Journal of Sound and Vibration 329 (2010) 1165–1176

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

Optimal bispectral detection of weak, quadratic nonlinearities in structural systems

J.M. Nichols [a,*], C.C. Olson [b]

[a] Naval Research Laboratory, 4555 Overlook Ave. Washington, DC 20375, USA
[b] National Research Council Postdoctoral Research Associate, Naval Research Laboratory, USA

ARTICLE INFO

Article history:
Received 13 January 2009
Received in revised form
22 October 2009
Accepted 26 October 2009
Handling Editor: M.P. Cartmell
Available online 20 November 2009

ABSTRACT

This work derives the expressions for the bispectrum and bicoherence functions for multi-degree-of-freedom spring–mass systems with quadratic nonlinearities subject to inputs described by a wide class of random processes. The derivation uses a Volterra series model for the system response and yields expressions for both "homogeneous spectra", where the output of only one degree of freedom is used, and "mixed spectra" where the bispectral density contains multiple response time series. This expression is then used to determine the optimal probability distribution of the input and the optimal bispectrum to compute for the goal of maximizing the probability of detecting the nonlinearity.

Published by Elsevier Ltd.

1. Introduction

Estimates of the auto-bispectral density obtained from system response data have been used in a number of applications to detect the presence of quadratic nonlinearities. Since the early work of Brillinger [1] and Hinich [2], the bispectrum has been recognized as a useful tool for detecting non-normality, often the hallmark of a nonlinear process. Applications have ranged from understanding natural physical processes [3] to detection of non-normality in radar signals [4]. Other works have used the estimated bispectrum to analyze structures. Nyffenegger et al. [5] used estimates of the bispectrum for discriminating among cylinders composed of different materials. Both Worden [6] and Hickey et al. [7] used estimates of the auto-bispectrum to detect different types of nonlinearity in a multi-degree-of-freedom (MDOF) system while Messina and Vittal [8] used estimates of the auto-bispectrum to detect nonlinear mode interaction in a power system [8]. In some cases structural damage will result in the presence of a nonlinearity. Rivola et al. [9] used the normalized auto-bispectrum to detect cracks in an experimental beam while Zhang et al. [10] focused on detecting gear faults, also using the auto-bispectrum. Subsequent work by Hillis et al. [11] used the bispectrum to detect cracks in a steel specimen. Both the auto-bispectrum and auto-trispectrum were used by Teng and Brandon in detecting the deterioration of jointed structures [12].

Although the bispectral density and bicoherence functions are frequently used as a signal processing tool in random vibration problems there exists few analytical works, particularly for multiple-degree-of-freedom (MDOF), nonlinear systems. For the single degree-of-freedom (SDOF) case, several authors have derived an expression for the bispectrum of a quadratically nonlinear system driven with Gaussian noise [1,13,14]. Additional works have used this expression in the context of structural dynamics to derive the Type-I and Type-II errors associated with bispectral detection schemes, again for SDOF, Gaussian driven structures [15]. Recent work by the authors has extended the analysis to include the possibility

* Corresponding author. Tel.: +1 202 404 5433; fax: +1 202 767 5792.
E-mail address: jonathan.nichols@nrl.navy.mil (J.M. Nichols).

0022-460X/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.jsv.2009.10.032

of an MDOF system [16] however this derivation was also limited to Gaussian distributed input. The special form of the characteristic function for a jointly Gaussian distribution allows one to invoke a valuable theorem in developing the analytical higher-order spectra. However, it was recently demonstrated that a proper transformation could be used to derive the bispectrum for a single degree-of-freedom system with a nearly arbitrary input distributions [17].

In this work, we derive a general expression for the auto-bispectral densities associated with a broad class of MDOF systems driven by any random process that can be described as a memoryless, third-order polynomial transformation of a Gaussian distributed random process. The formulation that follows is valid for any quadratically nonlinear, MDOF system forced at a single location. The resulting expressions consider both "homogeneous spectra", where the response of a single degree of freedom is being analyzed, and "mixed spectra", where the response at multiple points is considered in the estimate.

Once the expressions are obtained they are compared to estimated bispectral densities and bicoherence functions, performed using the direct method of estimation. It is shown how varying the location of the nonlinearity, and which of the bispectra is estimated can influence the detection problem. It is also shown that if the goal is nonlinearity detection using an uncorrelated (broadband) input sequence, the optimal input probability distribution for interrogating the structure is Gaussian.

2. Bispectrum and bicoherence for MDOF systems subject to non-Gaussian inputs

The auto-bispectral density function for a stationary system response is defined as the double Fourier transform (FT) of the third joint lagged cumulant. Denoting the response data from the ith degree-of-freedom as $y_i(t)$ and the de-meaned signal as $\tilde{y}_i(t)$ the expression for the bispectral density is

$$B_{y_k y_i y_j}(\omega_1,\omega_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}E[\tilde{y}_k(t)\tilde{y}_i(t+\tau_1)\tilde{y}_j(t+\tau_2)]\mathrm{e}^{-\mathrm{i}(\omega_1\tau_1+\omega_2\tau_2)}\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \tag{1}$$

Consider $y_i(t)$ to be the response of a linear structure to a stationary random process $x(t)$. It is easy to show that if $x(t)\sim\mathcal{N}(\mu,\sigma^2)$ i.e. is normally distributed with mean μ and variance σ^2, then $B_{y_k y_i y_j}(\omega_1,\omega_2)=0\ \forall\omega_1,\omega_2$. If $B_{y_k y_i y_j}(\omega_1,\omega_2)\neq 0$ there are two possibilities. One is that the structure obeys a nonlinear model such that a symmetric input distribution, like the Gaussian, leads to an asymmetric distribution on the output. The other possibility is that the structure is linear, but that the input distribution is asymmetric with a non-zero third moment. In the latter case the bispectrum is given by

$$B_{y_k y_i y_j}(\omega_1,\omega_2)=H_{1,i}(\omega_1)H_{1,j}(\omega_2)H_{1,k}(-\omega_1-\omega_2)B_{xxx}(\omega_1,\omega_2) \tag{2}$$

where $B_{xxx}(\omega_1,\omega_2)$ is the bispectrum of the input and $H_{1,i}$ is the linear system transfer function relating the input to the output at the "i"th degree-of-freedom. In order to help discriminate between these possibilities the bispectrum may be normalized by the product of linear transfer functions given in Eq. (2). Noting that the output power spectral density (PSD) for a linear system is $S_{y_i y_i}(\omega)=|H_{1,i}|^2 S_{xx}(\omega)$, a normalized version of the bispectrum is written as [1]

$$b_{y_k y_i y_j}(\omega_1,\omega_2)=\frac{B_{y_k y_i y_j}(\omega_1,\omega_2)}{\sqrt{S_{y_i y_i}(\omega_1)S_{y_j y_j}(\omega_2)S_{y_k y_k}(\omega_1+\omega_2)}} \tag{3}$$

and is referred to as the bicoherence function. For example, if the input consists of independent, identically distributed (iid) values chosen from a non-Gaussian distribution, and the structure is linear, the bicoherence will be the constant

$$b_{y_k y_i y_j}(\omega_1,\omega_2)=E[y_k(t)y_i(t)y_j(t)]\Delta_t^{1/2}/\sqrt{E[y_i(t)y_i(t)]E[y_j(t)y_j(t)]E[y_k(t)y_k(t)]}$$

where Δ_t is the sampling interval (the reason for the inclusion of $\Delta_t^{1/2}$ is given in the Appendix). A nonlinear structural response will result in non-constant values of the bicoherence, regardless of the distribution of the input. Thus, a test for nonlinearity may be realized by simply testing the constancy of the estimated bicoherence function (see e.g. Priestly [18]).

Analytical expressions for both $B_{y_k y_i y_j}(\omega_1,\omega_2)$ and $b_{y_k y_i y_j}(\omega_1,\omega_2)$ are now derived. In [17] this derivation was first presented but for single-degree-of-freedom (SDOF) systems only. The first step is to develop an analytical expression for $y_i(t)$ capable of capturing second-order nonlinearities. To this end we make use of the Volterra functional series approach described in detail by Schetzen [19]. This approach is essentially a Taylor series expansion for functions with memory [20]. Like the Taylor series, the Volterra approximation is only appropriate for systems with smooth (differentiable) nonlinearities and will not converge for arbitrary system parameters. Solutions far from linear (i.e. strongly nonlinear behavior) cannot be captured using this approach. For detection problems the focus is typically on weak nonlinearities for which the Volterra model is appropriate.

