David O. Smallwood
Mechanical and Thermal Environments Department
Sandia National Laboratories
P.O. Box 5800, MS-0865 Albuquerque NM 87185-0865, USA

Generation of Stationary Non-Gaussian Time Histories with a Specified Cross-spectral Density

The paper reviews several methods for the generation of stationary realizations of sampled time histories with non-Gaussian distributions and introduces a new method which can be used to control the cross-spectral density matrix and the probability density functions (pdfs) of the multiple input problem. Discussed first are two methods for the specialized case of matching the auto (power) spectrum, the skewness, and kurtosis using generalized shot noise and using polynomial functions. It is then shown that the skewness and kurtosis can also be controlled by the phase of a complex frequency domain description of the random process. The general case of matching a target probability density function using a zero memory nonlinear (ZMNL) function is then covered. Next methods for generating vectors of random variables with a specified covariance matrix for a class of spherically invariant random vectors (SIRV) are discussed. Finally the general case of matching the cross-spectral density matrix of a vector of inputs with non-Gaussian marginal distributions is presented.

INTRODUCTION

It is recognized that some dynamic and noise environments are characterized by time histories which are not Gaussian. Examples include radar clutter, sea-waves, high-intensity acoustic noise, and some surface transportation vibration. A better simulation of these environments can be generated if zero mean non-Gaussian time histories can be reproduced with a specified cross-spectral density (CSD) matrix and a specified marginal probability density function (pdf). The motivation for this paper is primarily non-Gaussian vibration. For some environments the non-Gaussian nature of the vibration response of a structure can have a significant effect on the fatigue life of a structure. Non-Gaussian vibration can also result in peak vibration levels significantly larger than would be expected from a Gaussian assumption.

If we can generate realizations of non-Gaussian vibration with a specified cross-spectral density matrix more accurate simulations of the environments will be possible. Note that the case of a single input and specifying the auto-spectral density is a special case of specifying the cross-spectral density matrix. Modern waveform reproduction techniques can be used to reproduce the realized waveforms on electrodynamic or electrohydraulic shakers to simulate the environments.

There has been a long interest in the generation of non-Gaussian random signals. Most of the interest has been in the simulation of sea waves (Hu

Received 8 July 1997; Revised 3 September 1997.
Shock and Vibration, Vol. 4, No. 5,6, pp. 361–377 (1997)
ISSN 1070-9622/97/$8.00

and Zhao (1993)) and radar clutter. Recent interest has developed for the simulation of surface transportation. Zero memory nonlinear (ZMNL) functions have been used for a long time to generate time sequences with a specified autospectrum (power spectrum) and a non-Gaussian probability density function. For example, Gujar and Kavanagh (1968) used an analog device to generate the zero memory nonlinear function. Other authors have also used this method for generating non-Gaussian sequences, examples include: Liu and Munson (1982), Li and Yu (1989), Ren et al. (1995), Winterstein (1988), Newsam and Wegener (1994), and Gordon and Ritcey (1995). Others have used a Poisson process to generate non-Gaussian random signals: Poirion (1993), and Smallwood (1996a). Spherically invariant random vectors (SIRV) have also been used, examples include: Rangaswamy et al. (1995), Conte and Longo (1987). Others have used ARMA (auto regressive moving average) models to generate sequences with the desired spectrum and then mixed sequences with a Gaussian and Laplace distribution to generate the desired non-Gaussian sequences: examples include: Walden (1992), and Wu and Cheng (1994). Hsueh and Hamernik (1990) and Steinwolf (1996, 1997) vary the phase spectrum to achieve a non-Gaussian distribution. Kim et al. (1987) use a quadratic function to generate non-Gaussian sequences.

Many of these methods will be discussed later in the paper. The paper is not a rigorous mathematical treatment of the problem, but rather a practical engineering approach. Methods for generating a single time history are discussed first, and then vectors of time histories are covered.

GENERALIZED SHOT NOISE FOR THE GENERATION OF NON-GAUSSIAN TIME HISTORIES

A probability density function can be specified by its moments, with higher order moments often becoming less important. In this section a method will be developed for producing realizations of a zero mean random process with a specified spectral density, mean square (defined by the spectrum), skewness, and kurtosis.

Skewness and kurtosis are defined for this paper as a normalized values given by

$$S^3 = \frac{E[x^3]}{E^{3/2}[x^2]} \quad \text{and} \quad K^4 = \frac{E[x^4]}{E^2[x^2]}, \tag{1}$$

where $x(t)$ is a stationary random process. Loosely, skewness is a measure of the symmetry of the process. If the positive side of the waveform has the same characteristics as the negative side of the waveform the skewness will be zero. If the positive peaks are greater than the negative peaks the waveform will have positive skewness. Kurtosis is a measure of the crest factor of the waveform. A Gaussian waveform will have a kurtosis of three. A waveform with an excess of peaks as compared with a Gaussian waveform will have a kurtosis greater than three.

This in essence specifies the first four moments of the distribution in addition to the spectral density. The solution is not unique. Many waveforms can have the same spectrum and the same first four moments.

A generalization of shot noise will be used for the realizations. Stationary shot noise is defined (Papoulis, 1965, p. 288) as a filtered sequence of impulses, governed by a Poisson process. If the filter impulse response is $h(t)$,

$$s(t) = \sum_{i=-\infty}^{\infty} h(t - \tau_i) \tag{2}$$

where τ_i are independent random times, governed by an exponential distribution with an average rate of λ occurrences per second. It is well known that as λ approaches infinity the process becomes Gaussian (Papoulis, 1965, p. 570). By taking advantage of the process for λ less than infinity, we can generate realizations with a specified skewness and kurtosis (or equivalently the third and fourth moments) different than for a Gaussian process. It can be shown that the first four moments of shot noise are given by (Papoulis, 1965; Lin, 1967)

$$\begin{aligned} E[s] &= \lambda \bar{h}_1, \\ E[s^2] &= \lambda \bar{h}_2, \\ E[s^3] &= \lambda \bar{h}_3, \\ E[s^4] &= \lambda \bar{h}_4 + 3\lambda^2 \bar{h}_2^2, \end{aligned} \tag{3}$$

where

$$\bar{h}_n = \int_{-\infty}^{\infty} h^n(t)\, dt. \tag{3a}$$

Further discussion will be restricted to the case where $\bar{h}_1 = 0$, which results in a zero mean process.

First the shot noise will be generalized with the addition of an amplitude parameter

$$x(t) = \sum_{i=-\infty}^{\infty} A h(t - \tau_i), \tag{4}$$

$$E\left[x^n(t)\right] = A^n E\left[s^n(t)\right]. \tag{5}$$

The above expression is not used directly but as a step to the following expression. The amplitude itself can be a stationary random process, in which case the generalized shot noise takes the form

$$x(t) = \sum_{i=-\infty}^{\infty} A_i h(t - \tau_i). \tag{6}$$

If A_i and τ_i are independent, and the mean of τ_i is long compared to the duration of $h(t)$, then A_i and $h(t-\tau_i)$ will be independent and

$$E[x^n(t)] = E[A^n]E[s^n(t)]. \tag{7}$$

An impulse response function will be chosen as the inverse Fourier transform of the square root of the required auto spectral density $G_{XX}(\omega)$

$$h(t) = \lambda^{-1/2} \int_{-\infty}^{\infty} G_{XX}^{1/2} e^{j2\pi ft}\, df. \tag{8}$$

If the spectrum of A_i is white with unity amplitude, the resulting spectrum of $x(t)$ will be the desired spectrum. The next step will be the generation of an amplitude function that can be used to control the 2nd through 4th moments of $x(t)$. In each case we will need three parameters, where λ will serve as one of the parameters. This leaves two parameters to define the amplitude function. Several options are available. One method would be to describe the amplitude as a random process of independent random variables, with a normal distribution described with a mean, μ_A, and standard deviation, σ_A, which gives

$$\begin{aligned} E[A] &= \mu_A, \\ E[A^2] &= \sigma_A^2 + \mu_A^2, \\ E[A^3] &= \mu_A(3\sigma_A^2 + \mu_A^2), \\ E[A^4] &= 3\sigma_A^4 + 6\sigma_A^2\mu_A^2 + \mu_A^4. \end{aligned} \tag{9}$$

A second method would describe a series of impulses with two fixed amplitudes, A_x and A_n, and a probability of 1/2 for each amplitude. Each amplitude A_i is chosen independently as one of these two amplitudes, which gives

$$\begin{aligned} E[A] &= \frac{1}{2}(A_x + A_n), \\ E[A^2] &= \frac{1}{2}(A_x^2 + A_n^2), \\ E[A^3] &= \frac{1}{2}(A_x^3 + A_n^3), \\ E[A^4] &= \frac{1}{2}(A_x^4 + A_n^4). \end{aligned} \tag{10}$$

Both the above methods generate sequences with the correct values for skewness and kurtosis for large values of kurtosis. As the kurtosis approaches 3, and λ gets larger, the values for μ_A, σ_A, A_x, and A_n from Eqs. (9) and (10) will over-predict the kurtosis as the independence assumption in Eq. (7) is violated.

