NOISE AND INTERFERENCE MODELING

Noise and interference are inevitable realities of collecting data or taking measurements in the real world. In some cases, the noise level may be so insignificant as to allow the engineer to ignore its effects. However, this situation generally only occurs in very controlled circumstances, such as those in the laboratory or when signal powers are exceptionally large relative to the noise.

Unfortunately, it is more generally the case that the noise and interference cannot be ignored. Rather, design and analysis must be done with careful attention to the corruptive effects of these disturbances. One way to ensure an effective final design is to have accurate models of the noise components of the signals of interest. Examples of the impact of high and low levels of noise on the observation of a sinusoid are shown in Fig. 1.

Modeling these effects can range from being relatively straightforward to being rather difficult if not impossible. To assist in this endeavor, it is the purpose of this article to describe various methods of characterizing noise and interference signals and to elaborate on some of the most popular models in practice today.

J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*. Copyright © 1999 John Wiley & Sons, Inc.

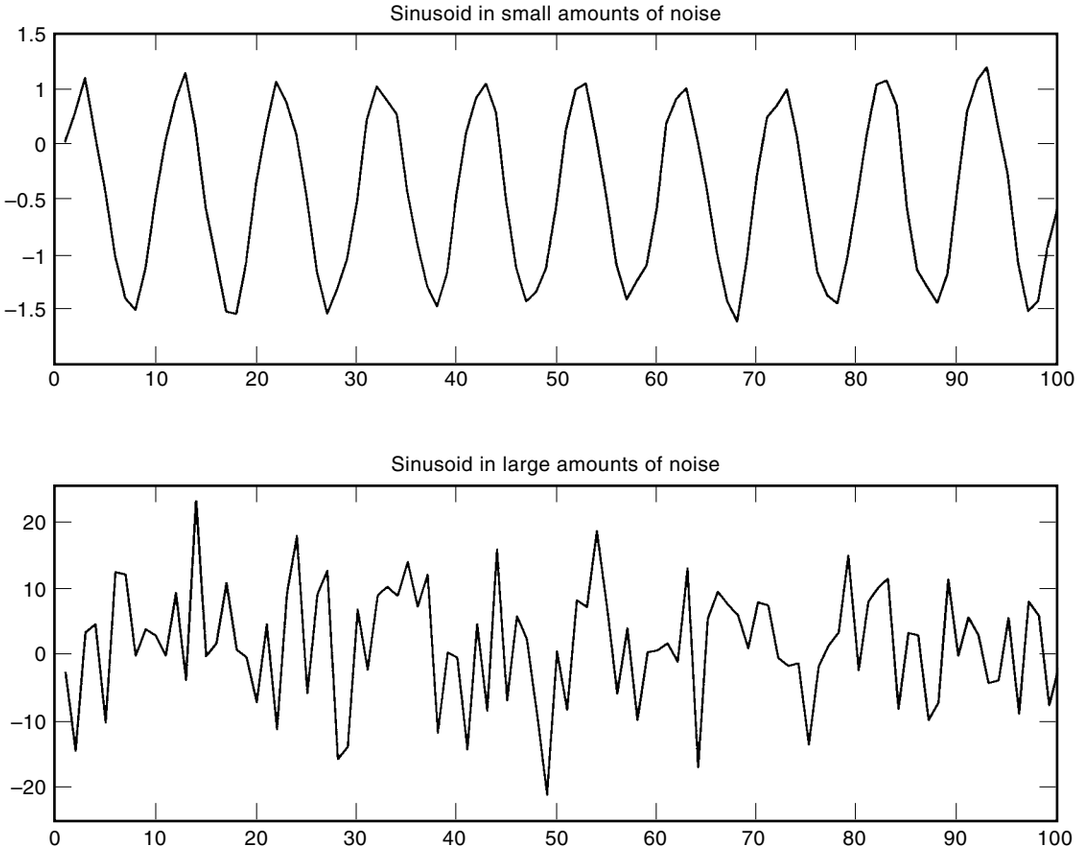

Figure 1. Plots of a common sinusoid embedded in low-level noise and high-level noise.

STATISTICAL DESCRIPTIONS OF NOISE

Because of its random nature, characterizing noise necessarily requires a variety of statistical methods. For applications involving the observation of continuous time or sampled signals, the most complete statistical description of the noise process is the so-called joint distribution of the process (1). Throughout this article we will denote an arbitrary noise or interference signal by $x(t)$.

Definition 1 The nth order joint distribution of $x(t)$, denoted $F_x(\boldsymbol{x}; \boldsymbol{t})$, is given by

$$F_{\boldsymbol{x}}(\boldsymbol{x};\boldsymbol{t}) = Pr[x(t_1) \le x_1 \qquad x(t_2) \le x_2, \ \ldots \qquad x(t_n) \le x_n] \quad (1)$$

where the length n vectors $\boldsymbol{x}$ and $\boldsymbol{t}$ represent the collections $\{x_1, x_2, \ldots, x_n\}$ and $\{t_1, t_2, \ldots, t_n\}$, respectively, and where, for example, $Pr[x(t_1) \le x_1]$ is the probability that the signal $x(t)$ at time t_1 is less than the constant x_1. For $F_x(\boldsymbol{x}; \boldsymbol{t})$ to be completely defined, it must be computed for all collections of times $t_1, t_2, \ldots, t_n$, vectors $\boldsymbol{x}$, and all integer values of n.

In modeling noise, one often wishes to know if the statistical characterization of the signal $x(t)$ changes with time. The most rigorous method for determining this important property is derived directly from the joint distribution.

If the joint distribution of the noise is shift invariant, that is, if

$$Pr[x(t_1) \le x_1;\ x(t_2) \le x_2, \ \ldots, \ x(t_n) \le x_n] = Pr[x(t_1 + \tau) \le x_1;\ x(t_2 + \tau) \le x_2, \ \ldots, \ x(t_n + \tau) \le x_n] \quad (2)$$

for all collections of times and vectors $\mathbf{x}$ and for all choices of n, then $x(t)$ is said to be *strictly stationary*. In this case, the entire statistical description of the process is a function of only the relative locations of the samples with regard to one another rather than the absolute locations of the time samples.

Unfortunately, in general, determining $F_x(\boldsymbol{x}; \boldsymbol{t})$ for large values of n is almost always impractical. However it may often be the case that one has sufficient information so as to be able to compute this important function for values of $n = 1$ or 2. In this case we may introduce a weaker form of stationarity that has many important applications and uses in practice. This type of characterization is often referred to as the so-called *second order* description of $x(t)$.