Using a two-term Volterra series, the signal model becomes

$$y_i(t)=\int_{-\infty}^{\infty}h_{1,i}(\tau)x(t-\tau)\,\mathrm{d}\tau+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h_{2,i}(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2$$

$$\bar{y}_i(t)=\int_{-\infty}^{\infty}h_{1,i}(\tau)E[x(t-\tau)]\,\mathrm{d}\tau+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h_{2,i}(\tau_1,\tau_2)E[x(t-\tau_1)x(t-\tau_2)]\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \tag{4}$$

where $h_{1,i}(\tau), h_{2,i}(\tau_1,\tau_2)$ are the linear and quadratic Volterra kernels, respectively. While more terms could be considered their contribution to the response is higher order and is not considered here. We could, for example, have added a cubic term to the model (4) which would better capture the quadratic nonlinearity and would also reflect the influence of smooth, third-order nonlinearities. The resulting solution for the bispectrum in this case will contain contributions from $h_{3,i}(\tau_1,\tau_2,\tau_3)$. However, in this work we are focused on the detection of weak, quadratic nonlinearities. Under these assumptions the additional terms turn out to be orders of magnitude smaller than those containing $h_{2,i}(\tau_1,\tau_2)$. Inclusion of the higher-order kernels in this case results in a minimal gain in model accuracy but comes at a huge cost analytically as will become apparent in the following derivation. We should also mention that if the goal is the detection of third-order nonlinearities the trispectrum is probably a more appropriate tool. For the trispectrum, third-order nonlinearities contribute to leading order as was recently demonstrated in [21].

Deriving the bispectrum involves substituting the signal model (4) into (1) and simplifying. In the original derivation of the bispectrum for MDOF systems, the assumption of a Gaussian distributed input was essential as it allowed for a theorem originally derived by Isserlis [22] to be used in the simplification. Isserlis's Theorem states

Theorem 2.1 (*Isserlis' Theorem*). *If* $\eta_1,\eta_2,\ldots,\eta_{2N+1}$ $(N=1,2,\ldots)$ *are normalized, jointly Gaussian random variables* (*i.e., for every i,* $E[\eta_i]=0$ *and* $E[\eta_i^2]=1$), *then*

$$E[\eta_1\eta_2\cdots\eta_{2N}]=\sum\prod E[\eta_i\eta_j]$$

and

$$E[\eta_1\eta_2\cdots\eta_{2N+1}]=0$$

where the notation $\sum\prod$ means summing over all distinct ways of partitioning $\eta_1,\eta_2,\ldots,\eta_{2N}$ into pairs (without regard to order). For example, consider the case of four Gaussian random variables $\eta_1,\eta_2,\eta_3,\eta_4$. Application of the above theorem gives the relation

$$E[\eta_1\eta_2\eta_3\eta_4]=E[\eta_1\eta_2]E[\eta_3\eta_4]+E[\eta_1\eta_3]E[\eta_2\eta_4]+E[\eta_1\eta_4]E[\eta_2\eta_3].$$

In the derivation one repeatedly encounters expectations of products of powers of the input $x(t)$ at various time-delays. For a jointly Gaussian input all odd-ordered products vanish as a consequence of Theorem 2.1 while all even ordered products can be broken into sums of auto- and cross-correlation functions. No analogous theorem exists for non-Gaussian joint distributions of $x(t)$.

In order to circumvent this problem, we consider the input to be a static (memoryless), nonlinear transformation of a jointly Gaussian distributed random process $w(t)\sim\mathcal{N}(0,\sigma_w^2)$ by choosing

$$x(t)=a_0+a_1w(t)+a_2w(t)^2+a_3w(t)^3. \tag{5}$$

This particular transformation appears to have been first used by Fleishman [23] in creating non-Gaussian distributions. Although the transformation is simply a third-order polynomial it can be used to generate a very large variety of distributions for $x(t)$, depending on how the parameters $a_{0\to3}$ are set. Note that this model does not assume that the input is spectrally white (iid). The resulting random process can be white or spectrally colored, depending on the process $w(t)$. This model is limited, however, to random processes with linear temporal correlations. For example, (5) could not be used to describe an input that is itself the output of a nonlinear system. However, one could conceivably replace (5) with a second Volterra series model (e.g., Eq. (4)) and carry the derivation forward. However, for this work we restrict ourselves to a spectrally colored input described by a broad class of probability density functions as specified by (5).

A derivation for the bispectrum subject to the class of non-Gaussian inputs generated by Eq. (5) can now be carried out. Substituting Eq. (5) into Eq. (4) and then into Eq. (1) yields a very large number of integrals, each involving the products of an even number of Gaussian distributed random variables $w(t)$ (all odd-ordered products are automatically zero by Theorem 2.1). Some of the products turn out to be quite large, in fact some terms involve the product of 16 $w(t)$'s. According to Isserlis's Theorem, the number of autocorrelation terms in the sum is $(2n)!/(n!\times2^n)$ where $2n$ is the number of Gaussian variables involved. For $2n=16$ this gives 2,027,025 terms, thus implementation of Theorem 2.1 was carried out using Mathematica software. Additionally, a number of the resulting terms turn out to be higher-order and can be discarded. Specifically, in the simplification one encounters a number of integrals where the following inequality has been assumed to hold:

$$\int H_2(\omega_1-\omega,\omega)S_{ww}(\omega)\,\mathrm{d}\omega \ll \int S_{ww}(\omega)\,\mathrm{d}\omega \tag{6}$$

for all choices of ω_1. For structural systems, one can get a sense of when (6) is expected to hold by considering a specific $H_2(\omega_1,\omega_2)$ kernel. For a single degree-of-freedom, quadratically nonlinear structure (see Eq. (14)), the second-order Volterra kernel can be shown to have the form

$$H_2(\omega_1-\omega,\omega)=-k_NH_1(\omega_1-\omega)H_1(\omega)H_1(\omega_1)$$

where $H_1(\omega)=1/(-m\omega^2+\mathrm{i}c\omega+k_L)$ is the familiar linear transfer function. Many structural systems are modeled as lightly damped and the denominator of $H_1(\omega)$ is dominated by the linear stiffness, k_L. The second-order kernel contains a product of three linear transfer functions, thus the ratio k_N/k_L^3 is expected to be a good predictor of when (6) can be expected to

hold. We have carried out the integral needed in (6) and have found that if $k_N/k_L^3 \ll 1$, the inequality (6) is indeed satisfied. We should point out that this is precisely what has already been assumed in modeling the system response i.e., a weakly nonlinear system.

The end result is the bispectral density function for any weakly quadratic nonlinear system subject to an input that can be described by the transformation (5)

