Mean and Variance of Estimates of the Bispectrum of a Harmonic Random Process—An Analysis Including Leakage Effects

Vinod Chandran, *Member, IEEE,* and Stephen L. Elgar, *Associate Member, IEEE*

***Abstract*—Analytical expressions are derived for the mean and variance of estimates of the bispectrum of a real-valued time series assuming a cosinusoidal model. The effects of spectral leakage, inherent in the discrete Fourier transform operation when the modes present in the signal have nonintegral number of wavelengths in the record, are included in the analysis. A single phase-coupled triad of modes can cause the bispectrum to have a nonzero mean value over the entire region of computation owing to leakage. The variance of bispectral estimates in the presence of leakage has contributions from individual modes and from triads of phase-coupled modes. Time-domain windowing reduces the leakage. The theoretical expressions for the mean and variance of bispectral estimates are derived in terms of a function dependent on an arbitrary symmetric time-domain window applied to the record, the number of data, and the statistics of the phase coupling among triads of modes. The theoretical results are verified by numerical simulations for simple test cases, and applied to laboratory data to examine phase coupling in a hypothesis testing framework.**

I. Introduction

In order to interpret bispectral values obtained from finite-length data records, the statistics of estimates of the bispectrum must be known. To achieve statistical stability, the time series is subdivided into short records, or blocks, for averaging. When there are a large number of short blocks, the estimate gains statistical stability at the expense of power spectral and bispectral resolution. Moreover, Fourier coefficients estimated from short records are susceptible to corruption from spectral leakage. This paper demonstrates that information about phase coupling between the true modes present in the signal can still be extracted from the bispectrum provided the statistics of the bispectrum are known, including the effects of leakage. Theoretical expressions are presented for the mean and variance of bispectral estimates calculated with Fourier-type methods. The results include the effects of phase-coupled triads, spectral leakage, and arbitrary symmetric data windows.

Brillinger and Rosenblatt [1] have investigated the asymptotic mean and variance of Fourier-type estimates of higher order spectra, and prove that under certain assumptions the kth order-spectral estimate is asymptotically unbiased and Gaussian distributed and that estimates of different order are asymptotically independent. They mention [2] that the pronounced ripples of the Dirichlet kernel (arising from spectral leakage for a rectangular window) may seriously distort the estimate and suggest tapering as a means of reducing the rippling. Several approximate expressions [3]–[8] for the mean and variance of the bispectrum have been proposed and used. All of these expressions, however, have ignored spectral leakage inherent in the discrete Fourier transform operation. The present study extends these results assuming a cosinusoidal model for the random process and relating the mean and variance of the direct Fourier-type of bispectrum estimate of the time series (assumed real) to the amplitude and phase statistics of the process, taking spectral leakage into consideration. For the sake of simplicity, it is assumed that statistical stability is provided by ensemble averaging, and there is no averaging over frequency bins in the bispectral domain. Assuming independence between the amplitude and phase for each true mode of oscillation of the process, the mean and variance may be related directly to the joint distributions of the phases. Hypothetical models with different assumptions about the phase statistics of the process under consideration (for example, model 1 may have no phase-coupled triads while model 2 may have a phase-coupled triad whose possibility is suggested by the presence of power spectral peak satisfying resonance conditions) may then be tested using the bispectrum estimate irrespective of whether the estimate is close to the true bispectrum or not. Thus, it is possible to quanitfy the statistical confidence in the interpretation of the conventional (i.e., direct Fourier transform) bispectrum estimate even in the presence of leakage.

Section II presents definitions and properties of the bispectral quantities needed for the subsequent statistical analysis. The mean and variance of bispectral estimates are derived in Appendixes A and B, respectively. The discrete Fourier transform (DFT) operation is expressed in terms of the discrete-time Fourier transform (DTFT) operation (z transform on the unit circle). A cosinusoidal model is assumed and knowing the DTFT's of the window and the time series, the DFT is expressed as a func-

Manuscript received October 26, 1989; revised September 12, 1990. This work was supported by the Coastal Sciences Branch of the Office of Naval Research.

The authors are with the Department of Electrical and Computer Engineering, Washington State University, Pullman, WA 991654-2752.

IEEE Log Number 9103197.

1053-587X/91$01.00 © 1991 IEEE

tion of the parameters (amplitudes, frequencies, and phases) of the model and a function dependent on the window, and thus includes leakage effects. The mean and the variance of bispectral estimates (Fourier type) involve products of discrete Fourier coefficients. These products are expanded (using MACSYMA[1]) and their phases are examined. Those terms that yield zero expected values under some reasonable assumptions are discarded. The remaining terms are separated into those that always contribute nonzero expected values (the phase-independent terms) and those that contribute nonzero expected values only in the presence of quadratic phase coupling. Important special cases of the expressions derived for the mean and the variance of bispectral estimates are discussed in Sections III and IV, respectively. Results from numerical simulations of random processes, including the effects of data windows, are compared to theory in Section V. The expressions are applied to laboratory data in Section VI. Conclusions follow in Section VII.

II. Bispectrum Estimate

Let a stationary, real-valued time series be subdivided into K subsections, each given by $g_i(n)$, $n = 0, 1, 2, \cdots, N-1$, with corresponding discrete Fourier transform

$$G_i(k) = \frac{1}{N} \sum_{n=0}^{N-1} g_i(n) \exp\left[-\frac{j2\pi kn}{N}\right] \tag{1}$$

for $k = 0, 1, 2, \cdots, N-1$. The estimate of the bispectrum is then defined as

$$\hat{B}(k_1, k_2) = \frac{1}{K} \sum_{i=1}^{K} G_i(k_1) G_i(k_2) G_i^*(k_1 + k_2) \tag{2}$$

in the triangular region $0 \leq k_1 \leq k_2 \leq k_1 + k_2 \leq N/2$ where it is unique. It is assumed that the data have been sampled at a rate fast enough to avoid spectral aliasing. If $w(n)$ is the time domain window applied to the data, the discrete Fourier transform (DFT) coefficient can be related to the discrete time Fourier transform (DTFT's) of the window and the signal as [9]

$$\begin{aligned} G(k) &= \text{DFT}\,[w(n)g(n)] \\ &= \frac{1}{N}\,\text{DTFT}\,[w(n)g(n)]\,(f)_{f=k/N} \\ &= \frac{1}{N}\,\mathcal{W}(f) \otimes \mathcal{G}(f)|_{f=k/N} \\ &= \frac{1}{N} \int_{-1/2}^{1/2} df\, \mathcal{W}\left(\frac{k}{N} - f\right) \mathcal{G}(f) \end{aligned} \tag{3}$$

where $\mathcal{W}(f)$ is the DTFT of the window function $w(n)$, $\mathcal{G}(f)$ is the DTFT of $g(n)$, $\otimes$ denotes convolution, and f denotes frequency normalized by the sampling frequency.

[1]A software system for symbolic and numerical mathematical manipulation distributed by Symbolics Inc., Cambridge, MA.

For a rectangular window

$$w(n) = \begin{cases} 1, & \text{for } n = 0, 1, \cdots, N-1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$\mathcal{W}(f) = \exp\,[-j\pi f(N-1)]\,\frac{\sin(N\pi f)}{\sin(\pi f)}.$$

Thus,

$$G(k) = \int_{-1/2}^{1/2} df \exp\,[-j\pi(k - fN)(1 - 1/N)] \cdot \frac{\sin\{\pi(k - fN)\}}{N \sin\{\pi(k - fN)/N\}}\,\mathcal{G}(f). \tag{4}$$

Assume a cosinusoidal model for the process that generates the real-time series $g(n)$, such that $g(n)$ is a superposition of P discrete cosinusoids

$$g(n) = \sum_{p=1}^{P} A_p \cos(2\pi f_p n + \phi_p). \tag{5}$$

A_p, ϕ_p denote the random amplitude and random phase, respectively, of the pth mode, whose frequency is f_p. Then the DTFT of $g(n)$ is given by

$$\mathcal{G}(f) = \sum_{p=1}^{P} \frac{A_p}{2}\,[\delta(f - f_p)e^{j\phi_p} + \delta(f + f_p)e^{-j\phi_p}] \tag{6}$$

where $\delta()$ denotes the Dirac delta function. Note that a harmonic model is a special case of the cosinusoidal model. The convolution integral in (4) can now be evaluated to yield an exact expression for the discrete Fourier coefficient

$$\begin{aligned} G(k) = \sum_{p=1}^{P} \frac{A_p}{2} \Bigg[&\exp\,[-j\pi(k - f_pN)(1 - 1/N)] \cdot \frac{\sin\{\pi(k - f_pN)\}}{N \sin\{\pi(k - f_pN)/N\}}\, e^{j\phi_p} \\ &+ \exp\,[-j\pi(k + f_pN)(1 - 1/N)] \cdot \frac{\sin\{\pi(k + f_pN)\}}{N \sin\{\pi(k + f_pN)/N\}}\, e^{-j\phi_p} \Bigg]. \end{aligned} \tag{7}$$

Hereafter, the Dirichlet kernel at $k - f_pN$ will be denoted by $D(k, f_p, N)$

$$D(k, f_p, N) = \frac{\sin\{\pi(k - f_pN)\}}{N \sin\{\pi(k - f_pN)/N\}}.$$

If the rectangular window is replaced by another symmetric window, $D(k, f_p, N)$ will be replaced by a different real function, and the expressions derived below will hold with the appropriate kernel substituted for $D(k, f_p, N)$.