A third method will describe impulses chosen with fixed amplitudes A_x or $-A_x$. The amplitudes A_i will be chosen independently with a probability of p for A_x and $(1-p)$ for $-A_x$. The expected values of A are

$$\begin{aligned} E[A] &= A_x p + (1-p)(-A_x) \\ &= A_x(2p-1), \\ E[A^2] &= A_x^2 p + (1-p)(-A_x)^2 = A_x^2, \\ E[A^3] &= A_x^3(2p-1), \\ E[A^4] &= A_x^4. \end{aligned} \tag{11}$$

The last method will be used in this paper because it leads to a simple closed form solution for the three parameters; λ, A_x, and p. The combinations of Eqs. (2), (7), and (11) gives

$$\begin{aligned} E[x] &= 0, \\ E[x^2] &= \lambda A_x^2 \bar{h}_2, \\ E[x^3] &= \lambda A_x^3 (2p-1)\bar{h}_3, \\ E[x^4] &= A_x^4(\lambda \bar{h}_4 + 3\lambda^2 \bar{h}_2^2). \end{aligned} \tag{12}$$

The solution for Eq. (12) for the unknown parameters is

$$\begin{aligned} \lambda &= \frac{\bar{h}_4}{\bar{h}_2^2}\left(\frac{1}{K^4 - 3}\right), \\ A_x &= \sqrt{\frac{E[x^2]}{\lambda \bar{h}_2}}, \\ p &= \frac{1}{2}\left(\frac{S^3\sqrt{\lambda \bar{h}_2^3}}{\bar{h}_3} + 1\right), \end{aligned} \tag{13}$$

where S^3 and K^4 are the normalized skewness and kurtosis for a zero mean process. For a Gaussian process the skewness and kurtosis are zero and three, respectively. Several observations can be made from the solution. First, as $K^4 \to 3$, $\lambda \to \infty$ as expected. Since λ must be positive, only a kurtosis greater than 3 can be achieved. Second, the product $A_x^2\lambda$ is equal to the constant $E[x^2]/\bar{h}_2$. Third, if we assume $\bar{h}_3 > 0$, if $S = 0$, then $p = 1/2$; if S is greater than 0, $p > 1/2$; and if S is less than 0, $p < 1/2$.

Since p is limited to the interval [0, 1], the skewness is limited by

$$\left|\frac{S^3\sqrt{\lambda \bar{h}_2^3}}{\bar{h}_3}\right| \leqslant 1. \tag{14}$$

Since as $K^4 \to 3$, $\lambda \to \infty$, $S^3 \to 0$, only a skewness of near zero is available as the kurtosis approaches 3. Fourth, as

$$\lambda \to 0, \quad K^4 \to \infty \tag{15}$$

which implies any kurtosis greater than 3 can be achieved. The spectrum does not approach zero because the amplitude of impulses will approach infinity.

It is conjectured that for any simulation with a specified kurtosis, a range of skewness values about zero are possible, and as the kurtosis approaches three, the range of possible skewness values approaches zero.

Digital Implementation

The impulse response function $h(t)$ can be approximated in a digital simulation by taking the inverse FFT of the square root of the auto spectral density of the random process to be simulated

$$\tilde{h}_i = \frac{1}{T^{3/2}\lambda}\,\mathrm{FFT}^{-1}(G_{xx}^{1/2}), \quad i = 0, \ldots, N-1, \tag{16}$$

where N is the block size and T is the duration of a frame. In this form samples of the impulse response for negative time will be in the upper half of the transform. A circular shift of $N/2$ will place the samples of the impulse response in a more convenient form. This will result in a delay in the output of $N/2$ samples, but will not change the final result. In theory the impulse response is not time limited. To reduce leakage caused by truncation of the impulse response, the sampled impulse response can be multiplied by a window and the mean removed.

$$h_i = \left(\tilde{h}_i - \frac{\mu_{hw}}{\mu_w}\right) w_i, \quad i = 0, \ldots, N-1, \tag{17}$$

where

$$\mu_{hw} = \frac{1}{N}\sum_{i=0}^{N-1} \tilde{h}_i w_i \quad \text{and} \quad \mu_w = \frac{1}{N}\sum_{i=0}^{N-1} w_i. \tag{18}$$

The times of the impulses are τ_m. The sample number of the mth impulse is given by k_m. A realization of the next delay is given by

$$k_m = \tau_m/\Delta t = k_{m-1} + \mathrm{round}(r_m/\Delta t), \tag{19}$$

where Δt is the sample interval, r_m are independent numbers generated from a exponential distribution with a mean of λ, and round( ) means round to the nearest integer. Since the impulse response function is non zero only in the interval 0 to $N-1$, the sum defining x (Eq. (6)) must extend only over the terms where $h(j-k_i)$ is non zero. This interval includes i's where $0 \leqslant j-k_i \leqslant N-1$. Samples of the random realization can the be generated using

$$x_j = \sum_i A_i h_{j-k_i}, \tag{20}$$

where i extends over the values

$$j \leqslant k_i \leqslant j + N - 1.$$

If the first impulse is at $k_1 = 1$, after a transient of N samples a stationary output will be generated. The frequency resolution (in this case we mean the bandwidth, B, if the smallest frequency feature which can be generated in the realization) of the expected spectral density of samples of $x(t)$, x_i will be determined by the block size, N, and the window used. As N increases the bandwidth of the smallest feature decreases such that product $B^*N\Delta t$ remains approximately a constant. Sometimes λ is so large that the quantization of τ_m can be a significant problem. For this case $h(t)$ can be sampled at a multiple of the sample rate for $x(t)$. An offset of the first sample of $h(t)$ and decimation can be used to improve the resolution of the delays.

Examples of this method are given by Smallwood (1996a). Poirion (1993) uses a variation of the Poisson shot noise to generate non-Gaussian waveforms.

POLYNOMIAL FUNCTIONS USED FOR THE GENERATION OF NON-GAUSSIAN WAVEFORMS

It has been suggested that a polynomial function of the form

$$y(t) = a_n x^n + \cdots + a_1 x + a_0 \tag{21}$$

could be used as a zero memory nonlinear function to transform a Gaussian variable into a variable with non-Gaussian characteristics.

Bendat (1990) and Merritt (1997) suggest three forms of a system to accomplish the desired goal as suggested by Fig. 1. Type 1: A is excluded and the system is reduced to a linear system followed by a zero memory nonlinear system. Type 2: B is excluded and the system is reduced to a zero memory nonlinear system followed by a linear system. And Type 3:

FIGURE 1 Generalized system using linear functions in series with a zero memory nonlinear function.

A zero memory nonlinear function is preceded and followed by a linear system. The linear systems A and B are used to shape the spectrum of the output. The zero memory nonlinear system is used to control the skewness and kurtosis.

Formulas for the skewness and kurtosis as a function of the parameters a_n can be developed (Merritt, 1995; Winterstein, 1988); n greater than 3 is seldom used. Winterstein has suggested that the form of Eq. (21) is appropriate for kurtosis greater than three, and that an inverse formula as below is appropriate for kurtosis less than three.

$$x(t) = a_n y^n + \cdots + a_1 y + a_0. \tag{22}$$

The three different forms have advantages and disadvantages: Type 1: A is excluded (a linear system followed by a zero memory nonlinear function) and the distribution of $y(t)$ will be the same as $v(t)$. If the distribution of $u(t)$ is Gaussian, the skewness and kurtosis of $v(t)$ can be predicted from the polynomial coefficients. However, the spectrum of $v(t)$ will be almost the same as the spectrum of $u(t)$, but not quite, as will be explained in a later section. Type 2: An alternate formulation is a zero memory nonlinear function followed by a linear system (B is excluded). The spectrum of $y(t)$ will be accurately predicted by A and the spectrum of $v(t)$, but the distribution will be changed (will move toward being closer to Gaussian) by A. This requires that the coefficients of the polynomial be chosen differently to give a distribution, which when distorted by A, will give the desired skewness and kurtosis. This typically requires an iterative solution. Winterstein typically uses a linear system followed by a zero memory nonlinear function, where Merritt seems to favor a zero memory nonlinear function followed by a linear system. Type 3: A type 3 system includes a linear system before and after (A and B) the zero memory nonlinear function.

In each case care must be taken in choosing the coefficients such that $y(x)$ is a monotonically increasing function of x over the range of the simulation. This implies that dy/dx is positive over the range of the simulation. Otherwise a nonrealizable cumulative distribution function (CDF) for y will result as will be explained in a later section.

