Constructions of Particular Random Processes

GLENN ERIC JOHNSON, MEMBER, IEEE

This paper reviews how to construct sets of random numbers with particular amplitude distributions and correlations among values. These constructions support both high-fidelity Monte Carlo simulation and analytic design studies. A variety of constructions is presented to free engineering models from "white or normal" limitations embodied in many current simulations. The methods support constructions of conventional stationary and normally distributed processes, nonstationary, nonnormal signal and interference waveforms, nonhomogeneous random scenes, nonhomogeneous volumetric clutter realizations, and snapshots of randomly evolving, nonhomogeneous scenes. Each case will have specified amplitude statistics, e.g., normal, log-normal, uniform, Weibull, or discrete amplitude statistics; and selected correlation, e.g., white, pink, or patchy statistics, clouds, or speckles. Sets of random numbers with correlation, nonstationarities, various tails for the amplitude distributions, and multimodal distributions can be constructed. The paper is essentially tutorial, emphasizing aspects of probability theory necessary to engineering modeling, and collecting example constructions.

I. Introduction

This paper develops constructions of general random processes. Since the measurement corruptions that set performance limits of detection, classification, and control systems are random processes, the constructions support faithful system performance evaluations. This paper shows how to construct sets of random numbers with particular amplitude distributions and correlations among values. Constructions of these random numbers support both high-fidelity Monte Carlo simulation, by devising realizations of the random numbers, and analytic design studies, by displaying example joint probability distributions. A variety of constructions is presented to free engineering models from "white or normal" limitations embodied in many current simulations. This paper is essentially tutorial, emphasizing aspects of probability theory necessary to engineering modeling, and collecting example constructions.

The suggested constructions provide joint probability distributions suitable for the study of systems driven by signals with any statistics, including nonwhite and nonnormal statistics. Using standard methods [1]–[5], together with the constructions, the effects of noise statistics and correlation on optimal detectors, demodulators, state estimators, and pattern classifiers can be evaluated for a wide variety of signal corruption. For example, the resulting joint probability distributions are required to evaluate the performance of locally optimum detectors operating in nonwhite, nonnormal, additive noise [6], [7].

The variety of processes that can be generated include conventional stationary and normally distributed processes in addition to nonstationary, nonnormal signal and interference waveforms, nonhomogeneous random scenes, nonhomogeneous volumetric clutter realizations, and snapshots of randomly evolving, nonhomogeneous scenes. Each case will have specified amplitude statistics, e.g., normal, log-normal, uniform, Weibull, or discrete amplitude statistics; and selected correlation, e.g., white, pink, or patchy statistics, clouds, or speckles. Although the methods reviewed in this paper have been developed for a considerable period ([7], [8], and earlier references therein), the more widely disseminated procedures for random number constructions [9]–[19] emphasize either normal random numbers or independent random numbers. And, although the methods for nonnormal and correlated random numbers are available, these more general methods are not commonly represented in scientific and engineering simulations. This paper motivates and describes techniques to construct more general sets of random numbers, random numbers with both particular correlations and first statistics.

The advantages of the techniques reviewed in this paper with respect to techniques commonly used in engineering simulations are threefold. The rigorously motivated random number generation techniques achieve:

—more accurate, reliable results from Monte Carlo simulation
—computationally efficient Monte Carlo simulation
—adequate simulation tests of processing algorithms

Techniques such as low-pass-filtering-independent random draws to achieve correlated random draws are often applied *ad hoc* in spite of the availability of discrete Karhunen–Loève expansions which precisely achieve intended correlations.

The conventional definition for random process refers to time evolution of a set of random features. However, since I am concerned only with sampled systems, I need not consider the potentially uncountable numbers of random draws required in the continuous time evolution of a system. Consequently, I discuss generation of random numbers

Manuscript received December 14, 1992; revised October 21, 1993.
The author is with The Analytic Sciences Corporation, Reston, VA 22090.
IEEE Log Number 9214860.

rather than random processes. Sampled time observations of a single random feature are considered distinct random numbers, and consequently, the methods for constructing finite sets of random numbers suffices for the sampled time evolution of random features. Nevertheless, one useful holdover from the theory of random processes is the concept of stationarity, and periodic stationarity [20]. These are symmetries of the first marginal distributions and correlations of pairs of features. In the present case of discretely indexed random features, these symmetries are manifest as equalities among the means for each feature, and matrix symmetries (e.g., Toeplitz or block Toeplitz) of covariance matrices.

Except for degenerate cases, the generation of three random features requires three random draws, three rolls of the dice. Mean values are obtained by rerolling the dice, in sets of three rolls, to achieve many sets of values for the three features, and accumulating average values. These average values are the ensemble average, the average over sets of random draws. One trial requires one generation of all three features.

Generally, two attributes of sets of random features are of greatest concern: the probability distributions for each random feature and the correlations of features. The first marginal probability distribution fixes the histogram of values for each random feature. Each feature achieves a particular value with a relative frequency set by this first marginal probability distribution. Interrelationships of feature values are indicated by correlations. Together with the first marginal distributions, the correlations identify the random features as samples of a stationary process with any particular power spectral density, or samples of a periodic stationary process, or as a more generally correlated set of random features. "Nonwhite" designates that the values for the random features are correlated. "Nonnormal" random draws include random features with non-Gaussian first marginal distributions in addition to sets of random features with Gaussian first marginal distributions that fail to be "jointly" normal. These two attributes, the first marginal distributions of each feature, and the correlations between features, are independent properties of a set of random features. The methods in Section III display constructions with independent specifications for the first marginal distributions and the correlations.

Monte Carlo digital simulations are often simply transformations from independent random draws into random draws with desired interrelationships and statistics. In a Monte Carlo simulation, a pseudorandom number generator is used in place of rolls of a die to achieve random numbers that are subsequently used to calculate sample values for the observable features of the stochastically described system.

Monte Carlo simulations can be greatly accelerated by directly generating required random results without repetitive front-end calculations. An example illustrates this result. Assert it is our job to evaluate the ability of a processor to recognize symbols transmitted over a channel corrupted by additive, mean zero, normally distributed noise. The processor evaluates linear combinations of received samples, and selects the symbol associated with the filter for which the largest output value is attained. A conventional simulation for this system might consist of generating many sets of corrupted symbol samples. Denote the number of samples processed for each symbol recognition as M. In the straightforward simulation, we would filter the same M samples in each one of N_s filters, and subsequently compare the output values to select the symbol represented by the corrupted samples. Accumulating a count of symbol errors over trials provides the estimate of symbol error rate. With $k = 1, \cdots N_s$, these output quantities can be represented

$$v_k = w_k^T r$$

where

r $= As_j + n$

A signal amplitude

s_j signal samples

n additive noise samples

R $= E[nn^T]$ noise covariance

w_k kth filter coefficients.

A more computationally efficient simulation is obtained by generating only the random outputs of the filters, v_k. We can calculate that the filter outputs, v_k, are normally distributed with means $Aw_k^T s_\ell$ and covariances $w_k^T R w_j$. We can just as well directly construct the correlated, normally distributed random numbers, v_k, as go through the procedure of filtering a set of randomly generated waveform samples. Statistically, there will be no difference between the numbers generated by the two procedures.

With M signal plus noise samples and N_s matched filters, generating the additive noise n requires M random draws, order M^2 operations for the linear combinations (Karhunen–Loève expansion) that generates correlated noise, and N_s filterings of length M of the M samples. Thus there are order $(N_s + M + 1)M$ operations per trial required in the conventional simulation: N_s filters requiring order M operations each to evaluate the output, plus the operations required to generate the corrupted symbol samples. Generating only the ouputs of the filters requires N_s random draws, and order N_s^2 operations for the Karhunen–Loève expansion that achieves the correct correlations of the outputs v_k. This requires only order $(N_s+1)N_s$ operations per trial to generate the correctly correlated filtered output values. This represents a substantial computational savings for a typical case of $N_s \ll M$. For example, we could be identifying one of 10 patterns in 512×512 pixel images. For this case, the number of computations is reduced by a factor of about 2^{18} for white noise and 2^{36} in generally correlated noise.

Results from the two simulation architectures are identical, yet the second approach provides significant computational savings. The result is that, within a simulation, correlations and first statistics of key normally distributed

random quantities need be calculated only once, and only for "output" features. Random realizations are subsequently constructed to these specifications. This efficiency can be combined with other Monte Carlo simulation enhancement techniques, such as importance sampling [21]. Similar gains are achieved in nonnormally distributed cases if a caveat is observed. Joint probability distributions are not uniquely defined by first marginal distributions and correlations, so only representative and not unique results are obtained. If the input statistics are characterized only to the level of first marginal distributions and correlations, then the efficient modeling approach can be justified.

The constructions also provide a practical approach to signal processing algorithm validation that improves upon much of the state-of-the-art for Monte Carlo simulation. A test suite with a variety of precisely targeted first statistics and correlations can be developed using the constructions. To validate operation of "robust" algorithms, a more substantial variation of trial statistics than that afforded by sampling normal or white processes is required. Suites of canonical noise environments with a wide variety of rigorously controlled statistics can be developed using the suggested constructions. Varying correlation, nonstationarities, varying tails of amplitude distributions, and multimodal distributions can all be included in test suites.

Despite the utility of the constructions, we cannot expect to sample "all" joint probability distributions for sets of random features using the constructions. Many distinct joint probability distributions share first marginal distributions and correlations. The constructions produce example realizations of joint probability distributions.

The first section and the appendices to this paper review well-established results from probability theory, to define notation, to motivate the constructions, and to provide a tutorial on random numbers as applied to engineering simulations. The constructions are presented in the second section. These constructions fulfill any requirement to generate particular, nonnormally distributed, correlated sets of random features. The final section displays the joint probability density functions for the constructed random numbers.