Equation (7) reveals that $G(k)$ comprises two terms, one with a random phase component $+\phi_p$, and the other with a random phase component $-\phi_p$ from each mode. Far away from the frequency f_p the magnitudes of these two terms are nearly equal and therefore the resultant phase will, on average, be close to zero or $+\pi$. Thus, although ϕ_p may be uniform random in $[-\pi, +\pi)$ for each mode, the phase of $G(k)$ is different as a result of leakage. In fact, the phase of $G(0)$ is always zero or $\pm\pi$ (the zero frequency DFT is the average value and hence positive or negative for a real-time series), and between 0 and f_p there is a transition of the probability density function of the phase from impulses at 0, $\pm\pi$ to uniform random, for a given single mode. Thus, the effect of leaked power at normalized frequency k/N is not identical to that of white noise.

Expressions for the mean and the variance of the bispectrum estimator are derived in Appendixes A and B, respectively. The central idea behind the expansions of terms leading to these expressions is to cast them in terms of the parameters of the model and their joint statistics. The resulting expressions are easily programmed to get numerical values. The order of complexity of these expressions is $O(P^3)$, where P is the number of true modes present in the model. The order of complexity for a brute force numerical simulation for the uncoupled, random-phase case is $O(MKN \log N)$, where N is the length of the FFT, K is the blocks averaged in the estimate, and M is the number of realizations averaged in the numerical simulation. Typically $MKN > P^3$.

III. Mean of the Bispectrum

Assuming independence between the statistics of the amplitudes and the statistics of the phases, the expected value of the estimate of the bispectrum is given by (Appendix A)

$$E[\hat{B}(k_1, k_2)] = \sum_{p,q,r=1}^{P} \frac{1}{8} E[A_pA_qA_r] \left\{ \sum_{a,b,c=0}^{1} \cdot \{D(k_1, (-1)^a f_p, N) D(k_2, (-1)^b f_q, N) \cdot D(k_1 + k_2, (-1)^{c+1} f_r, N) \cdot \exp[-j\pi(N-1)((-1)^a f_p + (-1)^b f_q + (-1)^c f_r)] \, E[\exp[j((-1)^a\phi_p + (-1)^b\phi_q + (-1)^c\phi_r)]]\} \right\}. \tag{8}$$

Equation (8) shows the dependence of the mean of the bispectrum on the joint distributions of triads of amplitudes and phases. In general, the mean can be computed from the following information.

1) *Estimation Parameters:* The length of the DFT N, and the window dependent function D.

2) *Model Deterministic Parameters:* The number of modes P, and the frequency f_p of each mode $p = 1, 2 \cdots P$.

3) *Model Statistical Parameters (Amplitude):* $E[A_pA_qA_r]$ for all triads of modes $p, q, r = 1, 2 \cdots P$, $p \le q \le r$.

4) *Model Statistical Parameters (Phase):* $E[\exp[j(\phi_p + \phi_q \pm \phi_r)]]$ for all triads of modes $p, q, r = 1, 2 \cdots P, p \le q \le r$.

A. Interpretation of the Expression for the Mean

a) Even when all true modes are statistically independent and have random phases it is theoretically possible for the expected value of the bispectrum to be nonzero. When the modes p, q, r are independent

$$E[\exp[j(\pm\phi_p \pm \phi_q \pm \phi_r)]] = E[\exp[j(\pm\phi_p)]] E[\exp[j(\pm\phi_q)]] E[\exp[j(\pm\phi_r)]]$$

for $p \ne q \ne r \ne p$. If two modes are identical, say $p = q$, then the right-hand side of the above equations is $E[e^{j(\pm 2\phi_p)}] E[e^{j(\pm\phi_r)}]$ or $E[e^{j(\pm\phi_r)}]$, and if all three modes are identical, then the right-hand side is $E[e^{j(\pm 3\phi_p)}]$ or $E[e^{j(\pm\phi_p)}]$. Therefore, even if $E[e^{j(\pm\phi_i)}] = 0$ for every mode i, there will be a nonzero contribution to the mean for every (k_1, k_2) if $E[e^{j3\phi_i}] \ne 0$ for any mode $i = 1, 2 \cdots P$.

b) For random phases uniformly distributed in $[0, 2\pi)$, however, the expected value of the bispectrum is zero even in the presence of leakage. In this case, $E[e^{j\phi_p}] = E[e^{j2\phi_p}] = E[e^{j3\phi_p}] = 0$ for all modes $p = 1, 2 \cdots P$. Note that $E[e^{j\omega\phi_p}]$ is the characteristic function of the phase ϕ_p and is equal to zero for all $\omega \ne 0$ for a uniform probability density.

c) Even a single phase-coupled triad of modes can cause the expected value to be nonzero in the entire region of computation, although the contribution decreases for points far away from the true triad.

d) The expected value can be computed for any symmetric window by replacing the function $D(k, f, N)$ with one that is appropriate for the particular window.

The expression for the mean is now evaluated for some simple test cases.

Case A (Random-phase triad without leakage): Let the model for the random process consist of 3 modes with independent, uniformly distributed random phases at frequencies f_1, f_2, and f_3 where $f_1 = k_1'/N$, $f_2 = k_2'/N$, and $f_3 = (k_1' + k_2')/N$ for some integers $k_1', k_2' \le N/2$ with $k_1' \le k_2'$. Then

$$E[\exp[j(\pm\phi_1 \pm \phi_2 \pm \phi_3)]] = E[\exp[j \pm\phi_2)]] E[\exp[j(\pm\phi_2)]] \cdot E[\exp[j(\pm\phi_3)]] = 0$$

and hence $E[\hat{B}(k_1, k_2)] = 0$ for all k_1, k_2 in the region of computation.

Case B (Random-phase triad with leakage: Let the model consist of 3 modes with independent, uniformly distributed random phases a frequencies f_1, f_2, and f_3 where $f_1 \ne k_1'/N$, $f_2 \ne k_2'/N$, and $f_3 \ne (k_1' + k_2')/N$ for any integers k_1', k_2'. Then, as in case A, $E[\hat{B}(k_1, k_2)] = 0$ for all k_1, k_2 in the region of computation. Thus, spectral leakage does not affect the expected value of the bispectrum for the case of uniformly distributed random phases with no phase coupling between modes.

Case C (Phase-coupled triad without leakage): Let the model consist of 3 modes with coupled phases such that $\phi_1 + \phi_2 = \phi_3$ at frequencies f_1, f_2, and f_3 where $f_1 = k_1'/N$, $f_2 = k_2'/N$, and $f_3 = (k_1' + k_2')/N$ for some integers $k_1', k_2' \leq N/2$ with $k_1' \leq k_2'$.

In this case, $E[\exp[j(\phi_1 + \phi_2 - \phi_3)]] = 1$.

Assume statistically independent amplitudes and let $E[A_1A_2A_3] = \overline{A}_1\overline{A}_2\overline{A}_3$. Then, for $i = 1, 2$,

$$D(k, f_i, N) = \begin{cases} 1, & \text{for } k = k_i'; \\ 0, & \text{otherwise} \end{cases}$$

and hence

$$E[\hat{B}(k_1, k_2)] = \begin{cases} \frac{1}{8}\overline{A}_1\overline{A}_2\overline{A}_3 & \text{when } k_1 = k_1' \text{ and } k_2 = k_2'; \\ 0, & \text{otherwise} \end{cases}$$

in the region of computation.

