MULTITAPER ESTIMATION OF BICOHERENCE

Yngve Birkelund and Alfred Hanssen

Electrical Engineering Group
Physics Department
University of Tromsø
N–9037 Tromsø, Norway

Edward J. Powers

Department of Electrical and Computer Engineering
The University of Texas at Austin
Austin, TX 78712-1084, USA

ABSTRACT

The statistical properties for bicoherence estimation are shown to be strongly connected to the properties of the power- and bispectral estimator used. Data tapering will reduce spectral leakage and frequency smoothing will reduce the variance. It is shown that correct normalization is essential to ensure unbiased results. The multitaper approach is shown to be superior to other non-parametric estimators for bicoherence estimation.

1. INTRODUCTION

The bicoherence is a useful tool for classification and quantization of non-linearity and non-Gaussianity in time series. Since a Gaussian linear processes theoretically has zero bicoherence, a non-zero constant valued bicoherence indicates non-Gaussianity. Furthermore, a frequency dependent bicoherence value indicates non-linearity [1]. In harmonic analysis, e.g. the classical three wave phase coupling example [2], the bicoherence provides the fraction of power carried by the non-linearity.

Since bicoherence is a normalized version of the bispectrum, higher order spectral analysis plays a significant role in the estimation of the bicoherence from real data. It is well known that the bispectrum estimates are prone to be noisy and statistically inconsistent. To reduce the impact this can have on the bicoherence values, segment averaging and/or frequency smoothing have been introduced in bicoherence estimation [2, 3]. In addition, data tapering has been suggested to reduce the spectral leakage for processes with large dynamical range to avoid biased bicoherence estimates. Recently, several publications report the use of multitaper technique in bispectral estimation [4, 5, 6, 7]. In both power- and bispectral estimation, the multitaper approach is superior to classical non-parametric estimators in terms of bias and variance.

In this paper, we will study the effect of data tapering and frequency smoothing in the case of bicoherence estimation. Some pitfalls in bicoherence estimation will be discussed. We will present the adaptive multitaper technique for bicoherence estimation, which provides full control over the smoothing bandwidth and leakage, and demonstrate its superiority over non-parametric estimators for bicoherence estimation.

The organization of this paper is as follows. In section 2 we present the bicoherence and some of its properties. In section 3 we discuss tapering and frequency smoothing of power- and bispectra, and we investigate the statistical properties of the resulting bicoherence estimators. Some illustrative numerical examples are shown in section 4. Finally, section 5 contains some recommendations and the conclusion.

2. BICOHERENCE

There exist several ways to normalize the bispectrum [3]. We will assume that the available data $x[n]$, $n = 0, 1, \ldots, N-1$, is from a real and stationary process with zero mean and Fourier transform $X(f) = 1/N \sum_{n=0}^{N-1} x[n] e^{-j2\pi fn}$. The bicoherence $b(f_1, f_2)$ can then be defined as

$$b(f_1, f_2) = \frac{M_3(f_1, f_2)}{\sqrt{M_2(f_1) M_2(f_2) M_2(f_1 + f_2)}}, \qquad (1)$$

where

$$M_2(f) = \mathrm{E}\left[X(f)X^*(f)\right] \qquad (2)$$

$$M_3(f_1, f_2) = \mathrm{E}\left[X(f_1)X(f_1)X^*(f_1 + f_2)\right]. \qquad (3)$$

For stochastic processes, the Fourier transform does not strictly exist. Still, the bicoherence definition in (1) is suitable, using the relationship to the process power spectrum $P(f) = NM_2(f)$ and its bispectrum $B(f_1, f) = N^2 M_3(f_1, f_2)$. Keeping these relationships in mind, we will for the rest of the paper denote $M_2(f)$ and $M_3(f_1, f_2)$ as the process power- and bispectrum, respectively. All estimators and statistical properties are related to $M_2(f)$ and $M_3(f_1, f_2)$, even if some references use $P(f)$ and $B(f_1, f_2)$.

