2$, $\lambda_6 = \lambda_4 + \lambda_5$, $\phi_1, \phi_2, \dots, \phi_5$ are all independent, uniformly distributed r.v.'s over $(0, 2\pi)$, and $\phi_6 = \phi_4 + \phi_5$. In Eq. 61 whereas $\lambda_1, \lambda_2, \lambda_3$ and $\lambda_4, \lambda_5, \lambda_6$ are at harmonically related positions, only the component at λ_6 is a result of phase coupling between the components at λ_4 and λ_5 ; additionally, λ_3 is an independent harmonic component. The power spectrum of the process consists of impulses at λ_i , $i = 1, 2, \dots, 6$, as illustrated in Fig. 11. Looking at the spectrum one cannot say if the harmonically related components are, in fact, involved in quadratic phase-coupling relationships. The third-order cumulant sequence $c_3^x(n, \ell)$ of $X(k)$ can be easily obtained as, [Raghuveer and Nikias, 1985]

$$\begin{aligned} c_3^x(n, \ell) = \frac{1}{4} \{ & \cos(\lambda_5 n + \lambda_4 \ell) + \cos(\lambda_6 n + \lambda_4 \ell) \\ & + \cos(\lambda_4 n + \lambda_5 \ell) + \cos(\lambda_6 n - \lambda_5 \ell) \\ & + \cos(\lambda_4 n - \lambda_6 \ell) + \cos(\lambda_5 n - \lambda_6 \ell) \} \end{aligned} \quad (62)$$

It is important to observe that in Eq. 62, *only the phase coupled components appear*. Consequently, the bispectrum evaluated in one of the first quadrant triangular regions of Fig. 7 shows an impulse only at (λ_5, λ_4) indicating that only this pair is phase-coupled (see Fig. 11). In the total absence of phase coupling, the third-order cumulant sequence and hence the bispectrum are both zero. Thus the fact that only phase coupled components contribute to the third-order cumulant sequence of a process is what makes the bispectrum a useful tool for detecting quadratic phase coupling and discriminating phase-coupling components from those that are not.

Any of the existing bispectrum estimation techniques can be used for the analysis of quadratic phase-coupling phenomena; however, each of those techniques will exhibit certain



12. A signal $Z(k)$ whose power spectrum and bispectrum are modeled by different linear systems. The noise processes $W(k)$, $V(k)$ are assumed independent.

advantages and a number of limitations. For example, the conventional techniques can serve as better quantifiers of phase coupling (i.e., degree of coupling) whereas the parametric methods (AR and ARMA) are better as detectors rather than quantifiers [Raghuveer and Nikias, 1985; 1986].

The use of conventional methods for bispectrum estimation in conjunction with the magnitude bicoherence index $|P_3^X(\omega_1, \omega_2)|$ has been used extensively for the detection and quantification of quadratic phase coupling. When the magnitude bicoherence index takes on a value close to unity at a frequency pair where phase coupling has occurred this indicates an almost 100-percent degree of phase coupling. On the other hand, a near-zero value of $|P_3^X(\omega_1, \omega_2)|$ at harmonically related frequency pairs will suggest an absence of phase coupling (an almost zero-percent degree of phase coupling). Certainly, one of the advantages of using the conventional approach to bicoherence index estimation is its ability to serve as a good quantifier by providing good estimates of the degree of phase coupling at harmonically related frequency pairs. For example, the conventional bicoherence estimate of the signal.

$$X(k) = \sum_{i=1}^4 \cos(\lambda_i k + \phi_i) \quad (63)$$

where $\lambda_3 = \lambda_1 + \lambda_2$, $\phi_3 = \phi_1 + \phi_2$, $\lambda_4 = \lambda_1 + \lambda_2$ and ϕ_1, ϕ_2, ϕ_4 are i.i.d., uniformly distributed over $[0, 2\pi]$, will show a peak at (λ_1, λ_2) of magnitude approximately 0.5 indicating only a 50-percent degree of phase coherence. Possible limitations for using conventional methods for the analysis of quadratic phase coupling are the high variance of bispectrum estimates and the low resolution when harmonically related frequency pairs are close to each other [Nikias and Raghuveer, 1986]. Let us note that the effect of these limitations is reduced as the number and/or length of the data segments increase.

Quadratic phase coupling was studied by Raghuveer and Nikias, [1985, 1986], using AR models for bispectrum estimation. The motivation to use AR techniques was to

take advantage of the high-resolution capability and low variance estimates associated with AR modeling. The justification for the use of AR methods for detection of quadratic phase coupling can be found in [Raghuveer and Nikias, 1985]. They have also shown that we cannot have just a single segment of data (i.e., just one set of fixed values for $\phi_1, \phi_2, \phi_3, \phi_4$ in Eq. 63) for detecting quadratic phase coupling between pairs of sinusoids. Segmentation of data into records is necessary to obtain consistent estimates of the third-order cumulants. The advantages of AR techniques arise in situations where the conventional estimators completely fail to resolve closely spaced frequency pairs. On the other hand, the limitation of the AR techniques lies on their inability to provide an accurate estimate of the degree of phase coherence when phase coupling does occur at harmonically related frequency pairs.

Swami and Mendel [1988] have shown that the trispectrum can be used to resolve cubic phase coupling, and, that if a signal contains components due to both quadratic and cubic phase coupling, the bispectrum of that signal is blind to the cubically-coupled components and can resolve the quadratically-coupled components, whereas the trispectrum of that signal is blind to the quadratically-coupled components and can resolve the cubically-coupled components. Hence, higher-order spectra can resolve different types of nonlinearities.

The linear model that can describe the bispectrum of a signal is generally different from the one that describes its power spectrum [Raghuveer and Nikias, 1985]. Consider the process, $Z(k)$, in Fig. 12. As we can see, $Z(k)$ is the sum of two signals, one which is the output of a linear system driven by white Gaussian noise $\{W(k)\}$ and the other, the output of a linear system driven by non-Gaussian white noise, $\{V(k)\}$. Further, $\{V(k)\}$ and $\{W(k)\}$ are independent, which implies that $X(k)$ and $Y(k)$ are also statistically independent, hence the bispectrum of $Z(k)$ is the sum of the bispectra of $X(k)$ and $Y(k)$. Since $X(k)$ is Gaussian, its bispectrum is zero and the bispectrum of $Z(k)$ is the bispectrum $Y(k)$, which means that the linear system, System II, of Fig. 12 models the bispectrum of $Z(k)$. In other words,

$$C_3^Z(\omega_1, \omega_2) = C_3^Y(\omega_1, \omega_2) = \gamma_3^Y T(\omega_1) T(\omega_2) T^*(\omega_1 + \omega_2). \quad (64)$$

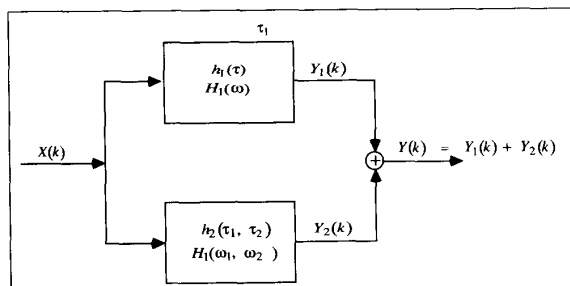
The power spectrum of $Z(k)$ is the sum of the power spectra of $X(k)$ and $Y(k)$. Since generally the contribution of $X(k)$ to the power spectrum of $Z(k)$ is significant, System II (alone) does not model the power spectrum. In this case,

$$C_2^Z(\omega) = C_2^X(\omega) + C_2^Y(\omega) = \gamma_2^X |H(\omega)|^2 + \gamma_2^Y |T(\omega)|^2. \quad (65)$$

Clearly, if $T(\omega)$ and $H(\omega)$ represent two different AR models, the power spectrum of $Z(k)$ is truly an ARMA whereas its bispectrum is AR-type.

Nonlinear processes can also be represented by general Volterra systems [Schetzen, 1989]. We only examine here the popular case of a second-order Volterra model.

