Journal of Sound and Vibration (2001) **246**(4), 723–735
doi:10.1006/jsvi.2001.3698, available online at http://www.idealibrary.com on IDEAL®

A CLASS OF NON-GAUSSIAN PROCESSES FOR MONTE CARLO SIMULATION

M. GRIGORIU

School of Civil & Environmental Engineering, Cornell University, 369 Hollister Hall, Ithaca, NY 14853-3501, U.S.A.

(*Received* 16 *June* 2000, *and in final form* 26 *March* 2001)

A new model is proposed to represent non-Gaussian stationary processes and develop Monte Carlo simulation algorithms for generating sample paths of non-Gaussian processes. The model is based on a class of non-Gaussian processes, the class of conditional Gaussian processes. Two representations are considered for these processes. The first representation is based on a randomized version of the classical spectral density function. The second representation uses the output of a linear filter with random coefficients subjected to Gaussian noise to define conditional Gaussian processes. The proposed model, its representations, and corresponding Monte Carlo simulation algorithms are illustrated by examples involving non-Gaussian random variables and processes. It is shown that the proposed model can match any second moment properties but, generally, can only fit approximately a specified marginal distribution.

1. INTRODUCTION

Most action on physical systems, material properties, stock prices, and other time series follow non-Gaussian distributions [1]. It is difficult to develop simple analytical models for a non-Gaussian process with an arbitrary probability law. Most current non-Gaussian models match only the second moment properties and the marginal distribution of a target non-Gaussian process, for example, the translation, the diffusion, and the filtered Poisson processes.

The translation process is a memoryless transformation of a Gaussian process with mean zero, variance one, and correlation function depending on the correlation function of the target non-Gaussian process. The translation process can follow any marginal distribution but not any correlation function. The correlation function of the underlying Gaussian process can be obtained by an iteration algorithm that converges if there exists a translation process with the required properties. Conditions for the existence of a translation process with a specified marginal distribution and covariance function are given in reference [1]. Diffusion processes can also be used to model non-Gaussian processes. Except for non-Gaussian processes with an exponential correlation function, the determination of the drift and the diffusion coefficients of a diffusion process matching a target marginal distribution and second moment properties is impractical [1, 2]. Filtered Poisson processes have been used less frequently in applications to model non-Gaussian processes [1].

The non-Gaussian models in this paper belong to the class of conditional Gaussian processes. The sub-Gaussian process, that is, a Gaussian process scaled by an α-stable random variable, is a member of this class [3]. The sub-Gaussian process is not ergodic

0022-460X/01/390723 + 13 $35.00/0

because its sample and ensemble properties differ. The lack of ergodicity is a common characteristic of the conditional Gaussian processes. The model in this paper may be viewed as a generalized version of the sub-Gaussian process. It is shown that the proposed model can have any second moment properties but cannot match exactly all marginal distributions. The model can only match a number of prescribed higher order moments. In contrast, the translation process can fit any marginal distribution but not an arbitrary correlation function. Examples are used to demonstrate the features of the proposed conditional Gaussian process; they illustrate the calibration of this process to target properties, and show its use in Monte Carlo simulation.

2. DEFINITION AND PROPERTIES

Let $\{G(\nu, \mathbf{z}),\ \nu \geqslant 0,\ \mathbf{z} \in D\}$ be a $[0, \infty)$-valued function defined on $[0, \infty) \times D$, where D is a subset of $\mathbb{R}^d$ it is assumed that $\int_0^\infty G(\nu, \mathbf{z})\, \mathrm{d}\nu < \infty$ for each $\mathbf{z} \in D$. Hence, $G(\cdot, \mathbf{z})$ can be the one-sided spectral density function of a weakly stationary process for each $\mathbf{z} \in D$.

Let $(\Omega_1, \mathscr{F}_1, P_1)$ be a probability space and

$$\mathbf{Z} : (\Omega_1, \mathscr{F}_1) \to (D, \mathscr{B}(D)) \tag{1}$$

denotes a measurable function, where $\mathscr{B}(D)$ is the Borel σ-field on D. Hence, $G(\cdot, \mathbf{Z}(\omega_1))$ has the properties of a one-sided spectral density function of a weakly stationary process for each $\omega_1 \in \Omega_1$. Consider another probability space, $(\Omega_2, \mathscr{F}_2, P_2)$, and the mapping

$$X : [0, \infty) \times (\Omega_1 \times \Omega_2) \to \mathbb{R} \tag{2}$$

defined by

$$X(t, \omega_1, \omega_2) = \int_0^\infty [\cos(\nu t)\, \mathrm{d}U(\nu, \omega_1, \omega_2) + \sin(\nu t)\, \mathrm{d}V(\nu, \omega_1, \omega_2)], \tag{3}$$

where

$$\mathrm{d}U(\nu, \omega_1, \omega_2) = \sqrt{G(\nu, \mathbf{Z}(\omega_1))}\, \mathrm{d}\tilde{U}(\nu, \omega_2), \quad \mathrm{d}V(\nu, \omega_1, \omega_2) = \sqrt{G(\nu, \mathbf{Z}(\omega_1))}\, \mathrm{d}\tilde{V}(\nu, \omega_2), \tag{4}$$

$\tilde{U}$ and $\tilde{V}$ are real-valued process on the probability space $(\Omega_2, \mathscr{F}_2, P_2)$ such that

$$\begin{aligned} &E[\mathrm{d}\tilde{U}(\nu, \omega_2)] = E[\mathrm{d}\tilde{V}(\nu, \omega_2)] = 0, \\ &E[\mathrm{d}\tilde{U}(\nu, \omega_2)\, \mathrm{d}\tilde{U}(\lambda, \omega_2)] = E[\mathrm{d}\tilde{V}(\nu, \omega_2)\, \mathrm{d}\tilde{V}(\lambda, \omega_2)] = \delta(\nu - \lambda)\, \mathrm{d}\nu, \\ &E[\mathrm{d}\tilde{U}(\nu, \omega_2)\, \mathrm{d}\tilde{V}(\lambda, \omega_2)] = 0 \end{aligned} \tag{5}$$

and $\delta(\xi) = 1$ for $\xi = 0$ and is zero otherwise.