THE USE OF PHASE TO CONTROL SKEWNESS AND KURTOSIS

A reasonable simulation of a random time history can be expressed in the frequency domain as

$$x(t) = \sum_{k=1}^{N} A_k \cos(2\pi k \Delta f t + \varphi_k). \tag{23}$$

If an FFT is used the above formulation results in one period of a pseudo-random signal. This is explained in detail by Smallwood and Paez (1991). The simulation can also be expressed in the form

$$x(t) = \sum_{k=1}^{N} a_k \cos(2\pi k \Delta f t) - b_k \sin(2\pi k \Delta f t). \tag{24}$$

For a Gaussian waveform, a_k and b_k are realizations of random variables, independent, normally distributed with a zero mean, and with a standard deviation given by

$$std(a_k) = std(b_k) = \sqrt{\Delta f S(k\Delta f)}, \tag{25}$$

where Δf is the spacing between frequency lines and $S(k\Delta f)$ is the desired auto spectral density at the frequency $k\Delta f$. Steinwolf (1996) and Steinwolf and Ibrahim (1997) give interesting formulas for the skewness and kurtosis for waveforms where a_k and b_k are not independent:

$$E[x^2] = \sum_{k=1}^{N} (a_k^2 + b_k^2), \tag{26}$$

$$\begin{aligned} S^3 E^{3/2}[x^2] &= E[x^3] \\ &= \frac{3}{4} \sum_{j=2k} \left\{ a_j (a_k^2 - b_k^2) + 2 b_j a_k b_k \right\} \\ &\quad + \frac{3}{2} \sum_{\substack{j-k=m \\ j<k}} \left\{ (a_j a_k - b_j b_k) a_m \right. \\ &\quad \left. + (a_j b_k + a_k b_j) b_m \right\}, \end{aligned} \tag{27}$$

$$\begin{aligned} K^4 = 3 + &\left\{ \sum_{k=1}^{N} (a_k^2 + b_k^2) \right\}^{-2} \left\{ -\frac{3}{2} \sum_{k=1}^{N} (a_k^2 + b_k^2)^2 \right. \\ &+ 2 \sum_{j=3k} \left[a_j a_k (a_k^2 - 3 b_k^2) - b_j b_k (b_k^2 - 3 a_k^2) \right] \\ &+ 6 \sum_{\substack{j=k+2n \\ k \neq n}} \left[(a_j a_k + b_j b_k)(a_n^2 - b_n^2) \right. \end{aligned}$$

$$
\begin{aligned}
&- 2(a_j b_k - a_k b_j) a_n b_n \big] \\
&+ 6 \sum_{\substack{j+k=2n \\ j<k}} \big[(a_j a_k - b_j b_k)(a_n^2 - b_n^2) \\
&+ 2(a_j b_k + a_k b_j) a_n b_n \big] \\
&+ 12 \sum_{\substack{j+k=n+m \\ j<k, n<m, j<n}} \big[(a_j a_k - b_j b_k)(a_n a_m - b_n b_m) \big] \\
&\times \big[(a_j b_k + a_k b_j)(a_n b_m + a_m b_n) \big] \\
&+ 12 \sum_{\substack{j+k+n=m \\ j<k<n}} \big[(a_j a_k - b_j b_k)(a_n a_m + b_n b_m) \big] \\
&\times \big[(a_j b_k + a_k b_j)(a_n b_m - a_m b_n) \big] \Bigg\}
\end{aligned} \tag{28}
$$

where the summation should be carried out only for those combinations of the indices j, k, m, n that satisfy the inequalities written for each summation.

Observe that the skewness will be zero when a_k and b_k are uncorrelated because all the terms involve a product $a_k b_j$ which has an expected value of zero when $j \neq k$. The kurtosis will approach 3 only as the number of terms, N, gets large. All the terms for the kurtosis except the first summation involve correlations of harmonics and sum and difference frequencies ($j = 3k$, $j = k + 2n$, $j + k = 2n$, $j + k = n + m$, $j + k + n = m$). Harmonic distortion caused by a nonlinear relation between x and y can clearly cause this correlation.

It is also clear that by introducing a correlation between some or all of the a_k and b_k coefficients, the skewness and kurtosis can be changed. It is also clear that if N is large, the solution is not unique; many different combinations of correlations can result in the same skewness and kurtosis. Hsueh et al. (1990) correlate the coefficients in a band of frequencies. Steinwolf (1996) uses an unspecified method to correlate the coefficients.

The generation of a continuous stationary waveform by this method will present some difficulty. The data are generated in frames depending on the choice of the sample rate and the number of coefficients, N. The joining of the blocks is not straightforward since discontinuities will exist at the frame boundaries. If the frames are long enough this may not be a significant problem. If a windowing and overlapping operation is attempted to remove the discontinuity, the distribution will be changed and moved toward a Gaussian distribution. The frames could be circularly shifted to a zero crossing to ease the discontinuity problem. A circular shift will not change the moments or the spectrum.

GENERATION OF A ZERO MEMORY NONLINEAR FUNCTION TO ACHIEVE A SPECIFIED DISTRIBUTION

For a Type 1 system with a single input, a Gaussian time history with the specified autospectral density (ASD) is generated. Smallwood and Paez (1991) and Tebbs and Hunter (1974) discuss an efficient way to accomplish this generation. A monotonic zero memory nonlinear function relating the Gaussian waveform to the desired realization is then established based on the cumulative distribution function (CDF) of the desired waveform and the known cumulative distribution function of a Gaussian waveform. The established function is used to transform the Gaussian waveform into a realization of the desired waveform. This method was suggested as early as 1968 by Gujar and Kavanagh.

If the probability density function of the desired waveform is known or can be estimated, the resulting cumulative distribution function can be used with the known distribution of a Gaussian waveform to establish a zero memory nonlinear transformation function. This function relates a random variable with a Gaussian distribution to a random variable with the desired distribution. To arrive at the required function, the formula for the change in variables using the cumulative distribution function will be used (Wirsching et al., 1995).

$$F_Y(y) = F_X\big(g^{-1}(y)\big), \tag{29}$$

where

$$y = g(x) \quad \text{or} \quad x = g^{-1}(y). \tag{30}$$

In summary

$$y = F_Y^{-1}\big(F_X(x)\big) = g(x) \tag{31}$$

is a monotonically increasing zero memory nonlinear function. The procedure for finding $g(x)$ from the known functions $F_Y(y)$ and $F_X(x)$ is best illustrated graphically in Fig. 2. $F_Y(y)$ and $F_X(x)$ will be restricted to functions which result in a monotonically increasing function $g(x)$. It is seen from Fig. 2 that a point in the $x - y$ plane (x_1, y_1) for which $F_X(x_1) = F_Y(y_1)$ is a point on the function $y = g(x)$. If both $F_Y(y)$ and $F_X(x)$ are known, $y = g(x)$ can be constructed. Usually for experimental data the probability density function is first estimated and the cumulative distribution function is estimated by integrating the probability density function. In this development $F_Y(y)$ is the target non-Gaussian distribution, and $F_X(x)$ is a Gaussian distribution.

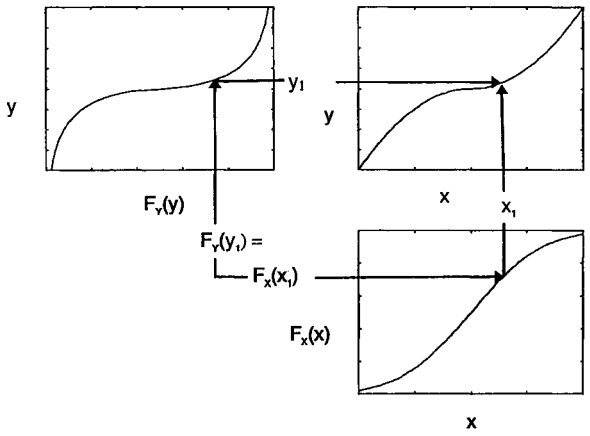

FIGURE 2 Generation of the zero memory nonlinear function transformation function, $y = g(x)$.

A sampled realization of a waveform with a Gaussian distribution and with the specified spectral density will then be generated. Each sample in the resulting realization will be transformed using the previously derived function, $y = g(x)$. If the function is "smooth" and monotonically increasing, the transformation will preserve the timing of all the zero crossings, minimums and maximums of the original waveform. Since most of the spectral information is contained in the zero crossings (Wise et al., 1977; Bendat and Piersol, 1986, Sec. 12.6.4), the spectrum will usually not be substantially changed. Some harmonic distortion will be introduced by the transformation. The spectrum of the distorted waveform can be estimated and an error spectrum generated. The error spectrum can be subtracted from the original spectrum if the subtraction does not result in a negative spectrum, and a new realization of the Gaussian waveform can be generated from the corrected spectrum in an iterative fashion. Often this iteration is not required as the distorted waveform spectrum is near enough to the target spectrum to be useful without correction.

If experimental data are used to generate the probability density function of the target spectrum, a procedure which results in a smooth continuous estimate of the probability density function is needed. A simple histogram will usually not be sufficient. The method used in this paper involves the kernel density estimator and is discussed by Silverman (1986). The actual probability density function estimation routine used in this paper was supplied by Norm Hunter of Los Alamos National Laboratories. The resulting probability density function estimate was integrated to estimate the cumulative distribution function.

The probability density function and the cumulative distribution function of the Gaussian distribution can be generated using standard functions. If the target distribution is known as one of the classical distributions the probability density function and cumulative distribution function can often be evaluated analytically.

In some cases sufficient data are not available to produce a smooth target cumulative distribution function of the desired data and, hence, a smooth zero memory nonlinear function. In this case a polynomial function can often be fitted to the zero memory nonlinear with good success. Often a least squares fit is used. If the normalized kurtosis is greater than three, a polynomial function of the form

$$y = \sum_{i=1}^{N} a_i x^i \tag{32}$$

can be used. If the normalized kurtosis is less than 3, a function of the form

$$x = \sum_{i=1}^{N} a_i y^i \tag{33}$$

can be used (Winterstein et al., 1994). N greater than 4 is seldom required. Care must be taken to restrict the coefficients to values which will result in a monotonic increasing function, $y = g(x)$, over the range of the data to be considered. Winterstein has suggested that often sufficient information about the distribution is not known for this method to be used, and that fitting the skewness and kurtosis using the polynomial fits of the previous section is the best that can be practically accomplished. However, it is the experience of the author and Merritt (1997) that the statistical uncertainties of estimating the skewness and kurtosis are so large that if sufficient data are available to estimate the skewness and kurtosis with a reasonable uncertainty, enough data are also available to make a reasonable estimate of the probability density function. By fitting a polynomial curve to the zero memory nonlinear function defined above, and by visually examining the fit, insight into the process can be gained.