Suppose a signal is represented by the 2nd-order Volterra model



13. Second-order Volterra system: linear and quadratic parts in a parallel structure [Nikias and Petropulu, 1993].

$$Y(k) = \sum_{\tau_1} h_1(\tau_1)X(k - \tau_1) + \sum_{\tau_1} \sum_{\tau_2} h_2(\tau_1, \tau_2)X(k - \tau_1)X(k - \tau_2) \quad (66)$$

where $X(k)$ is a stationary random process with zero mean. The identification problem is to determine the impulse response, $h_1(\tau)$ and the kernel $h_2(\tau_1, \tau_2)$. Note that Eq. 61 can be viewed as a parallel connection of a linear system $\{h_1(\tau)\}$ and a quadratic system $\{h_2(\tau_1, \tau_2)\}$, as illustrated in Fig. 13. Assuming that the input signal, $X(k)$, is a stationary, zero-mean Gaussian process, Tick [1961] has established the following fundamental relationships:

$$H_1(\omega) = \frac{C_{xy}(\omega)}{C_x^2(\omega)} \quad (67)$$

and

$$H_2(-\omega_1, -\omega_2) = \frac{C_{xy}(\omega_1, \omega_2)}{2C_x^2(\omega_1)C_x^2(\omega_2)} \quad (68)$$

where $C_{xy}(\omega)$, $C_{xy}(\omega_1, \omega_2)$ are the cross-spectrum and cross-bispectrum, respectively. Consequently, when we have access to the input and output of the system illustrated in Fig. 13, Eqs. 67, and 68 represent the system identification formulas. It is important to remember that Eqs. 67 and 68 are valid only for a Gaussian input signal. More general results, assuming a non-Gaussian input, have been obtained by Hinich [1985] and Powers, et. al. [1989]. Additional results on particular simple nonlinear systems have been reported by Brillinger [1977] and Rozario and Papoulis [1989].

Applications

There have been numerous applications of cumulants and polyspectra during the past 25 years. See [USS Comprehensive Bibliography] for more than 500 references. The applications can be grouped into three major categories: physics-related, 1D, and 2D/3D. The **physics-related applications** deal with oceans (internal waves, noise, shoaling gravity waves, wave coupling, wave interaction, wave phenomena, ship resistance to waves, and sea-surface temperature anomalies), earth (free oscillations), atmosphere

(pressure, turbulence), interplanetary (scintillation), wind (turbulence, currents), plasmas (wave interactions, nonlinear phenomena), electromagnetics (low frequency data), and crystallography (structures). The **1D applications** deal with: diagnosis (of surface roughness, machine faults, noisy mechanical systems), vibration analysis (signal pattern recognition, measurements, knock detection), speech (pitch detection, voiced/unvoiced decisions), noise (cancellation, ship radiated, from gears, bioelectric), system identification (nonminimum phase channels, input/output), detection, phase locking, FM signals, seismic deconvolution, range and doppler extraction, blind equalization, economic time series, brain potentials (eeg wave coupling) and biological rhythms. They also deal with a wide range of problems associated with: harmonic retrieval, nonlinear systems (Volterra, bilinear, phase coupling), array processing, sonar, and radar. The **2D/3D applications** deal with images (modeling, reconstruction, restoration, coding, motion estimation, sequence analysis), textures (model validation, discrimination), tomography (flow velocity field, 3D velocity field), speckle masking in astronomy, inverse filtering of ultrasonic images, and imaging photon-limited data.

In this section we give brief examples of how higher-order spectra have been applied to the following: array processing, classification, harmonic retrieval, time-delay estimation, blind equalization and interference cancellation. These examples are in no way exhaustive, but, instead, are meant to demonstrate the usefulness of higher-order spectra.

a. Array Processing

Array processing techniques play a very important role in the enhancement of signals in the presence of interference. Array processing problems include: direction of arrival (DOA) determination, determination of number of sources, beam forming, estimation of the source signal, source classification, sensor calibration, etc. See Van Veen and Buckley (1988) for an excellent introduction to array processing and its associated models.

So many novel and interesting array processing algorithms have appeared during the past ten years, why would one want to apply higher-order statistics to array processing problems? There are a number of answers to this question, including: (1) one of the most popular and important beamformers, namely, Capon's minimum-variance distortionless response (MVDR) beamformer, that has been the starting point for both signal enhancement and high-resolution DOA estimation, requires very specific and detailed information about the so-called array steering vector (e.g., source steering angles, array geometry, receiver responses), information that is often not available, or if available is not given to a high degree of accuracy; a technique recently developed by Dogan and Mendel (1992a), that is described below, shows how cumulants can be used to estimate the unknown steering vector, after which the MVDR beamformer can be used; (2) When additive noise is colored and Gaussian, a second-order statistics based high-resolution DOA algorithm, such as MUSIC, does not perform well; however a cumulant-based MUSIC

algorithm does perform well; (3) most second-order statistics based beamformers assume that the received signals are not coherent, which rules out the important case of multipath propagation; cumulant-based beamformers can work in the presence of multipath (e.g., Dogan and Mendel, 1992a). The following references apply higher-order statistics to array processing: Cardoso (1989, 1990, 1991a, 1991b), Chiang and Nikias (1989), Comon (1989), Dogan and Mendel (1992a,b, 1993a,b), Duvaut (1990), Forster and Nikias (1990, 1991), Gaeta and Lacoume (1991), Giannakis and Shamsunder (1991), Jutten, et al. (1991), Lacoume and Ruiz (1988), Lagunas and Vazquez (1991), Mohler and Bugnon (1989), Moulines and Cardoso (1991), Porat and Friedlander (1990), Ruiz (1991), Ruiz and Lacoume (1989), Scarano and Jacovitti (1991), and Shamsunder and Giannakis (1991a, 1991b).

Dogan and Mendel (1992a) use cumulants of received signals to estimate the steering vector of a narrowband non-Gaussian desired signal in the presence of directional Gaussian interferers with unknown covariance structure. They assume no knowledge of the array manifold or DOA information about the desired signal. The desired signal could be voiced speech, sonar, radar return or a communication signal. The example below is for a communications scenario and requires the use of fourth-order cumulants, because the third-order cumulants for the communication signals are identically zero.

Consider an array of M elements, with arbitrary sensor response characteristics and locations. Assume there are J Gaussian interference signals $\{i_j(t), j = 1, 2, \dots, J\}$, and a non-Gaussian desired signal $d(t)$, centered at frequency ω_0 . Sources are assumed to be far away from the array so that a planar wavefront approximation is possible. The additive noise is assumed to be Gaussian with unknown covariance. Consequently, the M measurements can be collected together to give the following model:

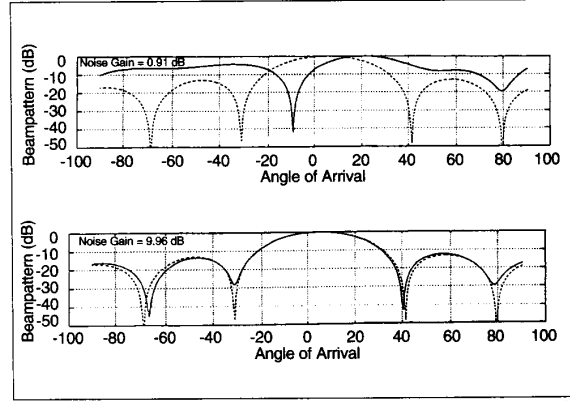
$$\underline{r}(t) = \underline{a}(\theta_d)d(t) + \underline{A}_I(\underline{\theta})\underline{i}(t) + \underline{n}(t) \quad (69)$$

where θ_d is the DOA of the desired signal; $\underline{a}(\theta_d)$ is the array steering vector of the desired signal; $\underline{A}_I(\underline{\theta})$ is the array steering matrix for the J interference sources $\underline{i}(t)$; $\underline{\theta}$ is a $J \times 1$ vector of DOA's for the interferers; $\underline{r}(t)$ is the $M \times 1$ vector of received signals; and $\underline{n}(t)$ is the $M \times 1$ vector of Gaussian noises. We let R denote the covariance matrix of $\underline{r}(t)$.

The output of an MVDR beamformer can be expressed as (Capon, 1969)

$$y(t) = \underline{w}^H \underline{r}(t) = [\beta_1 R^{-1} \underline{a}(\theta_d)]^H \underline{r}(t) \quad (70)$$

where constant β_1 maintains a specified response for the desired signal and \underline{w} denotes the weight vector of the processor. Clearly, MVDR beamforming requires knowledge of $\underline{a}(\theta_d)$. Without knowledge of the array manifold, it is not possible to determine $\underline{a}(\theta_d)$ even in the case of known θ_d . Consider the following vector of fourth-order cumulants $\underline{c} = \text{col}[c_1, c_2, \dots, c_M]$, where $c_j = \text{cum}\{r_1(t), r_1^H(t), r_1^H(t), r_j(t)\}$, $j=1,2,\dots,M$. Using the properties of cumulants stated in Table 2 and the receiver model in Eq.69, Dogan and Mendel (1992a) show that



14. Beam patterns and white-noise gains of processors in a single realization for SNR = 20 dB: (a) MVDR and (b) Cumulant-based MVDR.

$$\underline{c} = \beta_2 \underline{a}(\theta_d) \quad (71)$$

where β_2 is another scale factor. Using this result, it is now possible to estimate $\underline{a}(\theta_d)$ directly from the given data using estimated fourth-order cumulants. Observe that \underline{c} is a replica of the steering vector of the desired signal up to a scale factor. Substituting Eq. 71 into Eq. 70 we arrive at the following cumulant-based MVDR beamformer:

$$y(t) = \underline{w}_{\text{cum}}^H \underline{r}(t) = [\beta_3 R^{-1} \underline{c}]^H \underline{r}(t) \quad (72)$$

where $\beta_3 = (\underline{c}^H R^{-1} \underline{c})^{-1}$.

