Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Statistical Planning and Inference 130 (2005) 405–411

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Normalizing bispectra

Melvin J. Hinich[a,*], Murray Wolinsky[b]

[a]*Applied Research Laboratories, The University of Texas at Austin, P.O. Box 8029, Austin, TX 78713-8029, USA*
[b]*Los Alamos National Laboratory*

Received 24 September 2003; accepted 1 December 2003
Available online 29 July 2004

Abstract

Normalization of the bispectrum has been treated differently in the engineering signal processing literature from what is standard in the statistical time series literature. In the signal processing literature, normalization has been treated as a matter of definition and therefore a matter of choice and convenience. In particular, a number of investigators favor the Kim and Powers (Phys. Fluids 21 (8) (1978) 1452) or their "bicoherence" in Kim and Powers (IEEE Trans. Plasma Sci. PS-7 (2) (1979) 120) because they believe it produces a result guaranteed to be bounded by zero and one, and hence that it provides a result that is easily interpretable as the fraction of signal energy due to quadratic coupling. In this contribution, we show that wrong decisions can be obtained by relying on the (1979) normalization which is always bounded by one. This "bicoherence" depends on the resolution bandwidth of the sample bispectrum. Choice of normalization is not solely a matter of definition and this choice has empirical consequences. The term "bicoherence spectrum" is misleading since it is really a skewness spectrum. A statistical normalization is presented that provides a measure of quadratic coupling for stationary random nonlinear processes that has finite dependence.

Keywords: Bispectrum; Trispectrum; Spectrum; Chi-square

1. Introduction

In order to use an estimated bispectrum of a stochastic signal to detect nonlinearity or for phase-estimation, it is necessary to normalize the bispectrum by a product of the signal's spectrum. The proper normalization is the one used by Brillinger (1965, 1975), Hinich and

* Corresponding author. Fax: +1-490-4244.
E-mail address: hinich@mail.la.utexas.edu (M.J. Hinich).

0378-3758/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2003.12.022

Clay (1968), Brillinger and Rosenblatt (1967), Hinich (1982), Rosenblatt (1985) and Hinich and Messer (1995) rather than the one employed by Kim and Powers (1979). A number of investigators favor the Kim and Powers (1978) or their "bicoherence" in (1979) because they believe it produces a result guaranteed to be bounded by zero and one, and hence that it provides a result that is easily interpretable as the fraction of signal energy due to quadratic coupling (Elgar and Chandran, 1993).

We will first address the reason for normalization and then we will show that the Kim and Powers (1979) "bicoherence spectrum" is artificially bounded by one and depends on the resolution bandwidth of the sample bispectrum. As the bandwidth goes to zero the "bicoherence" goes to zero. This approach to normalization will destroy the evidence of nonlinearity in an application where the sample size is large enough to use a precise resolution bandwidth. Let us begin with the standard definition of a discrete-time *linear stochastic signal* $x(n\delta)$ where δ is a fixed sampling interval (Brillinger, 1975, p. 31; Priestley, 1981, p. 141). For simplicity, set $\delta = 1$ and $t = n$. A stochastic signal $x(t)$ is linear if it is of the form $x(t) = \sum_{k=-\infty}^{\infty} h(k)e(t-k)$ where $\{e(t)\}$ is a sequence of independent and identically distributed random variables with zero means. Such a sequence is called a *pure white noise* signal. In signal processing terminology, the signal is the output of a linear filter whose input is pure white noise. Assume that the impulse response $\{h(t)\}$ is square summable and the distribution of $e(t)$ has finite moments to ensure that all the cumulant spectra of the linear signal exist (Hinich, 1994).

The bicovariance of a linear signal is then

$$b_x(\tau_1, \tau_2) = Ex(t)x(t+\tau_1)x(t+\tau_2) = \mu_{3e} \sum_{k=-\infty}^{\infty} h(k)h(k+\tau_1)h(k+\tau_2), \quad (1.1)$$

where $\mu_{3e} = Ee^3(t)$. Thus, the signal's bispectrum is

$$\begin{aligned} B_x(f_1, f_2) &= \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} b(\tau_1, \tau_2) \exp[-\mathrm{i}2\pi(f_1\tau_1 + f_2\tau_2)] \\ &= \mu_3 H(f_1)H(f_2)H(-f_1-f_2), \end{aligned} \quad (1.2)$$

where $H(f) = \sum_{t=-\infty}^{\infty} h(t)\exp(-\mathrm{i}2\pi f t) = |H(f)|\exp[\mathrm{i}\phi(f)]$ is the filter's complex transfer function.