(1) *If G is measurable, that is, it is a random variable on $(\Omega_1, \mathscr{F}_1, P_1)$, the mapping $(t, \omega_1, \omega_2) \mapsto X(t, \omega_1, \omega_2)$ is a stochastic process on the product probability space $(\Omega_1 \times \Omega_2, \mathscr{F}_1 \times \mathscr{F}_2, P_1 \times P_2)$.*

Proof. The functions

$$(\omega_1, \omega_2) \mapsto \sqrt{G(\nu, \mathbf{Z}(\omega_1))}\cos(\nu t)\, \mathrm{d}\tilde{U}(\nu, \omega_2) \quad \text{and}$$

$$(\omega_1, \omega_2) \mapsto \sqrt{G(\nu, \mathbf{Z}(\omega_1))}\sin(\nu t)\, \mathrm{d}\tilde{V}(\nu, \omega_2)$$

are measurable in the arguments (ω_1, ω_2) for each (t, v) by the measurability of G, $\tilde{U}$, and $\tilde{V}$. Since summation and integration preserve measurability, X is measurable in (ω_1, ω_2) for each $t \geqslant 0$ so that it is a stochastic process on the product space. □

(2) *If* (1) $\tilde{U} = B_1$ *and* $\tilde{V} = B_2$, *where* B_1 *and* B_2 *are independent standard Brownian motions starting at zero*, (2) G *has continuous samples, and* (3) *the processes* B_i, $i = 1, 2$, *are independent of* G, *then the mapping* $(t, \omega_1, \omega_2) \mapsto X(t, \omega_1, \omega_2)$ *defined by equations* (2)–(5) *has a measurable version.*

Proof. Suppose that $\mathbf{Z}$ is a constant in $\mathbb{R}^d$ so that $g(v) = G(v, \mathbf{Z})$ is a $[0, \infty)$-valued deterministic function such that $\int_0^\infty g(v)\,\mathrm{d}v < \infty$. The function g can be the spectral density of a weakly stationary process. If this process is continuous in probability, then it has a measurable version, this is, a version, that is measurable jointly in v and ω_2 (reference [3], Exercise 9·15 (*i*), p. 443).

If $\tilde{U}$ and $\tilde{V}$ are two independent Brownian motions and G has continuous samples as assumed, the resulting version,

$$X(t, \omega_1, \omega_2) = \int_0^\infty \sqrt{G(v, \mathbf{Z}(\omega_1))}\,[\cos(vt)\,\mathrm{d}B_1(v, \omega_2) + \sin(vt)\,\mathrm{d}B_2(v, \omega_2)] \tag{6}$$

of X has continuous samples so that it is measurable in (t, ω_1, ω_2) (reference [5], Exercise 9·15 (ii), p. 443). □

(3) *The second moment properties of* X *are*

$$\mu = E[X(t)] = 0, \qquad c(\tau) = E[X(t)X(t+\tau)] = \int_0^\infty E[G(v, \mathbf{Z})]\cos(vt)\,\mathrm{d}v, \tag{7}$$

where the expectation $E[G(v, \mathbf{Z})]$ *is with respect to the* $\mathbb{R}^d$*-valued random variable* $\mathbf{Z}$.

Proof. The expectation, $E[X(t)] = E\{E[X(t)\,|\,\mathbf{Z}]\}$, is zero because $E[X(t)\,|\,\mathbf{Z}] = 0$ by the properties of U and V [equation (5)]. The covariance function, $E[X(t)X(t+\tau)]$, of this process can be calculated from $E\{E[X(t)X(t+\tau)\,|\,\mathbf{Z}]\}$ and

$$\begin{aligned}
&E[X(t)X(t+\tau)\,|\,\mathbf{Z}] \\
&= E\left[\int_0^\infty (\cos(vt)\,\mathrm{d}U(v) + \sin(vt)\,\mathrm{d}V(v)) \int_0^\infty (\cos(\lambda(t+\tau))\,\mathrm{d}U(\lambda) + \sin(\lambda(t+\tau))\,\mathrm{d}V(\lambda))\right] \\
&= \int_0^\infty (\cos(vt)\cos(v(t+\tau))\,E[\mathrm{d}U(v)^2] + \sin(vt)\sin(v(t+\tau))\,E[\mathrm{d}V(v)^2]) \\
&= \int_0^\infty G(v, \mathbf{Z})\cos(v\tau)\,\mathrm{d}v
\end{aligned} \tag{8}$$

by the properties of U and V in equation (5), the definition of the mean square integral, and properties of the expectation operator. □

(4) X *is a weakly stationary process that can have any second moment properties.*

Proof. This property is a direct consequence of equation (7). It is sufficient to select G such that its expectation matches the target spectral density function. □

(5) *If $\tilde{U}$ and $\tilde{V}$ are Gaussian processes, the conditional process $X|\mathbf{Z}$ is also Gaussian with mean zero and covariance function, $E[X(t)X(t+\tau)|\mathbf{Z}]$, equal to $\int_0^\infty G(v, \mathbf{Z})\cos(v\tau)\,dv$.*

Proof. The covariance function of $X|\mathbf{Z}$ is given by equation (8). That $X|\mathbf{Z}$ is Gaussian follows from equations (3)–(5). □

(6) *If $\tilde{U}$ and $\tilde{V}$ are Gaussian processes, the characteristic function of the random variable $X(t)$ is*

$$\varphi(u, t) = \varphi(u) = E\left\{\exp\left[-\frac{u^2}{2}\int_0^\infty G(v, \mathbf{Z})\,dv\right]\right\} \tag{9}$$

for each $t \geqslant 0$.