Under this characterization we seek to describe the mean and correlation of the noise signal rather than the instantaneous probabilities. Prior to presenting this definition, it is advantageous for us to supply the following definitions which will be used throughout the remainder of this article:

Definition 2 The amplitude probability density function of the signal $x(t)$ is given by

$$f_x(x;\, t) = \frac{\partial F_{\boldsymbol{x}}(x;\, t)}{\partial x} \quad \text{for all choices of } x \text{ and for all values of } t \tag{3}$$

while the joint amplitude density function of $x(t)$ is given by

$$f_{\boldsymbol{x}}(x_1,\, x_2;\, t_1,\, t_2) = \frac{\partial \partial F_{\boldsymbol{x}}(x_1,\, x_2;\, t_1,\, t_2)}{\partial x_1 \partial x_2} \quad \text{for all choices of } x_1 \text{ and } x_2 \text{ and for all } t_1 \text{ and } t_2 \tag{4}$$

These two definitions completely characterize all of the first- and second-order properties of $x(t)$ (e.g., mean values, signal energy and power, correlations and frequency content). In particular, the amplitude density function describes the statistical range of amplitude values of signal x at time t. Further, one should recall from basic probability theory (2) that the density functions $f_x(x;\, t)$ and $f_x(x_1,\, x_2;\, t_1,\, t_2)$ can be readily obtained from the nth joint distribution of the signal $x(t)$ given in Eq. (1) through simple limiting evaluations. Given these two density functions, we may define the second-order description of the noise or interference signal as follows:

Definition 3 The mean or average value of the noise as a function of time is defined as

$$\mu_x(t) = E[x(t)] \tag{5}$$

where the expectation is taken with respect to $f_x(x;\, t)$. The correlation function of the noise signal is defined as

$$R_x(t_1,\, t_2) = E[x(t_1)x^*(t_2)] \tag{6}$$

and where the above expectation is taken with respect to $f_x(x_1,\, x_2;\, t_1,\, t_2)$, and where x^* denotes the complex conjugate of $x(t)$.

If the mean function is a constant with respect to time (the average value of the noise does not change with time) and the correlation function of the noise is a function of only the difference of times and not their absolute locations, that is, $R_x(t_1,\, t_2) = R_x(0,\, t_2 - t_1)$, then the noise is said to be *wide-sense stationary*. In this case we simplify the notation by letting the mean be μ and the correlation function be given by $R_x(\tau)$ where $\tau = t_2 - t_1$.

A signal $x(t)$ being wide-sense stationary implies that the average statistical properties as measured by the mean and the correlation function do not depend on when one observes the noise and interference. This is a very beneficial property since designers attempting to combat this noise are not required to include an absolute clock in their design, rather a fixed design will always be optimal.

As an important observation, one should note that if the noise happens to be strictly stationary then the noise will also be wide-sense stationary. This is a direct consequence of the fact that the mean and correlation functions are expected values taken with respect to the first- and second-order joint density functions which, by supposition, are shift invariant. In general, the converse is not true. However, we will see later that among the most popular models for the noise, this will, in fact, be the case.

Power Spectral Density

Engineers are often more comfortable in describing or analyzing signals in the frequency domain as opposed to the statistical (or time) domain. This is equally true when dealing with random signals. However, random signals pose a few complications when one attempts to apply spectral techniques to them. This is a consequence of the fact that each time one makes an observation of this signal, the so-called noise *realization* (observation) is different and will therefore likely have a different spectrum. Moreover, many of these potential noise realizations might have infinite energy, and will therefore not have well-defined Fourier transforms.

Because of these complicating factors, we must consider the power spectrum (rather than the standard energy spectrum) averaged over all possible realizations to obtain a meaningful definition of the spectrum of a random signal. This approach leads to the well-known and oft-used *power spectral density* of the noise or interference as the basic frequency domain description of $x(t)$.

To begin, assume that we only observe the noise signal $x(t)$ from $t = -T$ to $t = +T$ (we will later let T approach infinity to allow for the observation of the entire signal). The Fourier transform of this observed signal is

$$X_T(\omega) = \int_{-T}^{T} x(t)e^{-j\omega t}\, dt \tag{7}$$

We may then write the squared magnitude of $X_T(\omega)$ as follows:

$$|X_T(\omega)|^2 = \int_{-T}^{T}\int_{-T}^{T} x(t_1)x^*(t_2)e^{-j\omega(t_1 - t_2)}\, dt_1\, dt_2 \tag{8}$$

As described above, for the definition to have any utility, we need to compute the average magnitude squared power spectrum, which requires that we must evaluate the expected value of $|X_T(\omega)|^2$ and then normalize this expectation by the length of the observation of $x(t)$ (this prevents the spectrum from blowing up with power signals). This leads to the following expression:

$$\frac{1}{2T}E[|X_T(\omega)|^2] = \frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T} R_x(t_1 - t_2)\, dt_1\, dt_2 \tag{9}$$

$$= \int_{-T}^{T}\left[1 - \frac{|\tau|}{2T}\right] R_x(\tau)\, d\tau \tag{10}$$

where the second equation arises from a transformation of variables. The final step in the definition is to allow the observation window to grow to $(-\infty,\, +\infty)$, that is, to take the limit of Eq. (10) as T tends to infinity. From this we arrive at the following definition which holds for all wide sense stationary random signals.

Definition 4 The power spectral density of a wide sense stationary random signal $x(t)$ is given by

$$S_X(\omega) = \int_{-\infty}^{\infty} R_x(\tau)e^{-j\omega\tau}\, d\tau \tag{11}$$

Note that as a consequence of the preceding analysis, the power spectral density turns out to simply be the Fourier transform of the correlation function. Thus the correlation

function of a random signal contains all the information necessary to compute the average spectrum of the signal and, furthermore, there is a one-to-one relationship between the average time dependency between samples (as measured by the correlation function) and the average frequency content in the signal.

The term "power spectral density" arises from the fact that $S_X(\omega)$ behaves like a probability density function. First, just like a standard density function, it is always non-negative, and second, when integrated over a certain frequency range one obtains the average power of the signal over that frequency range. That is, computing the integral

$$\int_{-\omega_2}^{-\omega_1} S_X(\omega)\,d\omega + \int_{\omega_1}^{\omega_2} S_X(\omega)\,d\omega \tag{12}$$

results in the total power in the signal $x(t)$ between the frequencies ω_1 and ω_2 (Note: we must consider both positive and negative frequencies.) This is just like the situation when one integrates a probability density function over a particular range to obtain the probability that the random variable will fall into that range.

Now given these few basic definitions, we may proceed to characterize various forms of noise or interference. From the statistical perspective, what differentiates various types of noise are, first and foremost, the specific joint distribution of the process. Since these joint density functions are exceedingly difficult to obtain, we are often relegated to comparing and classifying noise signals through $f_X(x;\, t)$ and $f_X(x_1,\, x_2;\, t_1,\, t_2)$, as well as the mean and correlation function or power spectral density.

NOISE MODELS

White Noise

White noise is a model used to describe any noise or interfering signal which has essentially equal power at all frequencies. More precisely, signals with a power spectral density given by

$$S_X(\omega) = \frac{N_0}{2} \qquad \text{for all frequencies} \tag{13}$$

are said to be white because, similar to white light, these signals have equal power at all (visible) frequencies. Through simple integration it is easily shown that the amount of power contained in white noise over an angular frequency range of B radians per second is N_0B. (Note that the factor of 2 in the definition of the power spectral density accounts for the two-sided nature of the Fourier transform.)

It is important to note that this definition of white noise is not an explicit function of either the amplitude density function or the joint density function given in Eqs. (3) and (4), but rather a function of only the power spectral density of the signal. Thus, any noise source, irrespective of the specific forms of $f_X(x;\, t)$ and $f_X(x_1,\, x_2;\, t_1,\, t_2)$ can be white noise, so long as the power spectral density is a constant.