$$\begin{aligned}
&B_{y_k y_i y_j}(\omega_1,\omega_2) = C(\omega_1,\omega_2)H_{1,i}(\omega_1)H_{1,j}(\omega_2)H_{1,k}(-\omega_1-\omega_2)\\
&+[C_2S_0^{(a)}+C_3S_1^{(a)}+C_4S_2^{(a)}+8a_2^4S_3^{(a)}]H_{1,j}(\omega_2)H_{1,k}(-\omega_1-\omega_2)H_{2,i}(-\omega_2,\omega_1+\omega_2)\\
&+[C_2S_0^{(b)}+C_3S_1^{(b)}+C_4S_2^{(b)}+8a_2^4S_3^{(b)}]H_{1,i}(\omega_1)H_{1,k}(-\omega_1-\omega_2)H_{2,j}(-\omega_1,\omega_1+\omega_2)\\
&+[C_2S_0^{(c)}+C_3S_1^{(c)}+C_4S_2^{(c)}+8a_2^4S_3^{(c)}]H_{1,i}(\omega_1)H_{1,j}(\omega_2)H_{2,k}(-\omega_1,-\omega_2)\\
&\quad+2\overline{x}C(\omega_1,\omega_2)[H_{1,j}(\omega_2)H_{1,k}(-\omega_1-\omega_2)H_{2,i}(\omega_1,0)\\
&+H_{1,i}(\omega_1)H_{1,k}(-\omega_1-\omega_2)H_{2,j}(\omega_2,0)+H_{1,i}(\omega_1)H_{1,j}(\omega_2)H_{2,k}(-\omega_1-\omega_2,0)]\\
&+2\overline{x}[C_2S_0^{(a)}+C_3S_1^{(a)}+C_4S_2^{(a)}+8a_2^4S_3^{(a)}]H_{1,k}(-\omega_1-\omega_2)H_{2,i}(-\omega_2,\omega_1+\omega_2)H_{2,j}(\omega_2,0)\\
&+2\overline{x}[C_2S_0^{(b)}+C_3S_1^{(b)}+C_4S_2^{(b)}+8a_2^4S_3^{(b)}]H_{1,k}(-\omega_1-\omega_2)H_{2,i}(\omega_1,0)H_{2,j}(-\omega_1,\omega_1+\omega_2)\\
&\quad+2\overline{x}[C_2S_0^{(c)}+C_3S_1^{(c)}+C_4S_2^{(c)}+8a_2^4S_3^{(c)}]H_{1,i}(\omega_1)H_{j,2}(\omega_2,0)H_{k,2}(-\omega_1,-\omega_2)\\
&+2\overline{x}[C_2S_0^{(a)}+C_3S_1^{(a)}+C_4S_2^{(a)}+8a_2^4S_3^{(a)}]H_{1,j}(\omega_2)H_{2,i}(-\omega_2,\omega_1+\omega_2)H_{2,k}(-\omega_1-\omega_2,0)\\
&+2\overline{x}[C_2S_0^{(b)}+C_3S_1^{(b)}+C_4S_2^{(b)}+8a_2^4S_3^{(b)}]H_{1,i}(\omega_1)H_{2,j}(-\omega_1,\omega_1+\omega_2)H_{2,k}(-\omega_1-\omega_2,0)\\
&\quad+2\overline{x}[C_2S_0^{(c)}+C_3S_1^{(c)}+C_4S_2^{(c)}+8a_2^4S_3^{(c)}]H_{1,j}(\omega_2)H_{i,2}(\omega_1,0)H_{k,2}(-\omega_1,-\omega_2)\\
&\quad+4\overline{x}^2C(\omega_1,\omega_2)[H_{1,k}(-\omega_1-\omega_2)H_{2,i}(\omega_1,0)H_{2,j}(\omega_2,0)\\
&+H_{1,j}(\omega_2)H_{2,i}(\omega_1,0)H_{2,k}(-\omega_1-\omega_2,0)+H_{1,i}(\omega_1)H_{2,j}(\omega_2,0)H_{2,k}(-\omega_1-\omega_2,0)]
\end{aligned} \tag{7}$$

where the functions $C(\omega_1,\omega_2), C_2, C_3, C_4, \overline{x}$, and a_2 describe the probability distribution of the input and the functions $S_{0,1,2,3}^{(a,b,c)}$ describe the input spectral properties. These functions are given in their entirety in the Appendix.

In order to derive the bicoherence an expression is also required for the power spectrum in terms of the coefficients a_0, a_1, a_2, a_3 and the spectral properties of the input $S_{ww}(\omega)$. This can be accomplished using the same general procedure as was used for the bispectrum. The expression for $S_{y_i y_i}(\omega_1)$ is also provided in the Appendix. An analytical expression for the bicoherence function can therefore be obtained by simply dividing the magnitude of Eq. (7) by the product of PSDs required by Eq. (3).

The final issue that needs to be addressed is how to obtain the linear and nonlinear Volterra kernels. These kernels will be dependent on the specific system under study. Expressions for $H_1(\omega)=\int_{-\infty}^{\infty} h_1(\tau_1)e^{-i\omega\tau_1}\,d\tau_1$ and $H_2(\omega_1,\omega_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2)e^{-i(\omega_1\tau_1+\omega_2\tau_2)}\,d\tau_1\,d\tau_2$, may be obtained via the harmonic probing technique described in detail in [24,25]. The cited works provide a clear presentation thus the material is not repeated here.

3. Estimation

Estimation of the bispectrum and bicoherence functions has been discussed at length in a number of references (see for example Nikias and Raghuveer [26] and Huber et al. [27]) and is briefly summarized here. Perhaps the most popular, and probably computationally most efficient approach, is accomplished in the frequency domain. It can be shown that Eq. (1) can also be written as

$$B_{y_k y_i y_j}(\omega_1,\omega_2) = E[dY_i(\omega_1)\,dY_j(\omega_2)\,dY_k(-\omega_1-\omega_2)] \tag{8}$$

where $dY_i(\omega)$ is the Fourier–Stieltjes representation (Cramér spectral representation) of the signal $y_i(t)$, i.e. $\tilde{y}(t)=(1/2\pi)\int_{-\infty}^{\infty} e^{i\omega t}\,dY(\omega)$ [18]. An estimator can be formed by averaging the discrete Fourier transform of the data over some number of segments, just as is commonly done for power spectrum estimation. Dividing the observed data $y(n),\ n=1\cdots N$ into S overlapping segments of length M gives $y_{s,i}(m)=y_i(m+sM-L),\ m=0\cdots M-1,\ s=0\cdots S-1$. The variable L denotes the degree of overlap ($0\le L<M$). The data segment is de-meaned (in accordance with the bispectrum definition), (possibly) windowed, and then Fourier transformed to give $Y_{s,i}(f_p)=\sum_m w(m)y_{s,i}(m)e^{-i2\pi pm/M},\ p=0,\ldots,M-1$ where $w(m)$ is the windowing function. The final estimator for Eq. (1) is then the average

$$\hat{B}_{y_k y_i y_j}(f_p,f_q) = \frac{\Delta_t^2}{SM}\sum_{s=0}^{S-1} Y_{s,i}(f_p)Y_{s,j}(f_q)Y_{s,k}^*(f_p+f_q) \tag{9}$$

for discrete frequencies f_p, f_q. Estimating the bicoherence simply involves dividing by the needed product of estimated power spectral densities

$$\hat{S}_{y_i y_i}(f_p) = \frac{\Delta_t}{SM}\sum_{s=0}^{S-1} Y_{s,i}(f_p)Y_{s,i}^*(f_p) \tag{10}$$

to give

$$\hat{b}_{y_k y_i y_j}(f_p, f_q) = \frac{|\hat{B}_{y_k y_i y_j}(\omega_1, \omega_2)|}{\sqrt{\hat{S}_{y_i y_i}(\omega_1)\hat{S}_{y_j y_j}(\omega_2)\hat{S}_{y_k y_k}(\omega_1+\omega_2)}} \tag{11}$$

Both bispectrum and bicoherence are complex normally distributed [1,27] with variance

$$\sigma_e^2 = \kappa \frac{M^2 \Delta_t}{N} S_{y_i y_i}(\omega_1) S_{y_j y_j}(\omega_2) S_{y_k y_k}(\omega_1+\omega_2) \tag{12}$$

and

$$\sigma_e^2 = \kappa \frac{M^2 \Delta_t}{N} \tag{13}$$

respectively. The pre-factor κ accounts for the possibility of overlapping segments ($L > 0$) and windowing and is discussed at length in Huber et al. [27].

4. Example system

For this work we consider the three degree-of-freedom, spring–mass–damper system governed by the equations

$$[\mathbf{M}]\ddot{\mathbf{y}}(t) + [\mathbf{C}]\dot{\mathbf{y}}(t) + [\mathbf{K}_L]\mathbf{y}(t) + \mathbf{g}(\dot{\mathbf{y}}, \mathbf{y}) = \mathbf{x}(t) \tag{14}$$

where

$$[\mathbf{M}] = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}, \quad [\mathbf{C}] = \begin{bmatrix} c_1+c_2 & -c_2 & 0 \\ -c_2 & c_2+c_3 & -c_3 \\ 0 & -c_3 & c_3 \end{bmatrix}, \quad [\mathbf{K}] = \begin{bmatrix} k_1+k_2 & -k_2 & 0 \\ -k_2 & k_2+k_3 & -k_3 \\ 0 & -k_3 & k_3 \end{bmatrix}$$

are constant coefficient mass, damping and stiffness matrices, respectively. The nonlinear function $\mathbf{g}(\cdot)$ provides quadratic coupling between masses. Here we consider both a quadratically nonlinear damping and restoring force between masses. For example, if the nonlinearity is located between masses 1 and 2 we have

$$\mathbf{g}(\dot{\mathbf{y}}, \mathbf{y}) = \begin{Bmatrix} -k_N(y_2(t)-y_1(t))^2 \\ k_N(y_2(t)-y_1(t))^2 \\ 0 \end{Bmatrix} + \begin{Bmatrix} -c_N(\dot{y}_2(t)-\dot{y}_1(t))^2 \\ c_N(\dot{y}_2(t)-\dot{y}_1(t))^2 \\ 0 \end{Bmatrix}$$

where k_N, c_N are the nonlinear stiffness and damping coefficients, respectively.