Proof. The conditional random variable $X(t)|\mathbf{Z}$ follows a Gaussian distribution with mean zero and variance $\int_0^\infty G(v, \mathbf{Z})\,dv$ at all times $t \geqslant 0$. Hence, its characteristic function is

$$E[e^{\sqrt{-1}uX(t)}|\mathbf{Z}] = \exp\left[-\frac{u^2}{2}\int_0^\infty G(v, \mathbf{Z})\,dv\right]$$

by properties of Gaussian variables. The expectation with respect to $\mathbf{Z}$ gives equation (9). □

(7) *If $\tilde{U}$ and $\tilde{V}$ are Gaussian processes, the marginal density of X is symmetric about zero so that its odd order moments are zero. The even order moments of X are*

$$E[X(t)^{2q}] = \frac{(2q)!}{2^q q!}E\left[\left(\int_0^\infty G(v, \mathbf{Z})\,dv\right)^q\right], \quad q \geqslant 1. \tag{10}$$

Proof. Since $X(t)|\mathbf{Z}$ is a Gaussian variable for each $t \geqslant 0$ with mean zero and variance $\int_0^\infty G(v, \mathbf{Z})\,dv$, its odd order moments are zero. Hence, the density of $X(t)$ must be symmetric about zero, consistent with equation (9) showing that the characteristic function of $X(t)$ is real-valued. The even order moments of $X(t)$ are

$$E[X(t)^{2q}|\mathbf{Z}] = \frac{(2q)!}{2^q q!}\left(\int_0^\infty G(v, \mathbf{Z})\,dv\right)^q, \quad q \geqslant 1. \tag{11}$$

by properties of Gaussian variables [1]. □

(8) *If $\tilde{U}$ and $\tilde{V}$ are Gaussian processes, X is strictly stationary.*

Proof. The conditional vector

$$(X(t_1), \ldots, X(t_n))|\mathbf{Z}$$

is Gaussian for any integer $n \geqslant 1$ and times $0 \leqslant t_1 < \cdots < t_n$ and has the same distribution as the vector

$$(X(t_1 + \tau), \ldots, X(t_n + \tau))|\mathbf{Z},$$

where τ denotes an arbitrary time shift. The stated property follows by eliminating the condition of $\mathbf{Z}$. □

The process X defined by equations (1)–(5) with $\tilde{U}$ and $\tilde{V}$ Gaussian processes defines the class of *conditional Gaussian process*. Processes in this class are used to model non-Gaussian

stationary processes defined by their second moment properties and marginal distribution and develop algorithms for generating samples of non-Gaussian processes.

3. CONDITIONAL GAUSSIAN PROCESSES

Two representations are considered for the members of the class of conditional Gaussian processes, processes with a random spectral density function and processes defined as the output of a linear filter with random coefficients subjected to a Gaussian input. These representations are consistent with the definition of the conditional Gaussian processes in the previous section and provide alternative implementations of the same concept.

3.1. RANDOM SPECTRAL DENSITY

Let

$$G(v, \mathbf{Z}) = \sum_{i=1}^{d} Z_i \varphi_i(v), \quad v \geqslant 0, \tag{12}$$

where φ_i are $[0, \infty]$-valued functions such that $\int_0^\infty \varphi_i(v)\,\mathrm{d}v < \infty$ and the random variables Z_i are positive so that the range, D, of $\mathbf{Z}$ in equation (1) must be a subset of $[0, \infty)^d$. Hence, $G(v, \mathbf{Z})$ is a positive random variable for all $v \geqslant 0$ and $\int_0^\infty G(v, \mathbf{Z})\,\mathrm{d}v < \infty$.

The function $G(v, \mathbf{Z}(\omega_1)) = \sum_{i=1}^{d} Z_i(\omega_1)\varphi_i(v)$ has the properties of the one-sided spectral density for each $\omega_1 \in \Omega_1$. For a fixed $\omega_1 \in \Omega_1$, define the conditional process $X \,|\, \mathbf{Z}(\omega_1)$ to be a stationary Gaussian process with mean zero and one-sided spectral density $G(v, \mathbf{Z}(\omega_1))$. This conditional process can be approximated by

$$X_n(t, \omega_1, \omega_2) = \sum_{k=1}^{n} \Gamma_k(\omega_1)\,[A_k(\omega_2)\cos(\nu_k t) + B_k(\omega_2)\sin(\nu_k t)], \tag{13}$$

where A_k, B_k are the independent standard Gaussian variables defined on the probability space $(\Omega_2, \mathscr{F}_2, P_2)$, that is, Gaussian variables with mean zero and variance one, ν_k denotes the central points of a partition, I_k, of the frequency axis, that is, $\cup_k I_k$ gives the frequency band of $G(v, \mathbf{Z})$ such that $I_k \cap I_l = \emptyset$ for $k \neq l$, and $\Gamma_k(\omega_1) = (\int_{I_k} G(v, \mathbf{Z}(\omega_1))\,\mathrm{d}v)^{1/2}$ are dependent random variables on $(\Omega_1, \mathscr{F}_1, P_1)$. For a fixed $\omega_1 \in \Omega_1$, X_n, that is, the conditional process $X_n \,|\, \mathbf{Z}(\omega_1)$, is Gaussian. However, X_n is not a Gaussian process.