Using the Fourier transform $X(f)$ of the whole data record directly in calculation of $M_2(f)$ and $M_3(f_1, f_2)$, known as the periodogram $M_2^{per}(f)$ and biperiodogram estimator $M_3^{per}(f_1, f_2)$, is however not recommended. This will always lead to a unity magnitude bicoherence since there is no expectation operations involved in the estimation of the spectra. Thus, we will in the following discuss the case where data $x[n]$ for $n = 0, 1, \ldots, N-1$ are divided into S segments, each with M data points and $N = SM$. For each segment the power- and bispectra are estimated, and the segment averaged results $\mathbf{A}[M_2^{per}(f)]$ and $\mathbf{A}[M_3^{per}(f_1, f_2)]$ are used in (1) to make a bicoherence-periodogram $\widehat{b}^{per}(f_1, f_2)$.

Since the expected value of the periodogram and biperiodogram is unbiased in the case of no spectral leakage [8, 7], the $\widehat{b}^{per}(f_1, f_2)$ will also be approximate unbiased. It is common to neglect the statistical variability of the denominator in (1), so that the variance can be approximated as [2]

$$\mathrm{Var}\left[\widehat{b}^{per}(f_1, f_2)\right] \simeq \frac{1}{S}\left(1 - b^2(f_1, f_2)\right). \qquad (4)$$

This approximation clearly shows that the variance is low for unity true bicoherence, while in the case of zero-level true bicoherence the variance is constant and independent of the Fourier transform amplitude $|X(f)|$.

3. IMPROVED POWER- AND BISPECTRAL ESTIMATES

Even if the segment averaged method provides a consistent estimator for the bicoherence, there are some problems that have to be taken into account. First, to obtain a reasonable zero level bicoherence in the white Gaussian case requires $S \gg 1$ [3]. Second, in the case of data with large dynamic range, spectral leakage can introduce severely biased bicoherence estimates.

It is possible to obtain lower variance in the bicoherence estimate if we introduce frequency smoothing in addition to the segment averaging. The improvement in variance properties is obtained at the cost of poorer frequency resolution [3]. Note that the power- and the bispectral estimate should be smoothed with the same bandwidth.

Spectral leakage can be reduced through data tapering [2]. A typical "bell-shaped" data window is the Hanning taper, which have significantly lower sidelobes than the rectangular data window. The improvement in leakage resistance is obtained at the cost of poorer frequency resolution since the main lobe is widened. We strongly recommend one to normalize the power- and bispectral estimates by the window energy and the bispectral window "energy" of the data window used [9, 7]. This will ensure unbiased estimates of constant valued bicoherence.

3.1. Multitaper estimates

Maximizing the spectral concentration λ of the taper $v[m]$ subject to a chosen half bandwidth f_B

$$\max_{V(f)} \lambda = \frac{\int_{-f_B}^{f_B} |V(f)|^2 df}{\int_{-1/2}^{1/2} |V(f)|^2 df} \tag{5}$$

where $V(f)$ is the Fourier transform of $v[m]$ and $m = 0, 1, \ldots, M-1$, leads to a set of M orthonormal data windows called the Discrete Prolate Spheroidal Sequences (DPSS) [10]. Normally, the $K = 2Mf_B$ tapers with the largest spectral concentrations are used in the multitaper approach. We denote these tapers and their corresponding spectral concentrations as $v_k[m]$ and λ_k, respectively.

For each data segment, a family of tapered data is obtained as $x_k[m] = x[m]v_k[n]$. Tapered periodogram $\widehat{M}_2^k(f)$ and tapered biperiodogram $\widehat{M}_3^{k_1k_2k_3}(f_1, f_2)$ can now be found as

$$\widehat{M}_2^k(f) = |X_k(f)|^2 \tag{6}$$

$$\widehat{M}_3^{k_1k_2k_3}(f_1, f_2) = X_{k_1}(f_1)X_{k_2}(f_2)X_{k_3}^*(f_1 + f_2), \tag{7}$$

where $X_k(f) = 1/M \sum_{m=0}^{M-1} x_k[m]e^{-j2\pi fm}$ is the Fourier transform of the data tapered by window no. k.