Liu et al. (1982) use a zero memory nonlinear function in a time domain technique where the Gaussian time history spectrum is shaped with a digital filter. Li (1989) uses a Type 1 system with a AR filter and a zero memory nonlinear function to model a Weibull distribution. Gordon and Ritcey (1995) also use a variation of the Type 1 method with a linear filter followed by a zero memory nonlinear function.

USING A SPHERICALLY INVARIANT RANDOM PROCESS (SIRP)

Rangaswamy et al. (1991, 1993, 1995) have written a series of papers which use a spherically invariant random process to simulate non-Gaussian random processes. let **H** be a linear manifold generated by some set of random variables $\{x_n\}$. If all the random variables in **H** have the same variance and have the same distribution function, we call the set $\{x_n\}$ *spherically invariant*. If the $\{x_n\}$ are arbitrary samples from a random process, this defines a *spherically invariant random process* (Blake and Thomas, 1968). An important property of a spherically invariant random process is that a linear operation on a spherically invariant random process is also a spherically invariant random process with the same distribution. A Gaussian process is a special case of a spherically invariant random process. But there are other distributions which are also spherically invariant including the Laplace, Cauchy, K-distribution, and the Student t-distributions (Rangaswamy et al., 1995). Another important property is that if a random vector is a spherically invariant random vector, then there exists a nonnegative random variable S such that the probability density function of the random vector conditioned on S is a multivariate Gaussian probability density function. These properties allow the generation of a random vector with a desired covariance matrix and a non-Gaussian probability density function in the following manner.

1) Generate a white zero-mean Gaussian random vector, **Z**, having an identity covariance matrix. It is not necessary for the mean to be zero, but it will be assumed in this paper.
2) Generate a random variable, V, from the probability density function $f_V(v)$, where f_V is the characteristic probability density function for the spherically invariant random vectors and is tabulated for several distributions by Rangaswamy et al. (1995).
3) Normalize the random variable, V, by a to obtain S. In other words generate $S = V/a$.
4) Generate the product given by $\mathbf{X} = \mathbf{Z}S$. At this step we have a white spherically invariant random vector having a zero mean and identity covariance matrix.
5) Perform the product given by

$$\mathbf{Y} = \mathbf{A}\mathbf{X}, \tag{34}$$

where

$$\mathbf{A}\mathbf{A}^{\mathbf{T}} = \Sigma, \tag{35}$$

where Σ is the desired covariance matrix and **Y** is the desired random vector.

This method has the advantage of keeping the desired covariance matrix and the desired probability density function independent. But the method also has some disadvantages. First, if the desired probability density function is not one of the tabulated probability density functions, the calculations can become quite tedious. Second, if the desired random vector is long, the desired covariance matrix gets quite large. For example, if it is desired to control the spectrum of a single random variable, the random vector needs to be at least as long as the autocorrelation function to properly identify the covariance matrix. Each row of the square covariance matrix is a time shifted autocorrelation function. The method will then generate frames of data as long as the random vector. If a stationary continuous sampled time history is desired, the problem of joining the frames is the same as discussed earlier.

Conte and Longo (1987) also use the spherically invariant random process method.

GENERALIZATION OF THE SMALLWOOD PAEZ METHOD FOR THE MULTIPLE INPUT CASE

In this section we want to discuss the generation of a vector of sampled time histories $\{x(t)\}$, where each element in the vector is a sampled time history. We would also like to specify the cross-spectral density matrix between the elements, $[G_{xx}]$, where the diagonal elements are the auto spectra of each element in the vector and the off diagonal elements are the cross-spectra between the elements.

In theory the method using a spherically invariant random process could be used to generate data for this case, but in practice it would be very awkward.

Smallwood and Paez (1991) showed how to generate a vector of normal stationary inputs with a specified cross-spectral density in an efficient manner. Briefly the cross-spectral density matrix is decomposed into the form

$$G_{xx} = HG_{ii}H', \tag{36}$$

where G_{xx} is the desired cross-spectral density matrix, H is a matrix of frequency response functions, H' is the complex conjugate transpose of H, and G_{ii} is a diagonal matrix with real components. Several methods can be used for this decomposition, the most common being a modified Cholesky decomposition. A vector of independent normal random sources is then synthesized with spectra given by the diagonal elements of G_{ii}. The resulting independent signals are generated in the frequency domain by setting the Fourier magnitude equal to the square root of the auto spectrum

and randomizing the phase. This is usually sufficient except for narrow band processes. For a more general case, or for the case of a narrow band process, the magnitude of the real and imaginary parts of the Fourier spectrum are picked from a normal distribution with zero mean and a standard deviation equal to the square root of the desired spectrum divided by 2. Typically a finite Fourier transform is used, given the spectrum at a set of $M/2 + 1$ frequency lines.

$$real\big(Z_i(f)\big) = \sqrt{G_{ii}(f)/2}\, normr$$

and

$$imaginary\big(Z_i(f)\big) = \sqrt{G_{ii}(f)/2}\, normr, \qquad (37)$$

where *normr* is a normal random variable with zero mean and unity standard deviation. Scaling of these vectors typically takes place later. This is very similar to the method described in the phase control section. This results in a vector of frames of data in the frequency domain, $\{Z(f)\}$. A vector is defined for each frequency line. The following discussion will be for a single frequency line. The calculations are repeated for each frequency line. The elements in the vectors are independent at this stage. The vectors are placed on the diagonal of a square matrix $[Z(f)]$. The square diagonal matrix is multiplied by the frequency response matrix to give a square matrix of drive signals

$$\big[D(f)\big] = \big[H(f)\big]\big[Z(f)\big]. \qquad (38)$$

The calculations are repeated for each frequency line. At this point each element in the matrix **D** is a frequency domain description of a frame of data. The results are transformed into the time domain using an inverse FFT.

$$\big[D(t)\big] = \mathrm{FFT}^{-1}\big[D(f)\big]. \qquad (39)$$

This results in an N by N matrix of frames of data for an N input problem. Each element in the matrix is a frame of data in the time domain M points long. The frames of data are then added as follows to produce a vector of outputs of length N, where each element in the vector is a frame M points long. The nth output is the sum of the data in the nth row of the matrix of drives, $\mathbf{D}(t)$. The mth column in $\mathbf{D}(t)$ represents the contribution of the mth independent noise source, Z_m. The output frames are windowed, overlapped, and added to previously generated frames to generate a continuous stationary normal output. The windowing controls leakage, and the overlap is chosen to produce a stationary output. The output vectors are then scaled to account for the sample rate, the block size, the gain from the overlapping of the frames, and the loss due to the window. The method is versatile and efficient. Details are contained in Smallwood and Paez (1991).

To extend this procedure to non-Gaussian waveforms, a zero memory nonlinear function is developed for each input using the desired cumulative distribution function for each individual input and the cumulative distribution function of a Gaussian distribution (Smallwood, 1996c). The synthesized Gaussian waveforms with the desired cross-spectral density are transformed into non-Gaussian waveforms using the developed zero memory nonlinear functions using a Type 1 system. As for the single input case the cross spectral density matrix is usually not seriously changed by this transformation, as will be illustrated by the following example.

Example

In this example a two input case is illustrated. The cross-spectral between the inputs is specified by specifying the coherence and phase between the inputs. The break points defining the cross-spectral density are given in Table 1. In the table only values at the break points are given. If a cell is blank it implies that the spectrum is a straight line at this frequency. Dashed lines are used to show these values on later plots. The required spectrum terminates at 2000 Hz. The spectra are calculated to 5000 Hz. Any spectral values above 2000 Hz are the result of leakage from lower frequencies.

Table 1. Break Points Defining the Cross-Spectral Density Matrix

Frequency (Hz)	$G_{11}(f)$	$G_{22}(f)$	Coherence	Phase
40	0.036	0.005	1	0
100	0.036		1	0
110	0.0036			
200		0.005		
210		0.009		
300	0.0036	0.009		
310	0.036	0.005		
500			0.5	
700			0.5	
1010	0.036	0.0025		
1020	0.0036	0.0009	0.1	
2000	0.0036	0.0009	0.1	
2020	0	0	1	
5000	0	0	1	0