3.2. LINEAR FILTER WITH RANDOM COEFFICIENTS

Let X be the stationary solution of the differential equation

$$\mathscr{L}[X(t)] = Y(t), \quad t \geqslant 0 \tag{14}$$

for some initial conditions, where

$$\mathscr{L} = \sum_{k=0}^{m} \xi_k(\mathbf{Z}) \frac{\mathrm{d}^{m-k}}{\mathrm{d}t^{m-k}} \tag{15}$$

is a differential operator of order m with random coefficients defined by the measurable mappings $\omega_1 \mapsto \xi_k(\mathbf{Z}(\omega_1))$, $k = 1, \ldots, m$. The input, Y, is a stationary Gaussian process with

mean zero and one-sided spectral density g_y. Conditions for the existence of a stationary solution of equation (14) can be found in references [4, 5].

Let $h(v, \mathbf{Z}(\omega_1))$ be the transfer function of $\mathscr{L}$ for a sample, $\mathbf{Z}(\omega_1)$, of Z. The corresponding one-sided spectral density of the stationary solution, $X(t) \mid \mathbf{Z}(\omega_1)$, is

$$G(v, \mathbf{Z}(\omega_1)) = |h(v, \mathbf{Z}(\omega_1))|^2 g_y(v), \tag{16}$$

provided that this solution exists. This process is Gaussian for a fixed ω_1. However, X is a non-Gaussian process if the condition on $\mathbf{Z}$ is eliminated.

3.3. MONTE CARLO SIMULATION

The generation of samples of X involves three steps:

(1) Generate a sample, $\mathbf{Z}(\omega_1)$, of $\mathbf{Z}$ based on the distribution of this random variable.
(2) Calculate the spectral density, $G(\cdot, \mathbf{Z}(\omega_1))$, corresponding to $\mathbf{Z}(\omega_1)$. The conditional process, $X \mid \mathbf{Z}(\omega_1)$, is Gaussian.
(3) Generate a sample path, $X(\cdot, \omega_1, \omega_2)$, of the conditional process $X \mid \mathbf{Z}(\omega_1)$. The generation of $X(\cdot, \omega_1, \omega_2)$, can be based on the spectral density, $G(v, \mathbf{Z}(\omega_1))$, given by equation (12) or equation (16) and the representation of X in equation (13). Current algorithms for generating sample paths of a stationary Gaussian process can be used to obtain $X(\cdot, \omega_1, \omega_2)$ [1]. If X is defined by equation (14), $X(\cdot, \omega_1, \omega_2)$ can also be obtained by integrating this equation for $\mathbf{Z}$ equal to its sample value, $\mathbf{Z}(\omega_1)$, and a sample path, $Y(\cdot, \omega_2)$, of Y. In this approach, the previous step is not needed.

Additional sample paths of X can be obtained by repeating the above steps as many times as needed. The ensemble of these samples has the required statistics.

As stated previously, the conditional process $X \mid \mathbf{Z}(\omega_1)$ is Gaussian but X is not a Gaussian process. The above Monte Carlo simulation algorithm also shows that X is not an ergodic process because its sample and ensemble statistics differ.

4. NUMERICAL EXAMPLES

Four examples are presented to demonstrate the calibration of conditional Gaussian models to target properties and the use of these models in Monte Carlo simulation. The examples include non-Gaussian random variable and stochastic process.

Example 1. Let

$$X = \sum_{i=1}^{d} Z_i Y_i \tag{17}$$

be a real-valued random variable, where Y_i are independent Gaussian variables with mean zero and variance one. The coefficients, Z_i, in the definition of X are independent random variables. There is no relationship between the collection of random variables Y_i and Z_i. Let $\mathbf{Z} = (Z_1, \ldots, Z_d)$. The conditional variable $X \mid \mathbf{Z}$ follows a Gaussian distribution with the first four moments

$$E[X \mid \mathbf{Z}] = \sum_{i=1}^{d} Z_i E[Y_i] = 0,$$

$$E[X^2 \mid \mathbf{Z}] = \sum_{i,j=1}^{d} Z_i Z_j E[Y_i Y_j] = \sum_{i=1}^{d} Z_i^2,$$

$$E[X^3 \mid \mathbf{Z}] = \sum_{i,j,k=1}^{d} Z_i Z_j Z_k E[Y_i Y_j Y_k] = 0,$$

$$E[X^4 \mid \mathbf{Z}] = \sum_{i,j,k,l=1}^{d} Z_i Z_j Z_k Z_l E[Y_i Y_j Y_k Y_l]$$

$$= \sum_{i,j,k,l=1}^{d} Z_i Z_j Z_k Z_l (\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) = 3 \sum_{i,j=1}^{d} Z_i^2 Z_j^2. \tag{18}$$

The fourth order moment could have been written directly because $X \mid \mathbf{Z}$ is a Gaussian random variable with mean zero and variance $E[X^2 \mid \mathbf{Z}] = \sum_{i=1}^{d} Z_i^2$, so that $E[X^4 \mid \mathbf{Z}] = 3(E[X^2 \mid \mathbf{Z}])^2$.

The first four moments of X,

$$E[X] = E[E[X \mid \mathbf{Z}]\} = 0, \qquad E[X^2] = E\{E[X^2 \mid \mathbf{Z}]\} = \sum_{i=1}^{d} E[Z_i^2],$$

$$E[X^3] = E\{E[X^3 \mid \mathbf{Z}]\} = 0,$$

$$E[X^4] = E\{E[X^4 \mid \mathbf{Z}]\} = 3 \sum_{i,j=1}^{d} E[Z_i^2 Z_j^2] = 3\left(\sum_{i=1}^{d} E[Z_i^4] + \sum_{i,j=1; i \neq j}^{d} E[Z_i^2]E[Z_j^2] \right) \tag{19}$$

result from equation (18) by eliminating the condition on $\mathbf{Z}$. The kurtosis coefficient,

$$\gamma_{4,x} = \frac{E[X^4]}{(E[X^2])^2} = 3\left(1 + \frac{\sum_{i=1}^{d} (E[Z_i^4] - (E[Z_i^2])^2)}{\sum_{i,j=1}^{d} E[Z_i^2]E[Z_j^2]} \right) \tag{20}$$

of X is larger than three because $(E[Z_i^2])^2 \leqslant E[Z_i^4]$ by the Cauchy–Schwarz inequality. Hence, the model can generate only random variables with symmetric densities about zero and heavier tails than the Gaussian distribution.