Many simulation results are given in Dogan and Mendel (1992a), not only for the incoherent source case but also for the coherent case (which requires a slight modification for the model in Eq. 69). Here we present results for the simplest scenario. The array is linear with 10 isotropic uniformly-spaced elements with uniform half-wavelength spacing. The record length is 1000 snapshots. This array is to be used for optimum reception of a BPSK signal which illuminates the array from 5 degrees broadside in the presence of temporally and spatially white, equal power, circularly symmetric sensor noise whose SNR is 20 dB. Fig. 14 compares the MVDR beamformer (assuming a priori knowledge of the DOA) and the cumulant-based MVDR beamformer. Plotted are the

beam pattern responses $P(\theta) = \left| \underline{w}^H \underline{a}(\theta) \right|^2$. Also shown on the plots are the white-noise gains of the beamformers. All responses are normalized to have a maximum value of 0 dB. For comparison purposes, the optimum beamformer response, calculated by using the true statistics and true steering vector information in Eq.70, is presented as the dashed curves. Observe that even when the MVDR beamformer is given true steering vector information it does not perform as well as the cumulant-based MVDR beamformer. The cumulant-based beamformer yields excellent performance without any knowledge of source DOA. Note, also, that for 100 Monte-Carlo runs the mean values and standard deviations of the white-noise gains for the MVDR beamformer are 0.179

± 1.360 , whereas the comparable results for the cumulant-based MVDR beamformer are 9.954 ± 0.015 . The theoretical white noise gain for this scenario equals 10.

Recently, Dogan and Mendel (1993a, b) have shown that there are now two additional reasons to use higher-order spectra in array signal processing: (1) HOS can increase the effective aperture of an array, and (2) HOS can not only eliminate the effects of additive Gaussian noise, but it can also eliminate the effects of additive non-Gaussian noise. This is accomplished by computing the cross-correlations that are needed by all high-resolution direction-of-arrival techniques (e.g., ESPRIT) using fourth-order cumulants.

b. Classification

Pattern or signal classification can be done working directly with the pattern or signal samples, or with attributes related to them. It is difficult to handle additive colored Gaussian noise with traditional approaches. A new approach (Giannakis and Tsatsanis, 1992), that is blind to additive colored or white Gaussian noise, works with a vector of cumulants or polyspectra, and extends correlation-based classification to HOS-based classification. It is based on the important fact that estimates of cumulants or polyspectra are asymptotically Gaussian. Consequently, one is able to begin with an equation like

$$\text{Estimate of HOS} = \text{HOS} + \text{estimation error} \quad (73)$$

in which the "estimation error" is asymptotically Gaussian, and extend traditional classification or detection procedures to this formulation. Working with a vector of higher-order statistics is in the spirit of using attributes which are derived from the original data. Consequently, higher-order statistics now provide new attributes to be used in pattern and signal classification (including detection) problems. Additional work on using higher-order statistics for detection and classification problems can be found in: Giannakis and Dandawate (1991), Giannakis and Tsatsanis (1990), Hinich (1990), Jouny, et al. (1991a,b), Kletter and Messer (1989, 1990a, 1990b), Sadler (1991), Sadler and Giannakis (1990), Shamsunder and Giannakis (1991), Swindelhurst and Kailath (1989), and Tsatsanis and Giannakis (1989, 1992).

c. Harmonic Retrieval

The estimation of the number of harmonics and the frequencies and amplitudes of harmonics from noisy measurements is frequently encountered in several signal processing applications, such as in estimating the direction of arrival of narrow-band source signals with linear arrays, and in the harmonic retrieval problem. For the latter problem and real signals, we begin with the model

$$Y(k) = \sum_{j=1}^p a_j \cos(k\omega_j + \phi_j) + N(k) = X(k) + N(k) \quad (74)$$

where the ϕ_j 's denote random phases which are i.i.d. and uniformly distributed over $[0, 2\pi]$, the ω_j 's are unknown deterministic frequencies, and the a_j 's are unknown deterministic amplitudes. The additive noise $N(k)$ is assumed to be

white or colored Gaussian noise with unknown spectral density. The problem is to estimate the number of signals p , the angular frequencies ω_j 's, and the amplitudes a_j 's.

This problem has been very widely studied. When the additive noise is white then second-order statistics-based high-resolution methods, like MUSIC or Minimum Norm, combined with SVD (to determine p), give excellent results. When the noise is colored these methods break down. They tend to overestimate the number of sinusoids by treating the colored noise as additional sinusoids. Higher-order statistics have no problem with any kind of Gaussian noise; hence, they have been successfully applied to this problem.

Third-order cumulants for $Y(k)$ equal zero; hence, this is an application where one must use fourth-order cumulants. The fourth-order cumulant of $Y(k)$ is, in general, a function of three lags (see Swami and Mendel (1991) or Mendel (1991) for the general expression); however, the diagonal slice of this cumulant is given as

$$c_4^y(\tau, \tau, \tau) \triangleq c_4^y(\tau) = -\frac{3}{8} \sum_{j=1}^p a_j^4 \cos(\omega_j \tau) \quad (75)$$

It is well-known (Marple, 1987) that the autocorrelation of $Y(k)$ is given by

$$c_2^y(\tau) = \frac{1}{2} \sum_{j=1}^p a_j^2 \cos(\omega_j \tau) \quad (76)$$

Comparing Eqs. 75 and 76 we see that (except for a difference in scale factors) $c_4^y(\tau)$ can be treated as an autocorrelation function of the following signal which is obviously related to $Y(k)$:

$$Y_1(k) = \sum_{j=1}^p a_j^2 \cos(k\omega_j + \phi_j) + N(k) \quad (77)$$

This means that already existing high-resolution methods, such as MUSIC and Minimum Norm, can be applied, just about as is, by replacing correlation quantities with $c_4^y(\tau)$. For details on exactly how to do this, see Swami and Mendel (1991). For an extensive comparison of correlation-based and cumulant-based methods for determining the number of harmonics when the amplitude of one harmonic decreases, when the frequency of one harmonic approaches the other, and when different lengths of data are used, for the case of two harmonics in colored Gaussian noise, see Shin and Mendel (1991, 1992). Additional work on applying higher-order statistics to harmonic retrieval problems can be found in: Anderson and Giannakis (1991), Ferrari and Alengrin (1991), Moulines and Cardoso (1991), Pan and Nikias (1988), Shi and Fairman (1991), Swami (1988), and Swami and Mendel (1988).

d. Time-Delay Estimation

One important application of time delay estimation methods is for source bearing and range calculation.

Let us assume that $\{X(k)\}$ and $\{Y(k)\}$ are two spatially separated sensor measurements that satisfy the equations

$$X(k) = S(k) + W_1(k) \quad (78)$$

$$Y(k) = AS(k - D) + W_2(k) \quad (78b)$$

where $\{S(k)\}$ is an unknown signal, $\{S(k - D)\}$ is a shifted and probably scaled version of $\{S(k)\}$, and $\{W_1(k)\}$ and $\{W_2(k)\}$ are unknown noise sources. The problem is to estimate the time delay D from finite length measurements of $X(k)$ and $Y(k)$. This situation arises in such application areas as sonar, radar, biomedicine, geophysics, etc. The basic approach to solve the time delay estimation problem is to shift the measurement sequence $\{X(k)\}$ with respect to $\{Y(k)\}$, and look for similarities between them. The best match will occur at a shift equal to D . In signal processing, "look for similarities" is translated into "taking the cross-correlation" between $\{X(k)\}$ and $\{Y(k)\}$. That is

$$\begin{aligned} c_{xy}(\tau) &= E\{X(k)Y(k + \tau)\} \\ &= Ac_2^s(\tau - D), \quad -\infty < \tau < \infty \end{aligned} \quad (79)$$

provided that $\{W_1(k)\}$ and $\{W_2(k)\}$ are zero-mean stationary signals, independent with each other and with $\{S(k)\}$. Note that

$$c_2^s(\tau) = E\{S(k)S(k + \tau)\} \quad (80)$$

is the covariance sequence of $\{S(k)\}$.