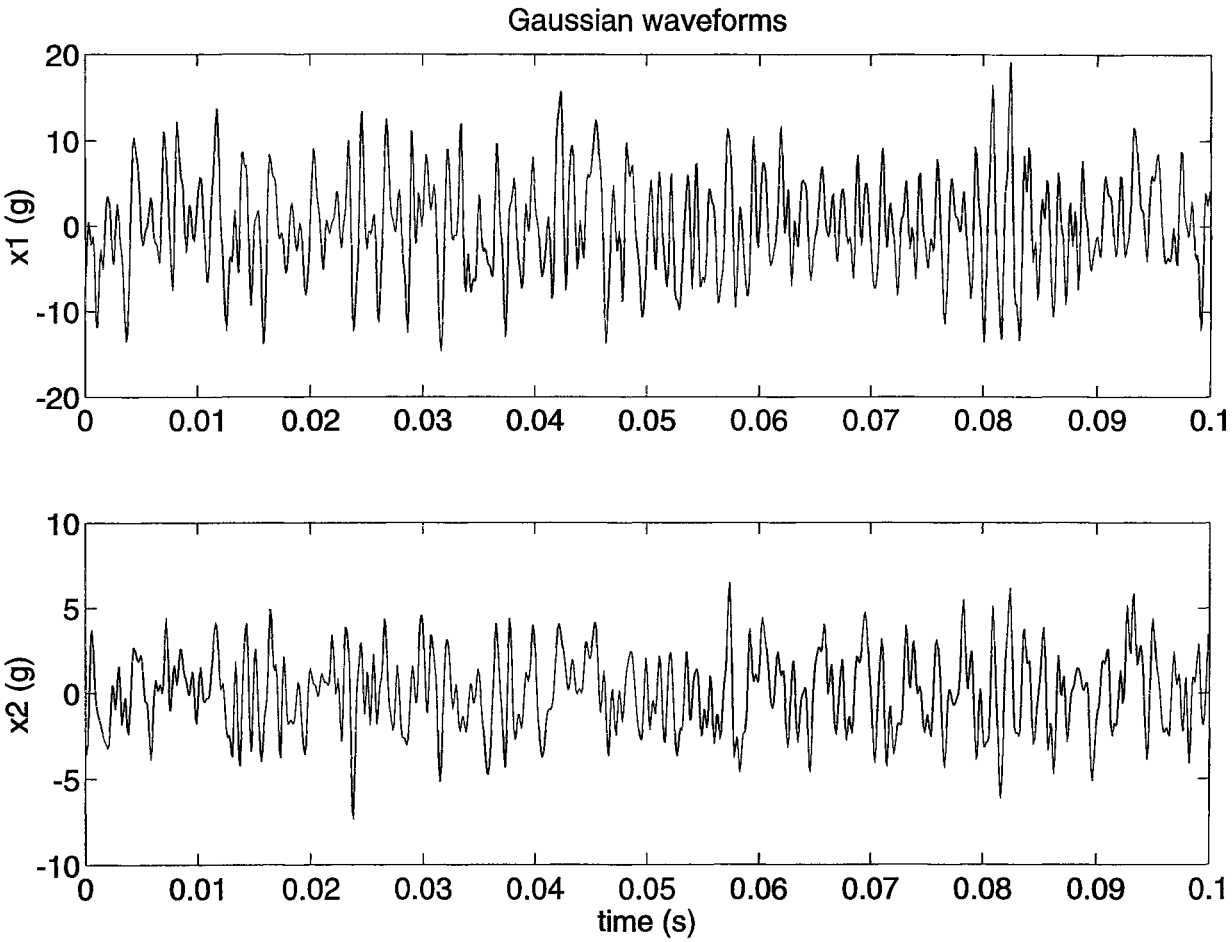

FIGURE 3 Sample Gaussian waveforms.

Realizations of waveforms for three cases are shown. The data were generated using the method of Smallwood and Paez, using a frame size of 2048, a sample rate of 10000 samples/s, and 10 frames of data. First, Gaussian waveforms are simulated. One thousand samples of the waveforms for the two inputs are shown in Fig. 3. The estimated spectra, coherence, and phase for the 20480 points are shown in Fig. 4. An estimate of the probability density function together with an equivalent Gaussian probability density function are shown in Fig. 5. The probability density function is shown as both a linear amplitude and a log amplitude. The log amplitude shows the structure of the tails better than the linear amplitude.

Second waveforms are generated with a uniform distribution. This is an extreme case with a normalized kurtosis of 1.8, requiring high distortion of the Gaussian data. The zero memory nonlinear functions are shown as Fig. 6. Notice the shape; the slope approaches zero as the absolute value of x gets large. It is apparent that a function of the form given by Eq. (33) would fit this function better than the form given by Eq. (32). A slight slope at the extremes of the zero memory nonlinear was allowed to assure a monotonic function. One thousand samples of the waveforms for the two inputs are shown in Fig. 7. The spectral estimates are shown as Fig. 8 and the estimated probability density function as Fig. 9.

The third example attempts to match a Student t-distribution with 5 degrees of freedom. This distribution has a normalized kurtosis of about 5.7. The zero memory nonlinear functions are shown in Fig. 10. Notice the different shape, as compared with Fig. 6; the slope increases as the absolute value of x increases. For this example the kurtosis is larger than 3. One thousand samples of the waveforms for the two inputs are shown in Fig. 11, the spectral estimates are shown as Fig. 12, and the estimated probability density functions together with an equivalent Student t-probability density function are shown as Fig. 13.

The results are generally acceptable. The distortion caused by the zero memory nonlinear function is not

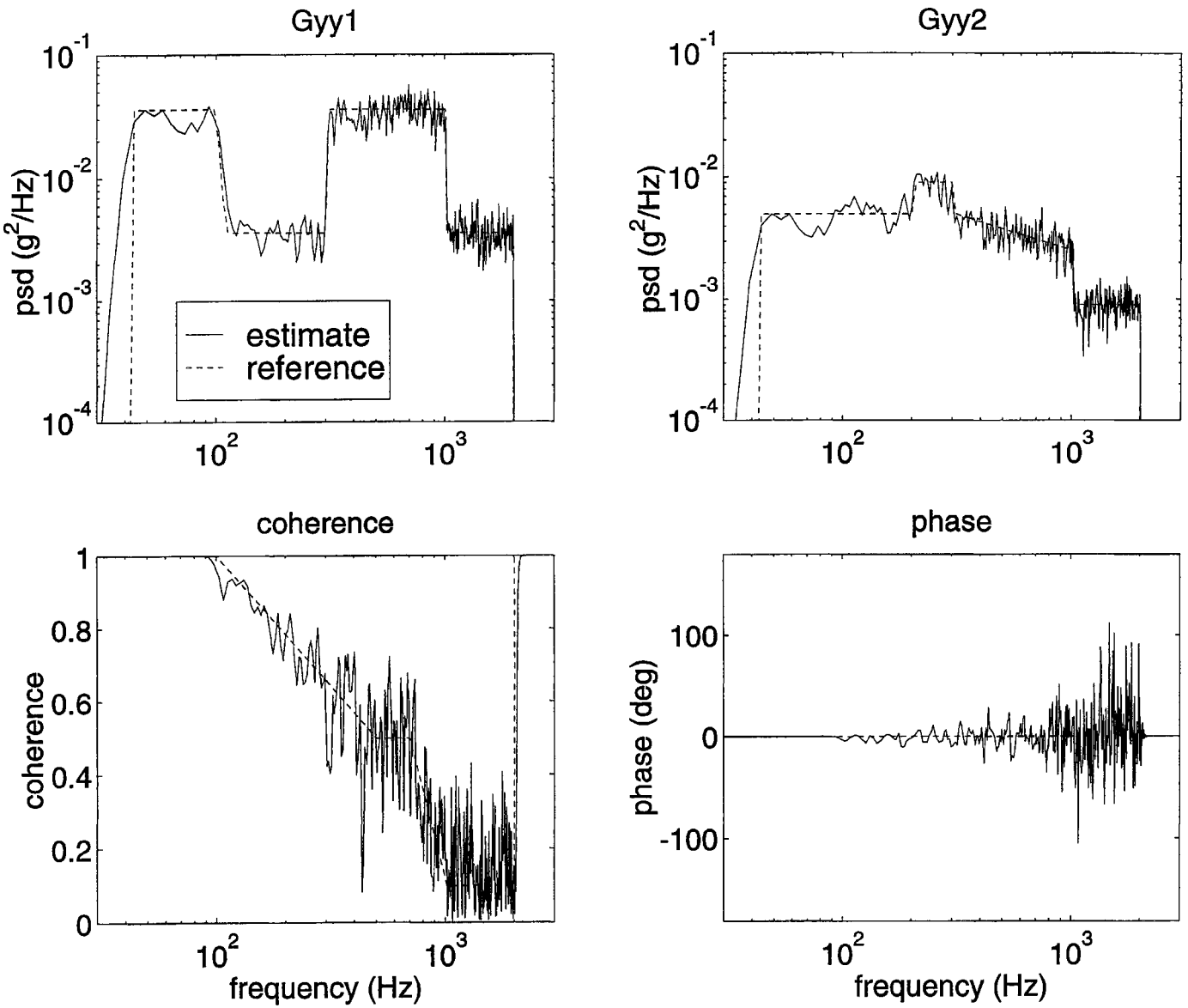

FIGURE 4 Power spectra and coherence for Gaussian waveform.

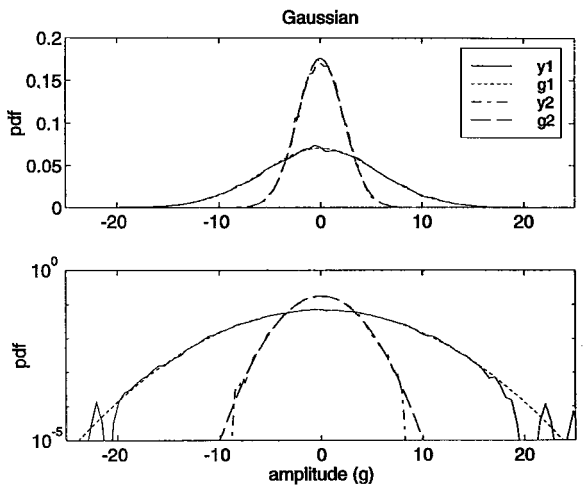

FIGURE 5 Probability density function of Gaussian waveforms.