Take, for example, Z_i to be independent random variables following an exponential distribution with parameter $\lambda > 0$. The moments of Z_1 are $E[Z_1^q] = \Gamma(q+1)/\lambda^q$, $q = 1, 2, \ldots$, [6] so that $E[X^2] = 2d/\lambda^2$, $E[X^4] = 12d(5+d)/\lambda^4$, and

$$\gamma_{4,x} = \frac{E[X^4]}{(E[X^2])^2} = 3(1 + 5/d). \tag{21}$$

The kurtosis coefficient decreases with d from 18 for $d = 1$ to 3 for $d \to \infty$. The limit, $\lim_{d \to \infty} \gamma_{4,x} = 3$, is consistent with the central limit theorem. Figure 1 shows a histogram of a scaled version of X with mean zero and variance one for $d = 10$, $\lambda = 1$, and the density of the standard Gaussian variable. The histogram, based on 10 000 samples of X, has heavier tails than the Gaussian density. The estimated and exact kurtosis coefficients of X are 4·7841 and 4·5, respectively, while this coefficient is 3 for a Gaussian variable. The generation of samples of X is based on the Monte Carlo algorithm outlined in the previous section. First, a sample, $\mathbf{Z}(\omega_1)$, of $\mathbf{Z}$ has been generated. Second, the expression of $X \mid \mathbf{Z}(\omega_1)$ has been determined. Third, a sample, $X(\omega_1, \omega_2)$, of X has been obtained from the expression of $X \mid \mathbf{Z}(\omega_1)$ and a sample, $(Y_1(\omega_2), \ldots, Y_d(\omega_2))$ of $(Y_1, \ldots, Y_d)$. These steps have been repeated 10 000 times to generate as many as 10 000 samples of X.

If Z_i's are independent gamma random variables with parameters (k, λ) and moments $E[Z_1^q] = \Gamma(k+q)/(\Gamma(k)\lambda^q)$, $q = 1, 2, \ldots$, then X has variance $\Gamma(k+2)d/(\Gamma(k)\lambda^2)$ and kurtosis

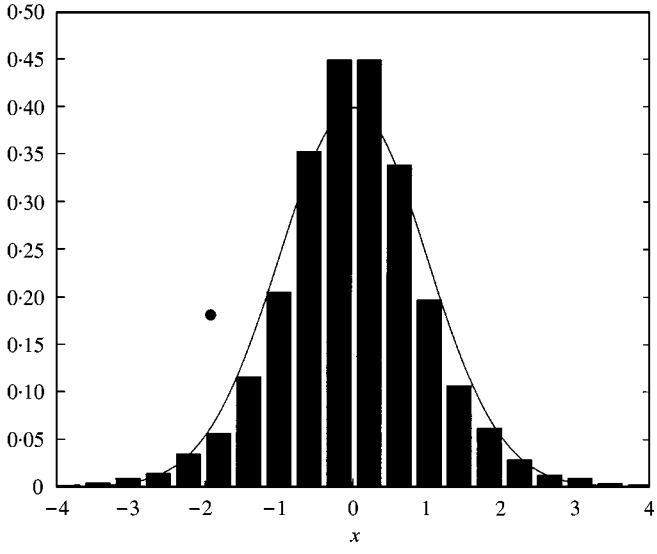

Figure 1. Histogram of X and the density of the standard Gaussian variable.

coefficient

$$\gamma_{4,x} = 3\left(1 + \frac{\Gamma(k)\Gamma(k+4) - \Gamma(k+2)^2}{d\Gamma(k+2)^2}\right). \tag{22}$$

As $d \to \infty$, $\gamma_{4,x}$ converges to 3 for each k indicating that X may approach a Gaussian random variable.

Suppose that the model of equation (17) needs to be calibrated to a random variable with mean $\mu = 10$, variance $\sigma^2 = 5$, and kurtosis coefficient $\gamma_4 = 6$ following a symmetric density about μ. The target mean can be matched by taking $X = \mu + \sum_{i=1}^{d} Z_i Y_i$ rather than its definition in equation (17). If Z_i are taken to be exponential variables, equation (21) implies $d = 5$ so that $\lambda = \sqrt{2}$ results from equation (19) and the value of the target variance. If the target kurtosis coefficient is assumed to be $\gamma_4 = 9$, equation (21) gives $d = 2{\cdot}5$ so that there is no solution because d must be a positive integer. Two options are possible to overcome this difficulty: assume correlation among the random variables Z_i or continue to assume independence among Z_i's but consider other distributions for these variables. The second option is illustrated.

If Z_i's are independent gamma variables with parameters (k, λ), three parameters are available to calibrate the target moments so that one of these parameters can be selected arbitrarily, for example, take $d = 10$. The condition that the target kurtosis be matched, that is, $\gamma_{4,x}$ given by equation (22) be equal to γ_4, gives $k = 0{\cdot}53$ and $0{\cdot}278$ for $\gamma_4 = 6$ and 9 respectively. The requirement that the variance of X coincide with the target variance provides the values of the other parameter of the gamma distribution, $\lambda = 0{\cdot}6367$ for $\gamma_4 = 6$ and $\lambda = 0{\cdot}4215$ for $\gamma_4 = 9$. Similar calculations give $k = 1$ and $\lambda = \sqrt{2}$ for $d = 5$ and $\gamma_4 = 6$, consistent with a previous result for exponential Z_i's.