Case D (Phase-coupled triad with leakage): Let the model consist of 3 modes with coupled phases such that $\phi_1 + \phi_2 = \phi_3$ at frequencies f_1, f_2, and $f_3 = f_1 + f_2$ where $f_1 \neq k_1'/N$, $f_2 \neq k_2'/N$ for any integers k_1', k_2'. Again, $E[\exp[j(\phi_1 + \phi_2 - \phi_3]] = 1$. Assume statistically independent amplitudes and let $E[A_1A_2A_3] = \overline{A}_1\overline{A}_2\overline{A}_3$. Then from (8)

$$E[\hat{B}(k_1, k_2)] = \frac{1}{8}\overline{A}_1\overline{A}_2\overline{A}_3 \sum_{(f_p, f_q, f_r)\in C_D} D(k_1, f_p, N) \cdot D(k_2, f_q, N) D(k_1 + k_2, f_r, N)$$

where C_D is a subset of the set of all possible triads chosen from $(\pm f_1, \pm f_2 \pm f_3)$ such that with the appropriate signs on the phases they reduce to $\pm(\phi_1 + \phi_2 - \phi_3)$. (See [10] for more details.) Unlike the random-phase case, spectral leakage does affect the estimates of the mean of the bispectrum in the case of phase-coupled triads. Note that the highest values of the mean of the bispectrum are for those k_1 and k_2 that are closest to f_1N and f_2N, respectively, because $[D(k_1, f_1, N) D(k_2, f_2, N) D(k_1 + k_2, f_1 + f_2, N)]$ is then close to its maximum value of unity, but all triads of Fourier components $(k_1, k_2, k_1 + k_2)$ will have nonzero bispectral values owing to spectral leakage.

Case E (White noise): Let the model be a discrete frequency approximation to white noise consisting of P statistically independent modes at frequencies f_p for $p = 1, 2 \cdots P$ uniformly spaced in the frequency interval [0, 0.5]. Note that P may be larger than N. For statistically independent amplitudes and statistically independent uniformly distributed phases

$$E[\exp[j(\pm\phi_p \pm \phi_q \pm \phi_r)]] = E[\exp[j(\pm\phi_p)]]E[\exp[j(\pm\phi_q)]] \cdot E[\exp[j(\pm\phi_r)]] = 0$$

and hence, as in case A, $E[\hat{B}(k_1, k_2)] = 0$ for all k_1, k_2 in the region of computation, as was shown by Brillinger and Rosenblatt [1].

IV. Variance of the Bispectrum

The variance of the bispectrum estimate is (Appendix B and [10])

$$\text{var}[\hat{B}(k_1, k_2)] = E[\hat{B}(k_1, k_2)\hat{B}^*(k_1, k_2)] - E^2[\hat{B}(k_1, k_2)] = \frac{1}{K}\{E[G^2(k_1)G^2(k_2)G^2(k_1 + k_2)] - E^2[\hat{B}(k_1, k_2)]\}. \tag{9}$$

Since the quantity within the braces is independent of K

$$\lim_{K\to\infty} \text{var}[\hat{B}(k_1, k_2)] = 0$$

and hence $\hat{B}(k_1, k_2)$ is a consistent estimator [1].

Recall that $G(k)$ contained a contribution from each mode present in the model. Therefore, the expansion of $[G^2(k_1)G_2(k_2)G^2(k_1 + k_2)]$ contains P^6 terms, each term representing the contribution from a sextuple of modes. The variance will thus depend, in general, on statistics of the amplitude and phase up to the sixth order. One or more of the six modes indicated above may be identical and assumptions can be made that considerably reduce the number of terms involved. First, assume that the amplitude statistics are independent of the phase statistics. The representative factor containing the phases in each term is then of the form $E[\exp[j(\pm\phi_p \pm \phi_q \pm \phi_r \pm \phi_{p'} \pm \phi_{q'} \pm \phi_{r'})]]$. Next, if there is no cubic or higher order coupling, all terms that involve four or more distinct modes can be eliminated. Note that each of the P^6 terms mentioned above actually expands into 2^6 subterms because of the positive and negative signs on each phase. For terms involving four or more distinct modes, the phases cannot cancel each other with positive and negative signs. A term for which the phases cancel contributes to the variance of bispectral estimates even in the absence of any phase coupling. Such terms are referred to here as phase-independent terms. Similar arguments reveal that the only subsets of terms that can contribute nonzero values are a) those that have six identical modes, say S_6; b) those that have four identical modes and another two identical modes, say $S_{4,2}$; and c) those that have three pairs of two identical modes, each say $S_{2,2,2}$. The terms in S_6 have nonzero contributions only from those subterms that have three positive and three negative signs so that the phases cancel. Let T_6 denote the sum of all such terms. Similarly, the terms in $S_{1,2}$ and $S_{2,2,2}$ contain subterms that contribute a nonzero value because the phases cancel. Their sums are denoted by $T_{4,2}^0$ and $T_{2,2,2}^0$, respectively. In addition, $S_{4,2}$ and $S_{2,2,2}$ contain some subterms which contribute nonzero values only in the presence of quadratic phase coupling, because they have factors such as $E[\exp[j2(2\phi_p \pm \phi_q)]]$ and $E[\exp[j2(\phi_p + \phi_q \pm \phi_r)]]$. Denote the sum of these as $T_{4,2}^\phi$ and $T_{2,2,2}^\phi$, respectively. Then

$$\text{var}[\hat{B}(k_1, k_2)] = T_6 + T_{4,2} + T_{2,2,2} - E^2[\hat{B}(k_1, k_2)] \tag{10}$$

where $T_{4,2} = T_{4,2}^0 + T_{4,2}^\phi$ and $T_{2,2,2} = T_{2,2,2}^0 + T_{2,2,2}^\phi$. Appendix B gives the full expressions and [10] provides the details of the derivation. With the above assumptions, the variance can be computed from the following additional information:

Model Statistical Parameters (Amplitude): $E[A_p^2A_q^2A_r^2]$ for all triads of modes $p, q, r = 1, 2 \cdots P$, $p \le q \le r$.

Model Statistical Parameters (Phase): $E[\exp[j2(\phi_p + \phi_q \pm \phi_r)]]$ for all triads of modes $p, q, r = 1, 2 \cdots P$, $p \le q \le r$.

The estimation parameters and the model deterministic parameters are the same as those for the mean.

The bispectrum is complex in general and the variance given above is the sum of the variances of the real and imaginary parts of the bispectrum. For large N and K, it has been shown [1], [5], [6] that the variances of the real part and the imaginary part are equal. Even without the assumption of large K it can be shown that the variances of the real part and imaginary part are equal (see Appendix C).

The same special cases that were discussed previously are now considered.

Case A (Random-phase triad without leakage): $\gamma(f_p, f_q, f_{r'}, f_{p'}, f_{q'}, f_r)$, given by (B4), is zero for all $f_p, f_q, f_{r'}, f_{p'}, f_{q'}, f_r$ and all k_1, k_2 in the region of computation except when $k_1 = k_1$ and $k_2 = k_2'$, for which $\gamma(f_1, f_2, f_3, -f_1, -f_2, -f_3)$ equals 1. Therefore $T_6 = 0$, $T_{4,2} = 0$, $T_{2,2,2} = \frac{1}{64}E[A_1^2A_2^2A_3^2]$, and since the mean value is zero

$$\text{var}[\hat{B}(k_1, k_2)] = \begin{cases} \dfrac{1}{K}\dfrac{1}{64}E[A_1^2A_2^2A_3^2] & \text{for } k_1 = k_1', k_2 = k_2' \\ 0 & \text{otherwise.} \end{cases}$$

Assuming statistically independent amplitudes,

$$E[A_1^2A_2^2A_3^2] = E[A_1^2]\,E[A_2^2]\,E[A_3^2] = \bar{A}_1^2\bar{A}_2^2\bar{A}_3^2 \propto P_1P_2P_3$$

where P_i is the power of mode i. Note that the variance is proportional to the triple product of power at each component of the triad, as originally shown by Brillinger and Rosenblatt [1], [2].

Case B (Random-phase triad with leakage): The mean value for this case was found to be zero. $E[\exp[j2(\pm 2\phi_i \pm \phi_j)]] = 0$ and $E[\exp[j2(\pm\phi_i \pm \phi_j \pm \phi_k)]] = 0$ for all modes $i \ne j \ne k \ne i$; $i, j, k = 1, 2, 3$. Thus

$$\text{var}[\hat{B}(k_1, k_2)] = \frac{1}{K}(T_6 + T_{4,2}^0 + T_{2,2,2}^0).$$

Consequently, spectral leakage for the case of random-phase triads increases the variance of the bispectral estimates for all k_1, k_2 relative to the case without leakage. Note that the case without leakage had zero variance for all $k_1 \ne k_1'$, $k_2 \ne k_2'$.