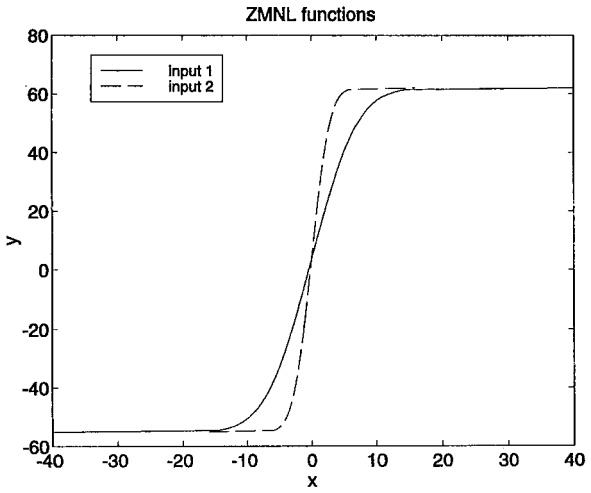

FIGURE 6 Zero memory nonlinear function for a uniform distribution.

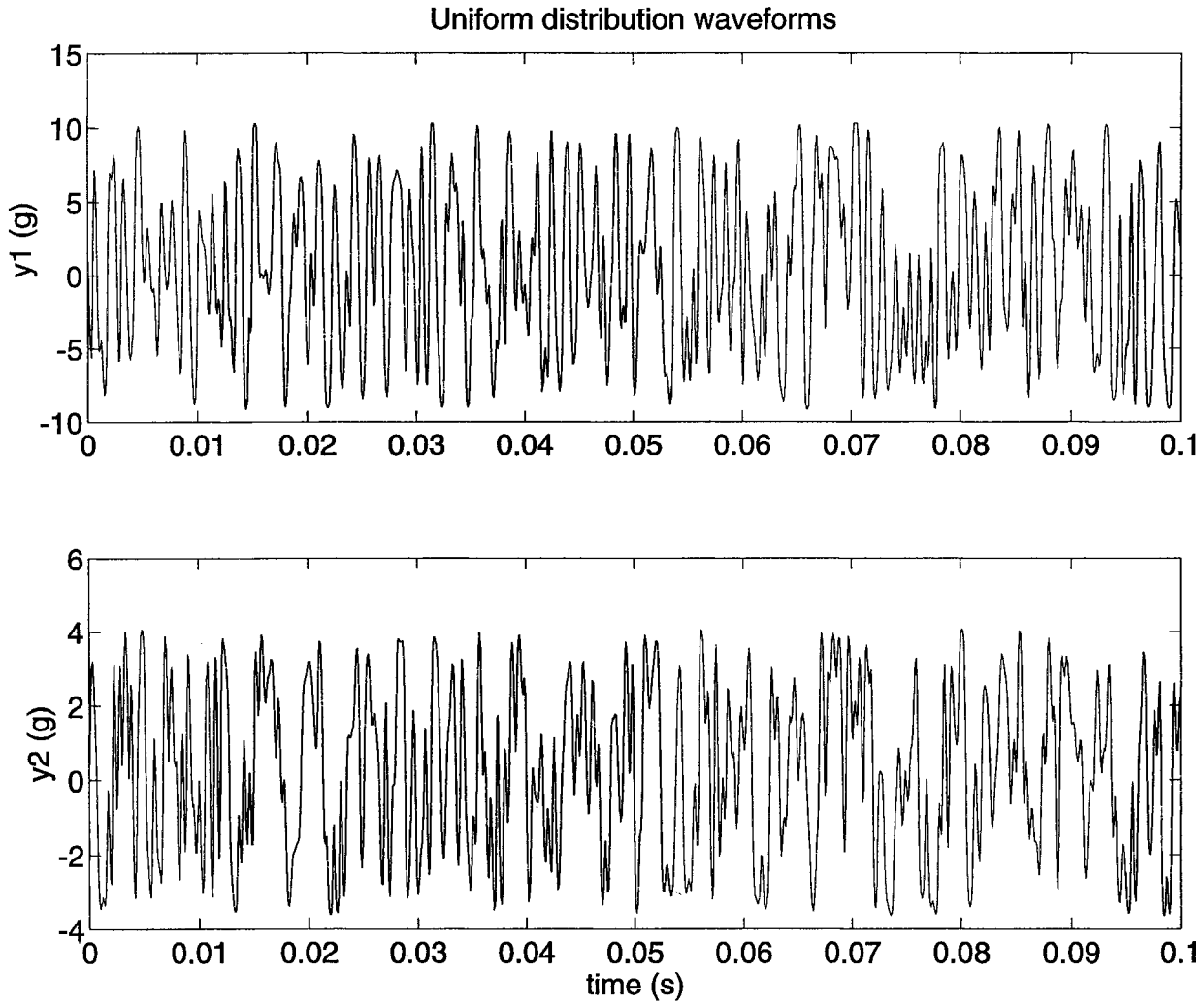

FIGURE 7 Sample of uniform waveforms.

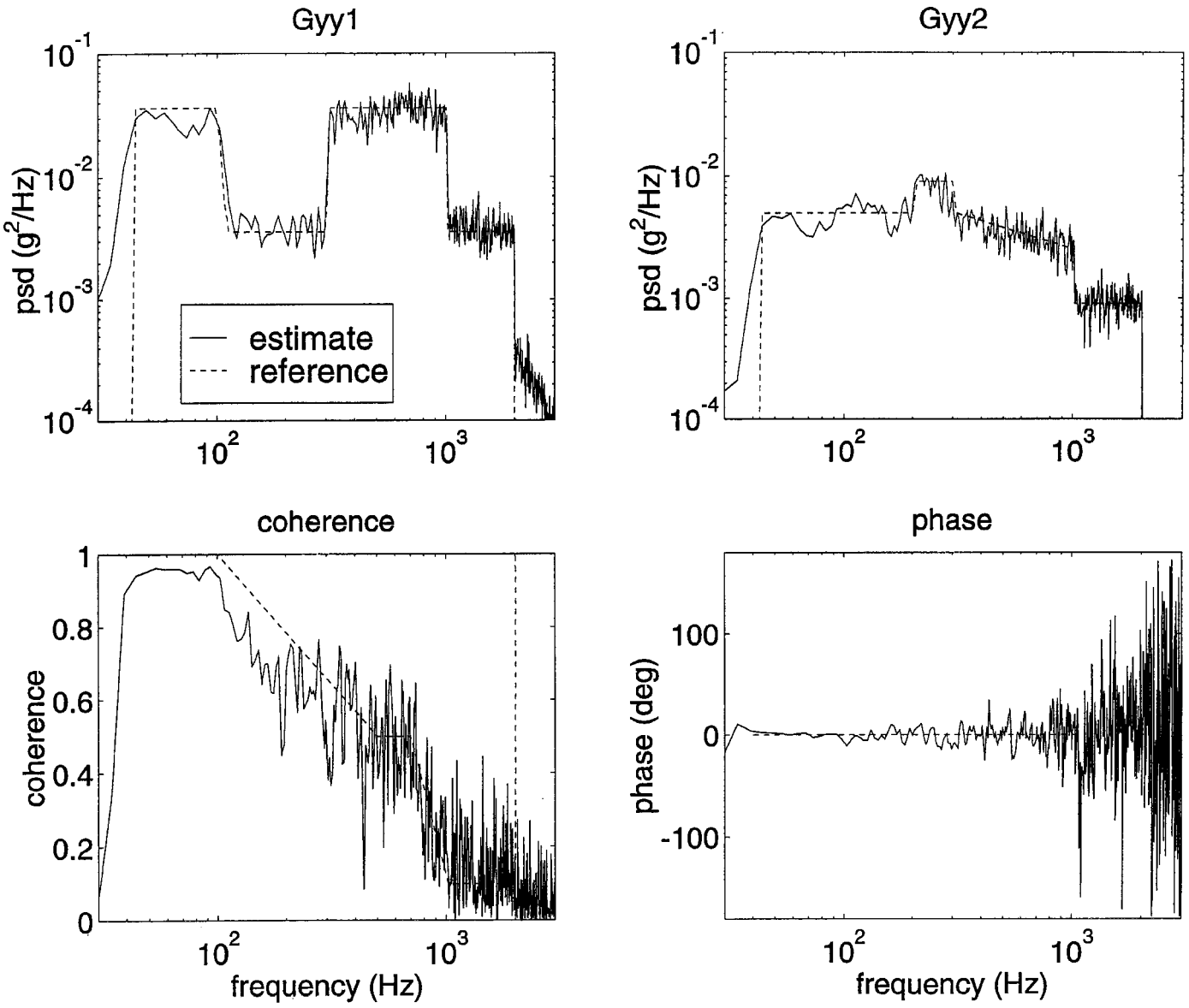

FIGURE 8 Spectra and coherence of uniform waveforms.

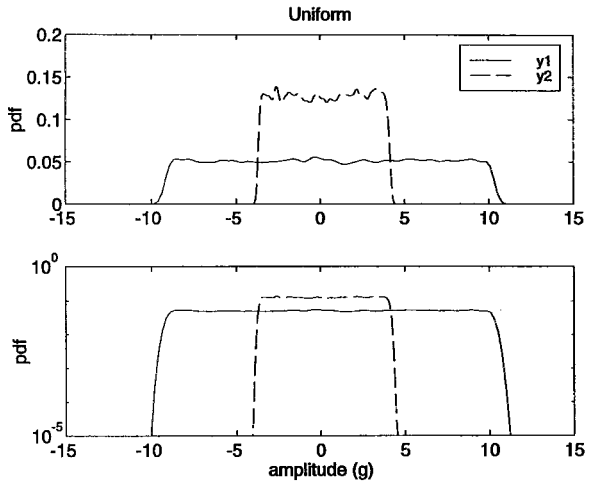

FIGURE 9 Probability density function of uniform waveforms.