The $c_{xy}(\tau)$ in Eq. 79 peaks at $\tau = D$. However, in practical situations, due to finite length data records and noise sources that are not exactly independent, the $c_{xy}(\tau)$ does not necessarily show a peak at the time delay D . Various window functions have been suggested to smooth the cross-correlation function in order to improve the quality of time delay estimates. These are ROTH, SCOT, PHAT, Eckart, and Hannan-Thompson (or maximum likelihood) just to name a few [Carter, 1987].

In practical application problems where the signal $\{S(k)\}$ can be regarded as a non-Gaussian stationary random process, and the noise sources are independent stationary Gaussian, the similarities between $\{X(k)\}$ and $\{Y(k)\}$ could also be "compared" in higher-order spectrum domains such as the bispectrum. Let us note that self-emitting signals from complicated mechanical systems contain strong quasi-periodic components and therefore can be regarded as non-Gaussian signals. The main motivation behind the use of higher-order spectra for time-delay estimation under the aforementioned assumptions is the fact that they are free of Gaussian noise. Assuming that $\{S(k)\}$ has also a nonzero measure of skewness, the following identities hold:

$$c_3^x(\tau, \rho) = E\{X(k)X(k + \tau)X(k + \rho)\} = c_3^s(\tau, \rho) \quad (81a)$$

$$c_{xyx}(\tau, \rho) = E\{X(k)Y(k + \tau)X(k + \rho)\} = c_3^s(\tau - D, \rho) \quad (81b)$$

where

$$c_3^s(\tau, \rho) = E\{S(k)S(k + \tau)S(k + \rho)\}$$

because the third-order cumulants of a Gaussian process are identically zero. Obtaining the bispectra of the third-order cumulants in Eqs. 81a and 81b we have

$$C_3^x(\omega_1, \omega_2) = C_3^s(\omega_1, \omega_2) \quad (82a)$$

$$C_{xyx}(\omega_1, \omega_2) = C_3^s(\omega_1, \omega_2) \exp\{j\omega_1 D\}. \quad (82b)$$

Assuming $C_3^s(\omega_1, \omega_2)$ is nonzero, the following ratio can be formed.

$$I(\omega_1, \omega_2) = \frac{C_{xyx}(\omega_1, \omega_2)}{C_3^x(\omega_1, \omega_2)} = \exp\{j\omega_1 D\}. \quad (83)$$

One way of computing the time delay D is to form the function

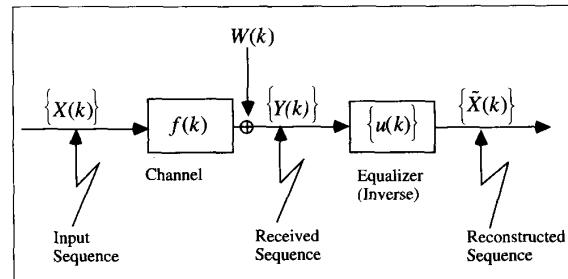
$$T(\tau) = \int_{-\pi}^{+\pi} \int_{-\pi}^{+\pi} \exp\{-j(\tau - D)\omega_1\} d\omega_1 d\omega_2 \quad (84)$$

which peaks at $\tau = D$. There are, of course, several other ways that have been developed based on parametric modeling of the third-order cumulants. It is important to note that Eq. 83 is "free" of Gaussian noise and thus better time delay estimates can be expected using $I(\omega_1, \omega_2)$. Figure 2 illustrates time delay estimation results in the presence of Gaussian noises obtained by 2nd-order and 3rd-order statistics-based methods. From this figure, it is apparent that the bispectrum-based method exhibits better noise reduction capability.

Time delay estimation methods based on higher-order statistics have been developed by Nikias and Pan [1988], Nikias and Liu [1990], Tugnait [1989], Zhang and Raghuvver [1991] and Oh et al. [1990].

e. Blind Deconvolution and Equalization

The blind deconvolution, or equalization problem, deals with the reconstruction of the input sequence given the output of a linear system and statistical information about the input. Blind deconvolution algorithms are essentially adaptive filtering algorithms designed in such a way that they do not need the external supply of a desired response to generate the error signal in the output of the adaptive equalization filter. In other words, **the algorithm is blind to the desired response**. However, the algorithm itself generates an estimate of the desired response by applying a non-linear transformation on sequences involved in the adaptation process. There are three important families of blind equalization algorithms depend-



15. Block diagram of a baseband communication system subject to additive noise.

ing on where the nonlinear transformation is being applied on the data. These are:

- (i) The Bussgang algorithms, where the nonlinearity is in the output of the adaptive equalization filter,
- (ii) The Polyspectra algorithms, where the nonlinearity is in the input of the adaptive equalization filter, and
- (iii) The algorithms where the nonlinearity is inside the equalization filter, i.e., nonlinear filters (e.g. Volterra) or neural networks.

Detailed discussion on blind equalization can be found in Haykin [1991] and Nikias and Petropulu [1993].

Let us consider a discrete-time linear transmission channel with impulse response, $f(k)$, which is unknown and possibly time-varying. The input data, $X(k)$, are assumed to be independent and identically distributed (i.i.d.) random variables with a non-Gaussian probability density function, with zero mean and variance, $E[X^2(k)] = \gamma_x^2$. Initially the noise will not be taken into account in the output of the channel. Then, the received sequence, $Y(k)$, (see Fig. 15) is

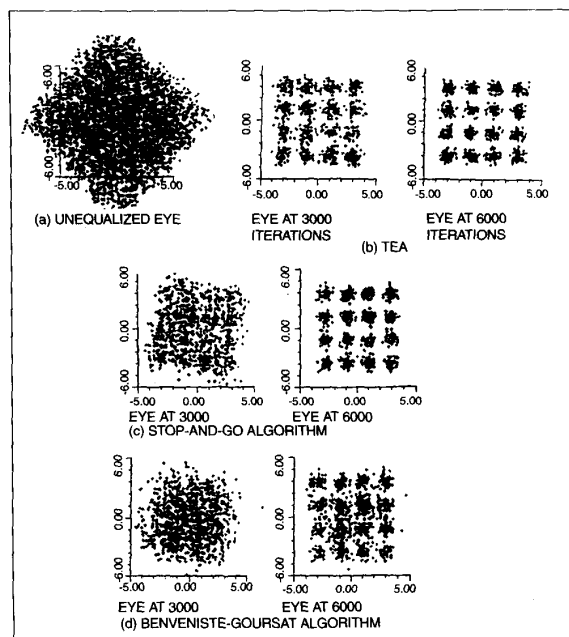
$$Y(k) = f(k) * X(k) = \sum_i X(k-i)f(i). \quad (85)$$

The problem is to restore $X(k)$ from the received sequence $Y(k)$, or equivalently, to identify the inverse filter (equalizer), $u(k)$, of the channel.

From Fig. 15 we see that the output sequence $\tilde{X}(k)$ of the equalizer is given by

$$\tilde{X}(k) = u(k) * Y(k) = u(k) * f(k) * X(k). \quad (86)$$

linear



16. (a) Discrete eye pattern of a 16-QAM constellation distorted by the channel example, (b) Equalized using TEA, (c) Equalized using Stop-and-Go algorithm, (d) Equalized using the Benveniste-Goursat algorithm (Hatzinakos and Nikias, 1991).

To achieve

$$\tilde{X}(k) = X(k-D)e^{j\theta} \quad (87)$$

where D is a constant delay and θ is a constant phase shift, it is required that

$$u(k) * f(k) = \delta(k-D)e^{j\theta}, \quad (88)$$

where $\delta(k)$ is the Kronecker delta function. Taking the Fourier transform of Eq. 88 we obtain

$$U(\omega)F(\omega) = e^{j(\theta-\omega D)}.$$

Hence, the objective of the equalizer is to achieve a transfer function

$$U(\omega) = \frac{1}{F(\omega)} e^{j(\theta-\omega D)} \quad (89)$$

In general, D and θ are unknown. However, the constant delay D does not affect the reconstruction of the original input sequence $X(k)$. The constant phase θ can be removed by a decision device.