The multitaper spectral estimates can thus be found as an average over all individual tapered estimates

$$\widehat{M}_2^{mt}(f) = \tfrac{1}{U_2} \sum_{k=0}^{K-1} \widehat{M}_2^k(f) \tag{8}$$

$$\widehat{M}_3^{mt}(f_1, f_2) = \tfrac{1}{U_3} \sum_{k_1\,k_2k_3=0}^{K-1} P_{k_1k_2k_3} \widehat{M}_3^{k_1k_2k_3}(f_1, f_2). \tag{9}$$

Here $P_{k_1k_2k_3} = \sum_{m=0}^{M-1} v_{k_1}[m]v_{k_2}[m]v_{k_3}[m]$ is a "bienergy" weighting function, and $U_2 = M \cdot K$ and $U_3 = M^2 \sum_{k_1,k_2,k_3=0}^{K} P_{k_1k_2k_3}^2$ are normalization constants to ensure approximately unbiased power- and bispectral estimates respectively.

In the case of signals with a large dynamic range, the adaptive multitaper approach introduces frequency dependent weight-functions $d_k(f)$ that will down-weight the result from each taper only where leakage is present. In case of no spectral leakage, the adaptive approach will perform similarly to the normal multitaper spectral estimates. The adaptive approach is discussed in detail in [6, 10].

The multitaper estimator for bicoherence $\widehat{b}^{mt}(f_1, f_2)$ is obtained using the segment averaged multitaper power- and bispectral estimates, $\mathbf{A}[\widehat{M}_2^{mt}(f)]$ and $\mathbf{A}[M_3^{mt}(f_1, f_2)]$ respectively, in (1).

3.2. Statistical properties

The bias and variance for multitaper spectral estimates can be described through the total spectral window $W_2(f)$ and total bispectral window $W_3(f_1, f_2)$, respectively. Similar results for both power- and bispectral estimates can be obtained for all conventional non-parametric techniques [9, 7, 11].

3.2.1. Expected value

The expected value of the multitaper power- and bispectral estimates can be written as [10, 7]

$$\mathrm{E}\left[\widehat{M}_2^{mt}(f)\right] = W_2(f) * M_2(f) \tag{10}$$

$$\mathrm{E}\left[\widehat{M}_3^{mt}(f_1, f_2)\right] = W_3(f_1, f_2) * M_3(f_1, f_2) \tag{11}$$

where $*$ denotes one- and two dimensional convolution, respectively. Here $W_2(f)$ and $W_3(f_1, f_2)$ can be written as

$$W_2(f) = \tfrac{1}{U_2} \sum_{k=0}^{K-1} |V_k(f)|^2 \tag{12}$$

$$W_3(f_1, f_2) = \tfrac{1}{U_3} \sum_{k_1\,k_2\,k_3=0}^{K-1} P_{k_1k_2k_3} W_{k_1k_2k_3}(f_1, f_2) \tag{13}$$

where $W_{k_1k_2k_3}(f_1, f_2) = V_{k_1}(f_1)V_{k_2}(f_2)V_{k_3}^*(f_1+f_2)$ and $V_k(f)$ is the Fourier transform of the k-th DPSS taper $v_k[m]$. Using DPSS tapers, $W_2(f)$ is the best approximation to a normalized ideal filter in terms of spectral concentration [10, 9]. Similarly, $W_3(f_1, f_2)$ corresponds to a normalized ideal hexagonal smoother in the bi-frequency domain [7].

We will approximate the expected value of the multitaper bicoherence estimator by

$$\mathrm{E}\left[\widehat{b}^{mt}(f_1, f_2)\right] \simeq \frac{\mathrm{E}\left[\widehat{M}_3^{mt}(f_1, f_2)\right]}{\sqrt{\mathrm{E}\left[\widehat{M}_2^{mt}(f_1)\right] \mathrm{E}\left[\widehat{M}_2^{mt}(f_2)\right] \mathrm{E}\left[\widehat{M}_2^{mt}(f_1+f_2)\right]}}. \tag{14}$$

Although this is an approximate expression, it has, however, proven very useful in interpreting the effect that difference power- and bispectral estimators have on bicoherence estimation.

For slowly varying processes, both power- and bispectral multitaper estimates are unbiased. Consequently, the bicoherence estimator is also unbiased in that case.

Near rapidly varying true bicoherence, the smoothing effects imply local bias in the bicoherence. Sharp maxima will be smoothed out, and since the maxima will be more spread out in the bispectral case, the resulting bicoherence estimate will be lower than the true value. If e.g. only the bispectral estimate is smoothed, sharp maxima in the true bicoherence will appear as a dip in an artificial raised smoothed platform.

