ERDC TR-18-7

US Army Corps
of Engineers®
Engineer Research and
Development Center

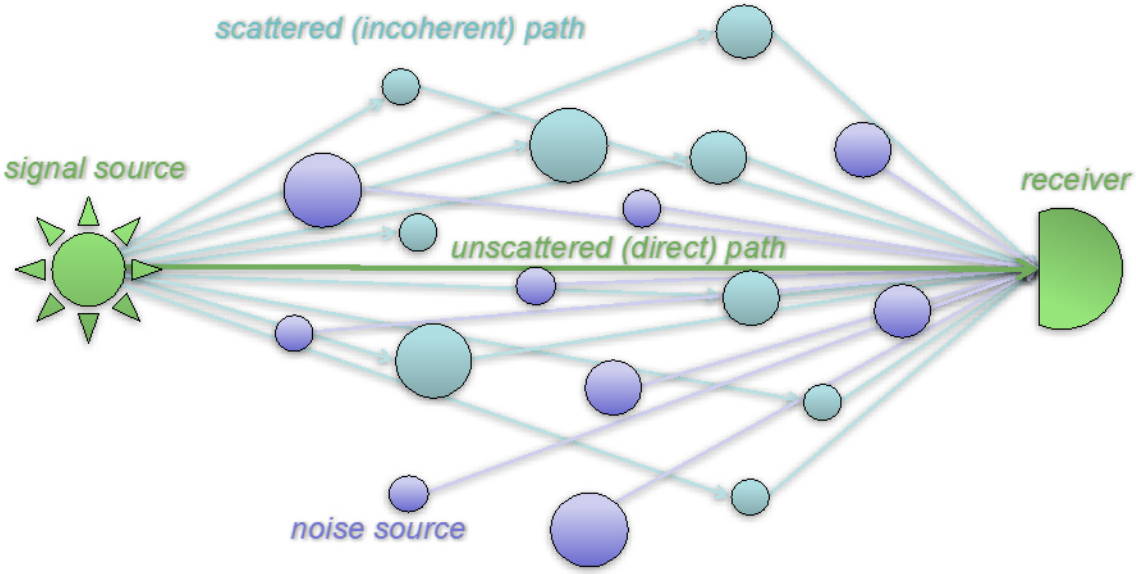
ERDC
INNOVATIVE SOLUTIONS
for a safer, better world

Environmental Quality/Installations and Geospatial Research and Engineering

Impact of Parametric Uncertainties on Scattered Signal Distributions and Receiver Operating Characteristics

D. Keith Wilson, Daniel J. Breton, Carl R. Hart, Chris L. Pettit,
Edward T. Nykaza, and Vladimir E. Ostashev

July 2018



Engineer Research and
Development Center

The U.S. Army Engineer Research and Development Center (ERDC) solves the nation's toughest engineering and environmental challenges. ERDC develops innovative solutions in civil and military engineering, geospatial sciences, water resources, and environmental sciences for the Army, the Department of Defense, civilian agencies, and our nation's public good. Find out more at www.erdc.usace.army.mil.

To search for other technical reports published by ERDC, visit the ERDC online library at http://acwc.sdp.sirsi.net/client/default.

Environmental Quality/Installations and
Geospatial Research and Engineering

ERDC TR-18-7
July 2018

Impact of Parametric Uncertainties on Scattered Signal Distributions and Receiver Operating Characteristics

D. Keith Wilson, Daniel J. Breton, Carl R. Hart, and Vladimir E. Ostashev

U.S. Army Engineer Research and Development Center (ERDC)
Cold Regions Research and Engineering Laboratory (CRREL)
72 Lyme Road
Hanover, NH 03755-1290

Edward T. Nykaza

U.S. Army Engineer Research and Development Center (ERDC)
Construction Engineering Research Laboratory (CERL)
2902 Newmark Dr.
Champaign, IL 61826-9005

Chris L. Pettit

U.S. Naval Academy
Aerospace Engineering Department
590 Holloway Rd.
Annapolis, MD 21402-5042

Final Report

Prepared for Headquarters, U.S. Army Corps of Engineers
Washington, DC 20314-1000

Under Environmental Quality/Installations (EQ/I), program element / project / task number 62272089604, "Real-Time Adaptation, Prediction, and Informatics for Dynamic Military Noise Environments (RAPID MNE)," and ERDC Geospatial Research and Engineering (GRE), program element / project / task number 611102T2400 "Statistical Evolution of Signatures Propagating in the Battlefield Environment"

Abstract

Many different distributions are used to model statistics of waves that have been randomly scattered in atmospheric and terrain environments. These distributions have varying analytical advantages and ranges of physical applicability. This report reviews several basic distributions and discusses how they can be extended to include spatial and temporal variability in the scattering process. For this purpose, a compound probability density function (pdf) can be introduced in which a basic pdf describing the underlying scattering process is modulated by a second pdf describing parametric uncertainties in the scattering. We describe some useful new formulations based on the compound pdf, including strong and Rytov (lognormal) scattering processes modulated by the environment. These new formulations lead to relatively simple marginalized signal power distributions (Lomax and lognormal, respectively). Furthermore, we show how the conditional scattered signal pdf may be viewed as a likelihood function in which the modulating pdf is the Bayesian conjugate prior. The parameters of the modulating process can thus be refined by simple sequential Bayesian updating. Finally, the impact of the parametric uncertainties on signal detection and receiver operating characteristic curves is discussed and shown to be a very important consideration in practical applications.

Contents

Figures and Tables

Figures

Tables

Preface

This study was conducted for the U.S. Army Corps of Engineers. Funding was provided by the Engineer Research and Development Center (ERDC) Environmental Quality/Installations (EQ/I) program under program element / project / task number 62272089604, "Real-Time Adaptation, Prediction, and Informatics for Dynamic Military Noise Environments (RAPID MNE)," and by the ERDC Geospatial Research and Engineering (GRE) program under program element / project / task number 611102T2400, "Statistical Evolution of Signatures Propagating in the Battlefield Environment." The EQ/I technical monitor was Mr. Alan Anderson of the ERDC Construction Engineering Research Laboratory (CERL). The GRE technical monitor was Dr. Randy Hill, ERDC Cold Regions Research and Engineering (CRREL).