The presentation seeks to emphasize the general utility of the models, collect useful results, provide examples for the main results, and develop the origins and limitations of the models. Rather than pursuing theoretical questions for generalized random processes, the discussion is directed toward discovery of obliging methods for finite sets of random features. These methods suffice for all multiple feature sampled systems, and hence, the suggested constructions are sufficient for all digital simulations. The description of finite sets of random numbers is more pointed than the mathematics of generalized random processes, yet the development is sufficiently powerful to solve nearly all problems of engineering interest. Subtle and difficult issues at the foundations of the theory of generalized random processes are simply excess baggage for sampled time Monte Carlo simulation.

II. Background on Random Number Construction

Two formal results greatly assist understanding the structure and limitations of random number constructions. These key results are the Central Limit Theorem (CLT), and the Problem of Moments (POM), briefly summarized in the Appendices. Since we will be summing independent random draws to achieve correlated random draws, we will run afoul of the CLT since it indicates that sums of independent random draws tend to become normally distributed. Our goal is to generate correlated nonnormally distributed draws. We must understand the convergence to normal indicated by the CLT in order to avoid it. The POM affects specification of random numbers: in general, statistical moments are an inadequate description of distributions of random numbers, and joint probability distributions must be specified. This result is not limited to a negligible class of pathological distributions. One result of the POM is that it is hopeless to accurately characterize the tails of a distribution by estimating moments.

The constructions evolve from joining two tasks: construction of arbitrarily distributed random draws from uniformly distributed random draws, and construction of correlated random draws from linear combinations of independent random draws.

A. Generating Independent Random Draws

All the constructions that follow require the ability to generate independent pseudorandom draws. The most widely disseminated approach to construction of independently distributed random draws is the linear congruential algorithm supplemented with a shuffle to mitigate correlations [9]. The linear congruential algorithm produces uniformly distributed random draws.

Given pseudorandom[1] numbers resulting from the linear congruential algorithm, the construction of independent random draws with arbitrary probability distributions becomes a problem of transforming independent, uniformly distributed random draws into random draws distributed according to the desired probability distribution. Algorithms for the generation of independent random draws with any desired distribution are well developed and there is an extensive literature [9]–[17] on the subject.

An arbitrarily distributed x is constructed as a transformation of independent random draws, u_k

$$x = f(u_1, u_2, \cdots u_n). \tag{1}$$

For example, the inverse probability method consists of transforming one uniformly distributed random variate, U,

[1] Sequences generated algorithmically that imitate random sequences are denoted pseudorandom. If the sequences are more purposefully tailored to be representative of random draws, they are denoted quasi-random. Whether a sequence is truly a random selection is a philosophical rather than a practical question. A sequence of all zeros is just as likely a sample of independent, uniformly distributed random numbers as any other sequence. However, although equally likely, it is not a representative sample. The statistics derived from this sample are far from the statistics derived from the majority of similarly described random sequences, and consequently, all zeroes is not a suitable sequence for use in a set of Monte Carlo simulation runs.

Table 1 Probability Distribution Transformations

Density Function	Transformation from $U(0, 1)$	Transformation from $N(0, 1)$
Uniform: 1	—	$e^{-\frac{1}{2}(N_1^2+N_2^2)}$
Normal: $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$	$\sigma\sqrt{-2 \ln U_1} \cos(2\pi U_2)$	—
Exponential: $\frac{1}{\bar{x}} e^{-\frac{x}{\bar{x}}}$	$-\bar{x} \ln U_1$	$\frac{\bar{x}}{2}(N_1^2 + N_2^2)$
Log Normal: $\frac{1}{\sqrt{2\pi}ax} e^{-\frac{1}{2} \ln^2 (x/x_m)/a^2}$	—	$x_m e^{aN_1}$
Rayleigh (Swerling I, II): $\frac{2x}{E[x^2]} e^{-2x^2/E[x^2]}$	$\sqrt{-E[x^2] \ln U_1}$	—
Gamma 2 (Swerling III, IV): $\frac{4x^2}{\bar{x}^2} e^{-\frac{2x}{\bar{x}}}$	$-r\frac{\bar{x}}{2}(\ln U_1 + \ln U_2)$	—
Sinusoidal: $\frac{1}{\pi\sqrt{A^2-x^2}}$	$A \sin(\pi U_1)$	—
Linear Envelope Ricean: $xe^{-\frac{1}{2}(x^2+s^2)} I_0(sx)$	—	$\sqrt{N_1^2 + (N_2+s)^2}$
Square-law Envelope Ricean: $e^{-(x+\frac{1}{2}s^2)} I_0(s\sqrt{2x})$	—	$\frac{1}{2}(N_1^2 + (N_2+s)^2)$
Weibull: $\beta\alpha(\alpha x)^{\beta-1} e^{-(\alpha x)^\beta}$	$\frac{1}{\alpha}(-\ln U_1)^{\frac{1}{\beta}}$	—
Chi-square, N dof: $\frac{1}{2^{N/2}\Gamma(N/2)} x^{N/2-1} e^{-x/2}$	—	$\sum_{k=1}^{N} N_k^2$
Gamma N: $\left(\frac{Nx}{\bar{x}}\right)^N \frac{1}{x\Gamma(N)} e^{-Nx/\bar{x}}$	$-\frac{\bar{x}}{N}\sum_{k=1}^{N} \ln U_k$	—
Largest of N Exponentials: $\frac{N}{\bar{x}} e^{-x/\bar{x}}(1-e^{x/\bar{x}})^{N-1}$	$-\bar{x} \ln(1 - U_1^{1/N})$	—
Cauchy: $\frac{1}{\pi w(1+(x/w)^2)}$	$w \tan(\pi(U_1 - \frac{1}{2}))$	—

into a random variate with distribution, $P(x)$

$$x = P^{-1}(U).$$

Transformations for several distributions commonly used in engineering analysis are collected in Table 1 and [17]. In Table 1, U_k denote independent, uniformly distributed random draws on (0, 1), and N_k denote independent, unity variance, mean zero, normal random draws. The transformations in Table 1 are used to construct nonnormally distributed random numbers from independent, uniformly or normally distributed random draws. Also defined for Table 1

$$\bar{x} = E[x] \qquad \sigma^2 = E[x^2] - E[x]^2$$
$$x_m = \text{median } x, \text{ i.e., } \int^{x_m} dP(x) = \frac{1}{2}$$
$$a = 2\ln(\bar{x}/x_m) \qquad (\alpha x_m)^\beta = \ln 2$$
$$\alpha\bar{x} = \Gamma\left(1 + \frac{1}{\beta}\right).$$

B. Generating a Sequence of Correlated Random Draws

Generation of correlated random draws from independent random draws is readily accomplished even for a general specification of covariance matrix [7], [11], [18], [19], [22].

A sequence of random features with any particular correlations is constructed using the discrete Karhunen–Loève expansion. This expansion constructs a sequence of N random draws v_n, with covariance R_{nm}, from N independent random draws u_m

$$v_n = E[v_n] + \sum_{m=1}^{N} H_{nm} u_m. \tag{2}$$

The u_m are independent, mean zero, unity variance random draws. The constructed random draws v_n have the desired correlations as long as the matrix H is a root of the covariance matrix

$$HH^T = R$$

with H^T denoting the matrix transpose of H.

It is sufficient that R be a covariance matrix. The condition for R to be a covariance matrix is given by the Bochner–Schwartz theorem [23]: there exists a positive measure μ, and points x_k, such that

$$R_{nm} = \int d\mu(k)\, e^{ik(x_n - x_m)}.$$

The sole challenge is to construct the root matrix H for the covariance matrix R, i.e., $HH^T = R$. Covariance matrices are nonnegative and therefore have a real root matrix H. Indeed, there are many such root matrices. A real, lower triangular root can be constructed by the Cholesky decomposition algorithm (similar to LU decomposition by Crout's algorithm [15] with $U = L^T$).

Since the root matrix calculated in Cholesky decomposition is lower triangular, each newly constructed correlated feature depends only on previously generated independent random draws and one newly generated independent random draw. Thus this particular selection for factorization of the covariance matrix is causal: the newly generated feature depends only on the previously generated features and one contemporaneous random draw. Since the inverse of a lower triangular matrix is lower triangular, the transformation is causally invertible as long as it is invertible.

Many other roots for the covariance matrix can be generated, for example, by diagonalizing the covariance matrix. If A diagonalizes the covariance matrix, and O is an arbitrary orthogonal matrix, then all

$$H = A\sqrt{D}O$$

are root matrices of R, i.e., $HH^T = R$ if $R = ADA^T$, D is diagonal, and $O^T = O^{-1}$. (Transpose can be extended to Hermitian transpose for complex matrices. As long as we can ensure that $v = Hu$ is real, complex roots can be employed.)

Perhaps the most common method of constructing correlated random features is particular to stationary processes and employs the fast Fourier transform (FFT) algorithm. This method is a numerically efficient implementation of the Karhunen–Loève expansion. The FFT algorithm performs the multiplication of the root matrix H times the independent random draws in a computationally efficient manner. However, this technique is limited to sets of random features which have symmetric Toeplitz covariance matrices, for example, uniformly sampled stationary random processes.

The computational efficiency is achieved only if the covariance matrix is diagonalized by the discrete Fourier transform (DFT) matrix. FFT algorithms achieve required matrix multiplications in order $N \ln(N)$ rather that N^2 operations, which more than compensates for a required augmentation of the covariance matrix.