Consider the system defined by Eq. (14) with $m_1 = m_2 = m_3 = 1.0\,\text{kg}$, $k_1 = k_2 = k_3 = 2000\,\text{N/m}$, $c_1 = c_2 = c_3 = 3.0\,\text{N}\cdot\text{s/m}$. The linear parameters will be fixed to these values for the remainder of this work. First, consider a restoring nonlinearity $k_N = 10^5\,\text{N/m}^2$ located between masses 1 and 2. Fig. 1 shows both the theoretical and estimated bispectrum ($\hat{B}_{\ddot{y}_2\ddot{y}_2\ddot{y}_1}(f_1,f_2)$, and $\hat{B}_{\ddot{y}_1\ddot{y}_2\ddot{y}_2}(f_1,f_2)$) associated with the acceleration response of this nonlinear system subject to Gaussian excitation ($a_0 = a_2 = a_3 = 0$, $a_1 = 1, \sigma_w = 1$) applied at the third mass. In this example we have used time-series from different points on the structure in order to illustrate the very different bispectra that can result. For the estimation, time-series of length $N = 2^{18}$ were generated by numerically integrating Eq. (14). The sampling interval was chosen to be $\Delta_t = 0.01\,\text{s}$ while the estimation parameters were chosen to be $M = 1024$, $L = 512$, and $w(m)$ was chosen based on a Hanning window. As a second example, consider the input to be governed by a highly non-Gaussian distribution (but with the same nonlinearity). Fig. 2 shows the distribution along with the predicted and estimated bispectrum. The same estimation parameters were used as for the previous example. The solution given by Eq. (7) serves as a good approximation for the bispectrum for a very large number of input distributions and works for any combination of time-series collected from the structure.

The solution can also be used to develop an expression for the analytical bicoherence function. This section concludes with a comparison between theory and estimate for the bicoherence function associated with a nonlinear system driven with an input conforming to the χ^2 distribution ($a_0 = a_1 = a_3 = 0$, $a_2 = 1$). Results of this comparison are shown in Fig. 3. For this example both quadratic damping and stiffness terms were placed between masses 1 and 2 with values $c_N = 75\,\text{N}\cdot\text{s}^2/\text{m}^2$ and $k_N = 10^5\,\text{N/m}^2$. Had the system been linear, theory predicts that the bicoherence would have been perfectly flat across the entire f_1, f_2 plane with the value $|b_{\ddot{y}_k\ddot{y}_i\ddot{y}_j}(f_1,f_2)| = 0.283$ due to the highly skewed nature of the input distribution. The nonlinearity, however, gives rise to various peaks, located along the combination resonances for this system, rising above this constant value. As was discussed previously, the estimator associated with the bicoherence has a much higher variance. For this reason the estimate was made with $M = 512$ point segment, sacrificing bias for variance. The estimate is still clearly more noisy than for the bispectrum yet is also clearly capturing the features predicted by the theory. We now turn our attention to a potential use of Eq. (7).

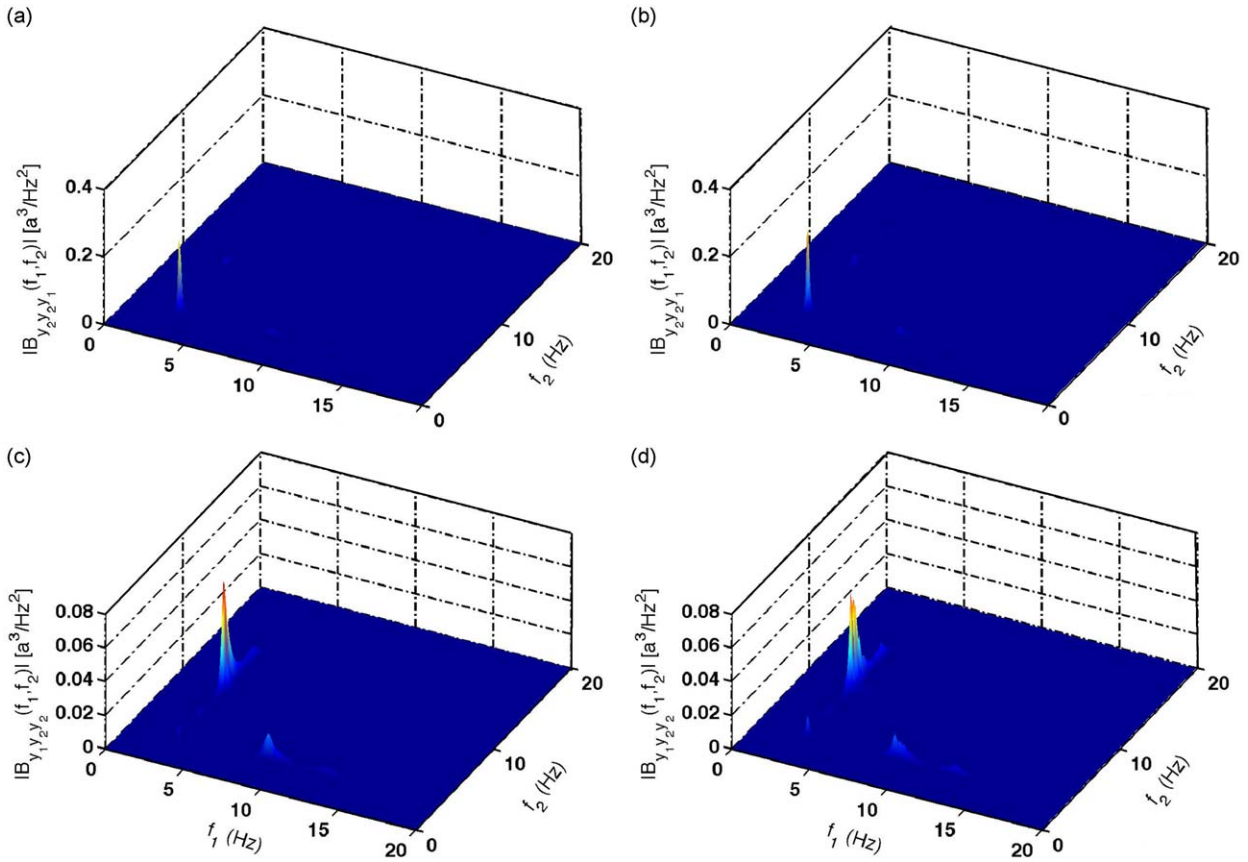

Fig. 1. (a) Theoretical and (b) estimated bispectrum magnitude based on multivariate acceleration response, $|B_{\ddot{y}_2\ddot{y}_2\ddot{y}_1}(f_1,f_2)|$. (c) Theoretical and (d) estimated bispectrum magnitude $|\hat{B}_{\ddot{y}_1\ddot{y}_2\ddot{y}_2}(f_1,f_2)|$. The forcing was a Gaussian distributed random process, obtained by setting $a_0=a_2=a_3=0$ and $a_1=1$. The power spectral density of the forcing was taken to be 0.01 N^2/Hz.

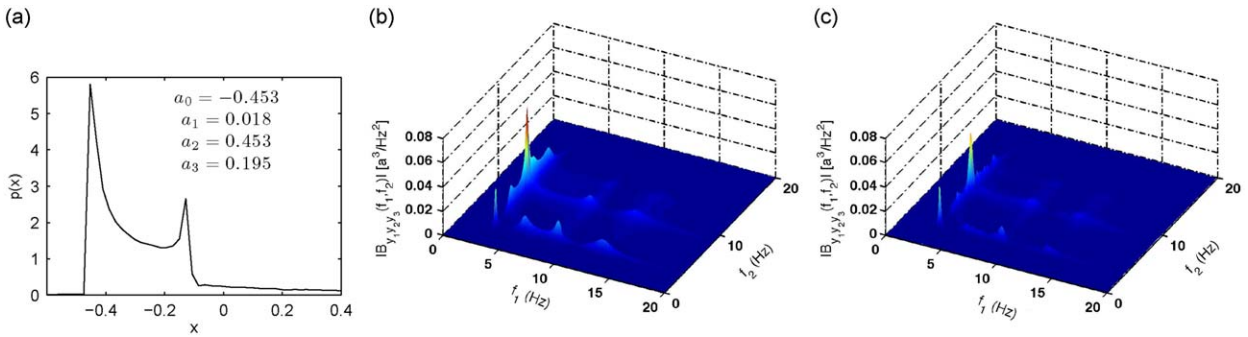

Fig. 2. (a) Non-Gaussian input distribution, obtained by setting $a_0=-0.453, a_1=0.018, a_2=0.453, a_3=0.195$, $\sigma_\omega=1$ in Eq. (5). (b) The theoretical and (c) the estimated output accelerance bispectrum $|B_{\ddot{y}_1\ddot{y}_2\ddot{y}_3}(f_1,f_2)|$ associated with this input applied at mass 3. A quadratic nonlinearity $k_N=10^5$ N/m^2 was placed between the first and second masses.