The work was performed by the Signature Physics Branch (CEERD-RRD) of the Research and Engineering Division (CEERD-RR), ERDC-CRREL; the Ecological Processes Branch (CEERD-CNN) of the Installations Division (CEERD-CN), ERDC-CERL; and the Aerospace Engineering Department, U.S. Naval Academy. At the time of publication, Dr. Marino A. Niccolai was Chief, CEERD-RRD; Mr. J. D. Horne was Chief, CEERD-RR; Dr. Chris Rewerts was Chief, CEERD-CNN; and Mr. Donald K. Hicks was Chief, CEERD-CN. The Deputy Director of ERDC-CRREL was Mr. David B. Ringelberg, the ERDC-CRREL Director was Dr. Joseph Corriveau, the Deputy Director of ERDC-CERL was Dr. Kirankumar Topudurti, and the ERDC-CERL Director was Dr. Lance D. Hansen.

COL Ivan P. Beckman was Commander of ERDC, and Dr. David W. Pittman was the Director.

Acronyms and Abbreviations

ATR	Automated Target Recognition
cdf	Cumulative Distribution Function
CERL	Construction Engineering Research Laboratory
CRREL	U.S. Army Cold Regions Research and Engineering Laboratory
dB	Decibel
EQ/I	Environmental Quality/Installations
ERDC	Engineer Research and Development Center
GRE	Geospatial Research and Engineering
pdf	Probability Density Function
RAPID MNE	Real-Time Adaptation, Prediction, and Informatics for Dynamic Military Noise Environments
RF	Radio Frequency
ROC	Receiver Operating Characteristic
rv	Random Variable
SNR	Signal-to-Noise Ratio

1 Introduction

1.1 Background

Acoustic, seismic, and electromagnetic waves are scattered by many types of phenomena and objects occurring in the atmosphere, ocean, and earth, including turbulence, internal (buoyancy) waves, particles, rocks, buried man-made objects, vegetation, hills, and buildings. The wave scattering impacts the performance of many systems used to detect signals, communicate, and remotely sense the environment. Since the scattering typically depends on fine and often dynamic details of the environment that cannot be resolved with most types of models and measurements, we typically view the scattering as a random process; that is, we endeavor to develop statistical models for the scattered signals.

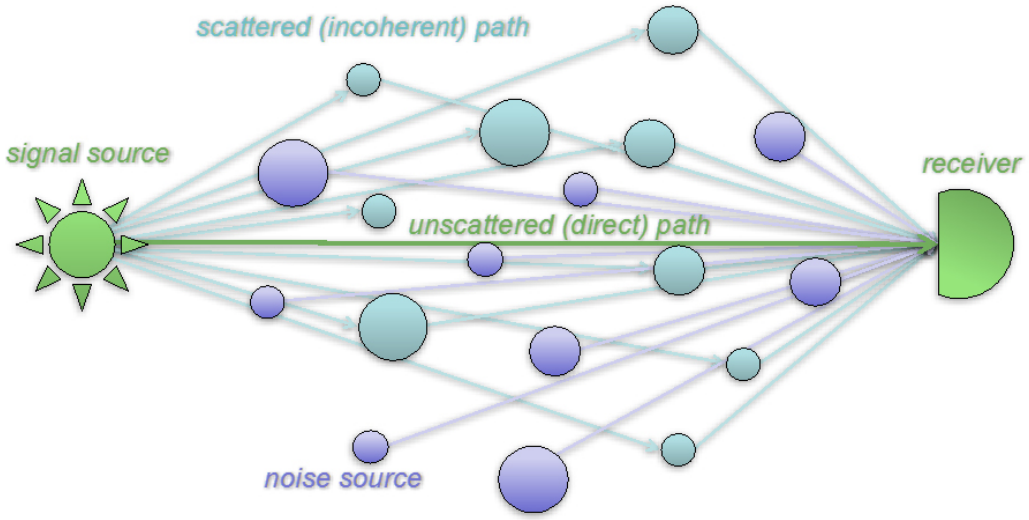
In this report, we address the problem of modelling the statistical distribution (more formally, the *probability density function*, or *pdf*) of the scattered signal power at one or more receivers after the signal has been randomly scattered by the environment. Many statistical models, physics based and empirical, have been previously formulated for such random signal variations caused by wave scattering.* The various models are intended for situations such as random scattering by turbulence or particles, multipath propagation, interference, scattering by rough surfaces, and superposition of multiple sources. However, models are available for only simple, idealized situations.

Figure 1 shows a conceptual diagram of the general problem. For present purposes, we simplify the discussion by assuming that the signals are harmonic or filtered into a narrow passband. The signal at a receiver consists of the signal of interest, which propagates along an unscattered (direct) path such that the signal amplitude and power undergo only small variations. The randomly scattered paths contribute incoherently, meaning that they arrive at the receiver with randomized amplitudes or phases. *Weak scattering* refers to situations where the energy propagated along the stable direct path dominates the received signal whereas *strong scattering*

* The following are among the main books written on the subject, to which the interested reader may refer for more background: Flatté et al. (1979), Rytov et al. (1989), Andrews and Phillips (2005), Ostashev and Wilson (2015).

refers to situations where the energy propagated along randomly scattered paths dominates.*

Figure 1. Conceptual diagram of the problem considered in this paper. The received signal consists of the signal of interest, which propagates along an unscattered (direct) path and multiple randomly scattered (incoherent) paths, plus noise originating from multiple, random sources.



We also depict in the figure noise originating from multiple, random sources, which, like the scattered paths, arrive at the receiver incoherently. The noise is an important consideration for problems involving signal detection. The random scattering and noise mechanisms lead to a probabilistic distribution for the signal and noise. Based on the signal and noise distributions, we may wish to furthermore calculate the probabilities of detection (for the signal of interest) and false alarm or other metrics describing the performance of a system.

Although we have framed the problem very generally in this report, wave scattering phenomena in terrestrial environments is of timely importance

* A more rigorous discussion of the difference between weak and strong scattering can be found in Ostashev and Wilson (2017). The definition between the two regimes is best based on the logarithm of the normalized signal amplitude (log-amplitude), designated by the symbol χ. Specifically, the variance of χ, $\langle\chi^2\rangle$, is less than one for weak scattering and close to one for strong scattering. Ostashev and Wilson (2017) suggest that $\langle\chi^2\rangle = 0.7$ is a useful boundary for distinguishing between weak and strong scattering. It is also important to keep in mind that while the signal may undergo fluctuations in the phase (designated by the symbol ϕ), the phase variance $\langle\phi^2\rangle$ impacts the coherence of the signal but not the strength of the scattering. Qualitatively, one may think of three distinct regimes for the scattered signals: weak scattering/strong coherence ($\langle\chi^2\rangle \ll 1$ and $\langle\phi^2\rangle \ll 1$), weak scattering/weak coherence ($\langle\chi^2\rangle \ll 1$ and $\langle\phi^2\rangle \gg 1$), and strong scattering/weak coherence ($\langle\chi^2\rangle \sim 1$ and $\langle\phi^2\rangle \gg 1$). These regimes can be conveniently depicted using strength-wave parameter diagrams, as shown by Ostashev and Wilson (2017).

in many Army applications. For example, operations in urban environments depend on effective communication between friendly forces and on detection of hostile forces. But the interaction of radio frequency (RF), acoustic, and other types of signals with the urban terrain results in complicated scattering effects, including multipath propagation and diffraction around buildings. These effects lead to "drop outs" (random fading) in the received signal energy as the emitter and receiver move through the urban environment.