Example 2. Suppose that the target stationary process is a band-limited white noise with mean zero, one-sided spectral density,

$$g(\nu) = g_0 \, 1_{[0,\bar{\nu})}(\nu), \quad 0 \leqslant \nu < \bar{\nu} \tag{23}$$

and a marginal density symmetric about zero with kurtosis $\gamma_4 = 5$, where $1_A(\xi) = 1$ for $\xi \in A$ and $1_A(\xi) = 0$ for $\xi \notin A$. The objective is to calibrate a conditional non-Gaussian model to this target process.

Consider the random spectral density function

$$G(\nu, \mathbf{Z}) = \sum_{i=1}^{d} Z_i 1_{[(i-1)\Delta\nu, i\Delta\nu)}(\nu), \tag{24}$$

where $d \geqslant 1$ is an integer, $\Delta\nu = \bar{\nu}/d$ is the size of a uniform partition of the frequency band of the target process, and the co-ordinates, Z_i, of the $\mathbb{R}^d$-valued random variable $\mathbf{Z}$ are such that $E[Z_i] = g_0$. The conditional Gaussian process, X, given by equation (13) with the random spectral density in equation (24) matches the second moment properties of the target process. The fourth moment of X is given by equation (10) with $q = 2$. To match the target fourth moment, it is necessary to specify higher order moments of $\mathbf{Z}$ in addition to its mean.

Suppose, for example, the dimension, d, of $\mathbf{Z}$ coincides with the number of harmonics, n, in the approximate representation, X_n, of X [equation (13)] and that

$$Z_i = pW + W_i, \quad i = 1, \ldots, d = n,$$

where W and W_i are independent exponential random variables with mean $1/\lambda$ and p is a real number that needs to be determined. The condition $E[Z_i] = g_0$ implied $p/\lambda + 1/\lambda = g_0$ or $\lambda = (p+1)/g_0$. The conditional second and fourth moments of X_n are

$$E[X_n(t)^2 \mid \mathbf{Z}] = \Delta\nu \sum_{i=1}^{d} Z_i, \qquad E[X_n(t)^4 \mid \mathbf{Z}] = (\Delta\nu)^2 \sum_{i,j=1}^{d} Z_i Z_j,$$

so that

$$E[X_n(t)^2] = \Delta\nu \sum_{i=1}^{d} g_0 = g_0\bar{\nu} = (p+1)\bar{\nu}/\lambda, \qquad E[X_n(t)^4] = (\Delta\nu)^2 \sum_{i,j=1}^{d} E[Z_i Z_j],$$

where

$$E[Z_i Z_j] = \begin{cases} E[(pW + W_i)^2] = (2p^2 + 2p + 2)/\lambda^2 & \text{if } i = j, \\ E[(pW + W_i)(pW + W_j)] = (2p^2 + 2p + 1)/\lambda^2 & \text{if } i \neq j \end{cases}$$

for the selected model of $\mathbf{Z}$. These expectations show that the correlation coefficient

$$\rho = \frac{E[(Z_1 - E[Z_1])(Z_2 - E[Z_2])]}{E[(Z_1 - E[Z_1])^2]} = \frac{p^2}{p^2 + 1}$$

of the equally correlated random variables Z_i is zero for $p = 0$ and approaches one as $p \to \infty$. The fourth moment of $X_n(t)$ is

$$E[X_n(t)^4] = (\Delta\nu)^2 \left[\sum_{i=1}^{d} (2p^2 + 2p + 2)/\lambda^2 + \sum_{i,j=1;\, i \neq j}^{d} (2p^2 + 2p + 1)/\lambda^2 \right]$$

$$= \frac{(\Delta\nu)^2}{\lambda^2} (2p^2 n^2 + 2pn^2 + n^2 - n) = \left(\frac{\bar{\nu}}{\lambda}\right)^2 (2p^2 + 2p + 1 - 1/n),$$

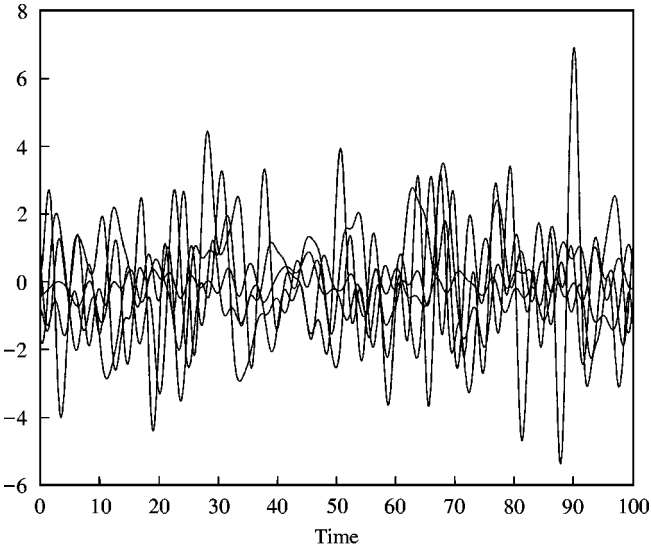

Figure 2. Five sample paths of X_n.

so that

$$\gamma_{4,X_n} = 3\frac{E[X_n(t)^4]}{(E[X_n(t)^2])^2} = 3\,\frac{2p^2 + 2p + 1 - 1/n}{(p+1)^2} \tag{25}$$

gives the kurtosis coefficient of $X_n(t)$. The range of this coefficient is $(3, 6)$.