5. Optimal bispectral detection

This section covers two of the basic issues associated with designing a bispectral nonlinearity detection scheme. First, we define the detection statistic to be used and show how this choice leads to a well-defined criteria for optimality in the detector. Next, we explore optimizing the detection scheme as a function of which bicoherence function to estimate, where to excite the structure, and which probability distribution to use in constructing the input signal.

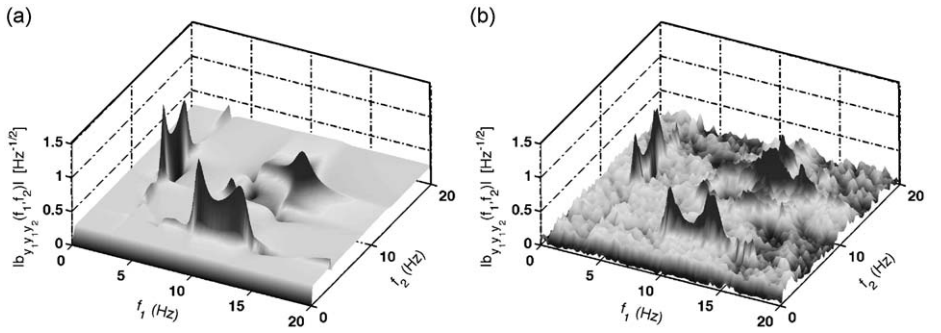

Fig. 3. (a) The theoretical and (b) estimated output accelerance bicoherence magnitude $|b_{\ddot{y}_1\ddot{y}_2}(f_1,f_2)|$ associated with the input applied at mass 3. The input distribution was χ^2 ($a_0=a_1=a_3=0, a_2=1, \sigma_w=1$). Quadratic nonlinearities $k_N=10^4\,\mathrm{N/m^2}$ and $c_N=75\,\mathrm{N\,s^2/m^2}$ were placed between the first and second masses.

5.1. Detection statistic

The first question to be answered is what to use as the detection statistic? Previous work by the authors has considered focusing on the detection probabilities of the estimated bispectrum peaks [15]. However, for MDOF systems there will be a large number of peaks making this detector a good deal more complicated. Additionally, the possibility of non-Gaussian inputs suggests a bicoherence-based test statistic rather than one based on the bispectrum. Following the work of Garth and Bresler [28,29], Richardson and Hodgkiss [3], Hinich and Wilson [4] we make use of the detection statistic

$$\mathcal{S}=\frac{2}{\sigma_e^2}\sum_{pq}|\hat{b}^r_{kij}(f_p,f_q)|^2 \tag{15}$$

where the sum is taken over all $D=M^2/16$ positive frequency bins in the estimated bicoherence function that satisfy $p,q\le M/4$. This constraint ensures that the sum frequency is below the Nyquist frequency i.e. $p+q\le M/2$. It was demonstrated in [28] that this test statistic was a sufficient statistic for the generalized likelihood ratio test (GLRT). The shorthand notation $\hat{b}^r_{kij}(f_p,f_q)$ has been adopted in place of $\hat{b}_{\ddot{y}_k\ddot{y}_i\ddot{y}_j}(f_p,f_q)$ as all results that follow are obtained using the acceleration output of Eq. (14). The superscript "r" will be used to denote the forcing location i.e. which degree of freedom is subject to $x(t)$.

As was mentioned earlier, the bicoherence estimates in each frequency bin are complex Gaussian distributed with common variance σ_e^2 thus the statistic S will follow a non-central chi-squared distribution with $2D$ degrees of freedom (each frequency bin contributes a real and imaginary part, hence the factor of 2). For large D, however, the non-central chi-squared distribution is well-approximated by a Gaussian distribution with mean and variance

$$\mu=2D+\Lambda$$

and

$$\sigma^2=2(2D+2\Lambda)$$

where $\Lambda=(2/\sigma_e^2)\sum_{p,q}|b^r_{kij}(f_p,f_q)|^2$ is the non-centrality parameter.

The nonlinearity detection problem is to discriminate between the hypotheses

$$\mathcal{H}_0:\mathcal{S}\sim\mathcal{N}(\mu_L,\sigma_L^2)$$

$$\mathcal{H}_1:\mathcal{S}\sim\mathcal{N}(\mu_N,\sigma_N^2) \tag{16}$$

where μ_L,σ_L and μ_N,σ_N are the linear and nonlinear mean and variance, respectively. The value for Λ is known analytically via Eqs. (7) and (A.4) for any combination of signals i,j,k and input locations r. The value for Λ is also given as a function of the input distribution via the parameter vector **a**. We are therefore in a position to find both the optimal bicoherence to compute and the optimal input distribution for nonlinearity detection.

For this detection problem, we chose the Neyman–Pearson criterion for optimality, that is to minimize the Type-II error (maximize probability of detection) for a given Type-I error (probability of false alarm). For the problem stated in (16), assuming similar values for the variance, this may be accomplished by maximizing the deflection coefficient [30]

$$d^2=\frac{|\mu_N-\mu_L|^2}{\sigma_N^2} \tag{17}$$

(even if the assumption of similar variance values is not made the results of the following optimizations were not changed). In what follows the goal is to maximize d w.r.t. various parameters of interest in this detection problem.

5.2. Which bicoherence to compute?

The goal of this section is to determine the optimal (in the sense described in the previous section) bicoherence to estimate. For example, if the nonlinearity is located between masses 1 and 2 should we base Eq. (15) on $\hat{b}^1_{123}$ or $\hat{b}^3_{211}$? Specifically, this optimization finds

$$\max_{i,j,k,r} d$$

For this optimization we considered the input $x(t)$ to be Gaussian, iid with a constant power spectral density of $S_{ww} = 0.01\,\mathrm{N}^2/\mathrm{Hz}$. For this particular choice of forcing $b^r_{kij}(f_p,f_q) = 0$ under the linear hypothesis, thus maximizing d strictly involves finding the largest value of μ_N. Fortunately for this problem we can exhaustively search the possibilities and find the test statistic that maximizes Eq. (17). There are three possible locations for both excitation and response giving a total of $3^4 = 81$ possible test statistics to estimate for a given nonlinearity location. First we consider the case where a quadratic stiffness nonlinearity, $k_N = 10^5\,\mathrm{N/m}^2$, is placed between masses 1 and 2. Fig. 4(a) plots the key component of μ_N (large μ_N equals large d) for all 81 possible bicoherence functions. A few of these are labeled in the figure and clearly show that some are more advantageous to use than others for detection purposes. It can be seen that the best bicoherence functions favor forcing at mass 3 and using the response at mass 1 as the last argument in the estimate

In order to test these predictions the following numerical experiment was performed. Ten separate realizations of Eq. (14), each consisting of $N = 32,768$ observations, were obtained from the system (14) using a Runge–Kutta numerical integration scheme. For each realization the test statistic $\mathcal{S}$ was estimated using the procedure outlined in Section 3. The parameters used were $M = 256$, $L = 128$ and a Hanning window was applied for smoothing. Using these parameters the adjustment factor in Eq. (13) becomes $\kappa = 0.273$. Fig. 4(b) shows the estimated receiver operator characteristic (ROC) curves associated with a few different choices of bicoherence function used in forming the test statistics. The ROC curve simply displays probability of detection vs. probability of false alarm associated with varying the detection threshold in a detection problem [31]. Large values of b^r_{kij} result in larger values μ_N which give rise to better probabilities of detection for a given probability of false alarm. These results clearly indicate that choice of signals used in computing the bispectrum and the choice of excitation point are of paramount importance. The ROC performance is exactly in line with that predicted. Using the test statistic based on b^1_{221}, for example, gives very poor detection performance.

Next, we consider changing the location of the nonlinearity to lie between masses 2 and 3 and repeat the above analysis with results shown in Fig. 5. Again, test statistics based on the predicted optimal choice of b^r_{kij} clearly yield the best detection performance. Regardless of where the damage is located one can obtain drastically different detection performance depending on which bicoherence function is used in computing $\mathcal{S}$.