Other important examples of wave scattering in Army applications include (1) the interaction of optical and acoustic waves with atmospheric turbulence, which degrades the performance of imaging and targeting systems; (2) obscuration of targets in forests and jungles; and (3) the scattering of seismic and radar waves by inhomogeneities in the soil (e.g., rocks), which complicates the ability to discern buried objects such as mines and unexploded ordinance.

1.2 Objectives

As mentioned in the preceding section, statistical modeling of random signal variations caused by wave scattering is already a mature area of research. The primary purpose is to show how such models can be extended to incorporate *uncertainties* in the scattering process and how algorithms can be developed that adapt to such uncertainties. The uncertainties are important because our knowledge of the terrestrial environment is limited such that we cannot make perfect predictions of wave scattering and its impacts on signal performance. The following are some important practical examples:

1. The scattering of acoustic, optic, and RF signals by the atmosphere depends on the local turbulence intensity, which varies randomly due to wind gusts, thermal plumes, dust clouds, and other phenomena. The variability of the turbulence in space and time (called *intermittency*) cannot be described deterministically and is therefore a source of uncertainty when predicting system performance.

2. Similarly, spatial variability of natural and man-made terrain features may lead to variations in surface and volumetric scattering. For example, the density and size of rocks, buildings, roadways, and trees, etc., will vary from one location to another. Many of the terrain features are too small to

be measured and incorporated directly into predictions; therefore, we must treat them as a source of uncertainty.

3. Predictions of signal propagation and sensor performance are often based on forecasts of the atmospheric state (i.e., numerical weather forecasts) at a future time. Since the forecast accuracy is limited by observations input to the model, imperfect model physics, and chaos, the atmospheric representation is a source of uncertainty in the signal predictions.

From the perspective of modeling wave scattering, these situations may all be regarded as examples of *parametric uncertainties*; that is, although we may have a specific model in mind for how the terrestrial environment scatters the signal, our knowledge of the parameters of that model is imperfect. We must assess these uncertainties and their impacts in order to put reasonable bounds on the range of sensor-system or communication-system performance.

1.3 Approach

This report is intended for a technical audience with some basic knowledge of statistics such as random variables (rvs) and modeling with probability density functions (pdfs). Some prior knowledge of wave scattering, signal processing, and/or Bayesian statistics is helpful but is not assumed. The report deals largely with statistical theory, with some discussion of the practical implications. Applications and comparisons to experimental data in various environments are planned for future studies.

The primary purpose of this report is to address the problem of parametric uncertainties in wave scattering. To this end, however, the report also provides an overview of single-variate (chapter 2) and multivariate (chapter 3) distributions for modeling randomly scattered signals. These chapters apply to signals along a single transmission path (one source, one receiver) and along multiple transmission paths (one or more sources and one or more receivers), respectively. While we do not provide here a comprehensive review of random signal distributions that have been used in the literature, the report provides a concise reference, which usefully consolidates the most important results. The discussion on multivariate distributions also incorporates some new results.

In chapter 4, we introduce the problem of modeling parametric uncertainties as based on compound pdfs. The essential idea is that the compound

pdf involves two distributions, one representing the wave scattering process and the other representing the uncertainties in the parameters of that process. This approach is shown to lead to a variety of possible distributions for modeling random signals. The common theme that emerges is how uncertainties dramatically impact the "tails" of the signal distributions, that is, the frequency of extreme variations.

Also in chapter 4, we show how the problem of modeling parametric uncertainties naturally relates to Bayesian inference of the wave scattering parameters. This relationship can be exploited by modeling the parametric uncertainties with Bayesian conjugate priors, which enables identification of statistical models for the parametric uncertainties with convenient analytical solutions. It also leads to sequential updating algorithms, which refine an initial prediction of the wave scattering parameters as new signal observations (i.e., observations of the random scattering process) become available. Referring to Figure 1, suppose we make an initial prediction for the signal statistics at the receiver (e.g., a prediction for the mean and variance). We transmit a signal from the source to the receiver. This transmission provides one random sample from the presumed distribution. On the basis of this sample, we can refine our prediction of the signal statistics. After repeating this process many times, we can provide a much more accurate characterization of the signal statistics than was possible initially.

In chapter 5, we explore how the parametric uncertainties impact the important practical problem of detecting a signal in noise. The problem is formulated in the classic manner as based on receiver operating characteristic (ROC) curves. We show how the parametric uncertainties have very dramatic effects on the ROC curves, thus indicating that conventional methods, which do not account for the uncertainties, potentially provide a very misleading assessment of system performance.

In chapter 6, we consider practical application of approaches described in chapters 2–4 to the modeling of signal transmissions in the presence of multiple sources of uncertainty regarding the source, receiver, and environmental parameters. The chapter also discusses formulation of multilevel models, which involve multiple levels of parametric uncertainties.

2 Single-Variate Distributions (Single Transmission Path)

In this section, we review distributions for single transmission paths (i.e., situations in which there is a single source and receiver). Such distributions are single variate. In general, we write the pdf for the scattered signal as $p(s|\boldsymbol{\theta})$, where s is the signal power at the receiver and $\boldsymbol{\theta}$ is a vector (set) of statistical parameters used to represent the pdf. (In this report, bolding is used to indicate vectors.) For a normal distribution, for example, $\boldsymbol{\theta}$ would typically consist of the mean m and variance σ^2. The vertical bar is used to separate the random variable (rv) (s in this case) from the distribution parameters.

In this chapter, we first consider the exponential pdf, which is applicable to strong scattering. Next, we consider the lognormal pdf, which is applicable to weak scattering. Then, we consider two distributions that can be used for both weak and strong scattering, namely the Rice and gamma pdfs. Lastly, we consider an extension of the gamma pdf called the generalized gamma pdf.

2.1 Exponential (strong scattering)

In conditions of strong scattering, the received signal consists of many independent, randomized contributions. The real and imaginary parts of the complex signal are then zero mean and normally distributed with equal variance. This leads to an exponential pdf for the signal power or, equivalently, a Rayleigh distribution for the amplitude (e.g., Burdic 1991 and Flatté et al. 1979).