The linear signal's spectrum is $\sigma_e^2 H(f)H(-f) = \sigma_e^2 |H(f)|^2$ where $\sigma_e^2 = Ee^2(t)$ is the variance of $e(t)$. The normalization used by Brillinger and Rosenblatt is

$$\begin{aligned} \Gamma(f_1, f_2) &= \frac{B_x(f_1, f_2)}{\sqrt{S_x(f_1)S_x(f_2)S_x(f_1+f_2)}} \\ &= \gamma_e \exp[\mathrm{i}(\phi(f_1) + \phi(f_2) - \phi(f_1+f_2))], \end{aligned} \quad (1.3)$$

where $\gamma_e = \mu_{3e}/\sigma_e^3$ is the *skewness* of $e(t)$. The magnitude $|\Gamma(f_1, f_2)|$ of the normalized bispectrum is called the *skewness function* in the statistical time series literature. In the engineering signal processing literature, it is called the *bicoherence spectrum* (Kim and Powers, 1978; Nikias and Raghuveer, 1987). This term is misleading since it implies that the bicoherence values will be bounded by one. This is not true for the skewness, and is true in an artificial and misleading way for the Kim–Powers (1979) definition.

Three important results are obvious from Eq. (1.3). First, the skewness function of the normalized bispectrum of a linear signal is $|\Gamma(f_1, f_2)| = |\gamma_e|$, which is constant for all bifrequencies independent of the shape of the signal's spectrum. This result is central to the Hinich (1982) test for linearity based on the estimated bispectrum.

Second, the skewness can be larger than one: for example $\gamma_e = 2$ for an exponentially distributed $e(t)$. It is therefore misleading to call the skewness a "bicoherence" for a random signal.

Third, the phase of the normalized bispectrum contains information about the phase of the complex linear transfer function that can be used to recover its phase (Lii and Rosenblatt, 1982), (Rosenblatt, Chapter VIII, 1985).

This standard normalization is defined for the true bispectrum and spectrum and does not depend on data. The Kim–Powers (1979) normalization is defined for an expected value of an infinite sum discrete Fourier representation of the signal $x(t)$. Elgar and Guza (1988) and Elgar and Chandran (1993) use a similar type of definition. The expected values in their paper can be properly interpreted using a probability measure function (Crame′r and Leadbetter, Chapter 6, 1967). To avoid measure theory, the normalization issue will be discussed using cumulants of the finite Fourier transform of the observed data.

2. Estimating the bispectrum

The spectrum and bispectrum, and thus the normalized bispectrum, can be estimated using conventional nonparametric methods (see Hinich and Clay, 1968). We use the frame averaging spectrum estimation method to illuminate the statistical issues of bispectrum normalization but the results will hold for any method that yields estimates that have similar asymptotic properties to the frame averaging method.

Consider a sample $\{x(t_1), \ldots, x(t_N)\}$ where $t_k = k\delta$. This sample is partitioned into $P = [N/L]$ non-overlapping frames of length $L\delta$ where the last frame is deleted if it has less than L observations. To simplify notation, normalize the time unit by setting $\delta = 1$. The *resolution bandwidth* is then $f_1 = 1/L$.

The pth frame is $\{x_p(1), \ldots, x_p(L)\} = \{x((p-1)L+1), \ldots, x(pL)\}$. The discrete Fourier transform of the pth frame is

$$X_p(k) = \sum_{t=1}^{L} x_p(t) \exp\left(-i2\pi \frac{kt}{L}\right)$$

and the periodogram of the mth frame is

$$\frac{1}{L} |X_p(k)|^2 = \frac{1}{L} X_p(k) X_p(-k).$$

Since $N \approx LP$, the frame-averaged estimate of the spectrum at frequency $f_k = k/L$ is (Hinich and Clay, 1968)

$$\hat{S}(f_k) = \frac{1}{N} \sum_{p=1}^{P} |X_p(k)|^2. \tag{2.1}$$

Then $E[\hat{S}(f_k)] = S(f_k) + O(1/L)$ where the error term of order $1/L$ is due to the frame windowing of the spectrum. The variance of the estimate for large values of L and P is $(1/P)S^2(f_k)$.

The frame-averaged estimate of the bispectrum at the bifrequencies (f_{k_1}, f_{k_2}) is (Hinich and Clay, 1968)

$$\hat{B}(f_{k_1}, f_{k_2}) = \frac{1}{N} \sum_{p=1}^{P} X_p(k_1) X_p(k_2) X_p(-k_1 - k_2). \tag{2.2}$$

Then $E[\hat{B}(f_{k_1}, f_{k_2})] = B(f_{k_1}, f_{k_2}) + O(1/L)$ and the variance for large L and P is $(L/P)S(f_{k_1})S(f_{k_2})S(f_{k_1} + f_{k_2})$.