3.2.2. Variance

Only considering the smoothing effect from the multitaper approach, the variance of the power- and bispectral estimates can be approximated by [9, 7]

$$\text{Var}\left[\widehat{M}_2^{mt}(f)\right] = W_2^2(f) * M_2^2(f) \quad (15)$$

$$\text{Var}\left[\widehat{M}_3^{mt}(f_1, f_2)\right] = W_3^2(f_1, f_2) * \text{Var}[\widehat{M}_3^{per}(f_1, f_2)], \quad (16)$$

where $\text{Var}[\widehat{M}_3^{per}(f_1, f_2)]$ is the variance of the biperiodogram. Since both $W_2(f)$ and $W_3(f_1, f_2)$ have an approximately flat smoothing support, a higher smoothing bandwidth f_B will provide wider supports and thereby lower variances.

Neglecting the variability of the denominator in (1), the variance of the multitaper bicoherence estimator can similarly be approximated by

$$\text{Var}[\widehat{b}^{mt}(f_1, f_2)] \simeq W_3^2(f_1, f_2) * \text{Var}[\widehat{b}^{per}(f_1, f_2)] \quad (17)$$

In the case of white noise, the variance reduction factor is $3(Mf_B)^2$ [4, 7]. For rapidly varying $M_2(f)$, $M_3(f_1, f_2)$ and/or $b(f_1, f_2)$, the bias in expected bicoherence values will effect the variance. This can partly be described using the estimators expected value instead of the true bicoherence value in (4)

Since multitaper estimators have been shown superior to other non-parametric estimators for both power- [10, 9] and bispectra [7] in terms of bias and variance, the multitaper bicoherence estimator will have the same superiority based on (14) and (17).

4. NUMERICAL EXAMPLES

4.1. Phase coupling

We will first investigate data with different degree of phase coupling between harmonic frequencies. The data are similar to the example given in [2], but we have chosen other frequencies and additive non-Gaussian zero-mean white noise. The test signal is

$$x[m] = \cos(\omega_b m + \theta_b) + \cos(\omega_c m + \theta_c) + \frac{1}{2}\cos(\omega_d m + \theta_d) + \cos(\omega_b m + \theta_b)\cos(\omega_c m + \theta_c) + n[m]$$

where $\omega_i = \pi f_i$ for $i = b, c$ and d, $f_b = 0.375$, $f_c = 0.5$ and $f_d = f_b + f_c$. The noise is defined as $n[m] = e[m] - 1$, where $e[m]$ is exponentially distributed with unity mean $\gamma_1 = 1$, unity variance $\gamma_2 = 1$, and skewness $\gamma_3 = 2$. Each segment contains $M = 128$ data points, and we average over $S = 4$ segments of data. The phases θ_b, θ_c and θ_d are independently drawn for each segment from a uniform distribution on $[-\pi, \pi)$.

The true bicoherence for this process contains a flat background noise with magnitude $\gamma_2/\sqrt{M\gamma_2^3} \simeq 0.177$. Sum and difference non-linear interaction are present at bifrequencies (f_c, f_b) and $(f_b, f_c - f_b)$, respectively. Except for these two points, the true bicoherence will have low values along the harmonic axes $f_1, f_2, f_1 + f_2 \in (f_c - f_b, f_b, f_c, f_b + f_c)$.

Estimates of the expected value and the variance for the bicoherence estimators are found using 10000 independent Monte Carlo simulations. The modulus of the expected value using four different bicoherence estimators are shown in Fig. 1. In Table 1, we have listed some important values for easier reference, and we have included the variance results.