The root of the stationary covariance matrix is selected to achieve the Karhunen–Loève expansion

$$v = E[v] + F\sqrt{D}u$$

with $FDF^* = R$. To achieve this specific diagonalization, the covariance matrix is augmented. Since Toeplitz matrices are not generally diagonalized by the DFT, the $N \times N$ symmetric Toeplitz matrix is embedded within a circulant, $2N \times 2N$ matrix

$$R_{nm} = \begin{pmatrix} \begin{pmatrix} R_0 & R_1 & R_2 \\ R_1 & R_0 & R_1 \\ R_2 & R_1 & R_0 \end{pmatrix} & \begin{matrix} 0 & R_2 & R_1 \\ R_2 & 0 & R_2 \\ R_1 & R_2 & 0 \end{matrix} \\ \begin{matrix} 0 & R_2 & R_1 \\ R_2 & 0 & R_2 \\ R_1 & R_2 & 0 \end{matrix} & \begin{matrix} R_0 & R_1 & R_2 \\ R_1 & R_0 & R_1 \\ R_2 & R_1 & R_0 \end{matrix} \end{pmatrix}.$$

A circulant matrix is diagonalized by DFT similarity transformation. The N-dimensional unitary DFT matrix is defined

$$F_{nm} = \frac{1}{\sqrt{N}} e^{2\pi i(n-1)(m-1)/N}.$$

Only N correctly correlated features are generated even though a $2N$-dimensional FFT is employed. The additional features are correlated with the desired features as if the features were on a ring. This can be understood by noting that the DFT implements a circular convolution of the root of the spectral density with the generated sequence of independent random draws. Generally, these additional features are of no interest and are discarded.

The independent random draws must have a conjugate symmetry, $u_k = (u_{2N+2-k})^*$, to construct real correlated features. If the independent random draws are chosen to be real, it follows that $u_{2N+2-k} = u_k$. For real independent draws, the generated sequence is symmetric.

The construction of correlated random numbers readily adapts to multidimensional situations. In multidimensional processes, the premium is on efficiency in generation.

In a general case, the root covariance matrix can be evaluated using the Cholesky decomposition algorithm. This algorithm takes on the order of $(LMN)^3/6$ passes through its inner loops for three dimensions of L, M, and N samples, respectively. The matrix multiplication in the generation of the correlated random draws takes on the order of $(LMN)^2/2$ multiplications and additions, so the calculation is dominated by the decomposition. If the covariance matrix factors

$$R(i,j,k;\ell,m,n) = R_x(i,\ell)R_y(j,m)R_z(k,n)$$

the Cholesky decomposition takes on the order of $(L^3 + M^3 + N^3)/6$ passes through the inner loops. The matrix multiplication now dominates, taking on the order of $LMN(L + M + N)/2$ multiplications and additions. Linear shift invariant (stationary in each dimension) covariance matrices have the symmetry that leads again to a computationally efficient diagonalization of the covariance matrix employing FFT algorithms. The operation count for stationary covariances is order $LMN\ \log_2(LMN)$.

C. First Marginal Distributions of the Correlated Features

The distributions of the correlated features generated in Section II-B are inherited from the distributions of the underlying independent random draws.

Since any particular correlated random feature has been expressed as a sum of independent random draws

$$v_n = \sum_{m=1}^{N} H_{nm}u_m$$

the first marginal probability distribution of each correlated feature is a distribution for the weighted sum of independent random draws.

Defining characteristic functions for each of the independent random draws, it is straightforward to manipulate the expression for the probability density of the correlated random draw into a single integral

$$p_n(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ikv} \prod_{m=1}^{N} \hat{C}_m(H_{nm}k)$$

where the characteristic function for the distribution of each independent random draw is defined

$$\hat{C}_n(k) = \int dP_n(x)\, e^{-ikx}$$

and the characteristic functions for the correlated random draws are

$$C_n(k) = \prod_{m=1}^{N} \hat{C}_m(H_{nm}k). \tag{3}$$

The key step is realizing that expectations over the independent random draws factor into products of expectations over each independent draw. The characteristic functions are bounded functions since the probability density measures have unit mass, and consequently, products of characteristic functions make sense [24]. In cases for which density functions are defined, the first marginal probability density function of the correlated feature is the multiple convolution integral

$$\begin{aligned} p_n(v) = {} & \frac{1}{\prod_{j=1}^{N} |H_{nj}|} \prod_{k=2}^{N} \int_{-\infty}^{\infty} ds_k \\ & \cdot \varphi_1\left(\frac{v - s_2}{H_{n1}}\right) \varphi_2\left(\frac{s_2 - s_3}{H_{n2}}\right) \\ & \cdots \varphi_{N-1}\left(\frac{s_{N-1} - s_N}{H_{n,N-1}}\right) \varphi_N\left(\frac{s_N}{H_{nN}}\right) \end{aligned} \tag{4}$$

where a density is defined for each probability distribution of the independent random draws

$$P_j(x) = \int^{x} dt\, \varphi_j(t).$$

III. Constructing Random Draws with Particular Correlations and First Marginal Distributions

A. Problem Background

Practical algorithms for constructing nonnormal, correlated sets of random features result from joining the methods of the previous sections.

If the joint probability distribution of the random features is known, then the extension to more than one dimension of the Inverse Probability Method constructs realizations of a general set of random numbers from independent, uniformly distributed random draws. This construction uses conditional, marginal distributions

$$\begin{aligned} x_1 &= P^{-1}(U_1; \infty, \cdots) \\ x_2 &= P^{-1}(U_2; \infty, \cdots | x_1) \\ &\vdots \\ x_N &= P^{-1}(U_N | x_1, \cdots x_{N-1}). \end{aligned} \tag{5}$$

That is, find the x_k that solve $P(x_k, \infty, \cdots | x_1, \cdots x_{k-1}) = U_k$ for uniformly distributed random draws U_k.

However, usually the joint probability distribution is not known, and our goal is to generate sets of random features given a description short of the full elaboration of the joint probability distribution. We wish to find a systematic way of constructing joint probability distributions constrained to have two properties, first marginal distributions and correlations, not readily imposed on guesses for joint distributions.

Control is exerted over the constructed first marginal distributions using the techniques of Section II-A, and over the correlations of the constructed features using the methods of Section II-B. However, the methods interfere; summing random draws modifies the distributions, and functions of random draws are correlated differently than the random draws themselves.

B. Construction Methods for Nonnormal, Correlated Random Numbers

Significant freedom in procedures to construct nonnormally distributed, correlated random numbers is anticipated since many joint distributions share first marginal distributions and correlations. Three approaches to the generation of nonnormal, correlated random features will be distinguished. In addition to the Inverse Probability Method, (5), the methods are designated:

DISTRIBUTION DISTORTION METHOD: Using the methods of Sections II-A and II-B, fix correlations of the random features generated from linear combinations of nonnormally distributed, independent random draws. Using the expressions from Section II-C, modify the probability distributions of the underlying independent random draws to achieve the desired first marginal distributions for the correlated features.

CORRELATION DISTORTION METHOD: Find a transformation from normal random draws to the desired first marginal distribution using a nonlinear transformation, i.e., the methods discussed in Section II-A. Generate a correlated, normally distributed set of random features from linear combinations of independent normally distributed random draws using the Karhunen–Loève expansion, (2). Fix the correlations for the underlying set of normal random features to produce the desired correlations for the nonnormal features.

These latter two approaches require only the two key attributes: first marginal distributions and correlations. When the joint distribution is unknown, yet there are estimates for the correlations and first statistics, these two approaches which do not require the joint distribution are the natural way to proceed.

The Distribution Distortion method fixes correlations among the generated features and seeks distributions for the independent random draws that produce the desired first marginal distributions for the correlated features. Distributions for the independent random draws must be selected to depart from the convergence to normal statistics indicated by the Central Limit Theorem. Tailoring distributions of the independent random draws involves solving for products of characteristic functions which equal the characteristic functions of the desired first marginal distributions.

The Correlation Distortion approach fixes output statistics and seeks to find correlations for a set of underlying random numbers to produce the correct correlations for transformed

random features. A nonlinear transformation is applied to construct the nonnormal correlated features from correlated normal random draws. This approach takes advantage of the stability of the normal distribution to generate correlated normally distributed random numbers from independent, normal random draws. Because of the relative ease of fixing the correlations and first statistics for each labeled feature, the Correlation Distortion approach is pursued below. This approach requires an evaluation of the correlations for the nonnormal features in terms of the correlations of the pretransformation, correlated, normally distributed random numbers.

The range of random processes constructable by any particular implementation of either the Distribution Distortion or Correlation Distortion methods may be limited. It is not always possible to find a solution for probability distributions of the independent random draws (Distribution Distortion method) or the covariance matrix of the underlying normal process (Correlation Distortion method). In the case of the Distribution Distortion method, the characteristic functions of infinitely divisible [26] distributions can always be factored, so infinitely divisible first marginal distributions are always achievable. However, if the roots of the characteristic functions indicated in (3) are not characteristic functions, the method must be abandoned. For the Correlation Distortion method, the derived covariance matrix must be nonnegative, and if it is not, alternative nonlinear transformations are sought. For example, in the case of correlated, uniformly distributed random features examined in Section III-D, negative correlations are not reachable with one selection for the nonlinear transformation, yet negative correlations are reached with a second nonlinear transformation.

The Distribution Distortion method is also denoted the "moving average" model in detection and estimation studies [6], [7]. I avoid this designation to prevent confusion with constant coefficient, moving average filter models. The Distribution Distortion method correctly models the correlation of all N constructed random features even if the features are, for example, samples of an autoregressive process. The Correlation Distortion method has also been employed in optimal detector studies [7], and is the method discussed in [8].

These constructions are suggestive of canonical or innovations [22], [25] representations for sets of random features. The canonical representation seeks a set of white (implying independent for normal random numbers), normal random numbers that transform, by causal and causally invertible transformation, into the desired set of random features. The transformation is not necessarily linear, as it must not be to achieve nonnormal random numbers. My discussion is less general than the canonical representations approach since I identified methods only for discretely indexed sets of random features. This discussion is also less restrictive than canonical representations, as any practical transformation, not necessarily invertible nor causal, is presented. Nevertheless, the construction methods presented are both causal and causally invertible, as long as the covariance matrix particular to the construction is nonsingular, the triangular factorization of the covariance is selected, and the marginal probability distributions are invertible functions, as required.