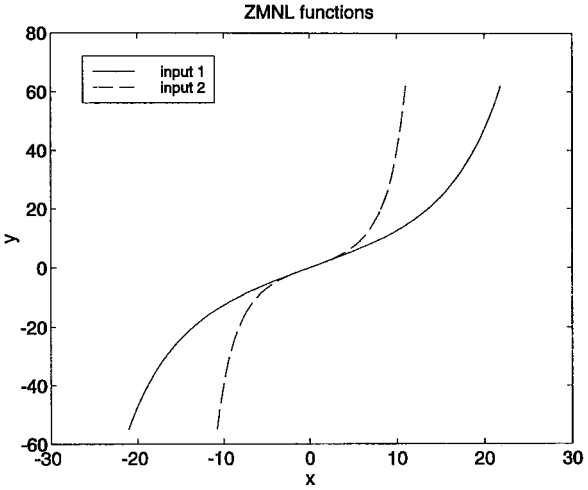

FIGURE 10 Zero memory nonlinear function of Student t-distributions.

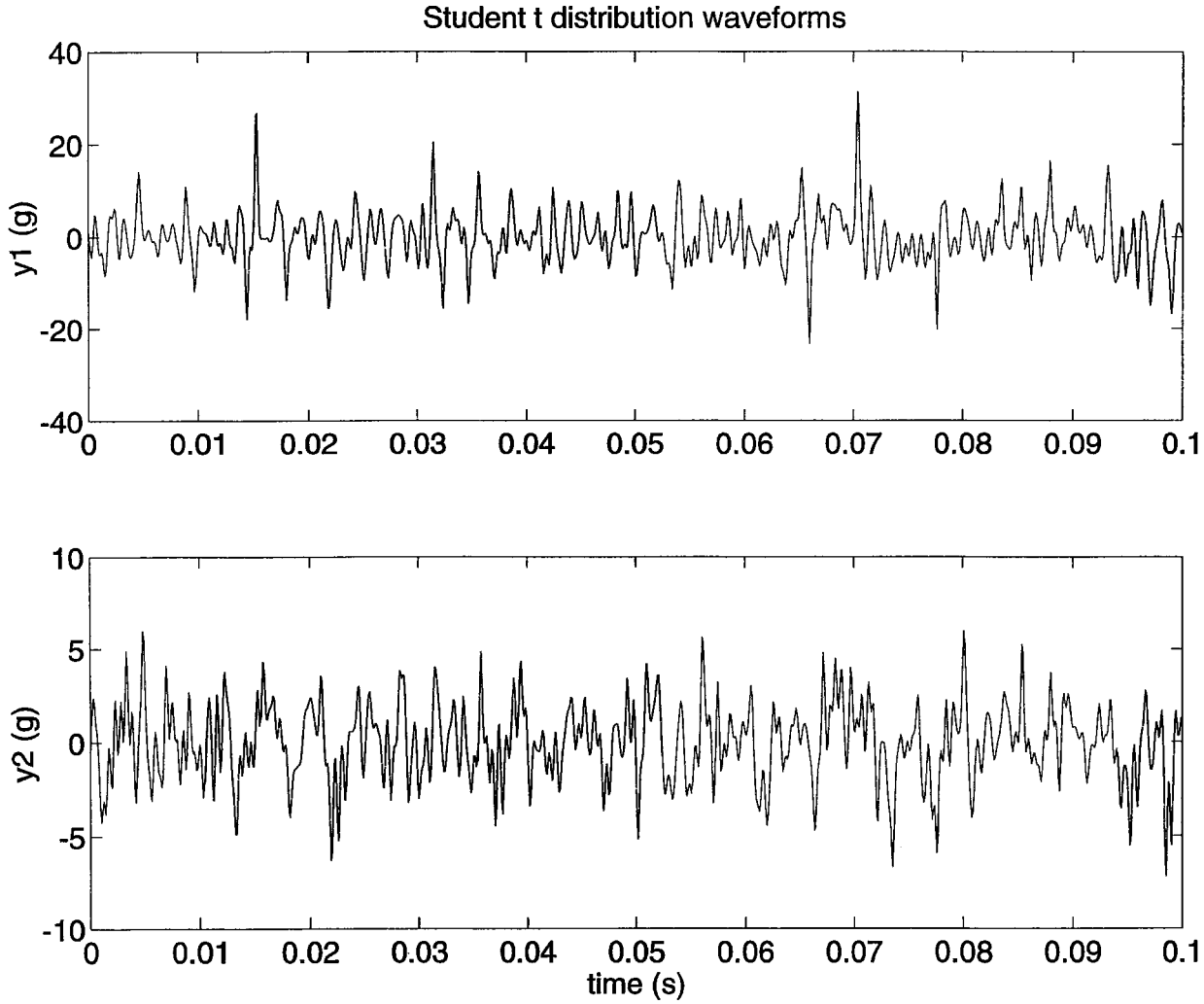

FIGURE 11 Sample of Student t-waveform.

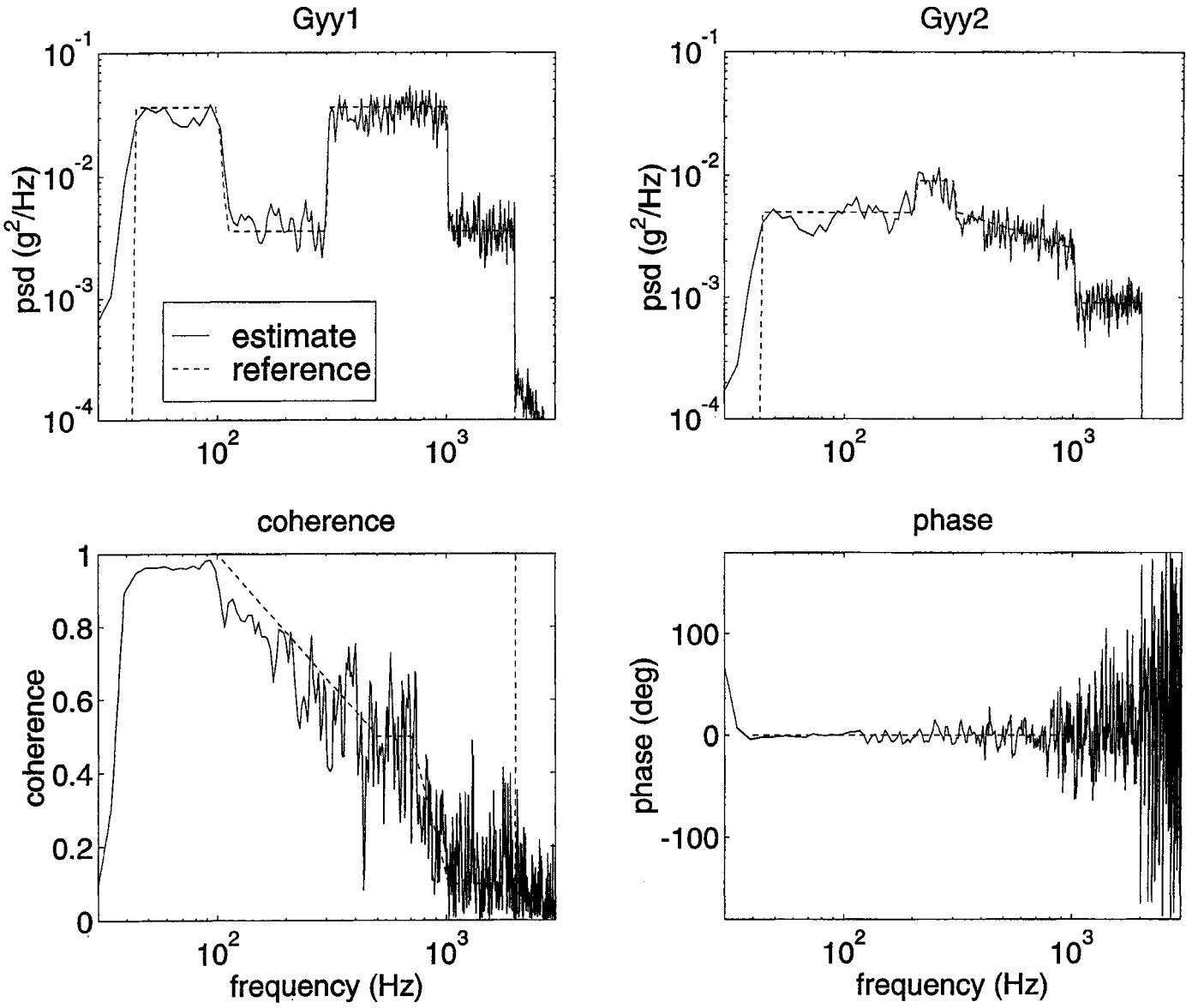

FIGURE 12 Spectra and coherence of Student t-waveform.