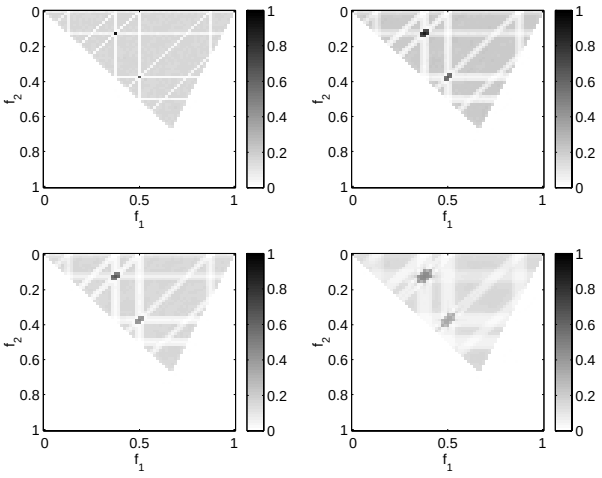

Fig. 1. Modulus of expected value for bicoherence estimators with $M = 128$ and $S = 4$. Upper left: Using periodogram and biperiodogram. Upper right: Using Hanning tapered periodogram and biperiodogram. Lower left: Using normalized Hanning tapered periodogram and biperiodogram. Lower right: Multitaper estimator with $Mf_B = 2$.

In the upper left corner of Fig. 1, is the expected magnitude for the bicoherence using periodogram and biperiodogram estimators. Since the harmonic frequencies match the discrete frequency bin available, the mean values shows clear identification of the degree of coupling. The problem is that the variance for the noise floor can be too high to distinguish between noise and true phase coupling.

The Hanning tapered bicoherence, whose expected magnitude is shown in the upper right corner of Fig. 1, introduces correlation between nearby frequencies. Near the bicoherence maxima, the neighboring frequencies are over estimated, as is the the noise floor level. Correct normalized Hanning tapers in the power- and bispectral estimates provide the correct noise level, but the peak values of the phase couplings are also lowered as shown in lower left corner of Fig. 1. We find that the variance of correct normalized bicoherence estimator is lower than in the untapered case. This variance reduction is connected to the frequency correlation introduced by the tapering itself [7].

The multitaper bicoherence expected magnitude is presented in lower right corner of Fig. 1. For slowly varying parts of values in the true bicoherence, this estimator provides unbiased estimates with very low variance. The variance result in the noise region, as indicated in right hand columb of Table 1, is close to the theoretical result $(1/3(Mf_B)^2 S = 0.021)$ from (17). Near the bicoherence peaks the hexagonal frequency smoothing effect is obvious, but still the different phase coupling can be identified clearly.

4.2. Tension Lag Platform (TLP) data

A TLP is a buoyant floating platform, tethered to sea bed by tensioned tendons. The resonant frequency of horizontal motion is designed to be below the expected frequency band of input seawave to avoid linear resonance. However, it is found that high-frequency incoming sea-waves can excite low-frequency drift oscillations in the horizontal plan due to second-order nonlinear ef-

Mean/ Var	$(f_c\text{-}f_b, f_b)$	(f_b, f_c)	$(0.25, 0.75)$
Theoretical	.943/ .027	.686/ .133	.177/ .250
(Bi-) periodogram	.924/ .023	.660/ .103	.162/ .252
Tapering	.892/ .032	.644/ .109	.217/ .258
Norm. tapering	.658/ .017	.475/ .059	.160/ .140
Multitaper	.466/ .007	.355/ .030	.162/ .024

Table 1. Mean and variance results for three bi-frequencies in the bicoherence. Theoretical variance results are obtained using (4).

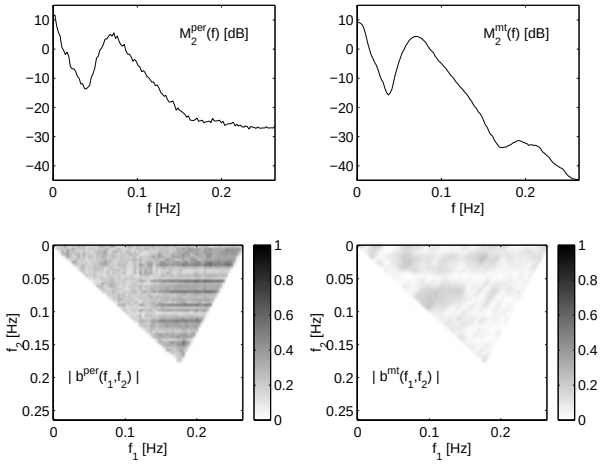

Fig. 2. Power spectrum (upper row) and magnitude bicoherence estimates (lower row) from TLP surge data, with $S = 22$, $M = 256$. Left: Using periodogram and biperiodogram. Right: Using adaptive multitaper estimation with $f_B = 4/M$.

fects [12]. The data we will use, comes from a model test of a TLP carried out in the wave basin of the Offshore Technology Research Center located at Texas A&M University. Using full scale measurements, the model had natural drift frequency at 0.007 Hz, while the random sea-waves were simulated as a Gulf of Mexico storm with most of their energy at frequencies around 0.07 Hz. The TLP surge motion data, after decimation, were divided into $S = 22$ segments each with $M = 256$ data points.