C. Approximate Methods

In addition to the three general, "exact" approaches, exact in the sense that the first marginal distribution and correlations are precisely realized, there are a wide variety of approximate constructions. The three methods distinguished above are more general in the sense that correlations are precisely achieved and are arbitrarily specified, e.g., the covariances need not be stationary (shift-invariant). The approximate methods are most readily applied to stationary (identically distributed and shift-invariant covariance) sequences of random numbers, with target correlations only approximated. Nonetheless, the approximate methods can be convenient when the goals are stationary sequences of random numbers and there is no stringent requirement to match correlations.

The approximate methods can be represented by the recursion

$$v_k = g_k\left(\sum_{j=1}^{N_1}\alpha_j v_{k-j} + \sum_{j=1}^{N_2}\beta_j u_{k-j} + u_k\right)$$

$$x_k = f_k(v_k)$$

with

u_k	independent random draws
v_k	correlated random draws
x_k	target random draws
α_k	autoregressive coefficients
β_k	moving average coefficients

With $g_k(x) = x$, the generation of the correlated random numbers v_k can be recognized as an ARMA, including pure AR and MA, filtering [27] of the independent draws u_k. This filtering corresponds to the more precise Karhunen–Loève expansion given in (2), Section II-B. The correlations resulting from this filtering are necessarily approximate if large numbers of random features are generated since the filter has a fixed number of degrees of freedom, the coefficients. The correlations are fixed by selection of ARMA model together with the filter coefficients.

One method [28] closely related to the Correlation Distortion method employs normally distributed independent random draws, resulting in normally distributed correlated draws ($g_k(x) = x$). A selected nonlinear transformation $f_k(x)$, as discussed in Section II-A, constructs the correlated random features with particular first marginal distributions. Another approach, represented by [29], constructs random features with infinitely divisible [26] distributions, and uses coefficients α_k that are discretely distributed random numbers, $\beta_k = 0$, and no nonlinear transformations. This approach is most closely related to the Distribution Distortion method. Another well developed method

[30], [31] employs the inverse probability transformation $P^{-1}(U)$ to fix the first marginal probability distributions from correlated, uniformly distributed random numbers. This method exploits the result that the fractional part of the sum of a uniformly distributed (0, 1) random number with any independent random number is also uniformly distributed (0, 1). With $\alpha_1 = 1$, $\alpha_k = 0$ for $k > 1$, and $\beta_k = 0$, the nonlinear transformation $g_k(x)$ is either the fractional part of x or of $1 - x$. Consequently, the correlated random numbers are uniformly distributed. In this case, the second transformation $f_k(x)$ consists of a uniform distribution preserving stitching transformation [30] followed by the inverse probability transformation to arrive at the desired first marginal distributions. The distribution of the independent random draws u_k, the stitching transformation, and the transformations $g_k(x)$ are manipulated to approximate target correlations.

D. Example Sets of Correlated Nonnormal Random Numbers

Log Normal: One tractable example of the Correlation Distortion method is construction of correlated, log-normally-distributed random features. In this case, it is relatively easy to solve for the required correlations of the underlying normal random numbers. This example has generally specified correlations, and log-normal first marginal distributions specified separately for each random feature.

Each correlated feature has a first marginal distribution given by the log-normal probability density function defined as

$$p_i(x) = \frac{1}{\sqrt{2\pi}\, a_i x} \exp\left(-\ln^2 (x/\mu_i)/(2a_i^2)\right)$$

where

μ_i median value for feature i

a_i log normal density shape parameter for distribution of feature i

This density is nonvanishing only for positive values $x > 0$. The shape parameter can be related to the mean to median ratio for the feature.

The construction begins by identifying the transformation from normal to log-normal features

$$x_i = \mu_i \exp(a_i v_i).$$

Now the methods of Section II-B are employed to generate correlated, mean zero, normal random numbers from independent normal random draws. For the correlated normal random numbers, define

$$r_{ij} = E[v_i v_j]$$

the covariance for normal features i and j.

The generated features x_i are jointly log-normal with means, medians, and second moments specified independently. The mean, median, and mode are in the same units as the features. The shape parameters are dimensionless. The average, median, mode (most likely value), variance, and any moment of a feature can be expressed with two parameters, herein selected as the shape parameter and the median.

$$E[(x_i)^n] = (\mu_i)^n \exp\left(\frac{1}{2} n^2 (a_i)^2\right)$$

are the moments, and $\mu_i \exp(a_i^2)$ is the mode.

The sole remaining task is to select the correlations of the underlying normal set of random numbers to produce the desired correlations in the constructed log normal random numbers. The covariance for the log-normal set of random features is evaluated, and then solved for r_{ij}

$$R_{ij} = E[x_i x_j] - E[x_i]E[x_j].$$

The mean of x_i is already available. The mixed second moments must be evaluated

$$\begin{aligned} E[x_i x_j] &= E[\mu_i \exp(a_i v_i)\mu_j \exp(a_j v_j)] \\ &= \mu_i \mu_j E[\exp(a_i v_i + a_j v_j)]. \end{aligned}$$

The indicated Gaussian integrals could be evaluated to find the covariances, but this Gaussian integration is most easily evaluated algebraically using Wick's theorem [32], [33]. Wick's theorem[2] for mean zero, normal random features provides

$$E[e^v] = e^{\frac{1}{2}E[v^2]}.$$

Simple manipulations result in the required mean values

$$E[x_i x_j] = \mu_i \mu_j \exp\left(a_i a_j r_{ij} + \frac{1}{2}(a_i^2 r_{ii} + a_j^2 r_{jj})\right).$$

We can select $r_{kk} = 1$. Now the covariances for the underlying normal set of random draws are determined in terms of specified covariances for the log normal random numbers

$$r_{ij} = \frac{1}{a_i a_j} \ln\left(1 + \frac{R_{ij}}{E[x_i]E[x_j]}\right).$$

Since we set $r_{ii} = 1$, use the parameters of the log-normal distribution to fix

$$R_{ii} = E[x_i]^2\left(e^{a_i^2} - 1\right).$$

This completes the construction of a log-normally distributed set of random features with specified correlations. This is an example construction of a non-shift-invariant, correlated, nonnormal set of random features. The construction began with independent normal features derived from independent uniformly distributed random draws, linear combinations of these normal features are correlated normal features, and then an individually selected nonlinear transformation for each feature produces the set of random features with particular first marginal distributions and correlations.

[2] Another result of Wick's theorem is that the expectation value of the product of $2n$ mean zero, jointly normal features equals the sum over all unordered pairings of products of expectations of pairs of features. The expectation of a product of an odd number vanishes. There are $(2n!)/(2^n n!)$ unordered pairings of $2n$ objects. This result is also known as Isserlis' formula [34].

Constant Modulus Waveforms: The second example constructs phase-modulated waveforms with any particular covariance. The constant modulus waveform is constructed

$$x_k = A \sin(\beta v_k) = \Re e\,(iAe^{-i\beta v_k}).$$

The v_k again denote real, jointly normal random variates with means and covariances

$$E[v_k] = 0$$
$$E[v_k v_j] = r_{kj}$$
$$r_{kk} = 1.$$

The modulus of this waveform is A. It should be noted that "constant modulus" as conventionally applied does not imply constant envelope. I adopt the convention that a phase-modulated waveform is constant modulus, i.e., the magnitude of the complex representation, $e^{i\beta v_k}$, is constant. The envelope of a phase modulated waveform is generally not constant, but is approximately constant in the narrowband approximation, which requires that the phase modulation be generated on a carrier frequency.

Wick's theorem provides the mean value

$$E[iAe^{-i\beta v_k}] = iAe^{-\frac{1}{2}\beta^2}.$$

The waveform is mean zero, $E[x_i] = 0$, although its imaginary component is not (but its Hilbert transform, the imaginary component in the analytic representation, is mean zero).

The first step in evaluating the covariance matrix is to find the expectation of the product

$$\begin{aligned} E[(A\sin(\beta v_k))(A\sin(\beta v_j))] &= -\frac{A^2}{4}E[e^{i\beta(v_j+v_k)} - e^{i\beta(v_j-v_k)} - e^{-i\beta(v_j-v_k)} + e^{-i\beta(v_j+v_k)}] \\ &= -\frac{A^2}{2}\left(e^{-\frac{1}{2}\beta^2 E[(v_j+v_k)^2]} - e^{-\frac{1}{2}\beta^2 E[(v_j-v_k)^2]}\right) \\ &= -\frac{A^2}{2}\left(e^{-\beta^2}e^{-\beta^2 r_{kj}} - e^{-\beta^2}e^{\beta^2 r_{kj}}\right). \end{aligned}$$

The covariance matrix for the constant modulus waveform is

$$R_{kj} = A^2 e^{-\beta^2}\sinh(\beta^2 r_{kj}).$$

The covariances of the underlying normal random draws v_k are fixed to achieve our selection for the particular covariance of the constructed constant modulus, phase modulated, waveform

$$r_{kj} = \frac{1}{\beta^2}\ln\left(R_{kj}\frac{e^{\beta^2}}{A^2} + \left(\left(R_{kj}\frac{e^{\beta^2}}{A^2}\right)^2 + 1\right)^{1/2}\right).$$

This completes construction of our constant modulus waveform with any particular covariance matrix. We are free to choose R_{kj} which specifies the construction of the underlying real, jointly normal set of random features that results in a phase-modulated sampled waveform with any desired set of correlations. If we select the correlations to be shift-invariant, we can define a PSD and have constructed a phase-modulated waveform with any required PSD.