Additionally, these results give a straightforward way to locate the nonlinearity. An input can be applied at both masses 1 and 3 and time-series recorded from each. We can then estimate b^1_{113} and b^3_{311}. If $b^1_{113} \gg b^3_{311}$, the nonlinearity is between masses 2 and 3 while if $b^3_{311} \gg b^1_{113}$, the nonlinearity is between masses 1 and 2.

5.3. What probability distribution to use for the input?

We also consider in the optimization the problem of determining the best input probability density function. Specifically, for a fixed input power, we seek the parameter vector $\mathbf{a} \equiv (a_0, a_1, a_2, a_3)$ that minimizes our Type-I and Type-II

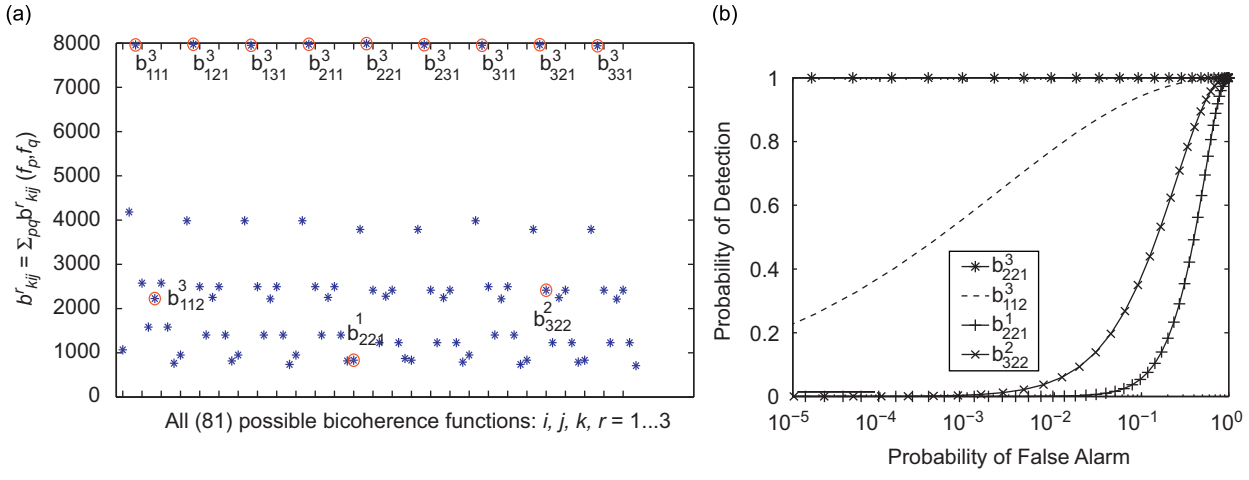

Fig. 4. (a) All 81 possible values of b^r_{kij} under $\mathcal{H}_1$ when the nonlinearity is located between masses 1 and 2 and (b) the ROC curves corresponding to different choices of test statistic $\mathcal{S}$ based on the associated bicoherence estimates.

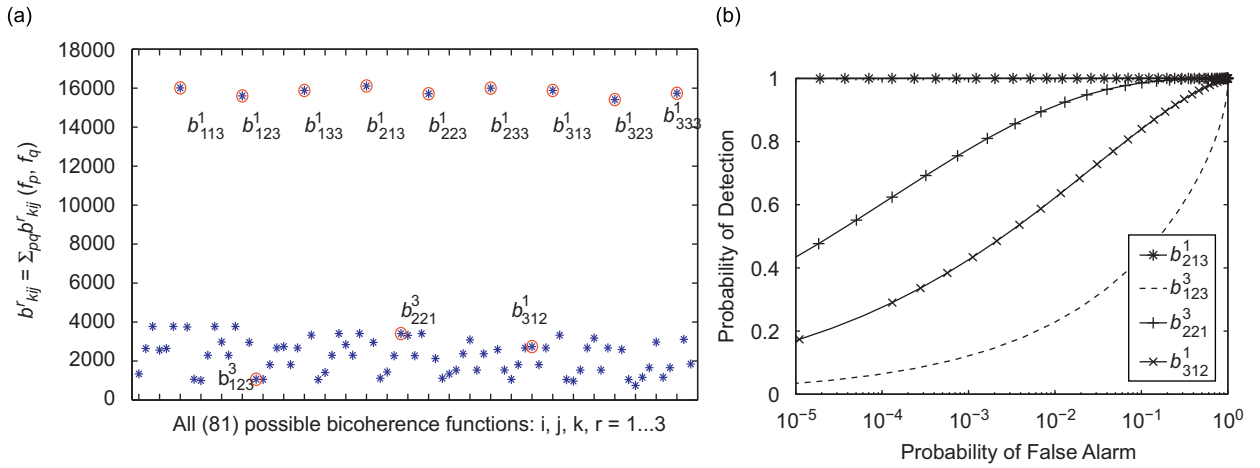

Fig. 5. (a) All 81 possible values of b^r_{kij} under $\mathcal{H}_1$ with the nonlinearity located between masses 2 and 3 and (b) the ROC curves corresponding to different choices of test statistic $\mathcal{S}$ based on the associated bicoherence estimates.

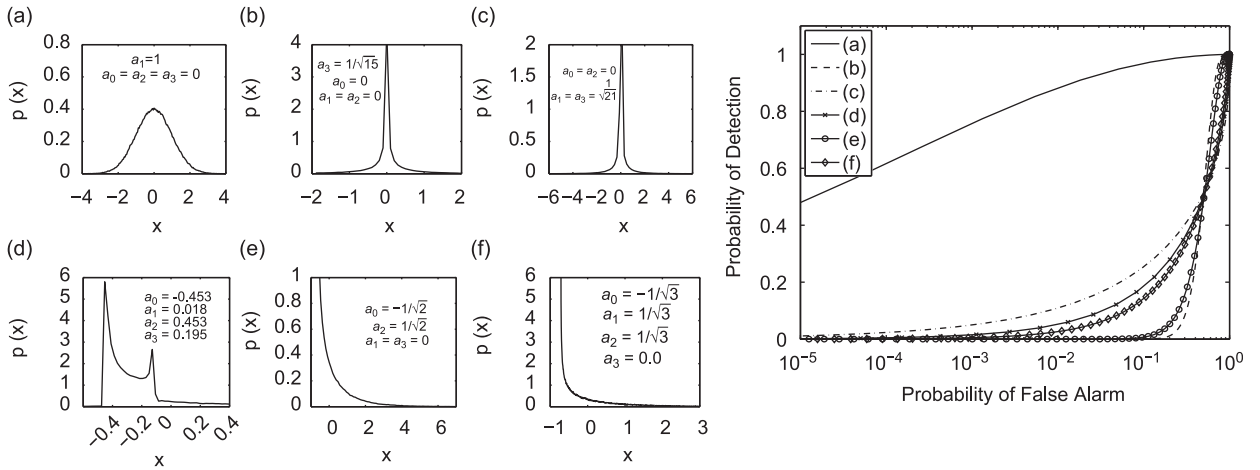

Fig. 6. Several different probability distributions for the input $x(t)$ along with the ROC curves associated with detecting a quadratic stiffness nonlinearity of $k_N = 30,000\,\mathrm{N/m^2}$. The Gaussian distribution was determined theoretically to be the optimal i.i.d. input probability distribution for detecting the nonlinearity. This is clearly seen to be the case in the estimated ROC curves.

error in making a detection. This optimization problem can also be written in terms of the deflection coefficient as

$$\max_{a_1,a_2,a_3} d$$

An exhaustive search is not possible in this case; for this reason differential evolution was used to search the parameter space. It should also be pointed out that only zero-mean distributions were considered in this work, i.e. $a_0 = -a_2\sigma_w^2$. This was done purely for practical reasons as the equipment used in vibration testing is often displacement limited. A large non-zero mean may help detection but may not be feasible experimentally. The optimization was also limited to iid inputs, that is to say there is no spectral coloring on the input signal. The goal here was to focus strictly on the form of the input probability density and not on where the signal energy should be concentrated.