The standard normalization of the estimated bispectrum is

$$\begin{aligned}\hat{\Gamma}(f_{k_1}, f_{k_2}) &= \frac{\hat{B}(f_{k_1}, f_{k_2})}{\sqrt{\hat{S}(f_{k_1})\hat{S}(f_{k_2})\hat{S}(f_{k_1} + f_{k_2})}} \\ &= \frac{\hat{B}(f_{k_1}, f_{k_2})}{\sqrt{S(f_{k_1})S(f_{k_2})S(f_{k_1} + f_{k_2}) + O(L^{-1})}}.\end{aligned} \tag{2.3}$$

Note that this normalization standardizes the variance of the bispectrum estimate using the estimated variance in place of the true variance. If $L = N^e$ where $0 < e < 0.5$, then $N^{-1/2+e}[\hat{\Gamma}(f_{k_1}, f_{k_2}) - \Gamma(f_{k_1}, f_{k_2})]$ has an asymptotically complex Gaussian distribution with a zero mean and unit variance as $N \to \infty$ (Hinich, 1982). Moreover, the $\Gamma(f_{k_1}, f_{k_2})$ are asymptotically independently distributed across the principal domain of the bifrequencies. A statistical normalization to the unit interval will be presented in Section 4 that uses these statistical properties.

Now let us turn to the problem with Kim and Powers (1979) normalization.

3. The Kim and Powers bicoherence spectrum

Now let us turn to the Kim and Powers (1979) bicoherence spectrum, which is

$$|\hat{b}(f_{k_1}, f_{k_2})| = \frac{|\hat{B}(f_{k_1}, f_{k_2})|}{\sqrt{\left[\frac{1}{N}\sum_{p=1}^{P} |X_p(k_1)X_p(K_2)|^2\right]\hat{S}(f_{k_1} + f_{k_2})}}. \tag{3.1}$$

It follows from the Schwartz Inequality and Eq. (2.1) that $|\hat{b}(f_{k_1}, f_{k_2})| \leqslant 1$ but we will now show that this bicoherence spectrum depends on the value of L used to estimate the bispectrum.

The denominator in Eq. (3.1) is

$$\left[\frac{1}{N}\sum_{p=1}^{P} |X_p(k_1)X_p(k_2)|^2\right]\hat{S}(f_{k_1} + f_{k_2}) = Z(k_1, k_2)\hat{S}(f_{k_1} + f_{k_2}), \tag{3.2}$$

where

$$Z(k_1, k_2) = \left[\frac{1}{N} \sum_{p=1}^{P} |X_p(k_1)X_p(k_2)|^2 \right]. \tag{3.3}$$

From Eq. (3.10) in Hinich (1994),

$$E|X_p(k_1)X_p(k_2)|^2 = L[T(f_{k_1}, f_{k_2}, -f_{k_1}) + LS(f_{k_1})S(f_{k_2})] + O(1), \tag{3.4}$$

where $T(f_{k_1}, f_{k_2}, -f_{k_1})$ is the trispectrum at the trifrequencies $f_{k_1}, f_{k_2}, -f_{k_1}$. Thus

$$E[Z(k_1, k_2)] = [T(f_{k_1}, f_{k_2}, -f_{k_1}) + LS(f_{k_1})S(f_{k_2})]. \tag{3.5}$$

The Kim–Powers bicoherence spectrum becomes

$$|\hat{b}(f_{k_1}, f_{k_2})| \approx \frac{|B(f_{k_1}, f_{k_2})|}{\sqrt{[T(f_{k_1}, f_{k_2}, -f_{k_1}) + LS(f_{k_1})S(f_{k_2})]S(f_{k_1} + f_{k_2})}} \tag{3.6}$$

for large values of N. It is obvious from Eq. (3.6) that the normalization depends on the frame length L and on the magnitude of the trispectrum.