To show how the exponential pdf arises, let $R = X + iY$, where R is the random complex amplitude of the narrowband signal, X is the random real part, and Y is the random imaginary part. By assumption, X and Y are drawn from independent normal distributions with zero mean and variance τ^2. That is, the pdf for the rv X is

$$p(x) = \mathcal{N}(x|0,\tau^2) = \frac{1}{\sqrt{2\pi}\tau} e^{-x^2/2\tau^2}$$

and likewise for Y. Here we have used $\mathcal{N}(x|m,\sigma^2)$ to indicate a normal pdf with mean m and variance σ^2. The probability that R has a magnitude less than some value r is thus

$$P(|R| \le r) = \frac{1}{2\pi\tau^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U\left(\frac{x^2+y^2}{r^2}\right) e^{-x^2/2\tau^2} e^{-y^2/2\tau^2}\, dx\, dy,$$

where $U(\xi)$ is a function that equals 1 when $0 \le \xi \le 1$ and zero otherwise. Converting the integral to cylindrical coordinates (ρ,θ), where $x = \rho\cos\theta$ and $y = \rho\sin\theta$, we have

$$P(|R| \le r) = \frac{1}{2\pi\tau^2}\int_0^{2\pi}\int_0^{r} e^{-\rho^2\cos^2\theta/2\tau^2} e^{-\rho^2\sin^2\theta/2\tau^2}\, \rho\, d\rho\, d\theta = \frac{1}{\tau^2}\int_0^{r} e^{-\rho^2/2\tau^2}\rho\, d\rho.$$

Changing the variable of integration to $\bar{s} = \rho^2/2\tau^2$, we readily find

$$P(S \le s) = \int_0^{s/2\tau^2} e^{-\bar{s}}\, d\bar{s} = 1 - e^{-s/2\tau^2},$$

where $s = r^2$ and $S = |R|^2$ is a random sample of the signal power. The preceding equation gives the cumulative distribution function (cdf) for s. To find the pdf, we differentiate with respect to s, the result being

$$p(s) = \frac{1}{2\tau^2} e^{-s/2\tau^2}. \tag{1}$$

In the statistical literature, the exponential pdf is often specified with the notation

$$\mathrm{Exp}(s|\lambda) = \lambda \exp(-\lambda s), \tag{2}$$

where λ is usually called the *rate parameter*. Comparing equations (1) and (2), we see that the former is an exponential pdf with rate parameter $\lambda = 1/2\tau^2$.

Since the parameter set $\boldsymbol{\theta}$ for the exponential pdf consists of just one parameter, λ, we thus indicate the signal pdf in the following equivalent forms:

$$p(s|\boldsymbol{\theta}) = p(s|\lambda) = \mathrm{Exp}(s|\lambda) = \lambda \exp(-\lambda s). \tag{3}$$

We define the mean and the variance of s as $m = \langle s \rangle$ and $\sigma^2 = \langle (s - \langle s \rangle)^2 \rangle$, respectively. The angle brackets indicate the expected value of an rv. For the exponential pdf, it can be shown that $m = \lambda^{-1}$ and $\sigma^2 = \lambda^{-2}$. Hence the normalized variance, σ^2/m^2, is 1 for the exponential pdf. Signals with a normalized variance of 1 are called *saturated*; such saturation is a defining characteristic of strong scattering. Furthermore, since $m = 2\tau^2$, for the exponential pdf, the mean scattered power equals the sum of the variances of the real and imaginary components of the signal.

Alternatively, we could parameterize the exponential pdf using the reciprocal of λ, which is designated here by the symbol θ (without bolding) and referred to as the *scale parameter*. That is,

$$p(s|\boldsymbol{\theta}) = p(s|\theta) = \mathrm{Exp}(s|\theta^{-1}) = \frac{1}{\theta} \exp\left(-\frac{s}{\theta}\right). \tag{4}$$

In this report, we will generally parameterize the exponential pdf using λ although in some cases, such as for describing turbulent intermittency in section 4.3, it will be more convenient to use θ.

Section 2.4 will provide example plots of the exponential pdf as it happens to be a special case of the gamma distribution, which is considered in that section.

2.2 Lognormal (weak scattering in the Rytov approximation)

For a lognormal pdf, the logarithm of the rv is normally distributed. This pdf is appropriate for situations involving weak scattering and follows from the Rytov approximation, which is discussed in many texts on wave propagation in random media (Flatté et al. 1979; Rytov et al. 1989; Andrews and Phillips 2005). In both RF and acoustical engineering, signals are often measured in decibels (dB), which are ten times the base-ten logarithm of the signal power, or twenty times the base-ten logarithm of the amplitude. Adoption of a lognormal pdf amounts to assuming that the signal in dB is normally distributed.

For the lognormal model, the logarithm of the signal, $\eta = \ln s$, is normally distributed:

$$\mathrm{p}(\eta|\mu, \phi) = \mathcal{N}(\eta|\mu, \phi) = \frac{1}{\phi\sqrt{2\pi}} \exp\left[-\frac{(\eta - \mu)^2}{2\phi^2}\right]. \tag{5}$$

Here, μ and ϕ^2 are the mean and variance of η. We call these parameters the log-mean and log-variance, respectively. By a transformation of variables to $s = \exp\eta$, we arrive at the lognormal pdf:

$$p(s|\mu,\phi) = \text{Lognorm}(s|\mu,\phi) = \frac{1}{s\phi\sqrt{2\pi}}\exp\left[-\frac{(\ln s-\mu)^2}{2\phi^2}\right]. \quad (6)$$

The mean of the signal s can be shown to be

$$m = e^{\mu+\phi^2/2}$$

whereas the variance is

$$\sigma^2 = \left(e^{\phi^2}-1\right)e^{2\mu+\phi^2} = \left(e^{\phi^2}-1\right)m^2.$$

Solving the two preceding equations for the log-mean, we find

$$\mu = \ln m - \frac{1}{2}\ln\left(1+\frac{\sigma^2}{m^2}\right), \quad (7)$$

and for the log-variance,

$$\phi^2 = \ln\left(1+\frac{\sigma^2}{m^2}\right). \quad (8)$$

Section 2.4 will provide example plots of the lognormal pdf, comparing it to the gamma pdf.

In some cases, the lognormal pdf can be approximated by a normal pdf. To see this, let us write $s = m + \sigma s' = m(1+\sigma s'/m)$, where s' is a fluctuation (order unity) in s as scaled by σ. If we assume that $\sigma/m \ll 1$, then $\eta = \ln s \cong \ln m + \sigma s'/m$. Since η is normally distributed, so too is s'. Based on this equation, the mean and variance η are $\mu = \ln m$ and $\phi^2 = \sigma^2/m^2$. (Note that these resuls are consistent with equations (7) and (8), respectively, when $\sigma/m \ll 1$.) Finally, it follows that s is normally distributed with mean $m = e^{\mu}$ and variance $\sigma^2 = \phi^2 e^{2\mu}$. Since the lognormal pdf is intended only for conditions of weak scattering, these relationships should be reasonable approximations.