Uniform: The next example constructs a correlated, uniformly distributed set of random features. If linear combinations of independent uniformly distributed random draws were formed to generate a correlated set of random features, the first statistics would be decidedly nonuniform. But, beginning with an independent, normally distributed set of random numbers, the correlated set of random features will have jointly normal distributions, and the transformation from normal to uniform statistics is elementary.

The key expression is again the transformation from normal random draws v_{ik} to uniform (0, 1) random draws x_i, in this case involving two independent normal random draws

$$x_i = \exp\left(-\frac{1}{2}(v_{i1}^2 + v_{i2}^2)\right).$$

For both independent realizations of the correlated normal random draws, define

$$r_{ij} = E[v_{ik}v_{jk}]$$

the covariance for normal samples i and j, independent realizations, $k = 1, 2$. For x_i to be uniformly distributed requires that $r_{ii} = 1$.

The only remaining task is to select the covariances of the two sets of underlying normal random numbers to reproduce the particular correlations for the uniformly distributed set of random features. Solve

$$R_{ij} = E[x_i x_j] - E[x_i]E[x_j]$$

for r_{ij}. Expanding and simplifying

$$\begin{aligned} R_{ij} &= E\left[\exp\left(-\frac{1}{2}(v_{i1}^2 + v_{i2}^2\right)\exp\left(-\frac{1}{2}(v_{j1}^2 + v_{j2}^2)\right)\right] \\ &\quad - E\left[\exp\left(-\frac{1}{2}(v_{i1}^2 + v_{i2}^2\right)\right] \\ &\quad \cdot E\left[\exp\left(-\frac{1}{2}(v_{j1}^2 + v_{j2}^2\right)\right]. \\ R_{ij} &= E\left[\exp\left(-\frac{1}{2}(v_{i1}^2 + v_{j1}^2)\right)\right]^2 - E\left[\exp\left(-\frac{1}{2}v_{i1}^2\right)\right]^2 \\ &\quad \cdot E\left[\exp\left(-\frac{1}{2}v_{j1}^2\right)\right]^2 \\ R_{ij} &= \frac{1}{\|1+r\|} - \frac{1}{(1+r_{ii})(1+r_{jj})}. \end{aligned}$$

Crucial use was made of the independence of the two realizations, that products of determinants equal the determinant of products, and that sums of nonnegative matrices are nonnegative. Also, $\|A\|$ denotes the determinant of A. The resulting Gaussian integrals were directly evaluated.

For example

$$E\left[\exp\left(-\frac{1}{2}(v_{i1}^2+v_{j1}^2)\right)\right]=\frac{1}{2\pi}\int dv_i \cdot\int dv_j\|r\|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}v^T r^{-1}v\right)\cdot\exp\left(-\frac{1}{2}(v_i^2+v_j^2)\right)$$

$$E\left[\exp\left(-\frac{1}{2}(v_{i1}^2+v_{j1}^2)\right)\right]=\|1+r^{-1}\|^{-\frac{1}{2}}\|r\|^{-1/2}.$$

Finally, elementary manipulations produce

$$r_{ij}=\sqrt{\frac{16R_{ij}}{1+4R_{ij}}}.$$

Since $r_{ii}=1$, we find that $R_{ii}=1/12$.

Rather than employ two independent, underlying sets of normal random numbers, the uniformly distributed set of random features can also be generated from a single set of normally distributed random numbers [8]. In this case, the transformation from normal to uniform is given by a normal error function $x=1-Q(v)$ ($Q(v)$ defined below in (6)) and the correlations are related: $r_{ij}=2\sin(2\pi R_{ij})$. This result can be achieved by direct evaluation of the indicated quadrature or by using Price's theorem [35].

Weibull: Another case evaluated by quadrature is provided by correlated Weibull [36]. In this case, the transformation from two independent sets of correlated normal draws to the Weibull distributed feature is

$$x_i=\frac{1}{\alpha}\left(\frac{v_{i1}^2+v_{i2}^2}{2}\right)^{\frac{1}{\beta}}.$$

The parameters of the Weibull distribution were introduced in Section II-A and Table 1. The relationship between the correlation of the Weibull distributed random features and the correlated $r_{ii}=1$ normal random draws is expressed in terms of Gauss hypergeometric and gamma functions [10]

$$R_{ij}=\frac{1}{\alpha^2}\Gamma^2(1+\beta^{-1})\left({}_2F_1(-\beta^{-1},-\beta^{-1},1;r_{ij}^2)-1\right).$$

Chi-Square: This example constructs a correlated, chi-square distributed set of random features. The chi-square distributions are another class of nonnormal distributions readily formed by this method. Again, the transformation from normal to the desired, chi-square statistics is elementary, employing N independent realizations of correlated normal random draws.

$$x_i=\frac{\bar{x}_i}{N}\sum_{k=1}^{N}v_{ik}^2.$$

The chi-square distribution is defined for positive values and parametrized by its mean and the number, N, of degrees of freedom

$$p(x)=\frac{N}{2\bar{x}\Gamma(N/2)}\left(\frac{Nx}{2\bar{x}}\right)^{(N-2)/2}\exp\left(-Nx/(2\bar{x})\right).$$

Since the mean is known, the only remaining work is to evaluate the expectation of mixed products of two features

$$E[x_ix_j]=\left(\frac{\bar{x}_i\bar{x}_j}{N^2}\right)E\left[\left(\sum_{k=1}^{N}v_{ik}^2\right)\left(\sum_{n=1}^{N}v_{jn}^2\right)\right]$$

$$E[x_ix_j]=\left(\frac{\bar{x}_i\bar{x}_j}{N^2}\right)\left(E\left[\sum_{k=1}^{N}v_{ik}^2\right]E\left[\sum_{k=1}^{N}v_{jk}^2\right]+\sum_{k=1}^{N}\left(E[v_{ik}^2v_{jk}^2]-E[v_{ik}^2]E[v_{jk}^2]\right)\right).$$

This is a model application of Wick's theorem. We find that

$$R_{ij}=E[x_ix_j]-\bar{x}_i\bar{x}_j=\frac{\bar{x}_i\bar{x}_j}{N}(E[v_{ik}^2v_{jk}^2]-1)=\frac{\bar{x}_i\bar{x}_j}{N}(2r_{ij}^2).$$

This is readily inverted to give the covariances of the N sets of independent realizations of correlated normal random draws used to construct the chi-square distributed set of random features

$$r_{ij}=\sqrt{\frac{N}{2}\frac{R_{ij}}{\bar{x}_i\bar{x}_j}}.$$

Again, since we set $r_{ii}=1$

$$R_{ii}=\frac{2}{N}(\bar{x}_i)^2.$$

A similar development relates covariances for the square-law envelope Rician model.

Two-Level Discrete: The closing example of the Correlation Distortion method constructs a set of random features distributed over two discrete levels. These random features have no conventional first density function; the probability measure is concentrated on two points, 0 and 1

$$p(x)=P\,\delta(x-1)+(1-P)\,\delta(x).$$

P can be interpreted as the probability that a feature will be observed with value unity.

As before, the key expression is the relationship between normal features and features with the desired first marginal distribution. One such relationship is

$$x_i=\begin{cases}1, & \text{if } Q(v_i)<P_i\\ 0, & \text{otherwise.}\end{cases}$$

The complementary normal distribution function is defined

$$Q(v)=\frac{1}{\sqrt{2\pi}}\int_v^\infty dt\;e^{-(1/2)t^2}. \tag{6}$$

For this relationship, we set $r_{ii}=1$.

The covariance of the discretely distributed set of random features is now

$$R_{ij}=\frac{1}{2\pi}\int_{A_i}^{\infty}dv_i\int_{A_j}^{\infty}dv_j\,\|r\|^{-1/2}\cdot\exp\left(-\frac{1}{2}v^Tr^{-1}v\right)-P_iP_j$$

where

$$Q(A_k)=P_k.$$

Table 2 Correlation Distortion Method Examples

Transformation from normal, v, to particular x	Covariance matrix to for underlying normal draws, r, given particular R
Log Normal: $x = \mu \exp(av)$	$r_{ij} = \frac{1}{a_i a_j} \ln\left(1 + \frac{R_{ij}}{E[x_i]E[x_j]}\right)$
Phase Modulated: $x = A\sin(\beta v)$	$r_{ij} = \frac{1}{\beta^2} \ln\left(R_{ij}\frac{e^{\beta^2}}{A^2} + \left(\left(R_{ij}\frac{e^{\beta^2}}{A^2}\right)^2 + 1\right)^{1/2}\right)$
Uniform: $x = \exp\left(-\frac{1}{2}(v_1^2 + v_2^2)\right)$ $x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{v} dt\, e^{-(1/2)t^2}$	$r_{ij} = \sqrt{\frac{16R_{ij}}{1+4R_{ij}}}$ $r_{ij} = 2\sin(2\pi R_{ij})$
Weibull: $x = \frac{1}{\alpha}\left(\frac{v_1^2+v_2^2}{2}\right)^{1/\beta}$	$R_{ij} = \frac{1}{\alpha^2}\Gamma^2(1+\beta^{-1})\left({}_2\mathbf{F}_1(-\beta^{-1}, -\beta^{-1}, 1; r_{ij}^2) - 1\right)$
Chi-Square: $x = \frac{\bar{x}}{N}\sum_{k=1}^{N} v_k^2$	$r_{ij} = \sqrt{\frac{N}{2}\frac{R_{ij}}{\bar{x}_i\bar{x}_j}}$
Square-law Envelope Rician: $x = \frac{1}{2}(v_1^2 + (v_2 + s)^2)$	$r_{ij} = \sqrt{R_{ij} + \left(\frac{s}{2}\right)^2} - \frac{s}{2}$