To show how this estimated "bicoherence spectrum" shrinks to zero as L increases, suppose that the signal is linear. Then the right-hand size of Eq. (3.6) is $|\gamma_e|/\sqrt{(\kappa_e + L)}$ where κ_e is the kurtosis (fourth-order standardized cumulant) of $e(t)$. By the Schwartz Inequality $|\gamma_e| \leqslant \sqrt{\kappa_e + 3}$ and thus

$$\frac{|\gamma_e|}{\sqrt{(\kappa_e + L)}} \leqslant \sqrt{\frac{\kappa_e + 3}{\kappa_e + L}}.$$

In most applications, L will be much larger than three and the magnitude of the "bicoherence" will lead the user of such normalization to believe that the signal is consistent with a null hypothesis that the signal is Gaussian. This dependence on the value of L of the "bicoherence" estimator essentially eliminates its value as a measure of Gaussianity and nonlinearity of the signal.

For a nonlinear stochastic signal the trispectrum term can be very larger than L and thus the "bicoherence spectrum" estimate can be shrunk by fourth-order nonlinearity.

We now present a statistical normalization that is useful for detecting nonlinear structure in a stationary random signal that has finite dependence.

4. A statistical normalization

Recall that the standard normalization is a skewness measure and thus is not bounded by one. The skewness function has no simple relationship with nonlinearity in the random signal. The asymptotic properties of the bispectral estimates provide a way to compute a normalization that provides more insight into the nonlinear structure of a random signal. For large N the statistics $Y(k_1, k_2) = 2N^{-1+2e}|\hat{\Gamma}(f_{k_1}, f_{k_2})|^2$ have (approximately) independently distributed non-central chi-square distributions with two-degrees-of-freedom and non-centrality parameters $\lambda(k_1,k_2)=2N^{-1+2e}|\Gamma(f_{k_1},f_{k_2})|^2$ for the array of bifrequency counters (k_1,k_2) in the principal domain. Thus if the signal is linear then the monotonically transformed statistics $P(k_1,k_2)=F_{\chi^2}[Y(k_1,k_2)|\lambda(k_1,k_2)]$ are approximately independently distributed uniform (0,1) variates where $F_{\chi^2}[y|\lambda(k_1,k_2)]$ denotes a non-central chi-square cumulative distribution function (cdf) with two-degree-of-freedom and non-centrality parameter $\lambda(k_1,k_2)$.

Recall that $|\Gamma(f_1,f_2)|=|\gamma_e|$ if the signal is linear, which implies that the non-centrality parameters are equal to $\lambda=2N^{-1+2e}\gamma_e^2$ for all bifrequencies in the principal domain. The statistical normalization then is to transform each $Y(k_1,k_2)$ using the cdf $F_{\chi^2}(y|\hat{\lambda})$ where $\hat{\lambda}$ is the mean value of all the estimates of the non-centrality parameters (Hinich, 1982). This type of probability normalization will now be used to show the difference between the Kim and Powers normalization and the standard one using some data from a wind tunnel test of a Learjet model obtained from Professor Ron Stearman, University of Texas at Austin.

5. An example

The bispectrum was calculated using the standard and Kim and Powers normalizations for a strain gauge signal from a test of a Learjet model in a wind tunnel. The data was low-pass filtered to avoid aliasing. The sampling rate is 2.0480 kHz. The sample size is 65,626 which permitted a resolution bandwidth of 2 Hz. The frame length is then $L=1024$ and the number of complete frames is 64. The spectrum and bispectra were computed for a bandwidth of 0 to 700 Hz. The spectrum of the signal for this band is in Fig. 1.

The support of the principal domain of the signal's bispectrum is the isosceles triangle defined by $\{0<f_1<700\,\text{Hz},\ 0<f_2<700\,\text{Hz},\ 0<f_1+f_2<700\,\text{Hz}\}$ (see Hinich and Messer (1995) for details about the principle domain). There are 1156 bifrequencies in this triangle.

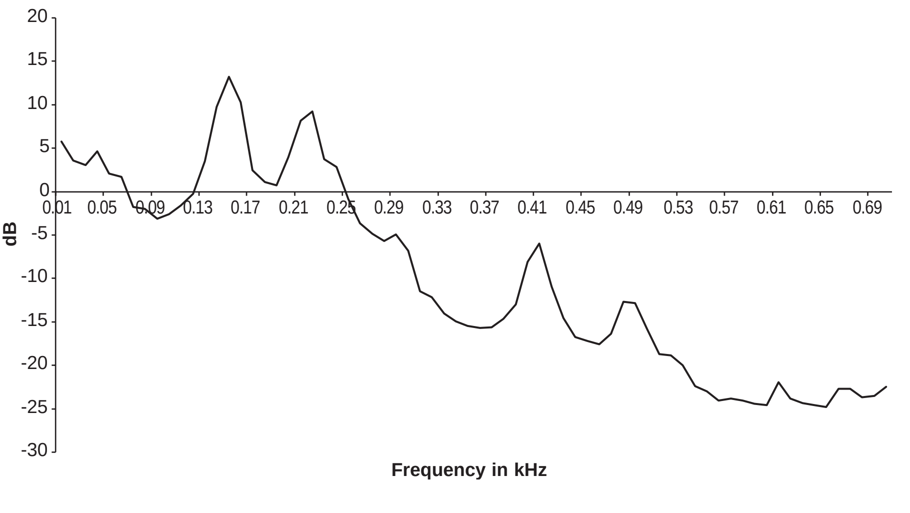

Fig. 1. Winglet signal spectrum.