The calibration of the conditional process X_n involves two steps. First, the value of p can be determined from the condition that γ_{4,X_n} is equal to the target kurtosis, $\gamma_4 = 5$. This condition gives $p = 4{\cdot}5$ for a relatively large value of n. Second, $E[Z_i] = g_0$ implies $(p+1)/\lambda = g_0$ or $\lambda = (p+1)/g_0$. Figure 2 shows five sample paths of X_n for $n = 200$ calibrated to a target process with spectral density function given by equation (23) for $\bar{\nu} = 2$ and $\gamma_4 = 5$.

Example 3. Suppose that

$$g(\nu) = \frac{g_0}{(\nu^2 - \nu_0^2)^2 + (2\nu_0\zeta\nu)^2}, \quad \nu \geqslant 0 \tag{26}$$

is the target one-sided spectral density function, where $\nu_0 > 0$, $g_0 > 0$, and $\zeta \in (0, 1)$. The approach of the previous example can be extended simply to account for the dependence of the spectral values on frequency.

An alternative approach is considered here. It is based on an approximation

$$\tilde{g}(\nu) = \sum_{i=1}^{d} a_i 1_{[0,\infty)}(\nu)\phi\left(\frac{\nu - \nu_i}{\sigma_i}\right), \quad \nu \geqslant 0 \tag{27}$$

of g, where $\phi(\xi) = \exp(-0{\cdot}5\xi^2)/\sqrt{2\pi}$ provides a collection of base functions, a_i are the yet undetermined constants, and σ_i, $\nu_i > 0$ are the specified constants. The unspecified constants a_i can be determined from the condition that the error

$$\mathrm{e} = \int_0^\infty (g(\nu) - \tilde{g}(\nu))^2\,\mathrm{d}\nu \tag{28}$$

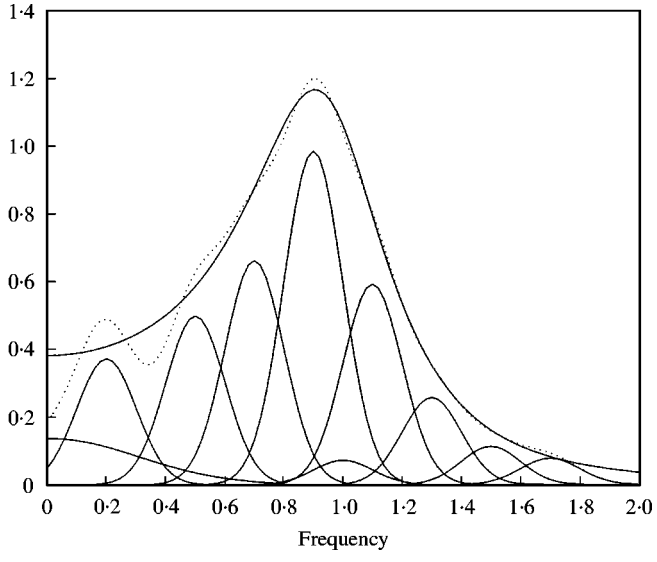

Figure 3. Target and approximate spectral density functions.

be minimized. The condition gives the system of linear equations

$$\sum_{i=1}^{d}\left[\int_0^\infty \phi\left(\frac{v-v_i}{\sigma_i}\right)\phi\left(\frac{v-v_j}{\sigma_j}\right)\right]a_i = \int_0^\infty g(v)\phi\left(\frac{v-v_j}{\sigma_j}\right), \quad j=1,\ldots,d \tag{29}$$

for the unknown constants a_i.

The top solid line in Figure 3 is the target spectral density $g(v)$ for $\zeta = 0{\cdot}3$, $v_0 = 1$, and $g_0 = 4\zeta v_0^3/\pi$ in the range $[0, 2]$. The other solid lines in the figure are the functions $a_i 1_{[0,\infty)}(v)\phi((v-v_i)/\sigma_i)$ for $d = 10$, frequencies, v_i, equal to 0·01, 0·2, 0·5, 0·7, 0·9, 1, 1·1, 1·3, 1·5, and 1·7 for $i = 1, \ldots, d$, $\sigma_1 = 0{\cdot}3$, and $\sigma_i = 0{\cdot}1$ for $i \geqslant 2$. The coefficients, a_i are 0·3434, 0·9321, 1·2485, 1·6578, 2·4667, 0·1830, 1·4841, 0·6433, 0·2832, and 0·1957 for $i = 1, \ldots, d$. The figure also shows with a dotted line the approximate spectral density, $\tilde{g}$, in equation (27). More accurate approximations can be obtained by refining the representation of the target spectral density given by equation (27).

Consider the random spectral density

$$\tilde{G}(v, \mathbf{Z}) = \sum_{i=1}^{d} Z_i 1_{[0,\infty)}(v)\phi\left(\frac{v-v_i}{\sigma_i}\right), \quad v \geqslant 0 \tag{30}$$

depending on an $\mathbb{R}^d$-valued random variable $\mathbf{Z} = (Z_1, \ldots, Z_d)$ such that $E[Z_i] = a_i$, where a_i are coefficients in the approximate representation of the target spectral density [equation (27)]. The process defined by equations (3) and (5) with the random spectral density, G, given by equation (30) matches the target second moment properties. The second order and higher order moments of $\mathbf{Z}$ need to be selected such that the marginal distribution of X or at least some of the higher order moments of X approximate satisfactorily the corresponding target properties.