This Gaussian integral over a semi-infinite region is not expressible in common elementary functions. As such, this example is more representative of general problems than the prior examples. This relationship between R_{ij} and r_{ij} can be inverted by standard numerical quadrature and root finding algorithms. Toward this end, it is convenient to express the relationship as a one-dimensional quadrature

$$R_{ij} = \frac{1}{\sqrt{2\pi}} \int_{A_j}^{\infty} dt e^{-(1/2)t^2} Q\left(\frac{A_i - r_{ij}t}{\sqrt{1 - r_{ij}^2}}\right) - P_i P_j.$$

Alternatively, elementary representations can be derived in the limits of small or large correlations. For the case of small correlation, $|r_{ij}| \ll 1$, $i \neq j$

$$\begin{aligned} R_{ij} &\approx \frac{1}{2\pi}\int_{A_i}^{\infty} dv_i \int_{A_j}^{\infty} dv_j \exp\left(-\frac{1}{2}(v_i^2 + v_j^2)\right) \\ &\quad \cdot (1 - r_{ij}^{-1} v_i v_j) - P_i P_j \\ R_{ij} &= \frac{r_{ij}}{2\pi}\int_{A_i}^{\infty} dv_i \int_{A_j}^{\infty} dv_j v_i v_j \exp\left(-\frac{1}{2}(v_i^2 + v_j^2)\right) \\ &= \frac{r_{ij}}{2\pi}\exp\left(-\frac{1}{2}(A_i^2 + A_j^2)\right) \\ r_{ij} &= 2\pi R_{ij}\exp\left(\frac{1}{2}(A_i^2 + A_2^2)\right). \end{aligned}$$

For large correlations, $r_{ij} = 1 - \epsilon$, $\epsilon \ll 1$

$$\begin{aligned} R_{ij} = &\int_{A_i}^{\infty} dv_i \int_{A_j}^{\infty} dv_j \frac{1}{2\pi\sqrt{2\epsilon}} \\ &\cdot \exp\left(-\frac{1}{4\epsilon}(v_i - v_j)^2 - \frac{1}{8}(v_i - v_j)^2 - \frac{1}{2}v_i v_j\right) \\ &- P_i P_j. \end{aligned}$$

One factor is a delta sequence

$$\delta(v_i - v_j) \approx \lim_{\epsilon \to 0} \frac{1}{\sqrt{4\pi\epsilon}} \exp\left(-\frac{1}{4\epsilon}(v_i - v_j)^2\right).$$

Then, using the fact that $Q(x)$ is a monotonically decreasing function

$$R_{ij} = Q(\max(A_i, A_j)) - P_i P_j = \min(P_i, P_j) - P_i P_j.$$

The example constructions are summarized in Table 2. The range of correlations achieved for each example transformation can be assessed using Table 3. The range of correlations achievable with the Correlation Distortion method and the particular transformations used in the examples is the image generated by evaluating the tabulated correlations for every nonnegative matrix r_{ij}. Note that for several of the example transformations, negative correlations cannot be reached.

A general feature of the Correlation Distortion method is that lack of correlation between the normal features implies independence of the constructed nonnormal features. That is, if the normal features used in the nonlinear transformation from normal to nonnormal statistics are uncorrelated, then the nonnormal features will necessarily be independent. By construction, the characteristic function will factor into expectations over each feature independently. However, the converse of this is not true; the nonnormal set of random features can be uncorrelated even though it is generated from a correlated normal set of random features. In this event, the nonnormal features will be uncorrelated, but not independent. One example suffices to establish this result.

The correlation of two features in the Correlation Distortion method is represented

$$\begin{aligned} R_{ij} = &\frac{1}{2\pi}\int dv_1 dv_2\, f_i(v_1) f_j(v_2) \\ &\times \left(\|r\|^{-1/2} e^{-1/2(v^T r^{-1} v)} - e^{-\frac{1}{2}(v_1^2 + v_2^2)}\right) \end{aligned} \quad (7)$$

where $f_i(v)$ is the transformation from a normally distributed random draw to the particularly distributed random

Table 3 Range of Correlations for Correlation Distortion Method Examples

Example	Correlation
Definition:	$\rho_{ij} = \frac{R_{ij}}{\sqrt{R_{ii}R_{jj}}}$
Log Normal:	$\rho_{ij} = \frac{e^{a_i a_j r_{ij}} - 1}{\sqrt{(e^{a_i^2}-1)(e^{a_j^2}-1)}}$
Phase Modulated:	$\rho_{ij} = \frac{\sinh(\beta^2 r_{ij})}{\sinh(\beta^2)}$
Uniform:	$\rho_{ij} = \frac{3r_{ij}^2}{4-r_{ij}^2}$ $\rho_{ij} = \frac{6}{\pi}\sin^{-1}(\frac{r_{ij}}{2})$
Weibull:	$\rho_{ij} = \frac{{}_2\mathbf{F}_1(-\beta^{-1},-\beta^{-1},1;r_{ij}^2)-1}{{}_2\mathbf{F}_1(-\beta^{-1},-\beta^{-1},1;1)-1}$
Chi-Square:	$\rho_{ij} = (r_{ij})^2$
Square-law Envelope Rician:	$\rho_{ij} = \frac{r_{ij}(s+r_{ij})}{(s+1)}$

feature. The covariance matrix is defined to be

$$r_{ij} = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}.$$

Factoring the nonnegative covariance matrix and changing integration variables yields

$$R_{ij} = \frac{1}{2\pi}\int dv_1 dv_2\, e^{-1/2(v_1^2+v_2^2)} \times\Big(f_1(\sqrt{1-r^2}v_1 + rv_2) - f_1(v_1)\Big) f_2(v_2).$$

The covariance of features i and j can be set to zero, without requiring $r = 0$, for example by an even f_1, odd f_2 pairing. This manipulation would provide an example of an uncorrelated, yet not independent, pair of draws generated by the Correlation Distortion method.

IV. Joint Probability Density Functions

The joint probability distributions of the set of random features constructed by the Distribution Distortion and Correlation Distortion methods are readily derived. These joint probability distributions are the unique descriptions of the constructed random features.

The joint probability distribution of sets of random features constructed from independent random draws by (2) is readily evaluated from the joint characteristic function

$$E\left[\exp\left(-i\sum_{j=1}^{N} k_j v_j\right)\right] = \prod_{m=1}^{N} E\left[\exp\left(-i\sum_{j=1}^{N} k_j H_{jm} u_m\right)\right] = \prod_{m=1}^{N} \hat{C}_m\left(\sum_{j=1}^{N} k_j H_{jm}\right)$$

where $\hat{C}_m(\cdot)$ are the characteristic functions for the distributions of the independent random draws. The inverse Fourier transform of this joint characteristic function gives the joint probability density. In the cases for which densities are defined, this can be written

$$p(x_1, \cdots x_N) = \frac{1}{\|H\|}\prod_{m=1}^{N} \varphi_m\left(\sum_{n=1}^{N} H_{mn}^{-1} x_n\right). \quad (8)$$

This is the expression for the joint probability density for sets of random features constructed by the Distribution Distortion method. That is, H is found as a root of the covariance R, and the distributions of the independent random draws are set to achieve the desired first marginal distributions, as expressed by (4).

The joint probability density for sets of random features constructed by the Correlation Distortion method is also readily derived. In this case, $x_m \equiv f_m(v)$ with v a mean zero, unity variance, normally distributed random draw. The joint probability density, (8), simplifies considerably using

$$\prod \exp(a_n) = \exp\left(\sum a_n\right)$$

and $HH^T = r$. The covariance r is for the underlying normal set of random draws, and the covariance R, for the nonnormal set of random features, is obtained from (7).

$$p(x_1, \cdots x_N) = \left(\frac{1}{(2\pi)^N \|r\|}\right)^{1/2} \cdot \exp\left(-\frac{1}{2} f^{-1}(x)_N^T r^{-1} f^{-1}(x)_N\right) \cdot \prod_{m=1}^{N} \frac{df_m^{-1}(x_m)}{dx_m} \quad (9)$$

where

$$f^{-1}(x)_N \equiv (f_1^{-1}(x_1), \cdots, f_N^{-1}(x_N)).$$

In this case, $f^{-1}(\cdot)$ must be defined to evaluate the density.

Equations (8) and (9) display a class of nonnormal, nonwhite joint densities. These densities can be substituted into analyses of statistical pattern recognition [5], state estimation [4], and optimal detection [1]–[3] procedures to explore the effects of noise distribution and noise correlation on probabilities of misclassification and probabilities of detection.

The constructions are summarized in Table 4. In Table 4, independent is abbreviated "i." In the case of the Distribution Distortion and Correlation Distortion methods, the covariances and first marginal distributions are independently specified. The Inverse Probability method requires knowledge of the joint probability distribution, from which the first marginal distributions and covariances are derived.

Appendix

A. The Description of Random Numbers

A set of random numbers results from drawing numbers from a random process, for example, the roll of a die, one draw for each random number. The value for each feature is not fixed, but rather the range of values and the relative

Table 4 Summary of Constructions

CORRELATION DISTORTION METHOD:

$u =$ i. normal $\quad HH^T = r \quad v = Hu \quad x_n = f_n(v_n)$

Distributions given by $x = f_n(v)$ where v is normally distributed

Covariance:

$$R_{ij} = \frac{1}{2\pi}\int dv_i dv_j\, f_i(v_i) f_j(v_j) \cdot \left(\|r\|^{-1/2} e^{-1/2(v^T r^{-1} v) - \epsilon^{-1/2}(v_i^2 + v_j^2)}\right)$$

DISTRIBUTION DISTORTION METHOD:

$v_n =$ i. general $\quad HH^T = R \quad x = Hv$

First marginal density, with $\varphi_n(v) =$ density of v_n, is

$$p_n(x) = \frac{1}{\prod_{j=1}^{N} |H_{nj}|} \prod_{k=2}^{N} \int_{-\infty}^{\infty} ds_k\, \varphi_1\left(\frac{x - s_2}{H_{n1}}\right) \varphi_2\left(\frac{s_2 - s_3}{H_{n2}}\right) \cdots \varphi_{N-1}\left(\frac{s_{N-1} - s_N}{H_{n,N-1}}\right) \varphi_N\left(\frac{s_N}{H_{nN}}\right)$$

Covariance: R

INVERSE PROBABILITY METHOD:

$$x_1 = P^{-1}(U_1, \infty, \cdots)$$
$$x_2 = P^{-1}(U_2, \infty, \cdots | x_1)$$
$$\vdots$$
$$x_N = P^{-1}(U_N | x_1, \cdots x_{N-1})$$

Marginal distribution of x_n: $\quad P_n(x_n) = P(\infty, \cdots, x_n, \infty, \cdots)$

Covariance: $\quad R_{nm} = E[x_n x_m] - E[x_n]E[x_m]$

frequency of values are specified. Each feature may assume any value within specified ranges. Each random draw for a feature may be discretely or continuously distributed, assuming either a finite of infinite number of distinct values. The sets of random numbers resulting from all possible sequences of draws is the ensemble of similarly described systems.