For this optimization, the nonlinearity was located between masses 1 and 2 such that the bicoherence feature b^3_{221} gives the best detection performance and was taken to be the detection statistic. The results of the optimization are not entirely surprising. It turns out that for the class of input distributions considered, a Gaussian distribution is optimal. In general, symmetric distributions performed well. This is due to the fact that for a symmetric input distribution the bicoherence is near zero if there is no nonlinearity. Because the variance of the test statistic increases with Λ, any input distribution that minimizes the value of the non-centrality parameter while maximizing the *difference* between linear and nonlinear non-centrality parameters is preferred. Highly asymmetric distributions performed poorly because even in the linear case they result in high values of Λ. Fig. 6 shows several different input distributions along with the generated ROC curves. The ROC curves were based on estimates of b^3_{221}, obtained from 20 independent realizations of Eq. (14). As predicted by theory, the

best performing ROC curve is clearly associated with a Gaussian input. Thus, we conclude that for detecting a quadratic nonlinearity in a structural system using iid. inputs a Gaussian distributed input is optimal.

6. Conclusions

A closed-form expression for both the bispectral density and bicoherence functions associated with a quadratically nonlinear, multi-degree-of-freedom system response was derived. The main assumption underlying the derivation is that the system be weakly nonlinear. This assumption is necessary both for the Volterra series model to hold, and for the simplification of the full bispectrum expression. Under this assumption, estimates of both the bispectrum and bicoherence functions were shown to be in close agreement with theory. These expressions were then used to derive both the optimal bicoherence to compute (including where to drive the structure) and the optimal probability distribution of the input for detecting the presence of a nonlinearity. The results indicate that sub-optimal choices of these parameters can significantly degrade detection performance. We conclude that for this example system, the forcing should be applied away from the nonlinearity and the signal in the last argument of the estimated bicoherence be taken close to the nonlinearity. As was mentioned, the very specific force/response combinations that lead to large values of the test statistic provide a convenient way to locate the nonlinearity. Additionally we recommend that for a bicoherence-based detector, a Gaussian distributed input results in the best possible detection performance over a wide range of zero-mean, broad-band (i.i.d.) inputs. It may be the case that deterministic inputs (e.g. sinusoid) or highly band-limited random inputs can produce still better detection performance. Future work, using the derived expression, will determine if additional detection gains can be realized by concentrating the input signal energy in certain frequency ranges (e.g. near resonance).

Acknowledgments

The authors would like to acknowledge funding from the Office of Naval Research, under contract No. N00014-08-WX-2-1107. We would also like to acknowledge Frank Bucholtz of the Naval Research Laboratory for providing a critical read of the manuscript.

Appendix A. Terms in bispectrum characterizing input probability distribution and power spectrum

The main analytical result of the paper, Eq. (7), is given in terms of coefficients describing both the probability distribution and spectral properties of the input. The input distribution properties are captured by the terms

$$C(\omega_1,\omega_2)=2a_1^2a_2S_0+6a_1a_2a_3(S_1+2\sigma_w^2S_0)+8a_2^3\tilde{S}+18a_2a_3^2(2S_2+\sigma_w^2S_1+\sigma_w^4S_0)$$

$$C_2=2a_1^4+24\sigma_w^2a_1^3a_3+108\sigma_w^4a_1^2a_3^2+216\sigma_w^6a_1a_3^3+162\sigma_w^8a_3^4$$

$$C_3=4a_1^2a_2^2+24\sigma_w^2a_1a_2^2a_3+36\sigma_w^4a_2^2a_3^2$$

$$C_4=12a_1^2a_3^2+72\sigma_w^2a_1a_3^3+108\sigma_w^4a_3^4$$

$$\overline{x}=E[x(t)]=a_0+\sigma_w^2a_2 \tag{A.1}$$

while the input spectral properties are governed by

$$S_0^{(a)}=S_{ww}(\omega_1+\omega_2)S_{ww}(\omega_2)$$

$$S_0^{(b)}=S_{ww}(\omega_1+\omega_2)S_{ww}(\omega_1)$$

$$S_0^{(c)}=S_{ww}(\omega_1)S_{ww}(\omega_2)$$

$$2\pi S_1^{(a)}=S_{ww}(\omega_2)\int S_{ww}(\omega_1+\omega_2+\omega_3)S_{ww}(\omega_3)\,\mathrm{d}\omega_3+S_{ww}(\omega_1+\omega_2)\int S_{ww}(\omega_2+\omega_3)S_{ww}(\omega_3)\,\mathrm{d}\omega_3$$

$$2\pi S_1^{(b)}=S_{ww}(\omega_1)\int S_{ww}(\omega_1+\omega_2+\omega_3)S_{ww}(\omega_3)\,\mathrm{d}\omega_3+S_{ww}(\omega_1+\omega_2)\int S_{ww}(\omega_1+\omega_3)S_{ww}(\omega_3)\,\mathrm{d}\omega_3$$

$$2\pi S_1^{(c)}=S_{ww}(\omega_1)\int S_{ww}(\omega_2+\omega_3)S_{ww}(\omega_3)\,\mathrm{d}\omega_3+S_{ww}(\omega_2)\int S_{ww}(\omega_1+\omega_3)S_{ww}(\omega_3)\,\mathrm{d}\omega_3$$

$$4\pi^2S_2^{(a)}=S_{ww}(\omega_1+\omega_2)\iint S_{ww}(\omega_2+\omega_3+\omega_4)S_{ww}(\omega_3)S_{ww}(\omega_4)\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4$$
$$+S_{ww}(\omega_2)\iint S_{ww}(\omega_1+\omega_2+\omega_3+\omega_4)S_{ww}(\omega_3)S_{ww}(\omega_4)\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4$$

$$4\pi^2 S_2^{(b)} = S_{ww}(\omega_1+\omega_2)\iint S_{ww}(\omega_1+\omega_3+\omega_4)S_{ww}(\omega_3)S_{ww}(\omega_4)\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4$$
$$+S_{ww}(\omega_1)\iint S_{ww}(\omega_1+\omega_2+\omega_3+\omega_4)S_{ww}(\omega_3)S_{ww}(\omega_4)\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4$$

$$4\pi^2 S_2^{(c)} = S_{ww}(\omega_1)\iint S_{ww}(\omega_2+\omega_3+\omega_4)S_{ww}(\omega_3)S_{ww}(\omega_4)\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4$$
$$+S_{ww}(\omega_2)\iint S_{ww}(\omega_1+\omega_3+\omega_4)S_{ww}(\omega_3)S_{ww}(\omega_4)\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4$$

$$4\pi^2 S_3^{(a)} = \iint S_{ww}(\omega_1+\omega_2+\omega_4)S_{ww}(\omega_2+\omega_3)S_{ww}(\omega_3)S_{ww}(\omega_4)\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4$$

$$4\pi^2 S_3^{(b)} = \iint S_{ww}(\omega_1+\omega_2+\omega_4)S_{ww}(\omega_1+\omega_3)S_{ww}(\omega_3)S_{ww}(\omega_4)\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4$$

$$4\pi^2 S_3^{(c)} = \iint S_{ww}(\omega_1+\omega_3)S_{ww}(\omega_2+\omega_4)S_{ww}(\omega_3)S_{ww}(\omega_4)\,\mathrm{d}\omega_3\,\mathrm{d}\omega_4$$

$$2\pi\tilde{S} = \int S_{ww}(\omega_1+\omega_3)S_{ww}(\omega_2+\omega_3)S_{ww}(\omega_3)\,\mathrm{d}\omega_3 \tag{A.2}$$

and $S_0 = S_0^{(a)}+S_0^{(b)}+S_0^{(c)}$, $S_1 = S_1^{(a)}+S_1^{(b)}+S_1^{(c)}$, $S_2 = S_2^{(a)}+S_2^{(b)}+S_2^{(c)}$.

It should be pointed out that each of the integrals in Eq. (A.2) extends over the real number line i.e. $-\infty\cdots+\infty$. This causes problems, however, when dealing with a random process that is both continuous and independently, identically distributed (iid). The reason for this, quite simply, is that a truly iid random process possesses a constant power spectral density i.e. $S_{ww} = const \equiv P$. By Parsevals relationship, the integral of the PSD gives the variance; however, this implies that an iid process has infinite variance. It also implies that the integrals in (A.2) are infinite. This leads to an apparent contradiction, that spectral analysis is not possible for a truly continuous iid random process. In fact, a sufficient condition for the polyspectra to exist is that

$$\sum_{\tau_1}\cdots\sum_{\tau_{n-1}}|C(\tau_1,\ldots,\tau_{n-1})| < \infty$$

where $C(\cdot)$ is the stationary joint cumulant of the random process under study [18]. For $n=2$ and 3 cumulants are equal to moments thus the power spectrum and bispectrum must possess limited second- and third-order correlations, respectively, to exist [18, pp. 213, 872]. In dealing with sampled data, however, the above criteria is automatically satisfied as the limits on the integrals are dictated by the sampling interval, Δ_t. For example, assuming an iid random process (constant PSD, P) the last equality in Eq. (A.2) becomes

$$2\pi\tilde{S} = P^3\int_{-\pi/\Delta_t}^{\pi/\Delta_t}\mathrm{d}\omega$$

$$\tilde{S} = P^3/\Delta_t \tag{A.3}$$

Thus, for an iid random process the integrals given by Eqs. (A.2) simplify considerably.