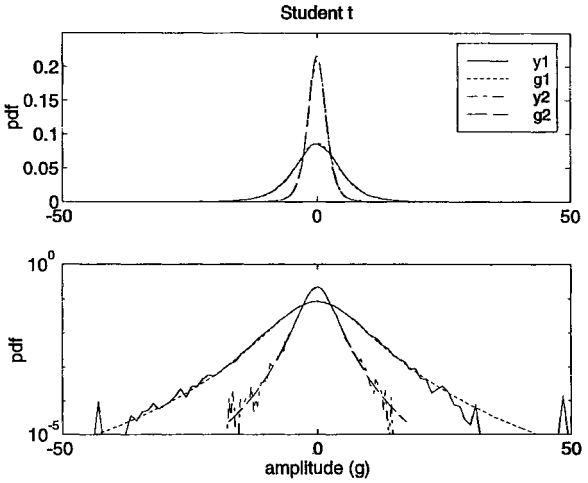

FIGURE 13 Probability density function of Student t-waveform.

severe. Evidence of high frequency distortion above 2000 Hz is evident in both the non-Gaussian waveforms. The distortion is worse for the more severe uniform distribution than for the Student t-distribution. Some loss of coherence is seen near a sharp decline in the spectral values. For example, compare the coherence from 100–200 Hz in Figs. 4, 8, and 12. As expected, the uncertainty in the estimated coherence and phase becomes large as the coherence gets small.

OTHER METHODS

A method for controlling the skewness and kurtosis was reported by Synergistic Technology Incorporated (1992). However the methods employed have not been published and will not be discussed here.

Kim et al. (1987) use a quadratic model based on a Volterra series to model the nonlinear response of moored vessels. The idea is to use a linear model in parallel with a quadratic model. The method is very similar to the models proposed by Bendat (1990).

CONCLUSIONS

Several methods are presented for generating non-Gaussian waveforms. It is seen that the solution is not unique, and that different methods will give results with different characteristics. The method chosen should depend on the application. A general method for the multiple input case is developed which has broad application where the cross-spectral density matrix of the inputs can be estimated.

– Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

REFERENCES

Bendat, J. S., 1990, *Nonlinear System Analysis and Identification*, John Wiley and Sons, New York, NY.

Bendat, J. S., and Piersol, A. G., 1986, *Random Data, Analysis and Measurement Procedures, 2nd Edition*, John Wiley and Sons, New York, NY.

Blake, I. F., and Thomas, J. B., 1968, "On a Class of Processes Arising in Linear Estimation Theory," *IEEE Transactions on Information Theory*, Vol. IT-14, No. 1, pp. 12–15.

Conte, E., and Longo, E., 1987, "Characterization of Radar Clutter as a Spherically Invariant Random Process," *IEE Proc.*, Vol. 134, Pt. F, No. 2, pp. 191–197.

Gordon, S. D., and Ritcey, J. A., 1995, "A Transformation Approach for Modeling and Detecting Non-Gaussian Signals," *IEEE 1058–6393/95.*

Gujar, U. G., and Kavanagh, R. J., 1968, "Generation of Random Signals with Specified Probability Density Functions and Power Spectral Density Spectra," *IEEE Transactions on Automatic Control*, pp. 716–719.

Hsueh, K. D., and Hamernik, R. P., 1990, "A Generalized Approach to Random Noise Synthesis: Theory and Computer Simulation," *J. Acoust. Soc. Am.* 87(3), pp. 12–1217.

Hu, S.-L. J., and Zhao, D., 1993, "Non-Gaussian Properties of Second-Order Random Waves," ASCE, *J. of Engr. Mech.*, Vol. 119, No. 2.

Kim, K. I., Powers, E. J., Ritz, C. P., Miksad, R. W., and Fischer, F. J., 1987, "Modeling of the Nonlinear Drift Oscillations of Moored Vessels Subject to Non-Gaussian Random Sea-wave Excitation," *IEEE J. of Oceanic Engineering*, Vol, OE-12, No. 4, pp. 568–575.

Li, G., and Yu, K.-B., 1989, "Modeling and Simulation of Coherent Weibull Clutter," *IEE Proceedings, Pt. F, Comm., Radar, and Signal Processing*, 136, 1, pp. 2–12.

Lin, Y. K., 1967, *Probabilistic Theory of Structural Dynamics*, McGraw-Hill, New York, NY.

Liu, B., and Munson, D. C., 1982, "Generation of a Random Sequence Having a Jointly Specified Marginal Distribution and Autocovariance," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-30, No. 6, pp. 973–983.

Merritt, R. G., 1995, "A Stochastic Model for the Pulse Method – Part 1: Large Deterministic Part," *Proceedings of the Institute of Environmental Sciences (IES)*, Mount Prospect IL.

Merritt, R. G., 1997, "A Stochastic Model for the Pulse Method – Part 2: Random Part," *Proceedings of the Institute of Environmental Sciences (IES)*, Mount Prospect IL.

Newsam, G. N., and Wegener, 1994, "Generating Non-Gaussian Random Fields for Sea Surface Simulations," *IEEE, Proc. of ICASSP'94*, Vol. 6, p. VI/195-8.

Papoulis, A., 1965, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, NY.

Poirion, F., 1993, "Numerical Simulation of Homogeneous Non-Gaussian Random Vector Fields," *J. of Sound and Vibration*, 160(1), pp. 25–42.

Rangaswamy, M., Weiner, D., and Ozturk, A., 1991, "Simulation of Correlated Non-Gaussian Interference for Radar Signal Detection," *IEEE, publication 1058-6393/91.*

Rangaswamy, M., Weiner, D., and Ozturk, A., 1993, "Non-Gaussian Random Vector Identification Using Spherically Invariant Random Processes," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 29, No. 1, pp. 111–124.

Rangaswamy, M., Weiner, D., and Ozturk, A., 1995, "Computer Generation of Correlated Non-Gaussian Radar Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 31, No. 1, pp. 106–116.

Ren, Y. J., Elishakoff, I., and Shinozuka, M., 1994, "Conditional Simulation of Non-Gaussian Random Fields for Earthquake Monitoring Systems," *Chaos, Solitons, & Fractals*, Vol. 5, No. 1, pp. 91–101.

Silverman, B. W., 1986, *Density Estimation for Statistics and Data Analysis,* Chapman and Hall, New York, NY.

Smallwood, D. O. and Paez, T. L., 1991, "A Frequency Domain Method for the Generation of Partially Coherent Normal Stationary Time Domain Signals," *62nd Shock and Vibration Symposium*, Vol. 2, pp. 373–382. Also, *Shock and Vibration*, Vol. 1, No. 1, pp. 45–53 (1993).

Smallwood, D. O., 1996a, "Generation of Time Histories with a Specified Auto Spectral Density, Skewness, and Kurtosis," *1996 Proceedings of the Institute of Environmental Sciences (IES)*, pp. 304–309, Mount Prospect IL.

Smallwood, D. O., 1996b, "Generation of Time Histories with a Specified Auto Spectral Density and Probability Density Function," *Proceedings of NOISE-Con 96,* Bellevue, Washington, pp. 763–768, Noise Control Foundation, P. O. Box 2469, Arlington Branch, Poughkeepsie, NY 12603.

Smallwood, D. O., 1996c, "Generation of Partially Coherent Stationary Time Histories with Non-Gaussian Distributions," *Proceedings of the 67th Shock and Vibration Symposium*, Vol. 1, pp. 489–498, SAVIAC, 2711 Jefferson Davis Highway, Suite 600, Arlington Virginia 22202-4158.

Steinwolf, A., 1996, "Shaker Simulation of Random Vibrations with High Kurtosis Value," 1996, *Proceedings of the Institute of Environmental Sciences (IES)*, pp. 310–317, Mount Prospect IL.

Steinwolf, A. and Ibrahim, R., 1997, "Analysis and Simulation of Asymmetric Probability Distributions During Fundamental and Environmental Random Vibration Tests," *Proceedings of the Institute of Environmental Sciences (IES)*, Mount Prospect IL.

Synergistic Technology Incorporated, 1992, "Software User's Manual for the STI-MEVCS Multiexciter Waveform Control System," Published by Synergistic Technology Incorporated, now a part of Spectral Dynamics Corp., 1983 Concourse Ave., San Jose, CA 95131.

Tebbs, J. D., and Hunter, N. F., 1974, "Digital Control of Random Vibration Tests Using a Sigma V Computer," *1974 Proceedings of the Institute of Environmental Sciences (IES)*, pp. 36–43, Mount Prospect IL.

Walden, A.T., 1993, "Simulation of Realistic Synthetic Reflection Sequences," *Geophysical Prospecting*, 41, 313–321.

Winterstein, S. R., Lange, C. H., and Kumar, S., 1994, "Fitting: A Subroutine to Fit Four-Moment Probability Distributions to Data," Sandia National Laboratories Technical Report, SAND94-3039, Albuquerque, NM.

Winterstein, S. R., 1988, "Nonlinear Vibration Models for Extremes and Fatigue," *J. of Engineering Mechanics*, Vol. 114, No. 10.

Wirsching, P. H., Paez, T. L., and Ortiz, K., 1995, *Random Vibration: Theory and Practice,* John Wiley and Sons, New York, NY.

Wise, G., Traganitis, A., and Thomas, J., 1977, "The Effect of a Memoryless Nonlinearity on the Spectrum of a Random Process," *IEEE Transactions on Information Theory*, IT-23, pp. 84–89.

Wu, W.-R., and Cheng, P.-P., 1994, "A Nonlinear IMM Algorithm for Maneuvering Target Tracking," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 30, No. 3, pp. 875–884.