A set of random numbers is fully described by a joint probability distribution. This joint distribution function is all there is to know about the set of random features. Once specified, the joint distribution determines all properties of a set of random features. The joint distribution specifies the relative frequency with which any resulting set of feature values, denoted $(r_1, r_2, \cdots)$, is observed with values less than $(x_1, x_2, \cdots)$.

$$P(x_1, x_2, \cdots) \equiv \text{ relative frequency } r_1 \le x_1, r_2 \le x_2, \cdots.$$

The concept of relative frequency is the intuitive notion of probability. The logically rigorous description of probability is the definition of a measure on subsets of feature values [26], [37]. Every subset of feature values is "weighted" by its relative frequency of observation within the ensemble. Measures rather than probability densities are required to describe probability distributions that apply finite weight to single points in configuration space in addition to finite weight to (Borel) subsets of configurations. If the measure is absolutely continuous with respect to Borel measure, then the density is the Radon–Nikodyn derivative of the probability distribution with respect to Borel measure [24].

Mean values of the features, averages over the ensemble of similarly described systems, are summations using the probability measure. Means are defined

$$E[f] \equiv \int f(x_1, x_2, \cdots)\, d^N P(x_1, x_2, \cdots)$$

with $f(\cdot)$ any summable function of the random features. In particular, the probability density[3] for the first marginal distribution of a feature, the correlations, and the statistical moments are readily defined given mean values

$$\begin{aligned} C_n(k) &= E[e^{-ikx_n}] \\ &= \text{ the characteristic function for } x_n \\ p_n(x) &= \frac{1}{2\pi}\int dk e^{ikx} C_n(k) \\ &= \text{ the probability density[3] of } x_n \\ \rho_{mn} &= \frac{(E[x_m x_n] - E[x_m]E[x_n])}{\sqrt{E[x_m^2] - E[x_m]^2}\sqrt{E[x_n^2] - E[x_n]^2}} \\ &= \text{ correlation between features } x_m \text{ and } x_n \end{aligned}$$

$$E\left[\prod_{k=1}^{N}(x_k)^{n_k}\right] = \text{ statistical moment of order } \sum_{k=1}^{N} n_k.$$

The characteristic function is the Fourier transform of the measure and a joint characteristic function is uniquely associated with each joint probability distribution.

The first marginal probability distribution specifies the distribution of a single feature without regard to any other feature

$$\begin{aligned} P_n(x_n) &= \text{ relative frequency } r_n \le x_n \\ &= P(\infty, \infty, \cdots, x_n, \infty, \cdots). \end{aligned}$$

Given ignorance, the normal (Gaussian) distribution is a natural choice for first marginal distribution. It is a natural choice both because it is the distribution to which sums of random draws tend, and because it is the minimum Boltzmann–Shannon entropy distribution. If a feature results from summing contributions from many constituent features, all of bounded variance and weakly correlated (strongly mixing) [26], [38], then the result is normal statistics. Also, among all distributions, the normal distribution assumes the least, i.e., has minimal information, among all distributions with a finite mean and variance.

Correlations are normalized second moments. The covariance of a set of random features is defined

$$\begin{aligned} R(j,k) &= E[(x_j - E[x_j])(x_k - E[x_k])] \\ &= E[x_j x_k] - E[x_j]E[x_k]. \end{aligned}$$

The correlations are obtained from the covariances by dividing the covariances by the standard deviations of the two features.

[3] In cases for which the density is defined.

The covariance of a finite set of random features is described by the matrix

$$R_{jk} \equiv R(j,k).$$

This matrix is nonnegative. For any choice of α_k

$$\sum_{k=1}^{N}\sum_{j=1}^{N} R_{jk}\alpha_j^*\alpha_k = E[\left|\sum_{k=1}^{N}\alpha_k(x_k - E[x_k])\right|^2] \geq 0.$$

Wide-sense-stationary (isotropic mean and shift invariant covariances) processes have power spectral densities. A generalization of stationary processes are harmonizable random processes [37] which are Fourier transforms of random measures. This random measure, a transform domain random process, also has correlations. The transform domain process is uncorrelated (has orthogonal increments) if and only if the original process is wide-sense-stationary [23]. In this event, the mean squared value of the transform domain random process is denoted the power spectral density. The transform itself is a random measure, denoted the spectral measure. Shift invariance of the means and covariances uniquely corresponds to the spectral measure lacking correlation between distinct regions in the transform domain. Accordingly, it is typical, rather than exceptional, to have "spectral" correlation in a random process. Only stationary processes lack this correlation. Strictly periodic cyclostationary processes exhibit correlations between pairs of transform domain features separated by multiples of a cycle frequency.

The characteristic function is the moment generating representation for a random number. The coefficients in the formal expansion of the characteristic function in powers of the transform variable k are the statistical moments

$$C(k) \approx \sum_{n=0}^{\infty} E[x^n]\frac{(-ik)^n}{n!}.$$

B. Central Limit Theorem

The Central Limit theorem implies that random draws, constructed as sums of independent random draws, for example by the methods of Section II-B, tend to be normally distributed. However, if nonnormal statistics is the goal of the construction, the conditions of the Central Limit theorem must be evaded. When nonnormal statistics are desired, a case with convergence to normal statistics so slow that the finite linear combinations of independent random draws do not exhibit normal statistics must be selected, or accessory nonlinear operations must be introduced.

There are closed classes of probability-density functions under convolution. As seen in Section II-C, random numbers with these density functions retain densities in the particular class as random numbers are summed. A function class is denoted closed under convolution if the convolution of two members of the class remains in the class. Probability distributions in a class all have the same functional form but differ by the values of parameters. For example, the normal distributions constitute a class of probability densities closed under convolution and parametrized by a mean and variance. Symbolically, denote each member of a class of densities as $p_c(x;\vec{a})$ with $\vec{a}$ an array of parameters. A class of densities is closed under convolution if

$$p_c(x;\vec{a}) * p_c(x;\vec{b}) = \int dy\, p_c(y;\vec{a})p_c(x-y;\vec{b}) = p_c(x;\vec{c}) .$$

In particular, for the class of normal densities

$$n(x;(m,\sigma)) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right)$$

$$n(x;(m_1,\sigma_1)) * n(x;(m_2,\sigma_2)) = n(x;(m_1+m_2,\sqrt{\sigma_1^2+\sigma_2^2})) .$$

Reference to the results presented in Section II-C finds that if the independent random draws used in constructions of correlated draws all come from a single, closed under convolution class, then the correlated feature must necessarily have a probability density in the class. We might conclude that if a nonnormal class is selected, that nonnormal statistics must necessarily result. Unfortunately, this construction method does not ensure that the resulting distribution is nonnormal. As large numbers of random draws are summed, their densities tend to converge on even more limited subclasses of densities, called stable [39] (or limit) densities. The closed under convolution classes contain stable densities. After many convolutions, the resulting parameters of the densities tend to converge upon certain limit values, and densities with these distinguished parameters are the stable densities.

The normal density class is distinguished by being the class of limit densities for all independent random numbers with finite second moments [26]. Thus the statistics of sums of sufficiently many independent, finite variance random draws will be normal, as long as the sum is not dominated by a few terms. This is the content of the Lindeberg–Feller Central Limit theorem.

Examples of closed under convolution distribution classes with the normal densities as limit distributions are the Poisson probability distribution given by the formal "densities"

$$p(x) = \sum_{k=0}^{\infty}\delta(x+a-k)\frac{a^k}{k!}e^{-a}$$

and the gamma densities

$$p(x) = a^{\nu}x^{\nu-1}e^{-ax}/\Gamma(\nu).$$

An example of a limit distribution with a divergent variance is the Cauchy probability density function

$$p(x) = \frac{1}{\pi w(1+(x/w)^2)}.$$

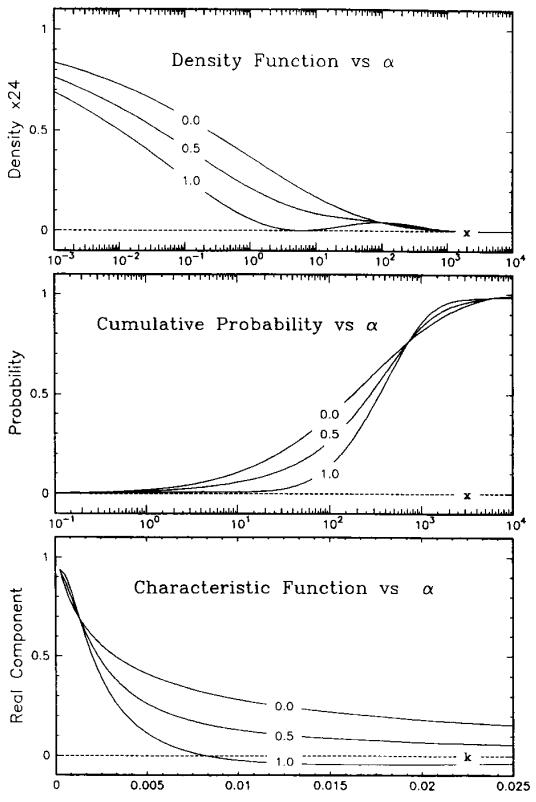

Fig. 1. Example distributions with identical moments.