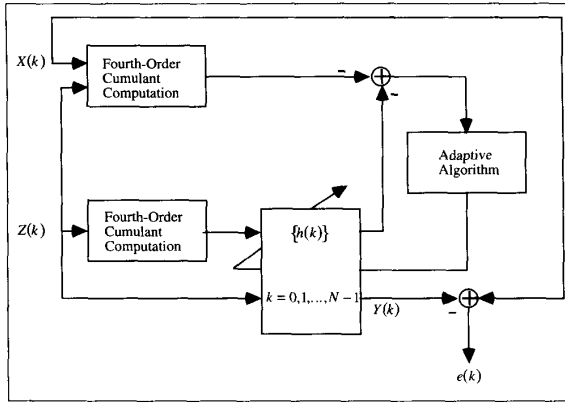
Blind equalization algorithms based on higher-order statistics (HOS) perform a nonlinear transformation on the input of the equalizer filter. This nonlinear transformation is a memory nonlinearity and it is identical to the generation of higher-order cumulants of the received channel data. One of the first blind equalizers based on HOS introduced in the literature is the Tricepstrum Equalization Algorithm (TEA) [Hatzinakos and Nikias, 1991] that estimates the equalizer impulse response by using the complex cepstrum of the fourth-order cumulants (tricepstrum) of the synchronously sampled received signal. Two extensions of the TEA have also been reported in the literature. The first one is the Power Cepstrum and Tricoherence Equalization Algorithm (POTEA) which recovers the Fourier magnitude of the equalizer using autocorrelations and its Fourier phase using fourth-order cumulants and the cepstrum of the tricoherence [Bessios and Nikias, 1991]. The second approach is an extension of TEA to the multichannel case using the cross-cumulants of the observed signals. It was thus designated as the Cross-Trispectrum Equalization Algorithm (CTEA) [Brooks and Nikias, 1991].

It has been well documented in the literature that the polyspectra-based blind equalizers achieve much faster convergence than the Bussgang-type algorithms at the expense of more computations per iteration. Fig. 16 illustrates the eye-diagrams of a 16-QAM constellation distorted by a channel and then equalized by TEA and two other algorithms that belong to the Bussgang family, namely, the Stop-and-Go and Benveniste-Goursat algorithms (Hatzinakos and Nikias, 1991). Clearly, the TEA algorithm opens the eye much earlier than the other two equalizers.

Blind equalizers based on HOS and parametric models have been developed by Porat and Friedlander [1991] and Tugnait [1991].

f. Interference Cancellation

When a signal of interest (SOI) is corrupted by an additive interference, and an auxiliary reference signal, which is



17. The configuration of the adaptive noise canceler using fourth-order statistics (ANC-FOS). The adaptive cancellation filter is $\{h(k)\}$.

highly correlated with only the interference, is also available, the elimination of the interference is accomplished by an adaptive noise canceling procedure. The reference signal is processed by an adaptive filter to match the undesired interference as closely as possible, and the filter output is subtracted from the primary input, which consists of the SOI and interference, to produce a system output. The objective of an adaptive noise canceller (ANC) is to produce a system output that best fits the SOI. Applications of ANC include the canceling of several kinds of interference in communications, speech, antenna sidelobe interference, and telephone circuits, just to name a few.

A conventional transversal ANC, which is denoted in this paper as ANC-SOS algorithm, utilizes the LMS algorithm and second-order statistics (SOS) [Haykin, 1991]. Applying the ANC-SOS algorithm in practice, we usually encounter two major difficulties. The first is that the ANC-SOS filter is affected directly by uncorrelated noises at the primary and reference inputs. The second is that ANC-SOS algorithm is problem-dependent; i.e., it is very sensitive to both the reference signal statistics and the choice of step size.

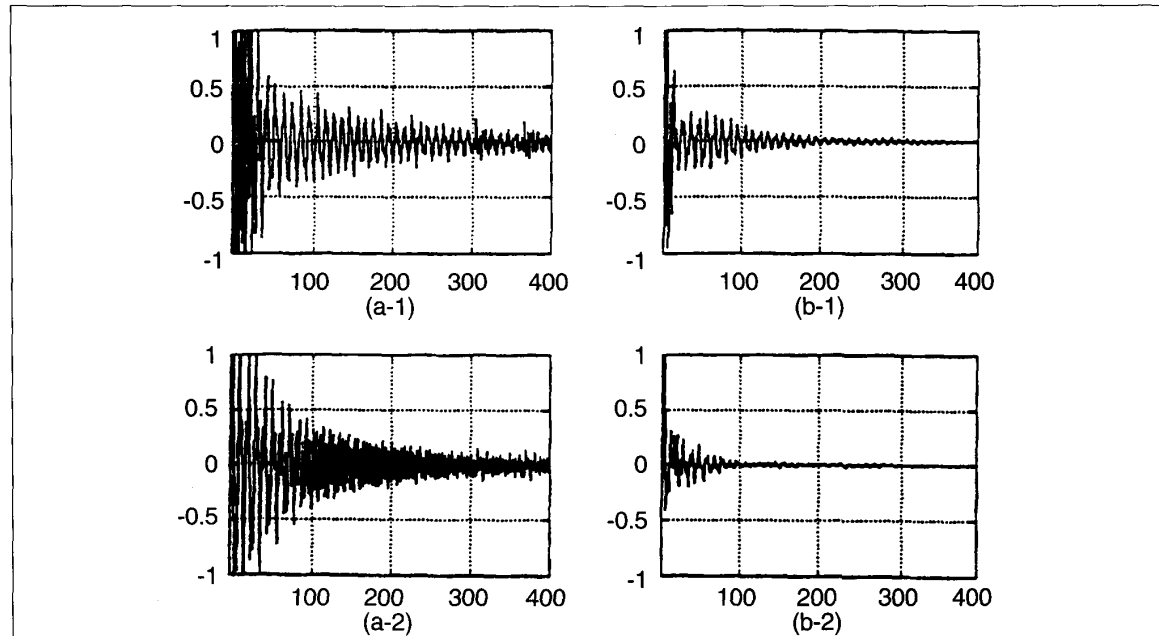
Let $\{X(k)\}$ and $\{Z(k)\}$ denote measurements of the primary and reference sensors, respectively, satisfying

$$X(k) = S(k) + I(k) + N_p(k) \quad (90)$$

$$Z(k) = W(k) + N_r(k) \quad (91)$$

where $\{S(k)\}$ is the signal of interest (SOI), $\{I(k)\}$ is the interference (narrowband or wideband), $\{W(k)\}$ is a reference signal highly correlated with the interference, and $\{N_p(k)\}$ and $\{N_r(k)\}$ are uncorrelated sensor noises. We assume that the SOI is zero-mean and any kind of a signal, i.e., deterministic or random, or combination of both, and uncorrelated with the interference and the reference signal. The reference signal is a stationary, zero-mean, non-Gaussian random process. The noises $\{N_p(k)\}$ and $\{N_r(k)\}$ are zero-mean, white or colored Gaussian, uncorrelated with each other and independent of the SOI, interference, and reference signal.

Shin and Nikias [1992] have developed a new ANC based on fourth-order statistics (ANC-FOS) and have shown that the ANC-FOS filter is independent of white or colored Gaussian uncorrelated noises and insensitive to both the reference



18. The error between the SOI and its reconstructed version as a function time. The first column shows results obtained by the ANC-SOS algorithm ($N = 32$) when $(a-1)A_1 = \sqrt{2}$ and $A_2 = 0.5$ and $(a-2)A_1 = 0.5$ and $A_2 = 1$; the second column shows results obtained by the ANC-FOS algorithm ($N = 16$) when $(b-1)A_1 = \sqrt{2}$ and $A_2 = 0.5$ and $(b-2)A_1 = 0.5$ and $A_2 = 1$.

signal statistics and the step size. Figure 17 illustrates the block diagram of the ANC-FOS algorithm.

We consider a typical example to compare the performance of the ANC-FOS algorithm with that of the ANC-SOS algorithm for eliminating interferences. Comparisons are presented in terms of the *error between the SOI and its reconstructed version by each ANC algorithm*. We assume that the SOI is deterministic BPSK having two states, and, satisfying

$$S(k) = \begin{cases} \cos(2\pi f_s T k), & \text{for } s_1 \\ -\cos(2\pi f_s T k), & \text{for } s_0 \end{cases} \quad (92)$$

where $f_s T = 0.43$ and the duration of one state is 20 samples. The reference signal is assumed to be a sum of real-valued sine waves ($W(k) = W_1(k) + W_2(k)$)

$$W_i(k) = A_i \sin(2\pi f_i T k + \phi_i), \quad i = 1, 2 \quad (93)$$

where A_i 's and f_i 's denote amplitudes and frequencies, respectively and ϕ_i 's are independent random variables uniformly distributed over $[-\pi, \pi]$. Note that $f_1 T = 0.1$ and $f_2 T = 0.25$. Each interference signal $\{I_i(k), i = 1, 2\}$ is generated through three MA(2) systems excited by a reference signal $\{W_i(k), i = 1, 2\}$. The corresponding MA coefficients equal $[1, 0.1, -0.3]$, $[1, 0.5, -0.1]$, and $[1, -0.2, 0.2]$, respectively. We assume that $I(k) = I_1(k) + I_2(k)$ with $A_1 = \sqrt{2}$, $A_2 = 0.5$ or $A_1 = 0.5$ and $A_2 = 0.1$. Figure 18 illustrates the results obtained by both ANC-SOS and ANC-FOS. From this figure, it is apparent that much faster convergence can be achieved with the ANC-FOS algorithm.