Example 4. Let X be the output of equation (14) with

$$\mathscr{L} = \frac{\mathrm{d}^2}{\mathrm{d}t^2} + 2Z_1 Z_2 \frac{\mathrm{d}}{\mathrm{d}t} + Z_2^2 \tag{31}$$

given by equation (15) for $m = 3$, $\mathbf{Z} = (Z_1, Z_2)$, $Z_1 \in (0, 1)$, $Z_2 > 0$, $\xi_0(\mathbf{Z}) = 1$, $\xi_1(\mathbf{Z}) = 2Z_1Z_2$, $\xi_2(\mathbf{Z}) = Z_2^2$, and Y a Gaussian white noise with mean zero and a one-sided spectral density of intensity $g_0 = 1/\pi_0$. The output, $X(t)|\mathbf{Z}$, is a stationary Gaussian process with mean zero, one-sided spectral density

$$G(v, \mathbf{Z}) = \frac{g_0}{(v^2 - Z_2^2)^2 + (2Z_1Z_2v)^2}, \quad v \geqslant 0 \tag{32}$$

and covariance function

$$E[X(t)X(t+\tau)|\mathbf{Z}] = \frac{\pi g_0}{2Z_1Z_2^3} e^{-Z_1Z_2\tau}\left[\cos(W\tau) + \frac{Z_1Z_2}{W}\sin(W\tau)\right], \tag{33}$$

where $W = Z_2\sqrt{1 - Z_1^2}$ and $\tau \geqslant 0$ [5]. The marginal distribution of X is given by

$$F_{X(t)}(x) = \int_0^\infty dz_1 \int_0^1 dz_2\, \Phi(x\sqrt{4z_1z_2}), \tag{34}$$

where $\Phi(\xi) = \int_{-\infty}^{\xi} \phi(u)\, du$.

The selection of the properties of $\mathbf{Z}$ such that X has specified second moment properties and marginal distribution is less simple than in the previous case. For example, suppose that the target one-sided spectral density is given by equation (26). The probability law of $\mathbf{Z}$ needs to be determined such that the condition $E[G(v, \mathbf{Z})] = g(v)$ is satisfied for each $v \geqslant 0$. Setting $E[Z_1] = \zeta$ and $E[Z_2] = v_0$ is not a solution because $G(v, \mathbf{Z})$ is a non-linear function of $\mathbf{Z}$. Iteration is needed to find moments of $\mathbf{Z}$ such that X has the desired second moment properties.

Simple results are possible if X is slightly non-Gaussian so that the random coefficients of the filter can be represented by $\mathbf{Z} = \boldsymbol{\mu} + \varepsilon\mathbf{R}$, where $E[\mathbf{Z}] = \boldsymbol{\mu}$, ε denotes a small parameter, and $\mathbf{R}$ is the random vector with mean zero. The first order approximation of $G(v, \mathbf{Z})$ is given by

$$G(v, \mathbf{Z}) \simeq G(v, \boldsymbol{\mu}) + \varepsilon \sum_{i=1}^{d} \theta_i(v)\, R_i, \tag{35}$$

where $\theta_i(v)$ is the partial derivative $\partial G(v, \mathbf{Z})/\partial Z_i$ evaluated at $\mathbf{Z} = \boldsymbol{\mu}$. If the mean value, $\boldsymbol{\mu}$, of $\mathbf{Z}$ is equal to the corresponding parameters in the target spectral density, that is, $E[Z_1] = \zeta$ and $E[Z_2] = v_0$, the approximate expectation of $G(v, \mathbf{Z})$ coincides with the target spectral density. A similar approximation can be developed for the marginal density of X. The approximation of equation (35) provides an alternative to equation (13). For example, the approximation of this equation to the order ε is

$$\begin{aligned} X_n(t) \simeq & \sum_{k=1}^{n} \sqrt{\alpha_k \Delta v}\,[A_k \cos(v_k t) + B_k \sin(v_k t)] \\ & + \frac{\varepsilon}{2} \sum_{k=1}^{n} \sqrt{\Delta v/\alpha_k}\left(\sum_{i=1}^{d} \theta_i(v_k, \boldsymbol{\mu})\, R_i\right)[A_k \cos(v_k t) + B_k \sin(v_k t)], \end{aligned} \tag{36}$$

where $\alpha_k = G(v_k, \boldsymbol{\mu})$. The approximation is based on the observation that the function $(a + \varepsilon b)^{1/2}$ can be approximated by $a^{1/2} + 0{\cdot}5\varepsilon b/a^{1/2}$ to the order ε. The term of order one in equation (36) is a Gaussian process with the target second moment properties and the term of order ε is a non-Gaussian process representing a correlation term.

5. COMMENTS

A new model has been proposed to represent non-Gaussian stationary processes and develop Monte Carlo simulation algorithms for generating sample paths of non-Gaussian processes. The model is based on a class of non-Gaussian processes, the class of conditional Gaussian processes. Two representations are considered for these processes. The first representation is based on a randomized version of the classical spectral density function. The second representation uses the output of a linear filter with random coefficients subjected to Gaussian noise to define conditional Gaussian processes. The proposed model, its representations, and corresponding Monte Carlo simulation algorithms have been illustrated by examples involving non-Gaussian random variables and processes. It was shown that the proposed model can match any second moment properties but, generally, can only fit approximately a specified marginal distribution.

REFERENCES

1. M. GRIGORIU 1995 *Applied Non-Gaussian Processes*: *Examples*, *Theory*, *Simulation*, *Linear Random Vibration*, *and MATLAB Solutions*. Englewood Cliffs, NJ: Prentice-Hall.
2. G. Q. CAI and Y. K. LIN 1995 *Physical Review E* **54**, 299–303. Generation of non-Gaussian stationary stochastic processes.
3. G. SAMORODNITSKY and M. S. TAQQU 1994 *Stable Non-Gaussian Random Processes. Stochastic Models with Infinite Variance*. New York: Birkhäuser.
4. J. L. MELSA and A. P. SAGE 1973 *An Introduction to Probability and Stochastic Processes*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
5. T. T. SOONG and M. GRIGORIU 1993 *Random Vibration of Mechanical and Structural Systems*. Englewood Cliffs, NJ: Prentice-Hall.
6. M. GRIGORIU 1980 *Journal of the Engineering Mechanics Division* **106**, 1423–1429. Table of dimensionless central moments.