The sampling interval is often omitted from discussions of higher-order spectral analysis and the units of the bispectrum are simply left as $[y]^3$. However, Eq. (1) is a density function with units $[y]^3/\mathrm{Hz}^2$. This becomes important when comparing observations to theory where one is typically working with spectral density functions. This is also the reason for the $\Delta_t^{1/2}$ found in the expression for the bicoherence discussed earlier. Treating both numerator and denominator of Eq. (3) as spectral density functions leads to the units $Hz^{-1/2}$ for the bicoherence. The bicoherence is clearly not bounded on the interval $0\to1$ and depends very much on the sampling interval used.

In order to derive the expression for the bicoherence, an expression for the power spectral density function is needed. Using the two-term Volterra model, Eq. (4), the power spectrum can be written as

$$S_{y_iy_i}(\omega_1) = [(a_1^2+6\sigma_w^2a_1a_3+9\sigma_w^4a_3^2)S_4^{(a)}+2a_2^2S_4^{(b)}+6a_3^2S_4^{(c)}]$$
$$\times\{H_{1,i}(\omega_1)H_{1,i}(-\omega_1)+2\overline{x}(H_{2,i}(0,-\omega_1)H_{1,i}(\omega_1)+H_{2,i}(0,\omega_1)H_{1,i}(-\omega_1))\} \tag{A.4}$$

where

$$S_4^{(a)} = S_{ww}(\omega_1)$$

$$2\pi S_4^{(b)} = \int S_{ww}(\omega_1+\omega_2)S_{ww}(\omega_2)\,\mathrm{d}\omega_2$$

$$4\pi^2 S_4^{(c)} = \iint S_{ww}(\omega_1+\omega_2+\omega_3)S_{ww}(\omega_2)S_{ww}(\omega_3)\,\mathrm{d}\omega_2\,\mathrm{d}\omega_3 \tag{A.5}$$

Combining Eqs. (7), (A.1), (A.2) and (A.4), (A.5) provides an analytical expression for the bicoherence function associated with the response of a multiple degree-of-freedom structure subject to colored, non-Gaussian excitation of the type defined by Eq. (5).

References

[1] D.R. Brillinger, An introduction to polyspectra, *The Annals of Mathematical Statistics* 36 (5) (1965) 1351–1374.
[2] M.J. Hinich, C.S. Clay, The application of the discrete fourier transform in the estimation of power spectra, coherence, and bispectra of geophysical data, *Reviews of Geophysics* 6 (3) (1968) 347–363.
[3] A.M. Richardson, W.S. Hodgkiss, Bispectral analysis of underwater acoustic data, *Journal of the Acoustical Society of America* 96 (2) (1994) 828–837.
[4] M.J. Hinich, G.R. Wilson, Detection of non-gaussian signals in non-Gaussian noise using the bispectrum, *IEEE Transactions on Acoustics, Speech, and Signal Processing* 38 (7) (1990) 1126–1131.
[5] R. Nyffenegger, M.J. Hinich, D. Ritter, S. Hansen, Material discrimination using bispectral signatures, *Journal of the Acoustical Society of America* 116 (3) (2004) 1518–1523.
[6] K. Worden, G.R. Tomlinson, Nonlinearity in experimental modal analysis, *Philosophical Transactions of the Royal Society A* 359 (2001) 113–130.
[7] D. Hickey, K. Worden, M.F. Platten, J.R. Wright, J.E. Cooper, Higher-order spectra for identification of nonlinear modal coupling, *Mechanical Systems and Signal Processing* 23 (4) (2009) 1037–1061.
[8] A.R. Messina, V. Vittal, Assessment of nonlinear interaction between nonlinearly coupled modes using higher order spectra, *IEEE Transactions on Power Systems* 20 (1) (2005) 375–383.
[9] A. Rivola, P.R. White, Bispectral analysis of the bilinear oscillator with application to the detection of cracks, *Journal of Sound and Vibration* 216 (5) (1998) 889–910.
[10] G.C. Zhang, J. Chen, F.C. Li, W.H. Li, Extracting gear fault features using maximal bispectrum, *Key Engineering Materials* 293–294 (2005) 167–174.
[11] A.J. Hillis, S.A. Neild, B.W. Drinkwater, P.D. Wilcox, Global crack detection using bispectral analysis, *Proceedings of the Royal Society A* 462 (2006) 1515–1530.
[12] K.K. Teng, J.A. Brandon, Diagnostics of a system with an interface nonlinearity using higher order spectral estimators, *Key Engineering Materials* 204 (2) (2001) 271–285.
[13] J.-M.L. Caillec, R. Garello, Asymptotic bias and variance of conventional bispectrum estimates for 2-d signals, *Multidimensional Systems and Signal Processing* 16 (2005) 49–84.
[14] P. Marzocca, J.M. Nichols, M. Seaver, S.T. Trickey, Second-order spectra for quadratic nonlinear systems by Volterra functional series: analytical description and numerical simulation, *Mechanical Systems and Signal Processing* 22 (2008) 1882–1895.
[15] J.M. Nichols, P. Marzocca, A. Milanese, On the use of the auto-bispectral density for detecting quadratic nonlinearity in structural systems, *Journal of Sound and Vibration* 312 (2008) 726–735.
[16] P. Marzocca, J.M. Nichols, A. Milanese, The analytical auto-bispectrum for multi-degree-of-freedom systems, *Journal of Engineering Mathematics*, doi:10.1007/s10665-009-9349-0.
[17] J.M. Nichols, C.C. Olson, J.V. Michalowicz, F. Bucholtz, The bispectrum and bicoherence for quadratically nonlinear systems subject to non-gaussian inputs, *IEEE Transactions on Signal Processing* 57(10).
[18] M.B. Priestly, *Spectral Analysis and Time Series, Probability and Mathematical Statistics*, Elsevier Academic Press, London, 1981.
[19] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, Wiley, New York, 1980.
[20] M. Schetzen, Nonlinear system modeling based on the Wiener theory, *Proceedings of the IEEE* 69 (12) (1981) 1557–1573.
[21] J.M. Nichols, P. Marzocca, A. Milanese, The trispectrum for gaussian driven multi degree-of-freedom nonlinear structures, *International Journal of Nonlinear Mechanics* 44 (2009) 404–416.
[22] L. Isserlis, On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables, *Biometrika* 12 (1/2) (1918) 134–139.
[23] A.I. Fleishman, A method for simulating non-normal distributions, *Psychometrika* 43 (4) (1978) 521–532.
[24] S.J. Gifford, G.R. Tomlinson, Recent advances in the application of functional series to non-linear structures, *Journal of Sound and Vibration* 135 (2) (1989) 289–317.
[25] K. Worden, G.R. Tomlinson, *Nonlinearity in Structural Dynamics*, Institute of Physics Publishing, Bristol, UK, Philadelphia, USA, 2001.
[26] C.L. Nikias, M.R. Raghuveer, Bispectrum estimation: a digital signal processing framework, *Proceedings of the IEEE* 75 (7) (1987) 869–891.
[27] P.J. Huber, B. Kleiner, T. Gasser, G. Dumermuth, Statistical methods for investigating phase relations in stationary stochastic processes, *IEEE Transactions on Audio and Electroacoustics* AU-19 (1) (1971) 78–86.
[28] L.M. Garth, Y. Bresler, A comparison of optimized higher order spectral detection techniques for non-Gaussian signals, *IEEE Transactions on Signal Processing* 44 (5) (1996) 1198–1213.
[29] L.M. Garth, Y. Bresler, The degradation of higher order spectral detection using narrowband processing, *IEEE Transactions on Signal Processing* 45 (7) (1997) 1770–1784.
[30] S.M. Kay, *Fundamentals of Statistical Signal Processing: Volume II, Detection Theory*, Prentice-Hall, New Jersey, 1998.
[31] R.N. McDonough, A.D. Whalen, *Detection of Signals in Noise*, second ed., Academic Press, San Diego, 1995.