As discussed earlier, decibel units are often used to represent the levels of RF and acoustic signals. Such units would correspond to the transformation $\xi = 10\log_{10} s$ rather than $\eta = \ln s$ as above. It can be readily shown that $\xi = (10/\ln(10))\eta$. Thus, the lognormal signal model described here is equivalent to treating the signal in decibels after a rescaling.

2.3 Rice (weak scattering in the Born approximation and strong scattering)

Like the exponential pdf, the Rice pdf applies to a signal with normally distributed real and imaginary parts of equal variance. However, the mean may be nonzero. This pdf is derived in Burdic (1991) and Flatté et al. (1979). Andrews and Phillips (2005) associate the Rice pdf with the Born approximation for weak scattering. The Rice pdf is exact in the limit of strong scattering (normalized variance close to 1).

Typically, the Rice pdf is written for the amplitude. For the signal power, the Rice pdf transforms to

$$p(s|\boldsymbol{\theta}) = p(s|\nu,\tau) = \frac{1}{2\tau^2}\exp\left(-\frac{s+\nu^2}{2\tau^2}\right) I_0\left(\frac{\sqrt{s}\nu}{\tau^2}\right). \quad (9)$$

Here, I_0 is the modified Bessel function of the first kind of zero order, ν is the mean of the signal amplitude (the amplitude along the direct path shown in in Figure 1), and τ^2 is the variance of the real and imaginary parts. The mean of s can be shown to equal $m = 2\tau^2 + \nu$ whereas the variance of s is $\sigma^2 = 4\tau^4 + 4\tau^2\nu^2$. Thus, given m and σ^2, we can determine the Rice pdf parameters as $\nu^2 = \sqrt{m^2 - \sigma^2}$ and $\tau^2 = (m - \nu^2)/2$.

Section 2.4 provides example plots of the Rice pdf and compares it to the gamma pdf.

2.4 Gamma (weak scattering approximation and strong scattering)

The gamma pdf can be derived from the Nakagami distribution, which has been found empirically to be a suitable distribution for signal amplitude in the presence of multipath propagation (Suzuki 1977). The Nakagami pdf is

$$p(r|\mu,\Omega) = \frac{2\mu^{\mu} r^{2\mu-1}}{\Gamma(\mu)\Omega^{\mu}} \exp\left(-\frac{\mu r^2}{\Omega}\right), \tag{10}$$

where $\Gamma(\cdot)$ is the gamma function, r is the signal amplitude, μ is a shape parameter, and Ω is a parameter controlling the spread of the distribution.

The gamma pdf follows directly from the Nakagami pdf when the latter is transformed from signal amplitude to power, $s = r^2$. For present purposes, we write the gamma pdf with the two parameters being defined as k, which is the same as μ in the Nakagami pdf and is called the *shape parameter*, and λ, which equals μ/Ω in the Nakagami pdf and is called the *rate parameter* (as in the exponential pdf). We thus have

$$p(s|\boldsymbol{\theta}) = p(s|k,\lambda) = \text{Gamma}(s|k,\lambda) = \frac{\lambda^k s^{k-1}}{\Gamma(k)} e^{-\lambda s}. \tag{11}$$

The mean of the gamma pdf is $m = k/\lambda$ whereas the variance is $\sigma^2 = k/\lambda^2$. Hence k and λ can be determined by setting $k = m^2/\sigma^2$ and $\lambda = m/\sigma^2$.

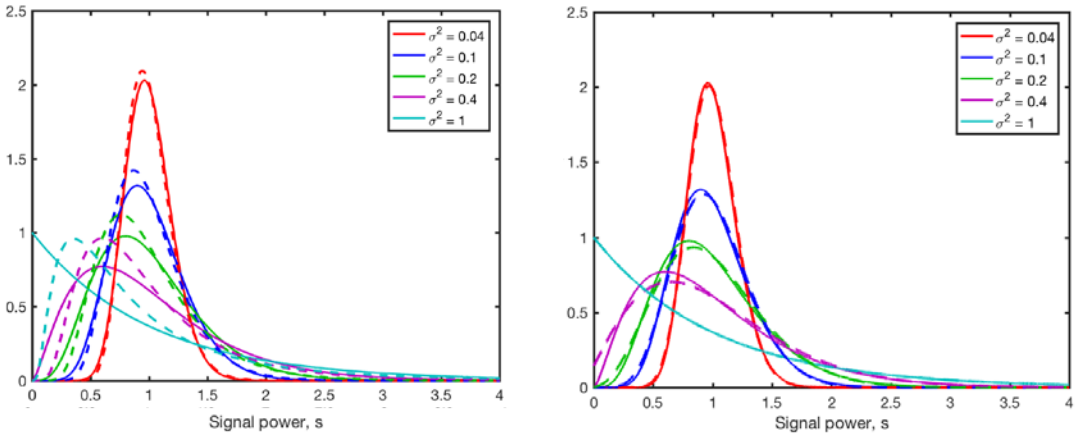
The gamma pdf, like the exponential, is sometimes parameterized using the inverse of λ, namely $\theta = \lambda^{-1}$, where θ is referred to as the *scale parameter*.

The gamma pdf, like the Rice pdf, reduces exactly to the exponential pdf in the limit of strong scattering, which corresponds to $k = 1$. Like the lognormal pdf, it converges to a normal pdf when the normalized variance is small and thus can also be regarded as suitable for weak scattering. Specifically, for large k, the gamma pdf approximates the normal pdf with mean $m = k\theta$ and variance $\sigma^2 = k\theta^2$. Using the relationships $m = e^{\mu}$ and $\sigma^2 = \phi^2 e^{2\mu}$ as derived in section 2.2 for the lognormal pdf in conditions of weak scattering, we find $k = \phi^{-2}$ and $\theta = \phi^2 e^{\mu}$.

Figure 2 compares the gamma distribution to the lognormal and Rice distributions (left and right parts of the figure, respectively). For these comparisons, the mean of all the pdfs has been set to one, and predictions for various values of the variance σ^2 are shown. Small values of the variance correspond to weak scattering whereas $\sigma^2 = 1$ corresponds to strong scattering. For weak scattering, which is the intended application of the lognormal distribution, the lognormal, gamma, and Rice pdfs are nearly

identical. All have a normal-like appearance in this regime. For strong scattering, the Rice and gamma pdfs are identical, and they are also very similar for intermediate values of σ^2. The main conclusions to be drawn are that the lognormal is useful only for weak scattering whereas the Rice and gamma are reasonable for situations ranging from weak to strong scattering.