The corresponding correlation function of white noise is obtained through the inverse Fourier transform as

$$R_x(\tau) = \frac{N_0}{2}\,\delta(\tau) \tag{14}$$

where $\delta(\tau)$ is the familiar Dirac delta function (3). Here one should observe from the above equation that white noise has zero correlation between all time samples of the signal, no matter how close the time samples are to one another. This does not imply that all samples of white noise are independent from one another, rather that they simply have zero correlation between them.

Unfortunately, by integrating the power spectral density over all frequencies we observe that the total power in white noise is infinite, which, of course, requires more energy than exists in the entire universe. Therefore, it is physically impossible for white noise to exist in nature. Yet, there are many noise processes which have essentially a constant power spectral density over a fixed and finite frequency range of interest, which suggests that white noise may be an appropriate and simplifying model. Yet, the real utility of the white noise model is in describing other "nonwhite" noise signals.

To demonstrate this, consider the output of a linear time-invariant system with frequency response $H(\omega)$ to a white noise input. Under very general conditions, it can be shown (2) that this output, given by $y(t)$, is a noise signal with power spectral density given by

$$S_Y(\omega) = \frac{N_0}{2}\,|H(\omega)|^2 \tag{15}$$

Now, conversely, let us assume that we observe some nonwhite noise signal with power spectral density $S_Y(\omega)$. If $\int \ln S_Y(\omega)/1 + \omega^2\,d\omega > -\infty$, then $y(t)$ can be modeled as the output of *some* linear time invariant system with a white noise input. That is, there will exist some well-defined system with frequency response given by $H(\omega)$ such that $N_0/2|H(\omega)|^2 = S_Y(\omega)$. These processes are, in fact, physically realizable and constitute all random signals seen in practice. Thus, from a spectral perspective we may replace essentially all random signals seen in nature by a linear system driven by a white noise signal.

Gaussian Noise

Gaussian noise is the most widely used model for noise or interference both in practice and in the scientific literature. There are two primary reasons for this: (1) many observations in nature follow the Gaussian law, and (2) the Gaussian density function is mathematically easy to analyze and manipulate.

The reason that so many natural observations are Gaussian arises from the so-called *Central Limit Theorem.* This theorem is stated as follows:

Theorem 1 Assume that Z_i are independent, and identically distributed random variables, each with mean μ_Z and finite variance σ_Z^2. Then the so-called "normalized sum" given by

$$Y_n \frac{1}{\sqrt{n}} \sum_{i=1}^{n-1} \frac{Z_i - \mu_Z}{\sigma_Z} \tag{16}$$

approaches a standard Gaussian random variable with mean equal to zero and unit variance, irrespective of the original distribution of the random variables Z_i.

What this very powerful theorem establishes is that if a particular observation of interest is a combination of infinitely many small and independent random components, then the combination, when properly normalized, is Gaussian Random variable with density function given by

$$f_{\boldsymbol{x}}(x;\, t) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{17}$$

This might suggest that all natural observations are Gaussian since most physical phenomena are impacted by many random events; however, extensive experimental results establish otherwise (4). Therefore, in addition to various forms of Gaussian noise, later in this article we will present a number of other "non-Gaussian" noise models.

Definition 5 A random noise signal is said to a Gaussian signal if all nth-order joint density functions are jointly Gaussian, that is, if all joint density functions are of the form

$$f_{\boldsymbol{x}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{det(\boldsymbol{K})}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{K}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right] \tag{18}$$

where the vector $\boldsymbol{\mu}$ and the matrix $\boldsymbol{K}$ represent the mean vector and covariance matrix (2), respectively, of the vector $\boldsymbol{x}(t_1), \boldsymbol{x}(t_2), \ldots, \boldsymbol{x}(t_n)$.

The important feature here is we see that for any value n, the joint statistics of a Gaussian signal are completely determined from the mean and correlation functions of $x(t)$. Furthermore, when the signal is in addition wide-sense stationary, then the joint statistics are shift invariant, and as such, the signal is also strictly stationary.

Gaussian noise is further differentiated by the nature of the mean functions and the correlation functions. When the correlation function is given by $R_X(\tau) = N_0/2\delta(\tau)$, then the signal is said to be *Gaussian White Noise,* which is the single most common model in noise analysis. In this case, each time sample will be independent from all others and have infinite variance. Other well-known Gaussian models which we describe below are Wiener noise, Thermal noise, and Shot noise.

Wiener Noise

The Wiener Process is the limiting case of a random walk signal. A random walk is a popular abstraction for describing the distance an individual is from home when this individual randomly steps forward or backward according to some probability rule. The model is as follows: Every T seconds an individual will take a length s step forward with probability 1/2 or a length s step backward with probability 1/2. Therefore, at time nT, the position of the random walk will be $x(nT) = ks - (n-k)s$, where n is the total number of steps taken, and k is the number of forward steps taken. For large values of time, it can be shown using the DeMoivre–Laplace Theorem (2) that

$$Pr(x(nT) = ks - (n-k)s) \approx \frac{1}{\sqrt{n\pi/2}} e^{-(2k-n)^2/2n} \tag{19}$$

or that the location of the walker is governed approximately by a Gaussian law.

Now, let us extend this by considering the limiting case where the length of time between steps tends to zero ($T \to 0$). It is easy to show that $E[x^2(nT)] = ns^2 = ts^2/T$ for $t = nT$. To ensure that we do not have a variance going to zero (the walker stops) or blowing up (not physically realizable), let us also add the condition that the size of the step be proportional to the square root of the time between steps, that is $s^2 = \alpha T$. Then the well-known Wiener Process is the signal $x(t)$, which is the limit of the random walk under these conditions. Examples of this are given in Fig. 2, where we have depicted realizations from three succeeding approximations to the Wiener process. In all figures, we have let the proportionality constant $\alpha = 1$. In Fig. 2(a) T and s equal 1, in Fig. 2(b) $T = 0.1$ and $s = \sqrt{0.1}$, and in Fig. 2(c) $T = 0.01$ and $s = 0.1$. One can see that as the approximation gets closer to the limiting value it becomes smoother (more continuous) and begins to exhibit long-term trends. This, of course, makes sense because the steps are getting much smaller and, as a consequence, it takes longer periods of time for the walker (signal) to change positions (values).

It should be pointed out that as $T \to 0$, the position of the random walk at any point in time is the sum of infinitely many random steps, and therefore by the Central Limit Theorem, the location at any point in time is a Gaussian random variable with zero mean and variance given by αt.

To compute the correlation function of the Wiener process, we note that $x(t_2) - x(t_1)$ is independent of $x(t_1)$ due to the fact that, by supposition, the random movements between times t_1 and t_2 are independent of the random movements up to time t_1. From this it must be that

$$E[(x(t_2) - x(t))x(t_1)] = E[x(t_2) - x(t_1)]E[x(t_1)] = 0 \tag{20}$$

from which it is easy to show that $R_X(t_1, t_2) = \alpha t_1$ whenever $t_2 > t_1$. Determining the correlation for the case that $t_1 > t_2$ results in the final value for the correlation function of the Wiener process as

$$R_X(t_1, t_2) = \alpha \min(t_1, t_2) \tag{21}$$

One can readily see that the Wiener process is not wide-sense stationary and therefore does not have a computable power spectral density. Nevertheless, it is a highly useful, physically motivated model for Gaussian data with relatively smooth realizations and which exhibits long-term dependencies.