Case C (Phase-coupled triad without leakage): Although $E[\exp[\pm j2(\phi_1 + \phi_2 - \phi_3)]] = 1$ for $k_1 = k_1'$ and $k_2 = k_2'$, $T_{4,2}^\phi$ and $T_{2,2,2}^\phi$ are zero because in this case the products of sinc functions involved in their coefficients is zero. As in case A, $\gamma(f_p, f_q, f_{r'}, f_{p'}, f_{q'}, f_r)$ is zero for all $f_p, f_q, f_{r'}, f_{p'}, f_{q'}, f_r$ and all k_1, k_2 in the region of computation except when $k_1 = k_1'$ and $k_2 = k_2'$, for which $\gamma(f_1, f_2, f_3, -f_1, -f_2, -f_3) = 1$. In fact, $\gamma(f_1, f_2, f_3, -f_1, -f_2, -f_3)$ appears only once in the expression for variance and this is in the expansion of $\tau_{2,2,2}(+f_1, -f_1, +f_2, -f_2, +f_3, -f_3)$ (see (B7)). Thus, $T_6 = 0$, $T_{4,2} = 0$, and

$$T_{2,2,2} = \begin{cases} \frac{1}{64}E[A_1^2A_2^2A_3^2] & \text{when } k_1 = k_1', k_2 = k_2' \\ 0 & \text{otherwise.} \end{cases}$$

Assuming statistical independence between the amplitudes,

$$E[A_1^2A_2^2A_3^2] = E[A_1^2]\,E[A_2^2]\,E[A_3^2] = \bar{A}_1^2\bar{A}_2^2\bar{A}_3^2.$$

Substituting these and the expression for the mean derived in the previous section into (10), yields

$$\text{var}[\hat{B}(k_1, k_2)] = \begin{cases} \dfrac{1}{K}\dfrac{1}{64}(\bar{A}_1^2\bar{A}_2^2\bar{A}_3^2 - \overline{A}_1^2\overline{A}_2^2\overline{A}_3^2) & \text{for } k_1 = k_1', k_2 = k_2' \\ 0 & \text{otherwise.} \end{cases}$$

If the amplitudes are constant,

$$E[A_1^2]E[A_2^2]E[A_3^2] = E^2[A_1]E^2[A_2]E^2[A_3]$$

and

$$\text{var}[\hat{B}(k_1, k_2)] = 0 \qquad \text{for all } k_1, k_2$$

in the region of computation.

Case D (Phase-coupled triad with leakage): $E[\exp[\pm j2(\phi_1 + \phi_2 - \phi_3)]] = 1$ in this case. Consequently, the expression for variance involves $T_{4,2}^\phi$ if the interaction involves a mode with itself or $T_{2,2,2}^\phi$ if it involves two different modes, as well as the phase-independent terms. From a phase-coupled triad of modes power leaks into all possible triads of frequencies in the region of computation, resulting in both an increased mean and an increased variance for the bispectrum relative to the mean and variance in the absence of the coupled triad.

Case E (White noise): The mean was found to be zero in the previous section. Only the phase-independent terms in $T_{4,2}$ and $T_{2,2,2}$ contribute to the variance and, therefore,

$$\text{var}[\hat{B}(k_1, k_2)] = \frac{1}{K}(T_6 + T_{4,2}^0 + T_{2,2,2}^0)$$

for all k_1, k_2 in the region of computation.

If N (length of the DFT) is increased for a given total number of points in the record, K (which is the total number of data divided by N) decreases and hence variance increases. As $N \to \infty$ while K remains constant, the number of modes P, which will be proportional to N for white noise (having constant power spectral density over the entire frequency range), also tends to infinity and hence the variance increases as shown previously [1], [5], [6].

A. Interpretation of the Expression for the Variance

Although the expression for the variance in (10) is unwiedly, it is quite easily programmed and computed, yielding quantitative values. In addition, (10) qualitatively indicates the following.

a) Every mode that leaks power increases the variance of bispectral estimates at all triads of frequencies in the region of computation.

b) Triads of coupled modes leak both coupled power and random power into all possible triads of frequencies in the region of computation, raising the mean and the variance of bispectral estimates.

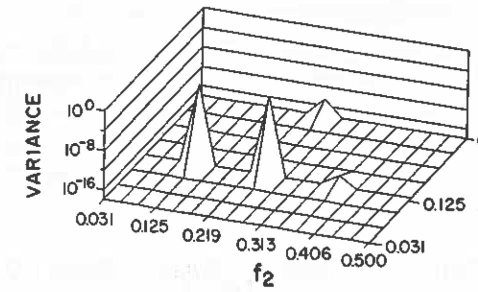
c) Given *a priori* knowledge about the modes involved and their amplitude and phase statistics, the statistics of the bispectrum, including the effects of leakage can be computed. This enables the time series to be divided into a larger number (K) of blocks yielding lower variances and hence statistically better estimates for verification of hypotheses.

V. Numerical Results and Simulations

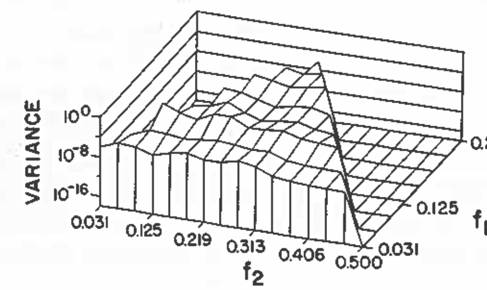
In this section the theoretical mean and variance of bispectral estimates for each of the cases discussed above are compared to corresponding values obtained from numerical simulations of random processes. Each subsection of the simulated time series was $N = 32$ points long and 256 realizations were averaged. Each realization was $K = 1$ subsection long. It suffices to show the correspondence for $K = 1$. For arbitrary K (K subsections in each realization) the variance gets divided by K while the mean remains unchanged. Unless otherwise stated, all Fourier amplitudes were equal to unity, and a rectangular data window was used. Further, the mean of the bispectrum refers to the real part of the mean for the test cases because the imaginary part is identically zero, and the variance of the bispectrum refers to the sum of the variances of the real and imaginary parts.

Case A (Random-phase triad without leakage): The model consisted of three modes with frequencies $f_a = 0.125$, $f_b = 0.250$, and $f_c = 0.375$, which correspond to exact integer number of wavelengths in a 32-point record (4/32, 8/32, and 12/32, respectively). Their phases were uniform random in $[0, 2\pi)$. Fig. 1(a) shows the variance of the bispectrum computed from (10). The values obtained from numerical simulations (not shown) are identical within the limits of numerical accuracy. The variance for the triad (0.125, 0.250, 0.375) is 1/64 as expected because $E^2[A_i] = 1$, $i = 1, 2, 3$, and $K = 1$ (see case A in Section IV). Note that (0.125, 0.125, 0.250) is also a triad of modes satisfying the resonance condition, and it is not phase coupled and the variance at ($f_1 = 0.125, f_2 = 0.125$) is also 1/64. Values at (0.250, 0.250) and (0.125, 0.375) are zero within the limits of numerical accuracy for 32-b floating point arithmetic. Theoretical and simulated values for the mean of the bispectrum (not shown) also agree to within these limits of accuracy.

Case B (Random-phase triad with leakage): If the frequencies of the modes in case A are shifted, there will be



(a)



(b)

Fig. 1. Variance of the estimate of the bispectrum for the test cases A and B. The frequencies (f_1, f_2) are normalized by the sampling frequency. $f_1 = k_1N$ and $f_2 = k_2N$ where $N = 32$ is the length of each record. (a) Random-phased triad with no leakage, (b) random-phase triad with leakage.

spectral leakage. In case B, the model consisted of three modes with frequencies $f_a = 0.1$, $f_b = 0.2$, and $f_c = 0.3$, which do not correspond to exact integer number of wavelengths in a 32-point record. Their phases were uniform random in $[0, 2\pi)$. Fig. 1(b) shows the variance of the bispectrum computed from (10). The values obtained from numerical simulations are identical. Note the spreading of the variance over the region of computation in Fig. 1(b).

Case C (Phase-coupled triad without leakage): The model consisted of the same three modes as defined for case A, except that now their phases were perfectly coupled. Figs. 2(a) and (b) show the theoretical mean and variance of the bispectrum computed from (8) and (10), respectively. The values obtained from numerical simulations are identical. The variance at (0.125, 0.250) is zero within the limits of numerical accuracy, as expected. Note that (0.125, 0.125, 0.250) is also a triad of modes satisfying the resonance condition, but it is not phase coupled and the variance at (0.125, 0.125) is therefore 1/64 (as pointed out in case A). Values at (0.250, 0.250) and (0.125, 0.375) are also zero within limits of accuracy.

Case D (Phase-coupled triad with leakage): The frequencies of the modes in case C can be shifted introducing spectral leakage, and thus, in case D, the model consisted of three modes with frequencies, $f_a = 0.1$, $f_b = 0.2$, and $f_c = 0.3$, which do not correspond to exact integer number of wavelengths in a 32-point record. Their phases were perfectly coupled. Figs. 2(c) and (d) show the theoretical mean and variance of the bispectrum. The values obtained from numerical simulations are identical. The variance near (0.1, 0.2) for the phase-coupled case (Fig.