Figure 2. Plots of various pdf models (*vertical axes*) vs. signal power. *Left*: Gamma (*solid lines*) and lognormal (*dashed lines*) pdfs for various values of the variance normalized by the squared mean. *Right*: Gamma (*solid lines*) and Rice (*dashed lines*) pdfs for various values of the variance normalized by the squared mean.

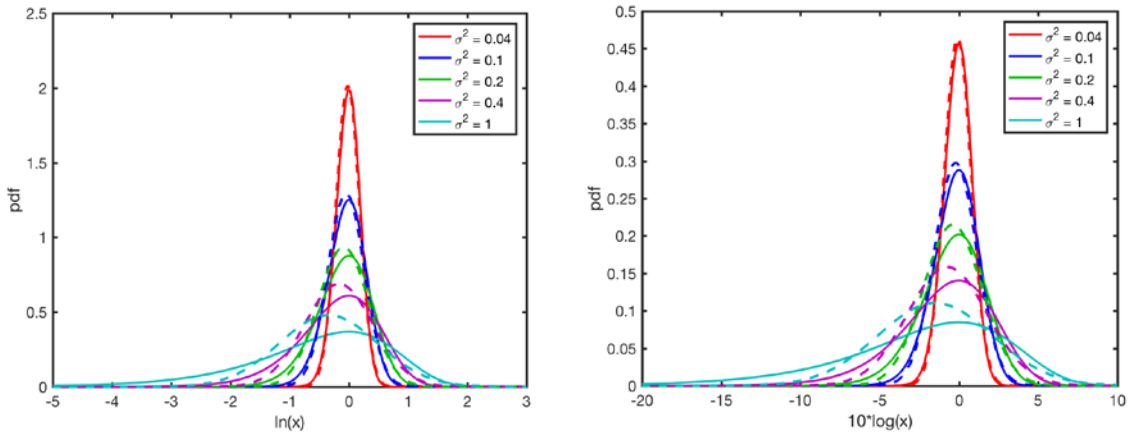


The pdfs shown in Figure 2 are for the untransformed signal power s. Since we often work with signal levels in decibels, it is interesting to compare the pdfs on a logarithmic scale as well. Figure 3 compares pdfs for the gamma and lognormal distributions using two different logarithmic abscissas (horizontal axes), namely the natural logarithm of the power ($\eta = \ln s$), and the level in decibels (i.e., $\xi = 10 \log_{10} s = 10\,\eta / \ln 10 \approx 4.343\eta$). With these logarithmic axes, of course, the lognormal pdf becomes a normal pdf. These plots also make apparent how the gamma pdf predicts more frequent "deep fading" (i.e., situations where the power is close to zero) than the lognormal pdf. Hence, a lognormal pdf will underestimate the occurrence of fading.

The distributions for the logarithms show that, with the gamma pdf, the logarithm has a significant negative skewness (non-normal behavior) as the variance approaches 1; that is, large, negative fluctuations in the level tend to be more frequent than large, positive fluctuations. This behavior is often observed in sound-level recordings. Dyer (1970) shows that, when x is given by an exponential pdf (i.e., $p(s) = \lambda \exp(-\lambda s)$), $\eta = \ln s$ has a mean of $-\ln \lambda - \gamma$ (where $\gamma = 0.577\ldots$ is Euler's constant) and a standard

deviation of $\pi/\sqrt{6}$. Hence (referring back to the discussion in the last paragraph of section 2.2), the standard deviation of the level in decibels is $(\pi/\sqrt{6})(10/\ln 10) = 5.57\ldots$ dB. Thus, when strongly scattered signals are analyzed in decibels, we expect to see a negatively skewed distribution with a standard deviation around 5.57 dB.

Figure 3. Comparison of gamma (*solid lines*) and lognormal (*dashed lines*) distributions for a mean of $m = 1$ and the variance set to five different values as shown in the legend. The horizontal axis for the plot on the left is the natural logarithm ($\ln x$) of the random variable (signal power); the horizontal axis on the *right* is in decibels ($10 \log x$).



2.5 Generalized gamma (weak and strong scattering with extreme events)

Ewart and Percival (1986) advocate for the *generalized* gamma pdf on the basis of its versatility and ability to approximate a wide variety of other commonly used pdfs for wave scattering. Unlike the pdfs considered thus far, which had either one parameter (the exponential) or two parameters (the lognormal, Rice, and gamma), the generalized gamma has three parameters. It retains k and λ from the gamma pdf but incorporates a new parameter, b. It is given by

$$p(s|\boldsymbol{\theta}) = p(s|k,\lambda,b) = \frac{b\lambda^{bk}s^{bk-1}}{\Gamma(k)}e^{-(\lambda s)^b}. \quad (12)$$

The generalized gamma pdf reduces to the ordinary gamma pdf when $b = 1$. Ewart and Percival demonstrated that the generalized gamma pdf agrees well with ocean acoustic scattering data from a variety of experiments. The additional parameter b provides flexibility in dealing with *dispersion* of the data (i.e., deviations from the number of extreme events as would be predicted from a normal or other baseline distribution).

Figures 4 and 5 show the generalized gamma distributions for various values of b for the cases $k = 1$ (strong scattering) and $k = 8$ (weak scattering), respectively. For both figures, the parameter λ was selected such that the mean of the signal power is 1 (using equations from Ewing and Percival 1986). The parameter b is seen to control the "tails" (extreme values) of the distribution. As b decreases, the pdfs change from a normal-like appearance to having tails exceeding the gamma distribution for the corresponding value of k. Based on their empirical fits to ocean acoustic data, Ewart and Percival found that b is *usually* less than 1; that is, in real data, elevated tails (overdispersion) are typically present.

Figure 4. Generalized gamma distribution for strong scattering ($k = 1$). Predictions are shown for various values of the parameter b. The *dashed line* is the ordinary gamma distribution with $k = 1$ (i.e., the exponential distribution). *Left* is a plot with linear axes whereas *right* is a plot with a logarithmic axis for the pdf.