C. The Problem of Moments

Joint probability distributions and joint characteristic functions are unique descriptions of a set of random numbers. However, there is not a unique correspondence of statistical moments and probability distributions. In particular, for common nonnormally distributed random numbers, probabilities of detection, and false alarm cannot be accurately estimated even if all the statistical moments are known. We cannot hope, in general, to determine a distribution by measuring its moments.

The "Problem of Moments" [40] is the study of correspondence of distributions and statistical moments. Even though a characteristic function uniquely defines a distribution and the characteristic function is formally a power series in the moments, distinct probability-density functions can have identical moments, identical to all orders. This result is a manifestation of the distinction between analytic and infinitely differentiable functions.

The lack of unique correspondence of moments and distributions is not restricted to pathological distributions. The one parameter family of probability density functions [39]

$$p_\alpha(x) = \frac{1}{24}(1 - \alpha \sin(x^{\frac{1}{4}}))\, e^{-x^{1/4}}$$

with $0 < \alpha < 1$, all have the same moments, $E[x^n] = (4n + 3)!/6$, which are independent of α. To every distribution uniquely defined by its moments, we can add a small proportion of $p_\alpha(x)$ and construct a distribution approximately equal to the original distribution, for which there is no longer a unique correspondence of moments with the distribution. Thus there are infinitely many distributions not uniquely associated with their moments near every distribution uniquely defined by its moments.

Nevertheless, the characteristic functions of probability-density functions are unique, i.e., the Fourier transform of $p_\alpha(\cdot)$ depends on α

$$C_\alpha(k) = \frac{1}{24ik} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(1 + m/4)}{m!\,(ik)^{m/4}} \cdot \left(1 + \frac{i\alpha(1-i)^m}{2} - \frac{i\alpha(1+i)^m}{2}\right).$$

This family of densities, their probability distributions, and characteristic functions are shown in Fig. 1, to emphasize that the example is not a pathological one.

The result is significant, from our constructive point of view, since it is not only lack of specification of the higher joint moments that prevents construction of a random number given a description of moments. For some distributions, even specification of all the moments is not sufficient to specify the distribution. As the plotted examples show, even within the two sigma points of the distributions, there can be significant variation among distributions that share moments. Without prior information, distributions must be estimated from histograms of observed values, and consequently, large numbers of samples will be required to accurately estimate the "tails" of the distributions.

As might be anticipated, unique correspondence of a distribution with a sequence of moments can be related to convergence of the expansion for the characteristic function in moments. A sufficient condition for unique correspondence limits the growth of the moments with order. The moments uniquely determine the density function if

$$\sum E[x^n]^{-1/2n} = \infty$$

i.e., the sum diverges. Necessary conditions also exist [40].

References

[1] D. Middleton, *An Introduction to Statistical Communication Theory*. New York: McGraw-Hill, 1960.

[2] ——, "Canonically optimum threshold detection," *IEEE Trans. Informat. Theory*, vol. IT-12, no. 2, pp. 230–243, 1966.

[3] H. L. van Trees, *Detection, Estimation, and Modulation Theory: Part I*. New York: Wiley, 1968.

[4] *Applied Optimal Estimation*, A. Gelb, Ed. Boston, MA: The M.I.T. Press, 1989.

[5] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles*. Reading, MA: Addison-Wesley, 1975.

[6] A. M. Maras, "Locally optimum detection in moving average non-Gaussian noise," *IEEE Trans. Commun.*, vol. 36, no. 8, pp. 907–912, 1988.

[7] A. B. Martinez, P. F. Swaszek, and J. B. Thomas, "Locally optimum detection in multivariate non-Gaussian noise," *IEEE Trans. Informat. Theory*, vol. IT-30, pp. 815–822, 1984.

[8] S. T. Li and J. L. Hammond, "Generation of pseudorandom numbers with specified univariate distributions and correlation coefficients," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-5, no. 5, pp. 557–561, 1975.

[9] D. E. Knuth, *The Art of Computer Programming, Seminumerical Algorithms* vol. 2. Reading, MA: Addison-Wesley, 1969, sec. 3.2.1.
[10] *Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun, Eds. Washington, DC: U.S. Gov. Printing Office, 1964, sec. 26.8.
[11] T. G. Lewis, *Distribution Sampling for Coupled Simulation*. Toronto, Ont., Canada: Lexington Books, 1975.
[12] R. Y. Rubinstein, *Simulation and the Monte Carlo Method*. New York: Wiley, 1981.
[13] P. Bratley, B. L. Fox, and L. E. Schrage, *A Guide to Simulation*. New York: Springer-Verlag, 1983.
[14] C. M. Rader, L. R. Rabiner, and R. W. Schafer, "A fast method of generating digital random numbers," *Bell Syst. Tech. J.*, pp. 2303–2310, 1970.
[15] W. H Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*. Cambridge, UK: Cambridge Univ. Press, 1988, ch.. 7.
[16] B. W. Schmeiser, "Random process generation," Tech. Rep. DTIC-ADA086758, June 1980.
[17] G. M. Dilliard, "Generating random numbers having probability distributions occuring in signal detection problems," *IEEE Trans. Informat. Theory*, pp. 616–617, 1967.
[18] G. Picci, "Stochastic realization of Gaussian processes," *Proc. IEEE*, vol. 64, no. 1, pp. 112–122, 1976.
[19] J. M. Geist, "Computer generation of correlated Gaussian random variables," *Proc. IEEE*, vol. 67, no. 5, pp. 862–863, 1979.
[20] L. E. Franks and W. A. Gardner, "Characterization of cyclostationary random signal processes," *IEEE Trans. Informat. Theory*, vol. IT-21, no. 1, pp. 4–14, 1975.
[21] P. G. Hoel, *Introduction to Mathematical Statistics*, 5th ed. New York: Wiley, 1984.
[22] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1984.
[23] I. M. Gel'fand, and N. Ya. Vilenkin, *Generalized Functions*, vol. 4. New York: Academic Press, 1964, ch. III.
[24] W. Rudin, *Real and Complex Analysis*. New York: McGraw-Hill, 1974.
[25] T. Kailath, "Innovations approach to detection and estimation," *Proc. IEEE*, vol. 58, no. 5, pp. 680–695, 1970.
[26] *Fundamentals of Mathematics*, vol. III, H. Behnke, F. Bachmann, K. Fladt, and W. Süss, Eds. Boston, MA: MIT Press, 1986, ch. 3a.
[27] R. A. Roberts and C. T. Mullis, *Digital Signal Processing*. Reading, MA: Addison-Wesley, 1987.
[28] B. Liu and D. C. Munson Jr., "Generation of a random sequence having a jointly specified marginal distribution and autocovariance," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, no. 6, pp. 973–983, 1982.
[29] L. S. Dewald and P. A. W. Lewis, "A new laplace second-order autoregressive time-series model-NLAR(2)," *IEEE Trans. Informat. Theory*, vol. IT-31, no. 5., pp. 645–651, 1985.
[30] D. L. Jagerman and B. Melamed, "The transition and autocorrelation structure of TES processes, Part I: General theory," *Commun. Statist.- Stochastic Models*, vol. 8, no. 2, pp. 193–219, 1992.
[31] ——, "The transition and autocorrelation structure of TES processes, Part II: Special cases," *Commun. Statist.- Stochastic Models*, vol. 8, no. 3, pp. 499–527, 1992.
[32] B. Simon, *The $P(\phi)_2$ Euclidean (Quantum) Field Theory*. Princeton, NJ: Princeton Univ. Press, 1974, ch. 1, pp. 9–11.
[33] G. C. Wick, "The evaluation of the collision matrix," *Phys. Rev.*, vol. 80, p. 268, 1950.
[34] L. Isserlis, "On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables," *Biometrika*, vol. XII, pp. 134–139, 1918.
[35] W. A. Gardner, *Introduction to Random Processes: with Applications to Signals and Systems*. New York: McGraw-Hill, 1990, pp. 115–117.
[36] W. J. Szajnowski, "The generation of correlated Weibull clutter for signal detection problems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES–13, no. 5, pp. 536–540, 1977.
[37] M. Loève, *Probability Theory*. New York: D. van Nostrand, 1963, ch. X.
[38] G. Jona-Lasino, "Probabilistic approach to critical behaviour," in *New Developments in Quantum Field Theory and Statistical Mechanics*, M. Lévy and P. Mitter, Eds. New York: Plenum, 1977.
[39] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume II*. New York: Wiley, 1966.
[40] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*. Providence, RI: Am. Math Soc., 1963.

Glenn E. Johnson (Member, IEEE) received the B.S. degree in physics from Harvey Mudd College, Claremont, CA, in 1972, and the M.S. and Ph.D. degrees in physics from the University of Maryland, College Park, in 1976 and 1978, respectively.

From 1979 to 1982, he was in the Advanced Radar Systems and Technology Department of Texas Instruments, Dallas, involved in design studies of airborne radar systems. From 1985 to 1991, he was a senior scientist in the Communications, Radars, and Sensors Division of Mission Research Corporation, Santa Barbara, CA, involved primarily in aircraft countermeasures and inserting machine intelligence into avionics. From 1982 to 1985 and since 1991, he has been on the technical staff of The Analytic Sciences Corporation, Reston, VA. His interests are mathematical physics, algorithms, and the design and digital simulation of radar and communication systems.