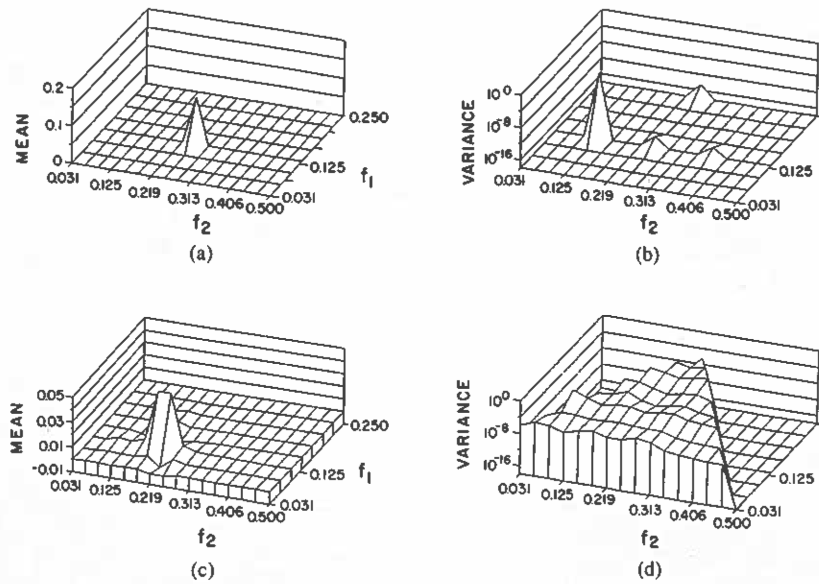


Fig. 2. Mean and variance of the estimate of the bispectrum for the test cases C and D. (a) and (b) are for phase-coupled triad with no leakage, and (c) and (d) are for phase-coupled triad with leakage.

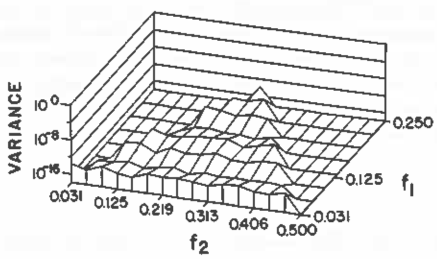


Fig. 3. Variance of the estimate of the bispectrum for the test case E (white noise).

2(d)) is lower than that for the random-phase case (Fig. 1(b)).

Case E (White noise): The model consisted of 16 modes. The ith mode had frequency randomly placed in the interval $[0.5(i - 1)/16, 0.5i/16]$, and had amplitude equal to 0.01. Fig. 3 shows the theoretical variance of the bispectrum computed from (10). The values obtained from numerical simulations are identical. Note that the variance is approximately constant in the region of computation as expected for white noise.

A. Hanning Window

As discussed above, the theoretical expressions for the mean and variance of estimates of the bispectrum can be applied to windowed data by replacing the Dirichlet kernel $[D(k, f_p, N)]$ appropriate for a rectangular window with the kernel corresponding to the window at hand. For example, consider a Hanning window

$$w_N(n) = \begin{cases} 0.5 \left[1 - \cos\left(\frac{2\pi n}{N-1}\right)\right] & \\ \quad\text{for } n = 0, 1, \cdots, N-1 & \\ 0, \quad \text{otherwise.} \end{cases}$$

The DTFT of this window, $\mathcal{W}_N(f)$, is given by

$$\mathcal{W}_N(f) = 0.5\mathcal{W}_R(f) - 0.25\left[\mathcal{W}_R\left(f - \frac{1}{N-1}\right) + \mathcal{W}_R\left(f + \frac{1}{N-1}\right)\right]$$

and

$$\mathcal{W}_N\left(\frac{k}{N} - f\right) = \exp\left[-j\pi(k - fN)(1 - 1/N)\right]N \cdot \left[0.5D(k, f, N) + 0.25D\left(k, f - \frac{1}{N-1}, N\right) + 0.25D\left(k, f + \frac{1}{N-1}, N\right)\right]$$

where $\mathcal{W}_R(f)$ is the DTFT of a rectangular window. Thus, the expressions derived for the mean and the variance of the bispectrum estimator can be used for a Hanning window by substituting the function

$$D_N(k, f_p, N) = \frac{1}{2}D(k, f_p, N) + \frac{1}{4}\left[D\left(k, f_p - \frac{1}{N-1}, N\right) + D\left(k, f_p + \frac{1}{N-1}, N\right)\right]$$

for $D(k, f_p, N)$. Similar functions may likewise be derived for other windows. Case D was repeated for a Hanning window in place of a rectangular window. Figs. 4(a) and (b) show the theoretical mean and variance of the bispec-

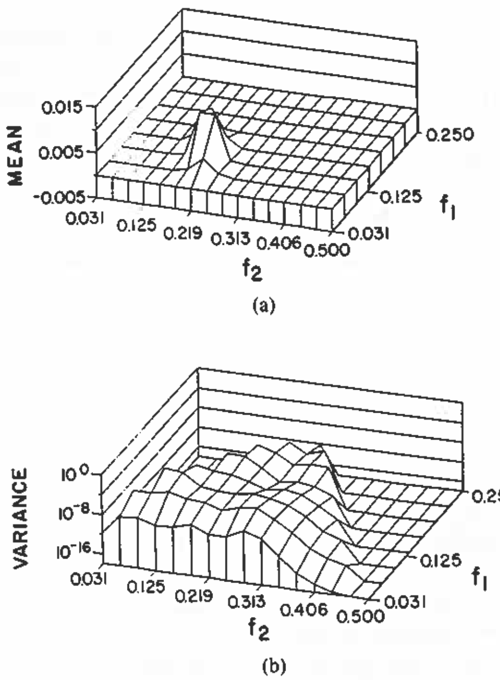


Fig. 4. Mean and variance of the estimate of the bispectrum for the test case D (phase-coupled triad with leakage), with a Hanning window in the time domain.

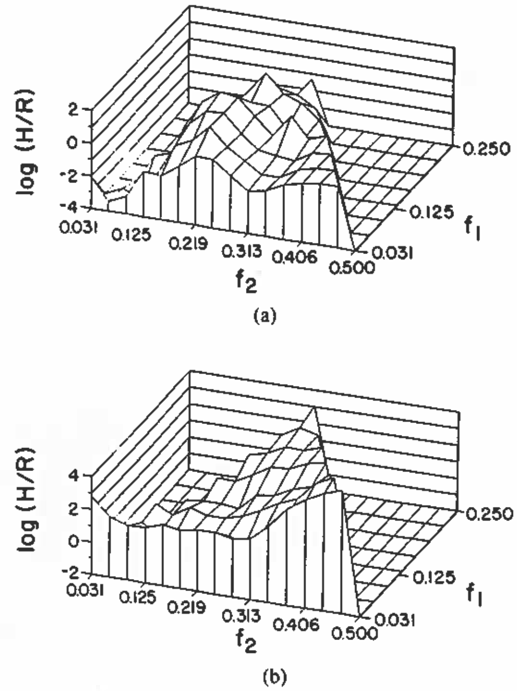


Fig. 5. H/R where H is the ratio of the mean of the real part of the estimate of the bispectrum to the standard deviation for the Hanning-windowed case and R is the same ratio for the rectangular-windowed case. A value of H/R greater than unity (i.e., $\log_{10}(H/R) > 0$) suggests a statistically improved estimate for the Hanning-windowed case. (a) Phase-coupled triad with leakage (test case D), (b) random-phase triad with leakage (test case B).

trum computed from (8) and (10), respectively, with D_N substituted for D, and the simulated results are identical.

In order to compare the two windows, the ratio of the mean to the standard deviation of the real part of the estimate of the bispectrum is computed for each window (the imaginary part is identically zero for the test cases). The ratio of this quantity for the Hanning window to that for the rectangular window is plotted in Figs. 5(a) and (b) for cases D (phase coupled) and B (not phase coupled), respectively. This ratio is independent of K. Note that for the coupled case (Fig. 5(a)) windowing results in improved estimates (i.e., ratio > 1) only in the vicinity of the coupled triad ($f_1 = 0.1, f_2 = 0.2$), while for the random case (Fig. 5(b)) the Hanning window results in improved estimates everywhere in the region of computation. The significance of this result is that if the observed bispectrum is compared with the bispectrum for phase-coupled and random-phase models in a hypothesis testing framework, the values in the vicinity of the phase-coupled triads serve as better estimates and they are improved by time-domain windowing. Although time-domain windowing reduces the relative variance of estimates of the bispectrum near phase-coupled triads in the presence of spectral leakage, windowing can decrease the resolution of bispectral estimates, similar to the effect on estimates of the power spectrum.