Thermal Noise

Thermal noise is one of the most important and ubiquitous sources of noise in electrical circuitry. It arises primarily from the random motion of electrons in a resistance and occurs in all electronic circuits not maintained at absolute zero degrees Kelvin. This resulting low-level noise is then scaled to significant levels by typical amplifiers used to amplify other low-level signals of interest.

To derive a model for Thermal noise, let us assume that we have a conducting rod of length L and cross-sectional area A for which we expect to measure low-level random voltages. Let $V_{x,k}(t)$ denote the component of velocity in the x direction of the kth electron at time t. The total current denoted by $I(t)$ is the sum of all electron currents in the x direction

$$I(t) = \sum_{k=1}^{nAL} i_k(t) = \sum_{k=1}^{nAL} \frac{q}{L} V_{x,k}(t) \tag{22}$$

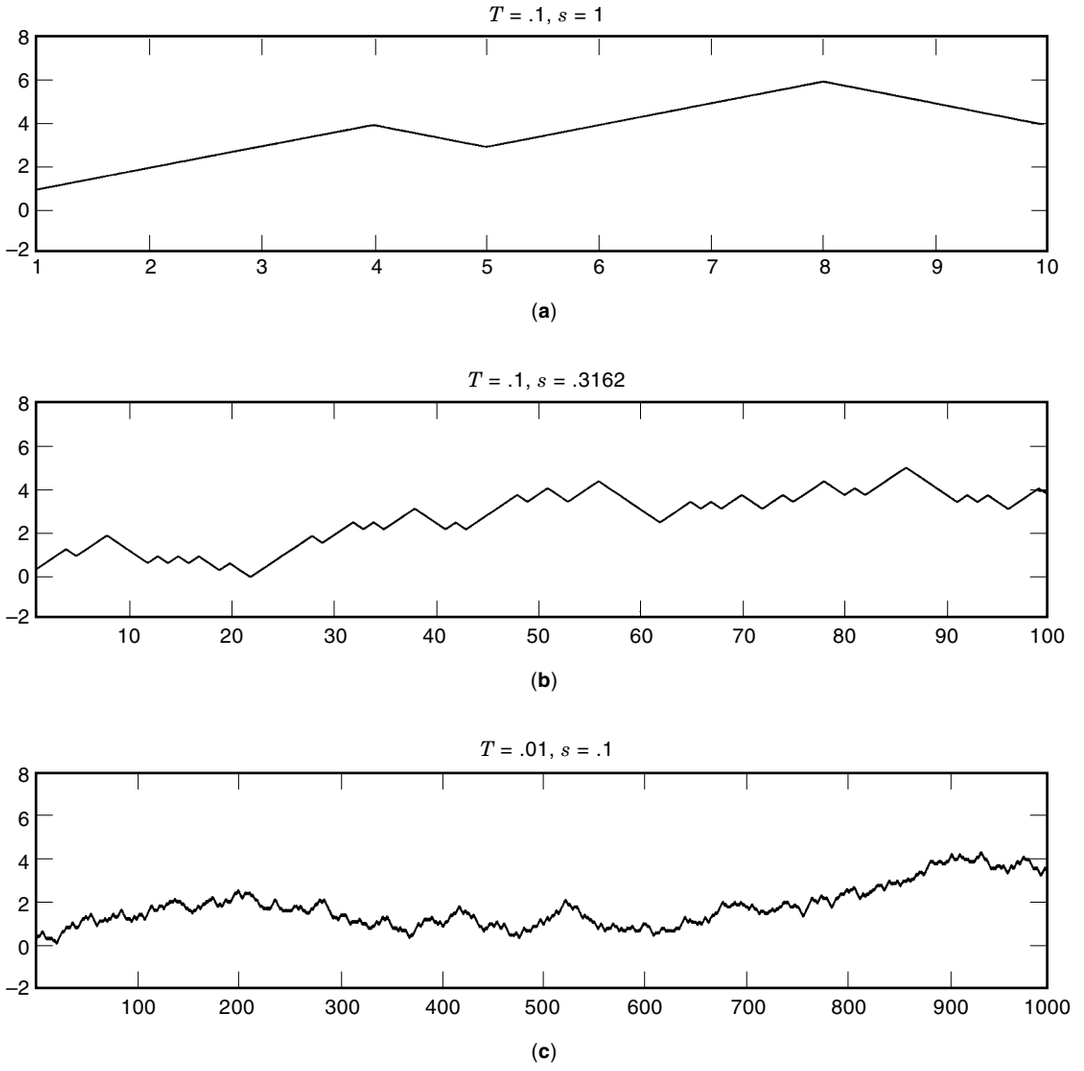

Figure 2. Plots of realizations from numerical approximations to the Wiener process for various values of s and T. In all cases, $\alpha = 1$. In (a), $T = 1$ and $s = 1$, in (b) $T = 0.1$ and $s = 0.3162$ and in (c) $T = 0.01$ and $s = 0.1$.

where n is the number of electrons per cubic centimeter and q is the charge of an electron. Let us assume that the average velocity of each electron is zero and that all electrons behave independently of one another and have an identical distribution. From this, it is easy to show that the correlation function of the current is given by

$$R_I(\tau) = \frac{nAq^2}{L} E[V_x(t)V_x(t+\tau)] \tag{23}$$

where we have dropped the subscript k for convenience. Assuming that electron collisions occur according to a Poisson law with intensity given by α (2) and that the average variance of the velocity of an electron is given by k_BT/m, where k_B is Boltzmann's constant and m is the mass of an electron (from the equipartition theorem for electrons), it can be shown that

$$R_I(\tau) = \frac{kT}{R} \alpha e^{-\alpha|\tau|} \tag{24}$$

where R is the resistance of the rod. Recognizing that $E(t) = RI(t)$, we can easily obtain the correlation function of the voltage as

$$R_E(\tau) = kT R\alpha e^{-\alpha|\tau|} \tag{25}$$

Note that as the intensity of the electron collisions α increases, the correlation function begins to approach a Dirac delta function, implying that the spectrum of thermal noise begins to better approximate white noise. Furthermore, we also observe that as we lower the temperature of the circuitry T, the power $R_E(0)$ in the thermal noise deceases proportionally. This, of course, suggests that if one wants to process very weak signals, for example, deep space signals, then the circuitry must be cooled using some form of coolant. The Central Limit theorem can be used to show that the voltage is very well approximated by a Gaussian process.

Shot Noise

Shot noise is used to model random and spontaneous emissions from dynamical systems. These spontaneous emissions are modeled as a collection of randomly located impulses (Dirac delta functions) given by $z(t) = \sum_i \delta(t - t_i)$ where the locations of the impulses in time (t_i) are governed by a Poisson law with intensity λ which is then driven through some system.