The averaged periodogram and a multitaper estimate of the surge power spectrum is shown in upper left and upper right corners of Fig. 2, respectively. The power spectra have two clear peeks, one near zero frequency (close to horizontal resonance frequency of 0.007 Hz) and the other around 0.07 Hz. The adaptive multitaper estimator [6] with bandwidth $f_B = 4/M$, provides a smoothed estimate without any ripples. Furthermore, since the leakage is reduced through the adaptive approach, it also shows some very low valued results near the Nyquist frequency ($f_N = 0.2635$ Hz).

In lower left and right corner of Fig. 2, the magnitude bicoherence estimate are shown for (bi-) periodogram ($M_3^{per}(f_1, f_2)$) and multitaper approach ($M_3^{mt}(f_1, f_2)$), respectively. The leakage produce wrong valued "stripes" in $M_3^{per}(f_1, f_2)$, and the remaining area is noisy. In the multitaper case, we see no "stripes" from leakage and the bicoherence magnitude can be found strongly frequency dependent.

5. CONCLUSIONS

Using the standard bicoherence estimator with periodogram and biperiodogram estimates, we have explained the statistical behavior involved with tapering and/or frequency smoothing for the bicoherence estimation case. To obtain an unbiased estimate, we have shown that it is extremely important to use correctly normalized power- and bispectral estimates. Tapering will reduce spectral leakage, and frequency smoothing will reduce the variance in bicoherence estimation.

We have suggested the use of a multitaper approach in bicoherence estimation to provide a clearly defined smoothing bandwidth. The multitaper approach has been shown to be superior to other non-parametric bicoherence estimators, and an adaptive approach can be used to reduce the leakage.

6. REFERENCES

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

[2] Y. C. Kim and E. J. Powers, "Digital bispectral analysis and its applications to nonlinear wave interactions," *IEEE Trans. Plasma Science*, vol. PS-7, no. 2, pp. 120–131, 1979.

[3] S. Elgar and R. T. Guza, "Statistics of bicoherence," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 36, no. 10, pp. 1667–1668, 1988.

[4] D. J. Thomson, "Multi-window bispectrum estimates," *Proc. IEEE Workshop on Higher-Order Statistics, Vail, Colorado*, pp. 19–23, 1989.

[5] Y. Birkelund and A. Hanssen, "Multitaper estimators for bispectra," *Proc. IEEE Workshop on Higher-Order Statistics, Caesarea, Israel*, pp. 207–211, 1999.

[6] Y. Birkelund and A. Hanssen, "Adaptive bispectral estimation using Thomson's multitaper approach," *Proc. IEEE Adaptive Systems for Signal Processing, Communication and Control Symposium, Lake Louise, Alberta, Canada*, pp. 283–288, 2000.

[7] Y. Birkelund and A. Hanssen, "Multiwindow bispectral estimation," *Proc. IEEE Workshop on Statistical Signal and Array Processing, Pocono, Pennsylvania*, pp. 640–644, 2000.

[8] V. Chandran and S. L. Elgar, "Mean and variance of estimates of the bispectrum of a harmonic random process - an analysis including leakage effects," *IEEE Trans. Signal Processing*, vol. 39, no. 12, pp. 2640–2651, 1991.

[9] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*, Cambridge University Press, Cambridge, U.K, 1993.

[10] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, no. 9, pp. 1055–1096, 1982.

[11] Y. Larsen and A. Hanssen, "Wavelet-polyspectra: Analysis of non-stationary and non-gaussian/non-linear signals," *Proc. IEEE workshop on Statistical Signal and Array Processing, Pocono, Pennsylvania*, pp. 539–543, 2000.

[12] B. Boashash, E. J. Powers, and A. M. Zoubir, *Higher-Order Statistical Signal Processing*, Wiley, New York, NY, 1995.