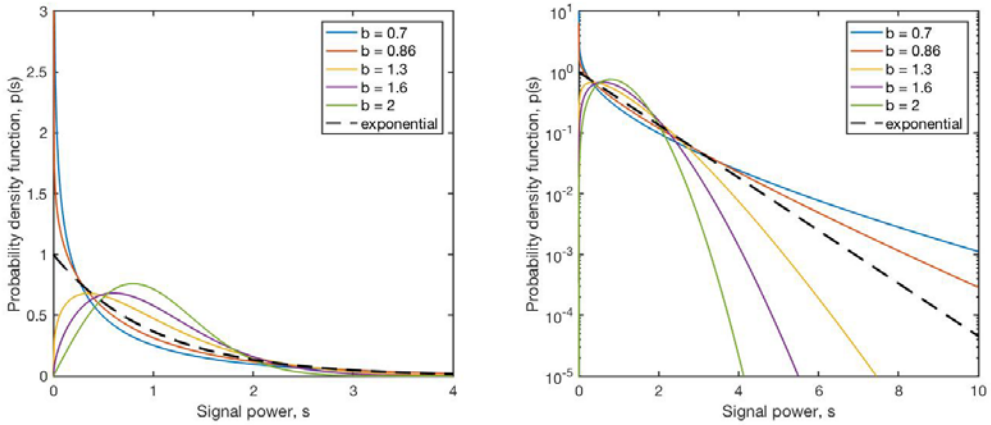
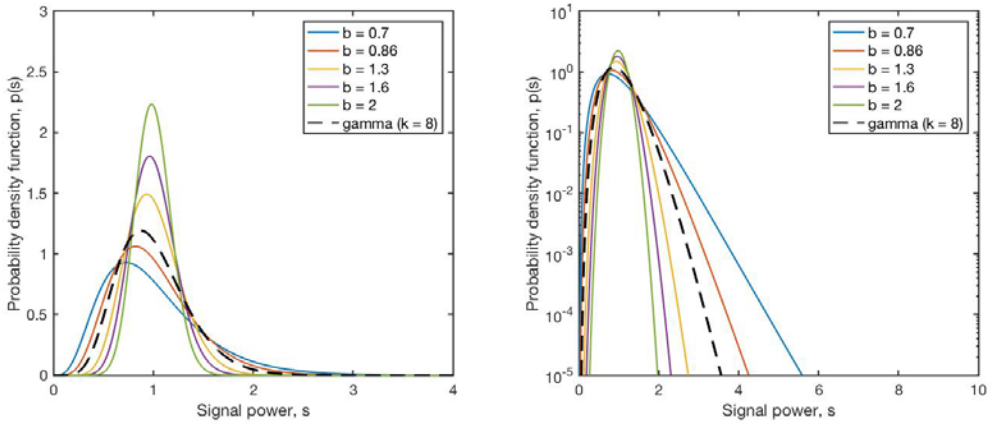


Figure 5. Same as Fig. 4 except that the distributions are for $k = 8$.



3 Multivariate Distributions (Multiple Transmission Paths)

Up until this point, we have considered signal transmissions along a single path. Let us now suppose there are multiple paths and that the transmissions along these paths may be correlated to some degree. In particular, we have in mind a situation where there are multiple source and receiver locations. Each source/receiver pair provides a different transmission path so that there is a total of $N = N_s \times N_r$ paths, where N_s is the number of sources and N_r the number of receivers. The multivariate distribution represents the received signal power along each of these N paths. The correlation of the received signal levels depends on the proximity and overlap of the paths. For now, we will not address how to model and calculate the path correlations.

Because of reflections or refraction, multiple transmission paths may also occur when there is a single source/receiver pair. Whether these paths are considered to be distinct, identifiable paths or part of the random scattering process depends on the modeling perspective. One practically important example is urban environments, for which acoustic and RF energy may reflect from buildings and other surfaces before reaching the receiver. Even in a flat, open environment, there may be direct- and ground-reflected ray paths, for which it is important to consider uncertainties in modeling the interference between the ray paths (e.g., Ostashev et al. 2011).

3.1 Lognormal (weak scattering in the Rytov approximation)

The lognormal model as described in section 2.2 applied to weak scattering along a single path. Since multivariate normal distributions have been well studied and lead to analytic solutions, it is straightforward to extend this treatment to multiple paths. The multivariate normal distribution for a column vector $\boldsymbol{\eta}$ of length N is given by

$$\mathcal{N}(\boldsymbol{\eta}|\boldsymbol{\mu},\boldsymbol{\Phi}) = \frac{1}{\sqrt{(2\pi)^K|\boldsymbol{\Phi}|}}\exp\left[-\frac{1}{2}(\boldsymbol{\eta}-\boldsymbol{\mu})^T\boldsymbol{\Phi}^{-1}(\boldsymbol{\eta}-\boldsymbol{\mu})\right]. \tag{13}$$

Here, $\boldsymbol{\mu}$ is the mean vector (length K), $\boldsymbol{\Phi}$ is the covariance matrix (size $N \times N$), the vertical lines indicate the determinant, the superscript T indicates transpose, and the superscript -1 indicates the matrix inverse.

As in the single variate case, we transform variables from $\boldsymbol{\eta}$ to $\boldsymbol{s} = \exp \boldsymbol{\eta}$. The exponentiation is applied to the vector on an element-by-element basis. The result is (see StackExchange 2017)

$$p(\boldsymbol{s}|\boldsymbol{\mu}, \boldsymbol{\Phi}) = \frac{1}{(s_1 s_2 \cdots s_K)\sqrt{(2\pi)^K |\boldsymbol{\Phi}|}} \exp\left[-\frac{1}{2}(\ln \boldsymbol{s} - \boldsymbol{\mu})^T \boldsymbol{\Phi}^{-1} (\ln \boldsymbol{s} - \boldsymbol{\mu})\right]. \quad (14)$$

To generate a multivariate lognormal vector, we simply generate normally distributed vectors for $\boldsymbol{\eta}$ and set $\boldsymbol{s} = \exp \boldsymbol{\eta}$.

3.2 Wishart (strong scattering)

Let us consider now strong scattering with multiple paths. This is not as straightforward as with the weak scattering case because there is no simple multivariate extension for the exponential pdf. However, we can consider the Wishart distribution, which is a generalization of the chi-squared distribution from single rvs to $N \times N$ random matrices.* Like the chi-squared pdf, the Wishart pdf has a specified number of degrees of freedom, which we indicate here by d. The degrees of freedom correspond to the number of normal variates that are squared and summed together. For two degrees of freedom ($d = 2$), the marginal distributions† of the diagonal elements of the matrix have exponential distributions. Thus, the Wishart distribution, with $d = 2$, appears to be appropriate for strong scattering along multiple paths.