Adaptive filter schemes based on higher-order statistics have also been published by Chiang and Nikias [1988, 1990] and Dandawate and Giannakis [1989].

Conclusion

During the past two decades spectrum estimation techniques have proved essential to the creation of advanced communication, sonar, radar, speech, biomedical, geophysical and imaging systems. These techniques only use second-order statistical information, which means that we have been assuming that the signals are inherently Gaussian. Most real-world signals are not Gaussian. It is no wonder, therefore, that spectral techniques often have serious difficulties in practice.

There is much more information in a stochastic non-Gaussian or deterministic signal than conveyed by its autocorrelation or spectrum. Higher-order spectra (i.e., polyspectra), which are defined in terms of the higher-order statistics of a signal, contain this additional information. In this tutorial, an overview of higher-order spectral analysis and its applications in signal processing has been presented.

Signal processing algorithms based on higher-order spectra are now available for use in commercial and military applications. The emergence of low cost very high speed hardware chips and the ever growing availability of fast computers now demand that we extract more information than we have been doing in the past from signals, so that better decisions can be made. All of the new algorithms that have been developed using higher-order spectra are application driven.

Acknowledgment

Most of the material presented in this tutorial paper is based on the textbook by C. L. Nikias and A. P. Petropulu, *Higher-Order Spectral Analysis*, Oppenheim Series of Signal Processing, Prentice-Hall, Inc. 1993, as well as the tutorials by C. L. Nikias and M. R. Raghuveer, "Bispectrum Estimation: A Digital Signal Processing Framework," *Proc. IEEE*, Vol. 75, July 1987 and J. M. Mendel, "Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications," *Proc. IEEE*, Vol. 79, March 1991. The plots in most of the figures were obtained using Hi-Spec™ which is a software package for signal processing with higher-order spectra. Hi-Spec™ is a trademark of United Signals & Systems, Inc. We thank Dae Shin and John Dogan for producing some of the figures in this paper.

Chrysostomos L. Nikias is Professor of Electrical Engineering and Director of the Center for Research on Applied Signal Processing (CRASP) at the University of Southern California (USC) at Los Angeles; Jerry M. Mendel is Professor of Electrical Engineering and the Director of the Signal & Image Processing Institute, also at USC.

References

- Anderson, J. and G. B. Giannakis, "Two-Dimensional Harmonic Retrieval Using Cumulants," *Proc. of 7th Workshop on Multidimensional Signal Processing*, p. 210, Lake Placid, NY, September 1991.
- Bartlett, M. S., *An Introduction to Stochastic Processes*, Cambridge University Press, UK, 1955.
- Bessios, A. G. and C. L. Nikias, "FFT-Based Bispectrum Computation on Polar Rasters," *IEEE Transactions on Signal Processing*, November 1991.
- Brillinger, D. R., *An Introduction to Polyspectra*, Ann. Math. Statist., **36**, pp. 1351-1374, 1965.
- Brillinger, D. R. and M. Rosenblatt, *Asymptotic Theory of Estimates of kth Order Spectra*, in *Spectral Analysis of Time Series*, B. Harris, ed., Wiley, New York, NY, pp. 153-188, 1967.
- Brillinger, D. R. and M. Rosenblatt, "Computation and Interpretation of kth - Order Spectra," in *Spectral Analysis of Time Series*, B. Harris, ed., John Wiley and Sons, New York, NY, pp. 189-232, 1967.
- Brillinger, D. R., "The Identification of a Particular Nonlinear Time Series System," *Biometrika*, **64**, 3, pp. 509-515, 1977.
- Brooks, D. H. and C. L. Nikias, "The Cross Bicepstrum: Properties and Applications for Signal Reconstruction and System Identification," *Proc. 1991 ICASSP*, pp. 3433-3436, Toronto, Canada, May 1991.
- Capon, J., "High-Resolution Frequency-Wavenumber Spectral Analysis," *Proc. of IEEE*, **57**, no. 8, pp. 1408-1418, August 1969.
- Cardosa, J.-F., "Source Separation Using Higher-Order Moments," *Proc. ICASSP*, Glasgow, Scotland, May 1989.
- Cardosa, J.-F., "Eigen-Structure of the Fourth-Order Cumulant Tensor with Applications to the Blind Source Separation Problem," *Proc. 1990 ICASSP*, pp. 2655-2658, Albuquerque, NM, 1990.
- Cardosa, J.-F., "Super-Symmetric Decomposition of the Fourth-Order Cumulant Tensor, Blind Identification of More Sources than Sensors," *Proc. 1991 ICASSP*, pp. 3109-3112, Toronto, Canada, May 1991a.
- Cardosa, J.-F., "Higher-Order Narrow-Band Array Processing," *Proc. of Int'l. Workshop on Higher-Order Statistics*, pp. 121-130, Chamrousse, France, July 1991b.