VI. Application to Laboratory Data

In this section laboratory data are examined using the theoretical results derived above. The data consist of a time series of water elevation in a long flume with a wavemaker (a computer controlled paddle) at one end and an absorbing beach (to reduce reflections) located far downstream. A water elevation sensor was placed one wavelength (3 m) downstream of the paddle, which was driven at approximately 0.3 Hz. Both the paddle motions [11] and the shallow water depths [12] result in quadratically phase-coupled harmonics. A 12288-point record (sampling frequency = 20 Hz) was analyzed. Power spectra for 512-point and 64-point records are shown in Fig. 6. Notice that leakage corrupts the power spectrum as the block size becomes smaller. The fundamental frequency was estimated as 0.303 Hz from a 2048-point block spectral analysis. The amplitudes of the fundamental mode and its first 15 harmonics were estimated from a 512-point Blackman windowed spectral analysis making use of the *a priori* information about the true frequencies. The model chosen consisted of 16 cosinusoids. $E[A_pA_qA_r]$ was assumed to be $A_pA_qA_r$, that is, randomness in the amplitudes was ignored. A 512-point bispectral analysis was performed and the biphases of the triads closest to (f_p, f_q, f_r) were used to compute $E[\exp[j(\phi_p + \phi_q - \phi_r)]]$, the statistical phase parameters required by the model for the coupled case. Notice that for the random model these expected values are zero. For computing the variance, $E[\exp[j2(\phi_p + \phi_q - \phi_r)]]$ must be known. For the random case these expected values are zero. For the coupled case, a value of 1 was assumed to get a worst case estimate of the variance. The mean and standard deviation were computed for both random-phase and phase-coupled cases for 64-point blocks so that $K = 192$. The phase-

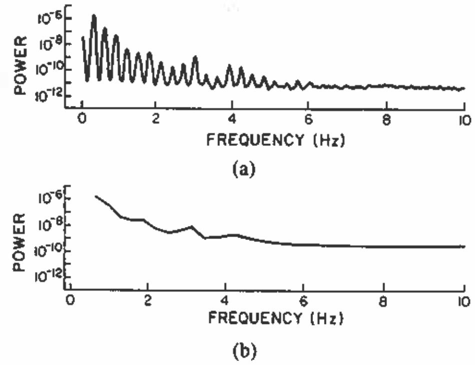


Fig. 6. Power spectrum of the laboratory data (the Nyquist frequency is 10 Hz): (a) 512-point blocks, Blackman windowed, average of 24 blocks, (b) 64-point blocks, rectangular windowed, average of 192 blocks.

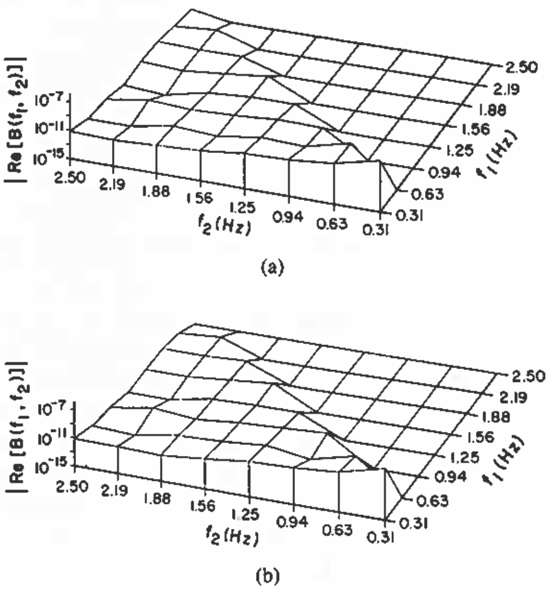


Fig. 7. The absolute value of the real part of the bispectrum. (a) Data, (b) expected value from the phase-coupled model.

coupled model assumed all possible triads of modes from the 16 harmonics that satisfy resonance conditions to be phase coupled. The absolute value of the real and imaginary parts of the bispectrum for the actual data and that expected from the phase-coupled model are plotted in Figs. 7 and 8. Note that for 64-point blocks the bispectrum does not have resolved peaks, and that the bispectrum computed from the phase-coupled model matches that computed from the data fairly well.

To determine, in a statistical sense, which model (random phase or phase-coupled) the data are consistent with, the ratio of the number of standard deviations that the real part of the observed bispectrum differs from the expected value for the random model (which is zero) to the number of standard deviations that the real part of the observed bispectrum differs from the expected value for the phase-coupled model is shown in Fig. 9(a). Only the ratios for triads consisting of the power spectral primary peak frequency (0.3125 Hz) and its harmonics are shown. Values greater than unity show deviation from the random model and conformity to the phase-coupled model. Fig. 9(b) shows the same ratio for the imaginary part of the bispectrum. Both figures indicate strong phase coupling in the data.

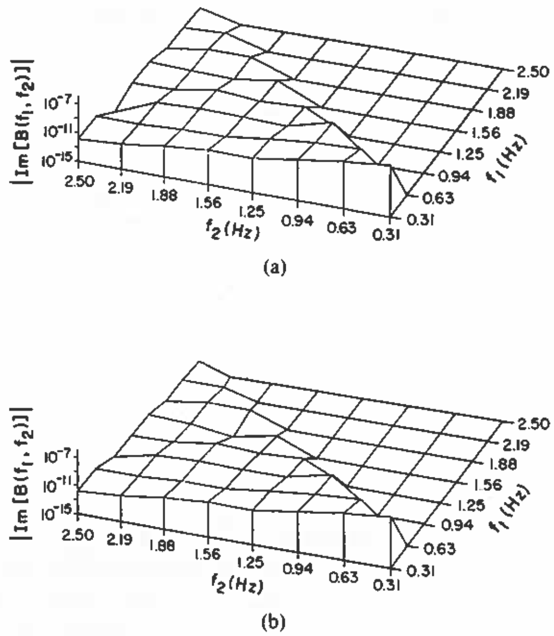


Fig. 8. The absolute value of the imaginary part of the bispectrum. (a) Data, (b) expected value from the phase-coupled model.

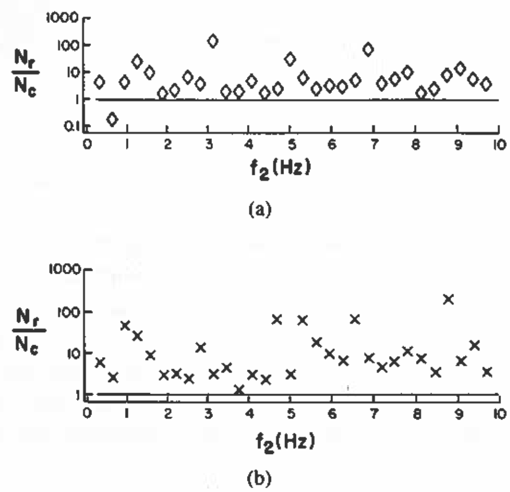


Fig. 9. Index of deviation from the random-phase model and conformity to the phase-coupled model versus frequency (f_2) of one of the components of triads consisting of f_1, f_2, $f_1 + f_2$, where $f_1 = 0.3125$ Hz is the power spectral primary peak frequency. The index N_r/N_c is the ratio of the number of standard deviations that the observed bispectrum differs from the expected value for the random model to the number of standard deviations that it differs from the expected value for the phase-coupled model. (a) Real part, (b) imaginary part.

Finally, to show that the theoretical results presented here can be used to divide the time series into a large number of blocks and achieve better statistical stability, the mean and standard deviation were computed for the same model with $K = 24$ and $K = 192$. These correspond to 512-point blocks (power spectral peaks are fairly well resolved as in Fig. 6(a)) and 64-point blocks (power spectrum is corrupted by leakage as in Fig. 6(b)). The differences between the expected values of the real part of the bispectrum for the phase-coupled model and the random-

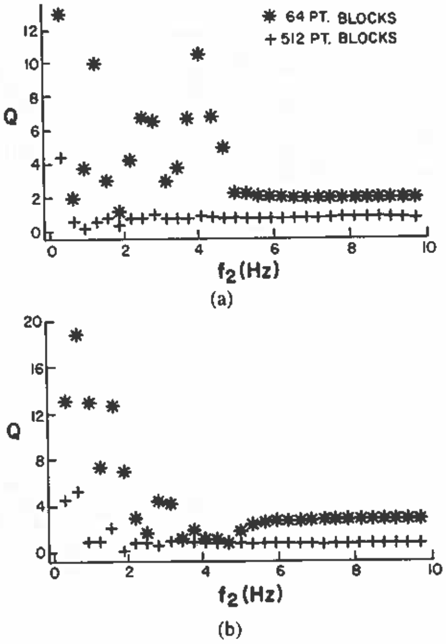


Fig. 10. Quality factor (Q) versus frequency (f_2) of one of the components of triads consisting of f_1, f_2, $f_1 + f_2$, where $f_1 = 0.3125$ Hz is the power spectral primary peak frequency. Q is the difference between the expected values of the bispectrum for the two models (random and phase coupled) normalized by the sum of the standard deviations for the two models. The asterisks are for $K = 192$ (64-point blocks) and the cross marks are for $K = 24$ (512-point blocks). (a) Real part, (b) imaginary part.

phase model (which is zero) normalized by the sum of the standard deviations for the two models are plotted in Fig. 10. The ratio of 64-point blocks is higher than that for 512-point blocks, indicating that better statistical reliability is possible by dividing the data into a larger number of blocks. By taking leakage into account in computing the statistics of the bispectrum, the reliable but low resolution estimates can be used to address questions of phase coupling in the hypothesis testing framework described above.