The output of a system with impulse response $h(t)$ to the input $z(t)$ is said to be "shot noise." From the convolution theorem, it is easily shown that this output is

$$x(t) = \sum_i h(t - t_i) \tag{26}$$

Using the properties of the Poisson law, is is easy to show that the mean value of shot noise is $\mu_x = \lambda H(0)$, where $H(\omega)$ is the frequency response (Fourier Transform) of the system. The power spectral density function of $x(t)$ is given by

$$S_X(\omega) = \lambda^2 H^2(0)\delta(\omega) + \lambda|H(\omega)|^2 \tag{27}$$

Thus the frequency content of the shot noise is determined entirely by the system $h(t)$ and the intensity of the Poisson process, that is, average rate of emissions. As before, because $x(t)$ at large values of time is composed of the superposition of many random elements [see Eq. (26)] $x(t)$ is often modeled as a wide-sense stationary Gaussian random process with power spectral density given by Eq. (27).

NON-GAUSSIAN MODELS

While the scope of the Central Limit Theorem might suggest otherwise, there are many data sets derived from the environment which do not conform well to the Gaussian model. Most of these data contain interfering signals emitted from a modest number of interferers or from interferers overlapping the signal of interest in the frequency domain in such a way that one could not easily remove these unwelcome elements by filtering.

Approaches to remedy this modeling problem fall into two categories: (1) generalize the Gaussian model to allow for statistical variation around the Gaussian process, or (2) attempt to derive new statistical models directly from the physics of the problem. In both cases, most non-Gaussian specifications typically do not go beyond the amplitude probability density function given in Eq. (3) and power spectral density. This is due, in large part, to the severe complexity of merely specifying valid and useful joint probability functions with the desired amplitude density.

These non-Gaussian models for noise and interference generally differ statistically from Gaussian noise in the rate at which the so-called "tail" of the amplitude density tends to zero. More specifically, for Gaussian noise the amplitude probability density function is approximated by $e^{-x^2/2\sigma^2}$ for large values of $|x|$. That is, the tails (the probability density function from some large value to $\pm$ infinity) of the amplitude density decay at an exponential rate with exponent equal to x^2. From a physical point of view, this rate translates into the relative frequency of observing large amplitude values from the noise signal. That is, amplitude density functions with tails which decay to zero faster than the Gaussian tail give rise to data sets with fewer large values in the noise than one might see with Gaussian data. On the other hand, when the tail of the amplitude density decays slower than that of the Gaussian, one is more likely to see large random values of the noise signal $x(t)$ than one would with Gaussian data. So, unlike differentiating noise models based on the spectrum of the noise, as seen in the preceding section, with non-Gaussian noise we typically differentiate various forms by the nature of the tail of the amplitude density function.

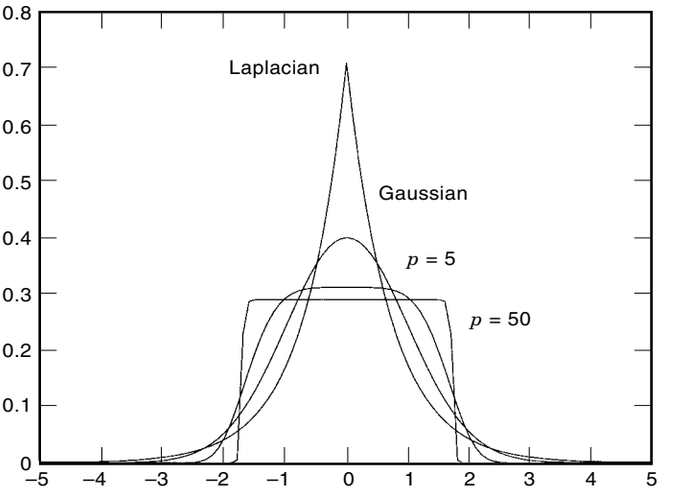

Figure 3. Plots of various specific amplitude density functions from the Generalized Gaussian model.

Generalized Gaussian Noise

The Generalized Gaussian model is the most straightforward extension of the standard Gaussian model. In this case, the amplitude probability density function is given by

$$f_X(x;\, t) = a(p)\exp\left(-\left[\frac{|x|}{A(p)}\right]^p\right) \tag{28}$$

where the constant p parameterizes the density function and where $A(p) = \sqrt{[\Gamma(1/p)/\Gamma(3/p)]}$, $a(p) = p/[2{*}\Gamma(1/p){*}A(p)]$ and where $\Gamma(z)$ is the gamma function (generalized factorial). It should be noted that, irrespective of the value of p, the variance of the amplitude density function is held constant (in this case the variance is arbitrarily set to one). Plots of the amplitude density function for $p = 1$, 2, 5, and 50 are depicted in Fig. 3. One can see that for small values of ($p < 2$), the tails of the density functions maintain relatively large values for large arguments. Conversely, for large values of ($p < 2$), the tails decay to zero quite rapidly.

From Eq. (28), it is easily recognized that for $p = 2$, the Generalized Gaussian model results in a standard Gaussian model. Alternatively, when $p = 1$, we obtain the well-known *Laplacian* model for noise. In this case, the tail of the amplitude density function decays at a rate of $e^{-|x|}$, thus implying that one will likely observe many more large values of noise than one would see with Gaussian noise. In general, for $p < 2$ we obtain tails with more mass (more probability), resulting in "impulsive" data, while for $p > 2$ we obtain tails