- Chiang, H. and C. L. Nikias, "Cumulant-Based Adaptive Time Delay Estimation," *Proc. 1988 Asilomar Conference*, Pacific Grove, CA, pp. 15-19, October 1988.
- Chiang, H. and C. L. Nikias, "The ESPRIT Algorithm with Higher-Order Statistics," *Proc. Workshop on Higher-Order Spectral Analysis*, pp. 163-168, Vail, CO, 1989.
- Chiang, H. and C. L. Nikias, "Adaptive Deconvolution and Identification of Nonminimum Phase FIR Systems Based on Cumulants," *IEEE Trans. Automatic Control*, AC 35, pp. 36-47, January 1990a.
- Chiang, H. and C. L. Nikias, "A New Method for Adaptive Time Delay Estimation for Non-Gaussian Signals," *IEEE Trans. Acoustics, Speech and Signal Processing*, ASSP 38, pp. 209-219, February 1990b.
- Comon, P., "Separation of Sources Using High-Order Cumulants," *SPIE Conf. on Advanced Algorithms and Architectures for Signal Processing, Real-Time Signal Processing Vol. XII*, pp. 170-181, San Diego, CA, August 1989.
- Dandawate, A. V. and G. B. Giannakis, "A Triple Cross-Correlation Approach for Enhancing Noisy Signals," *Proc. Workshop on Higher-Order Spectral Analysis*, Vail, CO, pp. 212-216, June 1989.
- Dogan, M. C. and J. M. Mendel, "Cumulant-Based Optimum Beamforming," USC-SIPI Report # 195, January 1992a; also accepted for publication in *IEEE Trans. on Aerospace and Electronic Systems* (to appear in 1994).
- Dogan, M. C. and J. M. Mendel, "Single Sensor Detection and Classification of Multiple Sources by Higher-Order Spectra," *Proc. of IEEE Statistical Signal & Array Processing Workshop*, Victoria, BC, Canada, October 7-9, 1992b.
- Dogan, M. C., and J. M. Mendel, "Joint Array Calibration and Direction-Finding With Virtual-ESPRIT Algorithm," *Proc. 1993 IEEE Signal Processing Workshop on Higher-Order Statistics*, Lake Tahoe, CA, June 7-9, 1993a.
- Dogan, M. C., and J. M. Mendel, "Antenna Array Noise Reconditioning by Cumulants," *Proc. 1993 IEEE Signal Processing Workshop on Higher-Order Statistics*, Lake Tahoe, CA, June 7-9, 1993b.
- Duvaut, P., "Principles of Source Separation Methods Based on Higher-Order Statistics," *Traitement du Signal*, 7, pp. 407-418, December 1990.
- Ferrari, A. and L. Alengrin, "Estimation of the Frequencies of a Complex Sinusoidal Noisy Signal Using Fourth Order Statistics," *Proc. 1991 ICASSP*, pp. 3457-3460, Toronto, Canada, May 1991.
- Forster, P. and C. L. Nikias, "Bearing Estimation in the Bispectrum Domain," *Proc. ASSP Workshop on Spectrum Estimation and Modeling*, pp. 5-9, Rochester, NY, October 1990; also, in *IEEE Transactions on Acoustics, Speech and Signal Processing*, 39, pp. 1994-2006, September 1991.
- Gaeta, M. and J. L. Lacoume, "Source Separation versus Hypothesis Testing," *Proc. of Int'l. Workshop on Higher-Order Statistics*, pp. 1269-1272, Chamrousse, France, July 1991.
- Giannakis, G. B., "Cumulants: a Powerful Tool in Signal Processing," *Proc. IEEE*, 75, pp. 1333-1334, 1987.
- Giannakis, G. B. and J. M. Mendel, "Cumulant-Based Order Determination of Non-Gaussian ARMA Models," *IEEE Trans. on Acoustics, Speech and Signal Processing*, 38, pp. 1411-1423, 1990.
- Giannakis, G. B. and M. K. Tsatsanis, "Signal Detection and Classification Using Matched Filtering and Higher-Order Statistics," *IEEE Trans. on Acoustics, Speech and Signal Processing*, 38, pp. 1284-1296, July 1990.
- Giannakis, G. B. and S. Shamsunder, "Modeling of Non-Gaussian Array Data Using Cumulants: DOA Estimation with Less Sensors than Sources," *Proc. of 25th Conf. on Info. Sciences and Systems*, pp. 600-606, The Johns Hopkins University, Baltimore, MD, March 1991.
- Giannakis, G. B. and A. V. Dandawate, "Detection and Classification of Non-Stationary Underwater Acoustic Signals Using Cyclic Cumulants," *Proc. Underwater Signal Proc. Workshop*, University of Rhode Island, October 1991.
- Giannakis, G. B. and M. K. Tsatsanis, "A Unifying Maximum-Likelihood View of Cumulant and Polyspectral Measures for Non-Gaussian Signal Classification and Estimation," *IEEE Trans. on Information Theory*, March 1992.
- Gratsteyn, I. S. and I. B. Ryzhik, *Table of Integrals, Series and Products*, New York Academic Press, 1980.
- Hassab, J. C., *Underwater Signal and Data Processing*, CRC Press, Inc., Boca Raton, FL, 1989.
- Hatzinakos, D. and C. L. Nikias, "Estimation of Multipath Channel Response in Frequency Selective Channels," *IEEE Journal Selected Areas in Communications*, 7(1), pp. 12-19, January 1989.
- Hatzinakos, D. and C. L. Nikias, "Blind Equalization Using a Tricepstrum Based Algorithm," *IEEE Trans. Communications*, 39, pp. 669-682, May 1991.
- Haykin, S., *Nonlinear Methods of Spectral Analysis*, 2nd edition, Berlin, Germany, Springer-Verlag, 1983.
- Hinich, M. J., "Testing for Gaussianity and Linearity of a Stationary Time Series," *J. Time Series Analysis*, 3(3), pp. 169-176, 1982.
- Hinich, M. J., "Identification of the Coefficients in a Non-Linear Time Series of the Quadratic Type," *J. of Economics*, 30, pp. 269-288, 1985.
- Hinich, M. J., "Detecting a Transient Signal by Bispectral Analysis," *IEEE Trans. Acoustics, Speech and Signal Processing*, ASSP 38, pp. 1277-1283, July 1990.
- Huber, P. J., B. Kleiner, T. Gasser and G. Dumermuth, "Statistical Methods for Investigating Phase Relations in Stationary Stochastic Processes," *IEEE Transactions Audio Electroacoustics*, AU 19, pp. 78-86, 1971.
- Jouny, I., R. Moses and F. Garber, "Classification of Radar Signals Using the Bispectrum," *Proc. 1991 ICASSP*, pp. 3429-3432, Toronto, Canada, May 1991a.
- Jouny, I. and E. K. Walton, "Applications of the Bispectrum in Radar Signatures Analysis and Target Identification," *Proc. of Int'l. Workshop on Higher-Order Statistics*, pp. 171-174, Chamrousse, France, July 1991b.
- Jutten, C., L. N. Thi, E. Dijkstra, E. Vittoz and J. Caelen, "Blind Separation of Sources: an Algorithm for Separation of Convolutional Mixtures," *Proc. of Int'l. Workshop on Higher-Order Statistics*, pp. 273-276, Chamrousse, France, July 1991.
- Kay, S. M., *Modern Spectral Estimation*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1988.
- Kim, Y. C. and E. J. Powers, "Digital Bispectral Analysis of Self-Excited Fluctuation Spectra," *Phys. Fluids*, 21(8), pp. 1452-1453, August 1978.
- Kletter, D. and H. Messer, "Detection of a Non-Gaussian Signal in Gaussian Noise Using High-Order Spectral Analysis," *Proc. Workshop on Higher-Order Spectral Analysis*, pp. 95-99, Vail, CO, 1989.
- Kletter, D. and H. Messer, "Suboptimal Detection of Non-Gaussian Signals by Third-Order Spectral Analysis," *IEEE Trans. on Acoustics, Speech and Signal Processing*, 36, pp. 901-909, 1990a.
- Kletter, D. and H. Messer, "Optimal Detection of a Random Multitone Signal and its Relation to Bispectral Analysis," *Proc. 1990 ICASSP*, pp. 2391-2394, Albuquerque, NM, April 1990b.
- Lacoume, J. L. and P. Ruiz, "Source Identification: a Solution Based on the Cumulants," *Proc. 4th ASSP Workshop on Spectral Estimation and Modeling*, pp. 199-203, August 1988.
- Lagunas, M. A. and G. Vazquez, "Array Processing from Third Order Functions," *Proc. of Int'l. Workshop on Higher-Order Statistics*, p. 21720, Chamrousse, France, July 1991.
- Lii, K.-S. and M. Rosenblatt, "Deconvolution and Estimation of Transfer Function Phase and Coefficients for Non-Gaussian Linear Processes," *The Annals of Statistics*, 10, pp. 1195-1208, 1982.
- Lii, K. S. and M. Rosenblatt, "Asymptotic Normality of Cumulant Spectral Estimates," *Theoretical Probability*, 1989.
- Lohmann, A. W., G. Weigelt and B. Wirmitzer, "Speckle Masking in Astronomy: Triple Correlation Theory and Applications," *Applied Optics*, 22, pp. 4028-4037, 1983.
- Lohmann, A. W. and B. Wirmitzer, "Triple Correlations," *Proc. IEEE*, 72, pp. 889-901, 1984.
- Marple, S. L., Jr., *Digital Spectral Analysis with Applications*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1987.
- Mendel, J. M., "Use of Higher-Order Statistics in Signal Processing and System Theory: an Update," *Proc. SPIE Conf. on Advanced Algorithms and Architectures for Signal Processing III*, pp. 126-144, San Diego, CA, 1988.
- Mendel, J. M., "Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications," *IEEE Proc.*, 79, pp. 278-305, March 1991.
- Mohler, R. R. and F. J. Bugnon, "A Second-Order Eigenstructure Array Processor," *Proc. Workshop on Higher-Order Spectral Analysis*, pp. 152-156, Vail, CO, 1989.