VII. Conclusions

Exact theoretical expressions are presented for the mean and the variance of estimates of the bispectrum of a harmonic random process in the presence of spectral leakage. In many practical situations where quadratic phase coupling and nonlinear mechanisms in the generation of time series are being investigated, it may be possible to deduce a harmonic model on theoretical considerations. Then the theoretical expressions derived here can be used to place the problem in a hypothesis-testing framework where the data may be compared to different models (with phase-coupled or random-phase triads). The theoretical expressions were verified by numerical simulations for a variety of test cases, and applied to laboratory data. Given *a priori* information about the modes involved in a random process and their amplitude and phase statistics (which may be obtained from a preliminary analysis), these expressions may be used to obtain the statistics of the bispectrum in the presence of leakage for arbitrary data windows.

Appendix A
Derivation of the Mean

The expected value of the bispectrum estimator, $\hat{B}(k_1, k_2)$, is

$$E[\hat{B}(k_1, k_2)] = E\left[\frac{1}{K}\sum_{i=1}^{K} G_i(k_1)G_i(k_2)G_i^*(k_1 + k_2)\right]$$

$$= \frac{1}{K}\sum_{i=1}^{K} E[G_i(k_1)G_i(k_2)\, G_i^*(k_1 + k_2)]$$

(linearity)

$$= E[G_i(k_1)G_i(k_2)G_i^*(k_1 + k_2)]$$

$$= E[G(k_1)G(k_2)G^*(k_1 + k_2)] \qquad \text{(stationary)}$$

$$= \sum_{p,q,r=1}^{P} \frac{1}{8} E\Bigg[A_pA_qA_r \sum_{a,b,c=0}^{1} \cdot \{D(k_1, (-1)^a f_p, N) D(k_2, (-1)^b f_q, N) \cdot D(k_1 + k_2, (-1)^{c+1} f_r, N) \cdot \exp[-j\pi(N-1)((-1)^a f_p + (-1)^b f_q + (-1)^c f_r)] \cdot \exp[j((-1)^a \phi_p + (-1)^b \phi_q + (-1)^c \phi_r)]\}\Bigg]. \tag{A1}$$

The expectation operator can be taken inside the summation by virtue of its linearity. When the amplitude and phase of each mode are independent of each other, their expected values can be separated.

Appendix B
Derivation of the Variance

The variance, var $[\hat{B}(k_1, k_2)]$, of the bispectrum estimator is

$$\text{var}\,[\hat{B}(k_1, k_2)] = E[\hat{B}(k_1, k_2)\hat{B}^*(k_1, k_2)] - E^2[\hat{B}(k_1, k_2)]$$

$$= \frac{1}{K}\{E[G^2(k_1)G^2(k_2)G^2(k_1 + k_2)] - E^2[\hat{B}(k_1, k_2)]\}. \tag{B1}$$

The details of the above derivation are given in [10]

$$E[G^2(k_1)G^2(k_2)G^2(k_1 + k_2)]$$

$$= \sum_{p,q,r=1}^{P} \sum_{p',q',r'=1}^{P} \frac{1}{64}\{E[A_pA_qA_rA_{p'}A_{q'}A_{r'} \cdot D(k_1, +f_p, N)D(k_2, +f_q, N)D(k_1 + k_2, +f_{r'}, N) \cdot D(k_1, -f_{p'}, N)D(k_2, -f_{q'}, N)D(k_1 + k_2, -f_r, N) \cdot \exp[-j\pi(N-1)(+f_p + f_q + f_r + f_{p'} + f_{q'} + f_{r'})] \cdot \exp[j(+\phi_p + \phi_q + \phi_r + \phi_{p'} + \phi_{q'} + \phi_{r'})]]$$

$+$ 63 other variations of this term with the signs of f_i, ϕ_i flipped for $i = p, q, r, p', q', r'$}. (B2)

$E[G^2(k_1)G^2(k_2)G^2(k_1+k_2)]$ depends upon the joint satistics of sextuples of amplitudes and phases in general. The following assumptions are made now to simplify the expression for variance.

1) The amplitude and phase are independent for each mode. Then the expectation operation can be separately applied to the amplitudes and the phases yielding $E[A_pA_qA_rA_{p'}A_{q'}A_{r'}]$ and $E[\exp[j(+\phi_p+\phi_q+\phi_r+\phi_{p'}+\phi_{q'}+\phi_{r'})]]$ within the summation in (B2).

2) There is no phase coupling higher than quadratic phase coupling and $E[e^{jn\phi_i}]=0$ for $n=\pm1,\pm2,\cdots,\pm6$; $i=1,2,\cdots,P$.

Depending upon the number of modes that are identical, the set of terms involved in the summation in (B2) can be split into the following disjoint subsets.

a) S_6 = all terms that have all six identical modes.

b) $S_{5,1}$ = all terms that have five identical modes.

c) $S_{4,1,1}$ = all terms that have four identical plus two distinct modes.

d) $S_{4,2}$ = all terms that have four identical modes and another two identical modes.

e) $S_{3,3}$ = all terms that have two groups of three identical modes each.

f) $S_{3,2,1}$ = all terms that have three groups of identical modes, having one, two, and three modes in each group.

g) $S_{2,2,2}$ = all terms that have three groups of two identical modes each.

h) S_* = all the remaining terms, that have four or more distinct modes.

Let T_i denote the sum of all terms in subset S_i. From assumption (2) it immediately follows that $T_*=0$. Enumerating the 64 subterms of a term in (B2) for each case, it can also be shown from assumptions (2) that $T_{5,1}=0$, $T_{4,1,1}=0$, $T_{3,2,1}=0$, and $T_{3,3}=0$. Therefore,

$$E[G^2(k_1)G_2(k_2)G_2(k_1+k_2)] = T_6+T_{4,2}+T_{2,2,2}. \quad \text{(B3)}$$

These terms are now developed by composition from simpler functions. Define

$$\begin{aligned}\gamma(f_p,f_q,f_{r'},f_{p'},f_{q'},f_r) &= D(k_1,+f_p,N)D(k_2,+f_q,N)D(k_1+k_2,+f_{r'},N)\\ &\cdot D(k_1,-f_{p'},N)D(k_2,-f_{q'},N)\\ &\cdot D(k_1+k_2,+f_r,N). \quad \text{(B4)}\end{aligned}$$

Then

$$T_6=\frac{1}{64}\sum_{i=1}^{P}\left\{E[A_i^6]\sum_{(a,b,c,x,y,z)\in C_6}\gamma((-1)^af_i,(-1)^bf_i,(-1)^cf_i,(-1)^xf_i,(-1)^yf_i,(-1)^zf_i)\right\} \quad \text{(B5)}$$

where C_6 is the set of all sextuples of binary numbers having three zeroes and three ones. It represents the 20 ways in which six identical phases can cancel each other. Define

$$\tau_{4,2}(f_{i1},f_{i2},f_{i3},f_{i4},f_{j1},f_{j2}) = \sum_{(f_a,f_b,f_c,f_d,f_x,f_y)\in P_{4,2}}\gamma(f_a,f_b,f_c,f_d,f_x,f_y) \quad \text{(B6)}$$

where $P_{4,2}$ is the set of possible combinations of $(f_{i1},f_{i2},f_{i3},f_{i4},f_{j1},f_{j2})$ in which the orders of $f_{i1},f_{i2},f_{i3},f_{i4}$ and f_{j1},f_{j2} are individually preserved. $P_{4,2}$ represents the ways in which two modes may be made identical from six given modes. Also define

$$\tau_{2,2,2}(f_{i1},f_{i2},f_{j1},f_{j2},f_{k1},f_{k2}) = \sum_{(f_a,f_b,f_c,f_d)\in P_{2,2}}\tau_{4,2}(f_a,f_b,f_c,f_d,f_{k1},f_{k2}) \quad \text{(B7)}$$

where $P_{2,2}$ is the set of possible combinations of $(f_{i1},f_{i2},f_{j1},f_{j2})$ in which the orders of f_{i1},f_{i2} and f_{j1},f_{j2} are individually preserved. $P_{2,2}$ represents the ways in which two modes may be made identical from four given modes. $\tau_{2,2,2}$ is thus a summation over the 90 possible combinations of six frequencies that form three groups of two identical frequencies each. Now $T_{4,2}$ and $T_{2,2,2}$ can be expressed in terms of these functions. Note below that they are sums of two groups of terms each. The first group of terms includes those terms that are phase independent because the phases cancel each other and contribute nonzero values to the variance irrespective of any phase coupling. These are denoted as $T^0_{4,2}$ and $T^0_{2,2,2}$, respectively. The second group of terms includes phases and contributes nonzero values to the variance only in the presence of quadratic phase coupling. These are denoted as $T^\phi_{4,2}$ and $T^\phi_{2,2,2}$, respectively.