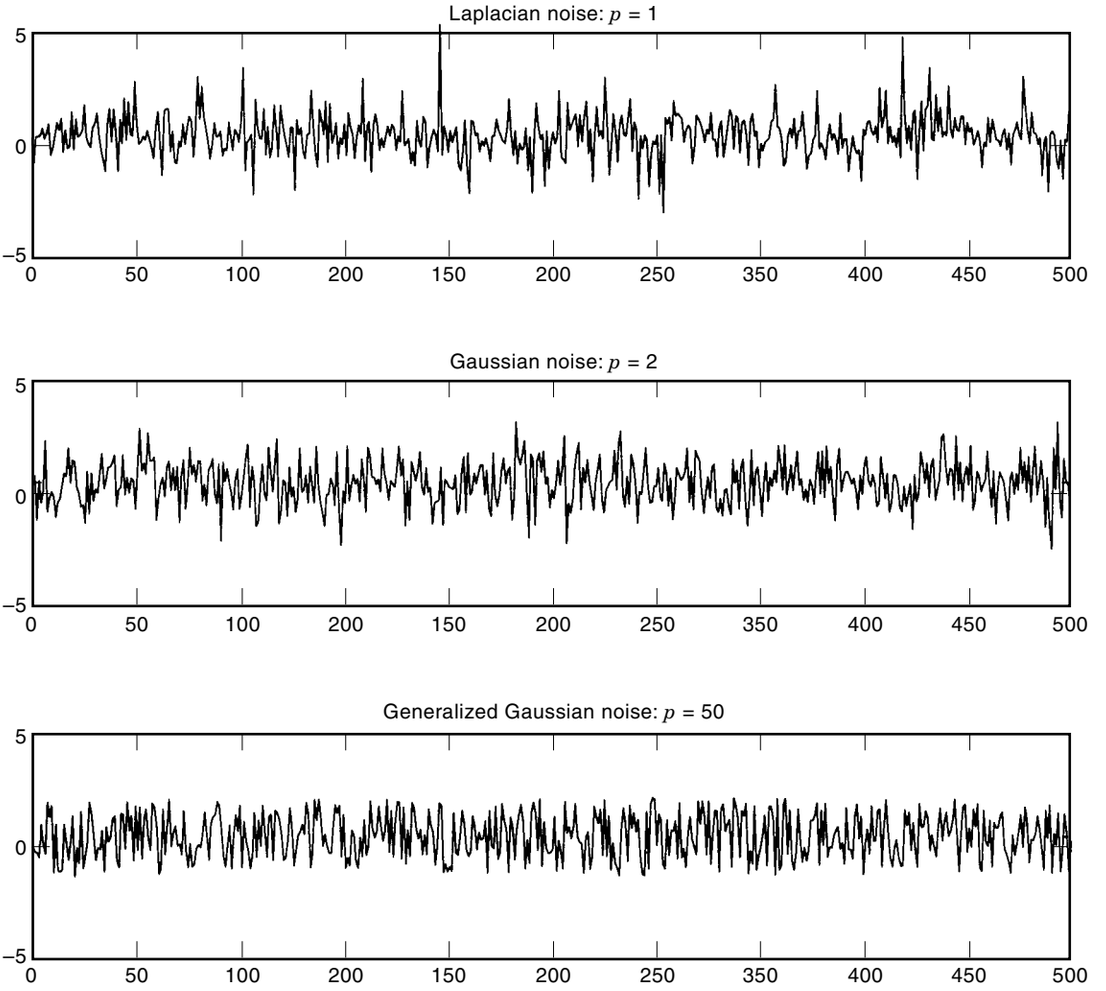

Figure 4. Plots of realizations from various density functions from the Generalized Gaussian family of density functions.

with less mass, resulting in data sets with very few outliers and more uniformly distributed over its range.

To demonstrate the variation in the data under this model, Fig. 4 shows sample realizations from Laplacian noise ($p = 1$), Gaussian noise ($p = 2$), and Generalized Gaussian noise $p = 50$. One can see that these data sets, all with a common variance (power), exhibit quite different behavior. The Laplacian noise better models interference which might contain spikes such as those produced by man-made devices, while $p = 50$ might better model interference arising from signals with well-contained powers.

Sub-Gaussian Noise

While in general one specifies a random signal through the joint distribution function, sub-Gaussian noise is specified through the joint characteristic function (2). As a reminder, this function is simply the Fourier transformation of the joint density function and, as such, there exists a straightforward one-to-one relationship between characteristic functions and joint density functions. As opposed to the Generalized Gaussian noise model, sub-Gaussian noise is parameterized in the exponent of the characteristic function rather than the exponent of the amplitude density function (5).

Definition 6 A noise or interference signal is said to be an α-sub Gaussian signal if for any positive integer n and time vector $\boldsymbol{t} = \{t_1, t_2, \ldots, t_n\}$, the characteristic function of the joint density function of $[x(t_1), x(t_2), \ldots, x(t_n)]$ is given by

$$\varphi(\boldsymbol{u}) = \exp\left(-\left[\frac{1}{2} \sum_{m,l=1}^{n} u_m u_l R(t_m, t_l) \right]^{\alpha/2} \right) \tag{29}$$

where $\alpha \in (1, 2]$ and where $R(t, s)$ is a positive definite function and where $\boldsymbol{u} = \{u_1, u_2, \ldots, u_n\}$.

Importantly, if the parameter $\alpha = 2$ then the sub-Gaussian process is simply a Gaussian process. Otherwise, the noise signal corresponds to some signal which has been parameterized away from the Gaussian signal. In all cases other than $\alpha = 2$, the corresponding tail of the amplitude density function decays at a polynomial rate (rather than an exponential rate) which translates into many more very large amplitude

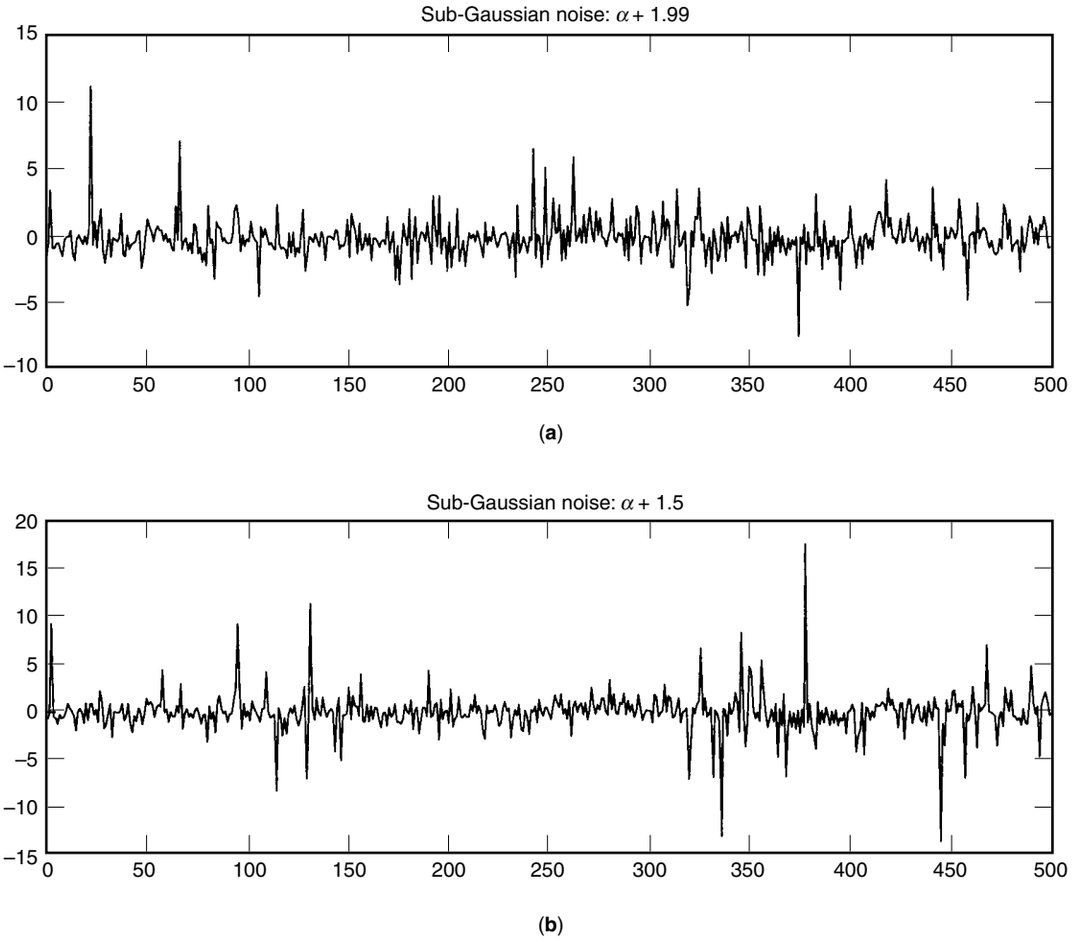

Figure 5. Plots of realizations from the sub-Gaussian noise family. (a) corresponds to $\alpha = 1.99$; (b) corresponds to $\alpha = 1.5$.

values than one would typically see with pure Gaussian noise. Therefore this model might be appropriate for characterizing data that contain many impulses.