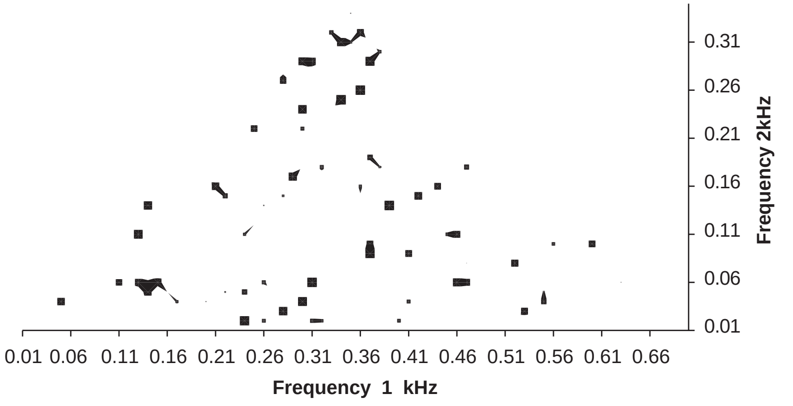

Fig. 2. Bispectrum probabilities > 0.9 of Winglet signal using standard normalization 1156 bifrequencies.

The bispectral probabilities greater than 0.9 are shown in Fig. 2 for the standard normalization. The Kim and Powers normalization shrinks the bispectral values so that there are *no* probabilities greater than 0.1 let alone 0.9. Thus the type of normalization matters.

References

Brillinger, D.R., 1965. An introduction to polyspectra. Ann. Math. Stat. 36, 1351–1374.

Brillinger, D.R., 1975. Time Series, Data Analysis and Theory. Holt, Rinehart & Winston, New York.

Brillinger, D.R., Rosenblatt, M., 1967. Computation and interpretation of k-th order spectra. in: Harris, B. (Ed.), Spectral Analysis of Time Series. Wiley, New York, pp. 189–232.

Crame′r, H., Leadbetter, M.R., 1967. Stationary and Related Stochastic Processes. Wiley, New York.

Elgar, S., Chandran, V., 1993. Higher-order spectral analysis to detect nonlinear interactions in measured time series and an application to Chua's circuit. Internat. J. Bifurcat. Chaos 3 (1), 19–34.

Elgar, S., Guza, R.T., 1988. Statistics of bicoherence. IEEE Trans. Acoust. Speech Signal Process. 36 (10), 1667–1668.

Hinich, M.J., 1982. Testing for Gaussianity and linearity of a stationary time series. J. Time Ser. Anal. 3, 169–176.

Hinich, M.J., 1994. Higher order cumulants and cuimulant spectra. Circ. Syst. Sign. Proc. 13 (4), 391–420.

Hinich, M.J., Clay, C.S., 1968. The application of the discrete Fourier transform in the estimation of power spectra, coherence and bispectra of geophysical data. Rev. Geophys. 6 (3), 347–363.

Hinich, M.J., Messer, H., 1995. On the principal domain of the discrete bispectrum of a stationary signal. IEEE Trans. Signal Process. 43 (9), 2130–2134.

Kim, Y.C., Powers, E.J., 1978. Digital bispectral analysis of self-excited fluctuation spectra. Phys. Fluids 21 (8), 1452–1453.

Kim, Y.C., Powers, E.J., 1979. Digital bispectral analysis and its applications to nonlinear wave interactions. IEEE Trans. Plasma Sci. PS-7 (2), 120–131.

Lii, K.S., Rosenblatt, M., 1982. Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes. Ann. Statist. 10, 1195–1208.

Nikias, C.L., Raghuveer, M.R., 1987. Bispectrum estimation: a digital signal processing framework. IEEE Proc. 75 (7), 869–891.

Priestley, M.B., 1981. Spectral Analysis and Time Series. Academic Press, London.

Rosenblatt, M., 1985. Stationary Sequences and Random Fields. Birkhauser, Boston.