$$\begin{aligned}T_{4,2}=\frac{1}{64}\sum_{\substack{i,j=1\\ i\neq j}}^{P}E[A_i^4A_j^2]\Bigg\{&\sum_{x=0}^{1}\sum_{(a,b,c,d)\in C_{2,2}}\{\tau_{4,2}((-1)^af_i,\\ &(-1)^bf_i,(-1)^cf_i,(-1)^df_i,(-1)^xf_j,(-1)^{x+1}f_j)\}\\ &+\sum_{x,y=0}^{1}\{\tau_{4,2}((-)^xf_i,(-1)^xf_i,(-1)^xf_i,\\ &(-1)^xf_i,(-1)^yf_j),(-1)^yf_j)\\ &\cdot\exp[-j2\pi(N-1)(2(-1)^xf_i+(-1)^yf_j)]\\ &\cdot E[\exp[j2(2(-1)^x\phi_i+(-1)^y\phi_j)]]\}\Bigg\}\end{aligned} \quad \text{(B8)}$$

where $C_{2,2}$ is the set of quadruples of binary numbers with two zeros and two ones and corresponds to the eight ways in which four identical phases can cancel each other with positive and negative signs. Finally,

$$\begin{aligned}T_{2,2,2}=\frac{1}{64}\sum_{\substack{i,j,k=1\\ i<j<k}}^{P}E[A_i^2A_j^2A_k^2]\Bigg\{&\sum_{a,b,c=0}^{l}\{\tau_{2,2,2}((-1)^af_i,\\ &(-1)^{a+1}f_i,(-1)^bf_j,(-1)^{b+1}f_j,\\ &(-1)^cf_k,(-1)^{c+1}f_k)\}\\ &+\sum_{a,b,c=0}^{1}\{\tau_{2,2,2}((-1)^af_i,(-1)^af_i,(-1)^bf_j),\\ &(-1)^bf_j,(-1)^cf_k,(-1)^cf_k)\\ &\cdot\exp[-j2\pi(N-1)((-1)^af_i\end{aligned}$$

$$+ (-1)^b f_j + (-1)^c f_k)] \cdot E[\exp [j2((-1)^a \phi_i + (-1)^b \phi_j + (-1)^c \phi_k)]]\bigg\}. \quad \text{(B9)}$$

Appendix C
Variances of the Real and Imaginary Parts of the Bispectrum

Brillinger and Rosenblatt [1], [2] show that the variances of the real and imaginary parts of the bispectrum are asymptotically (i.e., large K) Gaussian, and have equal variances. By decomposing the estimate of the bispectrum into a set of mutually uncorrelated terms (under the assumption of no cubic or higher order phase coupling) each of which makes equal contributions to the variances of the real and imaginary parts, it can be shown that the variances of the real and imaginary parts must be equal even if K is not large.

a) For the perfectly phase-coupled triad with no leakage it is shown in Section IV that the variance is zero. Therefore, the variances of the real and imaginary part must each be zero and thus they must be equal. This result extends to the case of several statistically independent phase-coupled triads because the variances sum to zero. The probability density functions of the real and imaginary parts are unit impulses at their respective values.

b) Equation (8) (ignoring the expected values) demonstrates that the estimate of the bispectrum is composed of a sum of terms of the form $A_p A_p A_r [f_D] [\cos s_1\phi_p + s_2\phi_q + s_3\phi_r) + j \sin (s_1\phi_p + s_2\phi_q + s_3\phi_r)]$, where $s_i = \pm 1$, $i = 1, 2, 3$ and f_D is a deterministic function. When the phase ϕ_i, $i = p, q, r$ is uniform random in $[0, 2\pi)$ and there is no phase coupling, it can be shown that $\cos (s_1\phi_p + s_2\phi_q + s_3\phi_r)$ and $\sin (s_1\phi_1 + p_2\phi_q + s_3\phi_r)$ have zero mean and equal variance. Therefore, there is an identical contribution to the variances from the real and imaginary parts of the estimate of the bispectrum from each such term. The total variance is a sum of these individual contributions because different triads of modes are mutually statistically uncorrelated in the absence of phase coupling. Thus, the variance of the real part is equal to the variance of the imaginary part for random-phase triads. When the number of triads is large, by the central limit theorem the real and imaginary parts of the estimate of the bispectrum will tend to have Gaussian [1], [2] probability distributions.

c) Perfectly phase-coupled triads with leakage will contribute nonzero values to the mean but they have zero variance, so the total variances of the real and imaginary parts still remain equal. (Phase-coupled triads (f_1, f_2, f_3) and (f_1, f_3, f_4) or any other pair with common modes will still be mutually uncorrelated because they will have at least four distinct modes between them and it is assumed that there is no cubic or higher order coupling between modes.)

d) Partial coupling (the case when the power at the component frequencies of a triad is not entirely owing to one phase-coupled mode but owing to another random-phase mode as well) is accommodated as a combination of perfectly phase-coupled and random-phase modes for which arguments a, b, or c hold.

References

[1] D. R. Brillinger and M. Rosenblatt, "Asymptotic theory of estimates of kth order spectra," in *Spectral Analysis of Time Series*, B. Harris, Ed. New York: Wiley, 1967, pp. 153-188.

[2] D. R. Brillinger and M. Rosenblatt, "Computation and interpretation of kth order spectra," in *Special Analysis of Time Series*, B. Harris, Ed. New York: Wiley, 1967, pp. 189-232.

[3] K. Hasselmann, W. H. Munk, and G. J. F. MacDonald, "Bispectra of ocean waves," in *Time Series Analysis*, M. Rosenblatt, Ed. New York, Wiley, 1963, pp. 125-130.

[4] R. A. Haubrich, "Earth noise, 5 to 500 millicycles per second. I.," *J. Geophys. Res.*, vol. 70, pp. 1415-1427, 1968.

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

[6] K. S. Lii, M. Rosenblatt, and C. Van Atta, "Bispectral measurements in turbulence," *J. Fluid Mech.*, vol. 77, pp. 45-62, 1976.

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

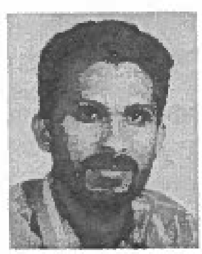
[8] K. S. Lii and K. N. Helland, "Cross-bispectrum computation and variance estimation," *ACM Trans. Math. Software*, vol. 7, pp. 284-294, 1981.

[9] D. F. Elliot, *Handbook of Digital Signal Processing—Engineering Applications*, D. F. Elliot, Ed. San Diego, CA: Academic, 1987, pp. 32-40.

[10] V. Chandran, "Two-dimensional bispectral analysis and leakage effects on the statistics of the bispectrum," Ph.D. dissertation, Washington State Univ., Aug. 1990.

[11] R. Flick and R. Guza, "Paddle generated waves in laboratory channels," *ASCE J. Waterway, Port, Coastal, and Ocean Div.*, vol. 106, pp. 79-97, 1980.

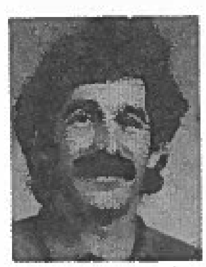
[12] M. Freilich and R. Guza, "Nonlinear effects on shoarling surface gravity waves," *Proc. Roy. Soc. Ser. A*, vol. 311, pp. 1-41, 1984.



Vinod Chandran (S'85-M'90) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, India, in 1982, the master's degree in electrical engineering from Texas Tech University, Lubbock, in 1985 and the Ph.D. degree in electrical and computer engineering from Washington State University, Pullman, in 1990.

Since August 1990, he has been a postdoctoral Teaching Associate in the Department of Electrical Engineering at Washingon State University. His current research interests include speech analysis, pattern recognition, and higher order spectral analysis.

Dr. Chandran is a member of the Optical Society of America, Tau Beta Pi, Phi Kappa Phi, and an associate member of Sigma Xi.



Stephen L. Elgar (A'87) received the bachelor's degree in mathematics and civil enginering from the University of Idaho, Moscow, in 1980, and the master's and Ph.D. degrees in oceanography from the Scripps Institution of Oceanography, University of California, San Diego, in 1981 and 1985, respectively.

In 1986 he joined the Department of Electrical Engineering at Washington State University, Pullman. His current research interests involve the study of nonlinear waves, and, in particular, ocean surface gravity waves and signal processing techniques necessary for analyzing ocean field measurements. His other research interests include nonlinear fluid mechanical systems, and higher order spectra of chaotic processes, including fluid flows and mechanical vibrations.

Dr. Elgar is a member of the American Geophysical Union, the American Physical Society, and the American Meteorological Society.