There are a number of important properties of sub-Gaussian processes (6,7). From a modeling point of view, the most significant is that any sub-Gaussian process can be expressed as the product of a Gaussian process scaled by a so-called $\alpha/2$ stable random variable. That is, if $x(t)$ is an α-sub Gaussian signal, then

$$x(t) = S^{1/2}y(t) \tag{30}$$

where $y(t)$ is a Gaussian random process with zero mean and correlation function given by $R(t, s)$ and S is an independent, positive $\alpha/2$-stable random variable. (A $\alpha/2$-stable random variable is a random variable with characteristic function given by $\varphi(u) = \exp[-\gamma|u|^{\alpha/2}]$.) One interesting property of sub-Gaussian signals is that all samples from this signal are always dependent on one another. In addition, for $\alpha \neq 2$, it can be shown that samples from $x(t)$ will have infinite variance, thus making this model somewhat problematic in terms of accurately representing the average power of an interference signal. Nevertheless, the sub-Gaussian noise model has been used extensively to model noise or interference with a significant number of outliers or impulses and, in some applications, it serves well as a replacement for shot noise as an efficient model.

Middleton Noise and Interference

Unlike both the Generalized Gaussian model and the sub-Gaussian model, Middleton models have been derived directly from physical arguments. These models attempt statistically to characterize the amplitude density of the envelope (8) of a signal in the presence of various types of man-made interference after it has passed through the front end of receiver. Importantly, these models are canonical, in the sense that they do not depend on the specific source of the interference—only the relative bandwidth of the interferers with respect to the receiver. Therefore, interferers such as power lines, atmospheric noise, and various mechanical machines are all accommodated well by these very powerful models.

There are two basic forms of Middleton noise (9–11). Class A noise represents the statistical interference when the bandwidth of the interfering signals are on the order of or smaller than the bandwidth of the front end of the receiver. In this case, transient effects from the receiver can be ignored.

Alternatively, Class B noise considers the case where the receiver bandwidth is substantially smaller than the interfer-

ence bandwidth and thus transient effects must be accounted for. In addition to Class A and B noise, Middleton has introduced the notion of Class C noise, which is simply a linear combination of both A and B noise (12).

Physical Model. For both Class A and Class B noise, it is assumed that there are an infinite number of potential sources of interference within reception distance of the receiver. The individual interfering signals are assumed to be of a common form, except for such parameters as scale, duration, and frequency, among others.

The locations and parameters of the interferers in the source field are assumed to be randomly distributed according to the Poisson law. The emission times for each source are also assumed to be random and Poisson-distributed in time. Physically speaking, this implies that the sources are statistically independent both in location and emission time.

Class A Noise. As described above, the physical model for Class A noise assumes that the bandwidth of the individual interfering signals is smaller than the bandwidth of the receiver (data-collection system.) This allows for the simplification of ignoring all transient effects (ringing) in the output of the receiver front end.

Avoiding the tedious analysis and simply stating the result, the (approximate) amplitude probability density function under these assumptions was shown by Middleton (9) to be

$$f_{\boldsymbol{x}}(x;\, t) = e^{-A} \sum_{m=0}^{\infty} \frac{A^m}{m!} \frac{1}{\sqrt{2\pi\sigma_m^2}} e^{-\frac{x^2}{2\sigma_m^2}} \tag{31}$$

where

$$\sigma_m^2 = \frac{m/A + \Gamma}{1 + \Gamma} \qquad \text{and} \qquad \Gamma = \sigma_G^2/\Omega \tag{32}$$

One can see from the above equation that the Class A model is parameterized in three parameters (A, Γ, Ω), each with physical significance. The parameter A is a measure of the impulsiveness of the interference and is given by the product of the average number of impulses emitted per unit time with the average duration of the an impulse. Small values of A indicate a highly impulsive signal, since this would imply a small number of interferers each with very short pulses. For large values of A we have many interferers each with long duration and, thus, by the Central Limit theorem, the interference is nearly a Gaussian process. The constant Γ quantifies the ratio of the intensity of the independent Gaussian component arising from thermal or ambient noise (given by σ_G^2) to the mean intensity of the interference (given by Ω).

Class B Noise. Class B noise assumes that the bandwidth of the interfering signals is larger than the front-end bandwidth of the receiver. This assumption significantly complicates the analysis. In particular, one is required to use two separate models for small and large signal levels instead of a single model as in the Class A model. Because of the level of detail required to describe this model, it is recommended that the interested reader seeking a precise definition be directed to the original work found in (9).

ϵ-Mixture Noise

One might observe from the above Middleton models that the amplitude density functions are infinite linear combinations of zero mean Gaussian models. Interestingly, it has been well known in the statistics literature that appropriately scaled Gaussian density functions can be combined together to obtain nearly all valid amplitude density functions, so it is not surprising that this combination appears as a canonical representation of signal interference.

However, working with infinite sums of density functions requires much careful analysis. It is because of this that researchers have introduced a simplifying approximation to the Middleton Class A model, which is referred to as the ϵ-mixture model.

In this simplification, the amplitude probability density function of interference plus background or thermal noise is approximated by a combination of just two Gaussian density functions:

$$f_{\boldsymbol{x}}(x;\, t) = (1-\epsilon) \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} e^{-\frac{x^2}{2\sigma_\epsilon^2}} + \epsilon \frac{1}{\sqrt{2\pi\gamma\sigma_\epsilon^2}} e^{-\frac{x^2}{2\gamma\sigma_\epsilon^2}} \tag{33}$$

where the constant ϵ determines the fraction of impulsivity found in the data and where γ represents the ratio of intensities of the impulsive component to nominal ambient noise. In order to maintain a fixed power level of σ^2 in the noise and interference model for all choices of ϵ and γ, the parameters must satisfy the following power constraint:

$$\sigma_\epsilon^2 = \frac{\sigma^2}{1 - \epsilon + \epsilon\gamma} \tag{34}$$

Sample realizations from the ϵ-mixture model are depicted in Fig. 6. The top figure corresponds to purely Gaussian noise ($\epsilon = 0$), while the middle and bottom figures correspond to $\epsilon = .01$ and .1, respectively. As one can see, as ϵ increases the average number of "impulses" (spikes in the data) increases, while the average power of background thermal noise decreases in accordance with Eq. (34). Therefore, this model offers some of the flexibility required by the Middleton models without the requisite computational complexity.

CONCLUSIONS

This article has described the basic techniques for characterizing noise and interference from both a statistical and a signal processing perspective. These approaches have varied from the most analytic and fundamental (nth-order joint density or distribution functions) to more practical and data-oriented (means, correlation functions, and power spectral density.)

The primary differentiation between various noise models has been based on the statistical characteristics of the signal. Of particular importance was the "Gaussianity" or non-Gaussianity of the signal. In many important applications, the Central Limit Theorem implies that the corruptive effects of noise will be well approximated by a Gaussian process. However, basic physical arguments, such as those found in Middleton's work, also suggest that much of what we consider to be Gaussian noise is, in fact, substantially non-Gaussian.