- Moulines, E. and J.-F. Cardoso, "Second-Order versus Fourth-Order MUSIC Algorithms: an Asymptotical Statistical Analysis," *Proc. of Int'l. Workshop on Higher-Order Statistics*, pp. 221–224, Chamrousse, France, July 1991.
- Nikias, C. L. and M. R. Raghuveer, "Bispectrum Estimation: A Digital Signal Processing Framework," *Proceedings IEEE*, **75**(7), pp. 869–891, July 1987.
- Nikias, C. L. and R. Pan, "Non-Minimum Phase System Identification via Cepstrum Modeling of Higher-Order Moments," *Proc. ICASSP-87*, pp. 980–983, Dallas, TX, 1987.
- Nikias, C. L. and R. Pan, "Time Delay Estimation in Unknown Gaussian Spatially Correlated Noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 7(3), pp. 291–325, 1988.
- Nikias, C. L. and F. Liu, "Bicepstrum Computation Based on Second- and Third-Order Statistics with Applications," *Proc. ICASSP'90*, pp. 2381–2386, April 1990.
- Nikias, C. L. and A. P. Petropulu, *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework*, Prentice-Hall, Inc., 1993.
- Oh, W. T., S. B. Kim and E. J. Powers, "The Squared Skewness Processor for Time Delay Estimation in the Bispectrum Domain," *Signal Processing V: Theory and Applications*, L. Torres, E. Masgrau and M. A. Lagunas (eds.), Elsevier Science Publishers, pp. 111–114, 1990.
- Oppenheim, A. V. and R. W. Schaffer, *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
- Pan, R. and C. L. Nikias, "The Complex Cepstrum of Higher-Order Cumulants and Nonminimum Phase System Identification," *IEEE Transactions Acoust., Speech, and Signal Processing*, ASSP **36**(2), pp. 186–205, February 1988a.
- Pan, R. and C. L. Nikias, "Harmonic Decomposition Methods in Cumulants Domains," *Proc. ICASSP'88*, pp. 2356–2359, New York, NY, April 1988b.
- Papoulis, A., *Probability, Random Variables and Stochastic Processes*, 3rd ed., McGraw-Hill, New York.
- Petropulu, A. P. and C. L. Nikias, "The Complex Cepstrum and Bicepstrum: Analytic Performance Evaluation in the Presence of Gaussian Noise," *IEEE Trans. Acoust., Speech and Signal Processing*, **38**(7), pp. 1246–1256, July 1990.
- Petropulu, A. P. and C. L. Nikias, "Signal Reconstruction from the Phase of the Bicepstrum," *IEEE Transactions on Acoust., Speech and Signal Processing* **40**(3), pp. 601–610, March 1992.
- Pflug, A. L., G. E. Ioup and R. L. Field, "Properties of Higher-Order Correlation and Spectra for Band Limited, Deterministic Transients," *J. Acoustics, Soc. Am.*, **91**(2), pp. 975–988, February 1992.
- Porat, B. and B. Friedlander, "Direction Finding Algorithm Based on High-Order Statistics," *Proc. 1990 ICASSP*, pp. 2675–2678, Albuquerque, NM, 1990.
- Powers, E. J., C. P. Ritz, C. K. An, S. B. Kim, R. W. Miksad and S. W. Nam, "Applications of Digital Polyspectral Analysis to Nonlinear Systems Modeling and Nonlinear Wave Phenomena," *Workshop on Higher-Order Spectral Analysis*, pp. 73–77, Vail, CO, June 1989.
- Raghuveer, M. R. and C. L. Nikias, "Bispectrum Estimation: A Parametric Approach," *IEEE Trans. on Acous., Speech and Signal Processing*, ASSP **33**(5), pp. 1213–1230, October 1985.
- Raghuveer, M. R. and C. L. Nikias, "Bispectrum Estimation via AR Modeling," *Signal Processing*, **10**, pp. 35–48, 1986.
- Rao, T. S. and M. M. Gabr, *An Introduction to Bispectral Analysis and Bilinear Time Series Models*, Lecture Notes in Statistics, **24**, Springer-Verlag, New York, NY, 1984.
- Rao, S. S. and C. Vaidyanathan, "Estimating the Number of Sinusoids in Non-Gaussian Noise Using Cumulants," *Proc. 1991 ICASSP*, pp. 3469–3472, Toronto, Canada, May 1991.
- Rosenblatt, M. and J. W. Van Ness, "Estimation of the Bispectrum," *Ann. Math. Statist.*, **36**, pp. 1120–1136, 1965.
- Rosenblatt, M., *Stationary Sequences and Random Fields*, Birkhauser, Boston, MA, 1985.
- Rozario, N. and A. Papoulis, "The Identification of Certain Nonlinear Systems by Only Observing the Output," *Workshop on Higher-Order Spectral Analysis*, pp. 78–82, Vail, CO, June 1989.
- Ruiz, P. and J. L. Lacoume, "Extraction of Independent Sources from Correlated Inputs: a Solution Based on Cumulants," *Proc. Workshop on Higher-Order Spectral Analysis*, pp. 146–151, Vail, CO, 1989.
- Ruiz, P., "Sources Identification Using Cumulants: Limits and Precautions of Use," *Proc. of Int'l. Workshop on Higher-Order Statistics*, pp. 257–264, Chamrousse, France, July 1991.
- Sadler, B. and G. B. Giannakis, "On Detection with a Class of Matched Filters and Higher-Order Statistics," *Proc. of 5th ASSP Workshop on Spectrum Estimation and Modeling*, pp. 222–226, Rochester, NY, October 1990.
- Sadler, B., "Sequential Detection Using Higher-Order Statistics," *Proc. of Int'l. Conf. on Acoustics, Speech and Signal Processing*, pp. 3525–3528, Toronto, Canada, May 1991.
- Scarano, G. R. and G. Jacovitti, "Sources Identification in Unknown Coloured Noise with Composite HNL Statistics," *Proc. 1991 ICASSP*, pp. 3465–3468, Toronto, Canada, May 1991.
- Schetzen, M., *The Volterra and Wiener Theories on Nonlinear Systems*, updated edition, Krieger Publishing Company, Malabar, FL, 1989.
- Shamsunder, S. and G. B. Giannakis, "Detection and Parameter Estimation of Multiple Sources via HOS," *Proc. of Int'l. Workshop on Higher-Order Statistics*, pp. 265–268, Chamrousse, France, July 1991a.
- Shamsunder, S. and G. B. Giannakis, "Wideband Source Modeling and Localization: a HOS-Based Approach," *Proc. 25th Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, November 1991b.
- Shi, Z. and F. W. Fairman, "Cumulant Based Approach to Harmonic Retrieval Problem Using a State Space Approach," *Proc. 1991 ICASSP*, pp. 3505–3508, Toronto, Canada, May 1991.
- Shin, D. C. and J. M. Mendel, "Comparison Between Correlation-Based and Cumulant-Based Approaches to the Harmonic Retrieval and Related Problems," USC-SIPI Report #177, University of Southern California, May 1991.
- Shin, D. C. and J. M. Mendel, "Assessment of Cumulant-Based Approaches to Harmonic Retrieval," *Proc. 1992 IEEE ICASSP*, San Francisco, CA, March 1992a.
- Shin, D. C. and C. L. Nikias, "Adaptive Noise Canceller for Narrowband and Wideband Interferences Using Higher-Order Statistics," USC-SIPI Report #220, September 1992b.
- Swami, A. and J. M. Mendel, "Cumulant-Based Approach to the Harmonic Retrieval Problem," *Proc. of IEEE Conf. on ASSP*, pp. 2264–2267, New York, NY, 1988a.
- Swami, A., "System Identification Using Cumulants," Ph. D. Dissertation, USC-SIPI Report #140, University of Southern California, Department of Electrical Engineering-Systems, Los Angeles, CA, 1988b.
- Swami, A. and J. M. Mendel, "ARMA Parameter Estimation Using Only Output Cumulants," *Proc. IV IEEE ASSP Workshop on Spectrum Estimation and Modeling*, pp. 193–198, Minn., MN, 1988; also, *IEEE Trans. on Acoustics, Speech and Signal Processing*, **38**, pp. 1257–1265, July 1990.
- Swami, A. and J. M. Mendel, "Cumulant-Based Approach to the Harmonic Retrieval and Related Problems," *IEEE Trans. on Signal Processing*, **39**, pp. 1099–1109, 1991.
- Swindlehurst, A. L. and T. Kailath, "Detection and Estimation Using the Third Moment Matrix," *Proc. ICASSP 1989*, pp. 2325–2328, Glasgow, Scotland, May 1989.
- Tick, L. J., "The Estimation of Transfer Functions of Quadratic Systems," *Technometrics*, **3**(4), pp. 562–567, November 1961.
- Tsatsanis, M. K. and G. B. Giannakis, "Object and Texture Detection and Classification Using Matched Filtering and Higher-Order Statistics," *Proc. of 6th Workshop on Multidimensional Signal Processing*, pp. 32–33, Monterey, CA, September 1989.
- Tsatsanis, M. K. and G. B. Giannakis, "Object and Texture Detection and Classification Using Higher-Order Statistics," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1992.
- Tugnait, J. K., "Time Delay Estimation in Unknown Spatially Correlated Gaussian Noise Using Higher-Order Statistics," *Proc. 23rd Asilomar Conf. Signals, Systems, Computers*, pp. 211–215, Pacific Grove, CA 1989.
- United Signals & Systems, Inc., "Comprehensive Bibliography on Higher-Order Statistics (Spectra), Culver City, CA, 1992.
- Van Ness, J. W., "Asymptotic Normality of Bispectral Estimates," *Ann. Math. Statist.*, **37**, pp. 1257–1272, 1986.
- Van Veen, B. and K. Buckley, "Beamforming: a Versatile Approach to Spatial Filtering," *IEEE ASSP Mag.*, pp. 4–24, 1988.
- Zhang, W. and M. R. Raghuveer, "Non-Parametric Bispectrum-Based Time-Delay Estimators for Multiple Sensor Data," *IEEE Trans. Signal Processing*, **39**(13), pp. 770–774, March 1991.