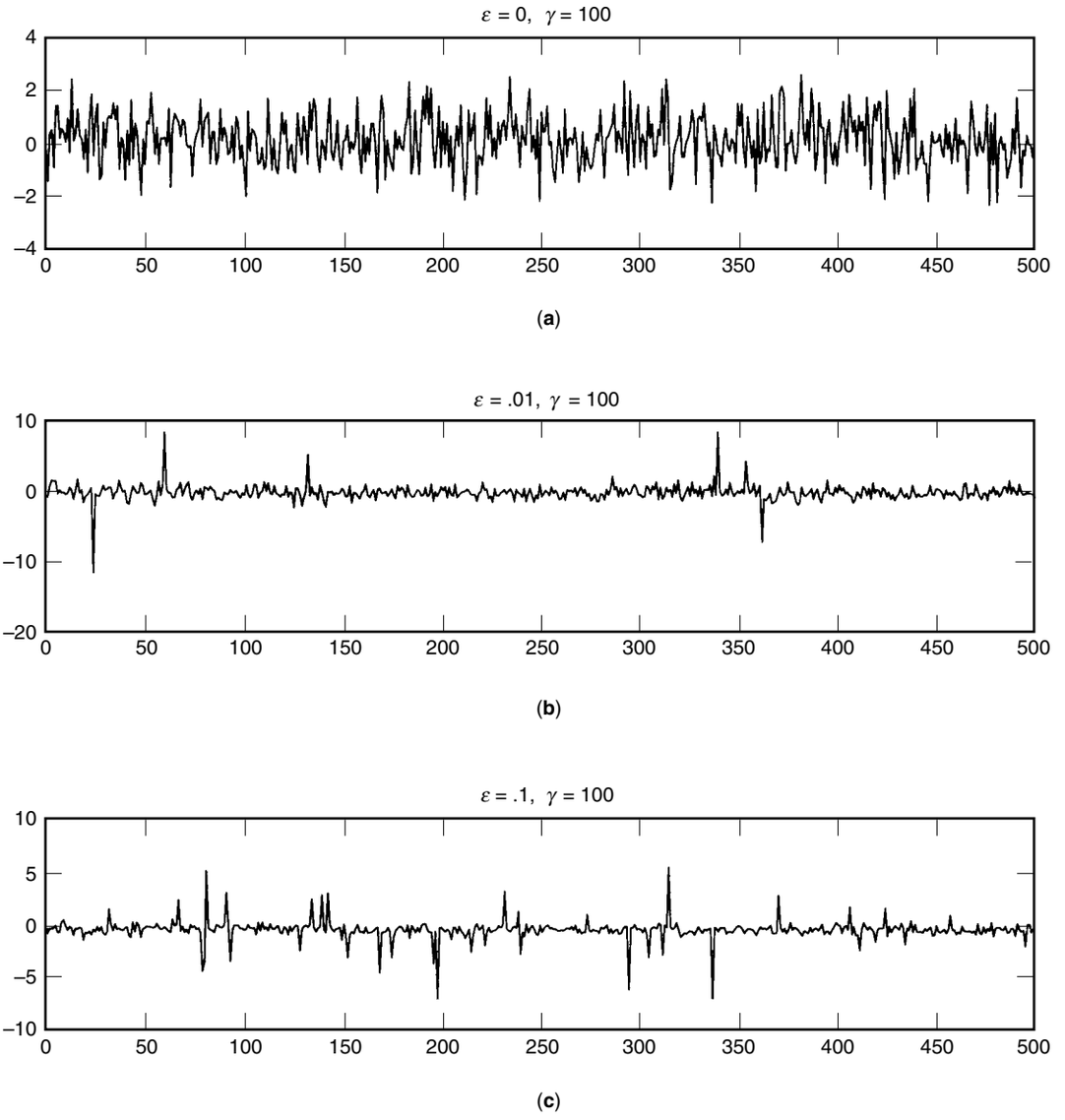

Figure 6. Plots of realizations from the ϵ-mixture noise family. (a) corresponds to $\epsilon = 0$ (pure Gaussian noise,) (b) corresponds to $\epsilon = 0.01$ (a 1% chance of observing interference at any point in time) and (c) corresponds to $\epsilon = .1$ (a 10% chance of observing interference at any point in time.) One should observe that as ϵ increases, the empirical frequency of "impulses" arising from interference increases proportionally.

Moreover, it has been shown in much research that imposing the Gaussian model on data which are decidedly or marginally non-Gaussian can have detrimental effects on the final designs. It is therefore critical that when one is dealing with data, particularly arising from interfering sources, that the appropriate model be used and used accurately.

For readers interested in pursing these topics further, one should begin with the important original work found in (13). Explicit models governing a wide variety of random phenomena can be found in (14–16). For an advanced treatment of random process in general, books by Doob (17) and Wong and Hajek (18) are excellent if not somewhat advanced treatments of the theory of random signals.

BIBLIOGRAPHY

1. R. M. Gray and L. D. Davisson, *Random Processes: A Mathematical Approach for Engineers,* Englewood Cliffs, NJ: Prentice-Hall, 1986.
2. A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* New York: McGraw-Hill, 1984.
3. A. V. Oppenheim, A. S. Willsky, and I. T. Young, *Signals and Systems,* Englewood Cliffs, NJ: Prentice-Hall, 1983.
4. E. J. Wegman, S. C. Schwartz, and J. B. Thomas (eds.), *Topics in Non-Gaussian Signal Processing,* New York: Springer-Verlag, 1989.
5. C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications,* New York: Wiley, 1995.

6. Y. Hosoya, Discrete-time stable processes and their certain properties, *Ann. Probabil.,* **6** (1): 94–105, 1978.
7. G. Miller, Properties of certain symmetric stable distribution. *J. Multivariate Anal.,* **8**: 346–360, 1978.
8. J. G. Proakis, *Digital Communications,* 3rd ed., New York: McGraw-Hill, 1995.
9. D. Middleton, Canonical non-gaussian noise models: Their implication for measurement and for prediction of receiver performance, *IEEE Trans. Electromagn. Compatibil.,* **EMC-21**: 209–220, 1979.
10. D. Middleton, Statistical-physical models of electromagnetic interference, *IEEE Trans. Electromagn. Compatibil.,* **EMC-19**: 106–127, 1977.
11. D. Middleton, Procedures for determining the parameters of the first-order canonical models of class a and class b electromagnetic interference, *IEEE Trans. Electromagn. Compatibil.,* **EMC-21**: 190–208, 1979.
12. D. Middleton, First-order non-gaussian class c interference models and their associated threshold detection algorithms, Tech. Rep. NTIA Contractor Rep. 87-39, Natl. Telecommun. Inf. Administration, 1987.
13. N. Wax, ed., *Noise and Stochastic Processes,* New York: Dover, 1954.
14. N. L. Johnson, S. Kotz, and A. W. Kemp, *Univariate Discrete Distributions,* 2nd ed., New York: Wiley, 1992.
15. N. L. Johnson and S. Kotz, *Continuous Univariate Distributions,* Vol. 1, Boston: Houghton-Mifflin, 1970.
16. N. L. Johnson and S. Kotz, *Continuous Univariate Distributions,* Vol. 2, Boston: Houghton-Mifflin, 1970.
17. J. L. Doob, *Stochastic Processes,* New York: Wiley, 1953.
18. E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems,* New York: Springer-Verlag, 1985.

GEOFFREY C. ORSAK
Southern Methodist University

NOISE, CIRCUIT. See CIRCUIT NOISE.