The derivation of the Wishart pdf for $d = 2$ proceeds similarly to the derivation of the exponential pdf. Let $R_n = X_n + iY_n$, where R_n is the complex amplitude of the signal along path n, X_n is the random real part, and Y_n is the random imaginary part. By assumption, X_n and Y_n are drawn from independent normal distributions with zero mean and variance τ_n^2. Following the procedure in section 2.1, we can readily show that $S_n = R_n^2$ has an exponential pdf with rate parameter $\lambda_n = 1/2\tau_n^2$. But what is the distribution of the cross terms (i.e., $R_m R_n^*$, when $m \neq n$)?

Writing out the real and imaginary parts of the cross term, we have $R_m R_n^* = (X_m + iY_m)(X_n - iY_n) = X_m X_n + iX_n Y_m - iX_m Y_n + Y_m Y_n$. Let us assume that we have removed the deterministic phases of the signals relative

* Note that the multivariate lognormal distribution in section 3.1 was formulated in terms of a random *vector*. The Wishart distribution is formulated in terms of a random *matrix*.

† A *marginal distribution* results from removing other variables from the joint distribution by integrating over them. One is thus left with a single-variate pdf for the variable of interest.

to one another; for example, we could multiply each signal by e^{-ikd_n}, where d_n is the length of the propagation path. Then the cross-correlations must have zero phase; that is, the expected value of the cross terms X_mY_n and X_nY_m is zero. Furthermore, X_mX_n and Y_mY_n must have the same distribution because the X_n and Y_n have the same distributions. Since the X_n and Y_n are normally distributed, we can conveniently define cross-correlation coefficients such that

$$\rho_{mn} = \frac{\langle X_m X_n \rangle}{\tau_m \tau_n} = \frac{\langle Y_m Y_n \rangle}{\tau_m \tau_n}.$$

Or, defining the $N \times 1$ vectors $\boldsymbol{X} = [X_1, X_2, \ldots, X_N]$ and $\boldsymbol{Y} = [Y_1, Y_2, \ldots, Y_N]$, we have the $N \times N$ covariance matrix

$$\boldsymbol{V} = \langle \boldsymbol{X}\boldsymbol{X}^T \rangle = \langle \boldsymbol{Y}\boldsymbol{Y}^T \rangle = \begin{bmatrix} \tau_1^2 & \rho_{12}\tau_1\tau_2 & \cdots & \rho_{1N}\tau_1\tau_N \\ \rho_{12}\tau_1\tau_2 & \tau_2^2 & \cdots & \rho_{2N}\tau_2\tau_N \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1N}\tau_1\tau_N & \rho_{2N}\tau_2\tau_N & \cdots & \tau_N^2 \end{bmatrix}. \quad (15)$$

(The superscript T indicates the matrix transpose.) Hence, we are interested in the pdf for a matrix

$$S = \boldsymbol{X}\boldsymbol{X}^T + \boldsymbol{Y}\boldsymbol{Y}^T,$$

where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are N-variate normal distributions with zero mean and covariance $\boldsymbol{V}$. But this is simply the definition for the Wishart distribution with $d = 2$. (The number of degrees of freedom corresponds to the terms in the preceding summation for S.) The Wishart distribution, for $d > N - 1$, is written

$$p(\boldsymbol{S}|d, \boldsymbol{V}) = \frac{|\boldsymbol{S}|^{(d-N-1)/2}}{2^{dN/2}\Gamma_N\left(\frac{d}{2}\right)|\boldsymbol{V}|^{d/2}} \exp[-\mathrm{tr}(\boldsymbol{V}^{-1}\boldsymbol{S})/2], \quad (16)$$

where $\boldsymbol{S}$ is the $N \times N$ random matrix, $\boldsymbol{V}$ is an $N \times N$ positive definite matrix called the *scale matrix* (in our case, the covariance matrix as defined by equation (15)), and

$$\Gamma_p(a) = \pi^{p(p-1)/4} \prod_{j=1}^{p} \Gamma[a + (1-j)/2].$$

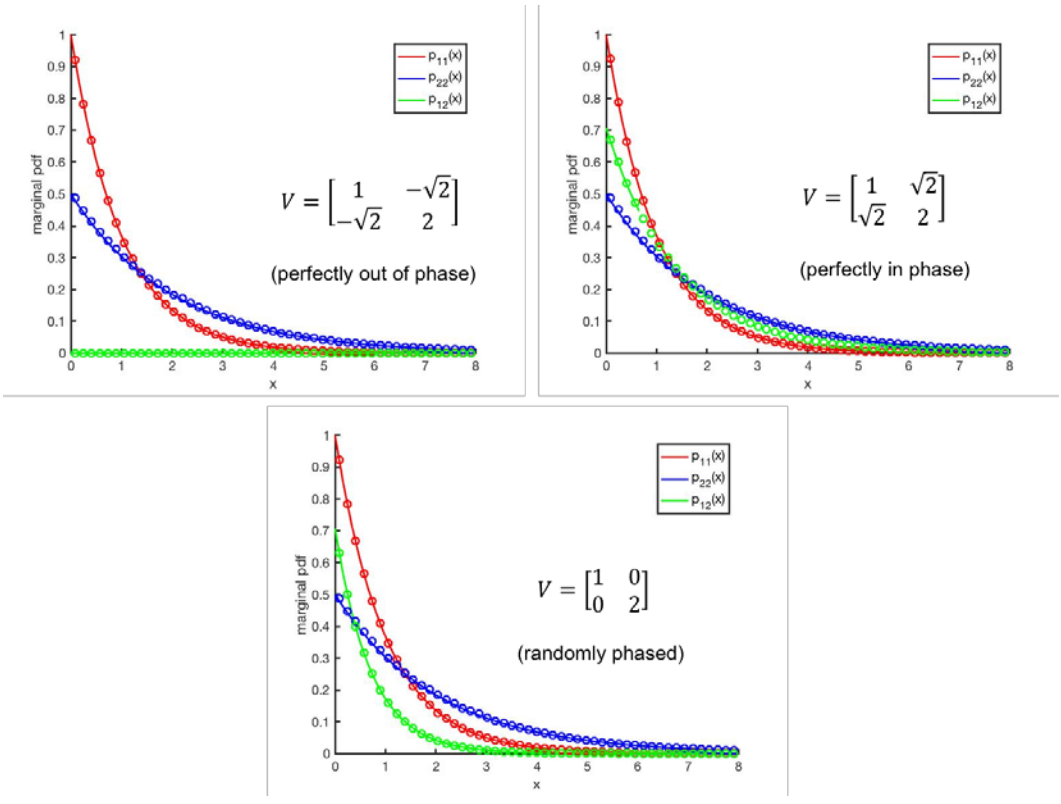
As indicated earlier, the marginals of the diagonal elements of the Wishart distribution are exponentially distributed. The off-diagonal marginals can be shown to have a variance-gamma distribution (Pearson et al. 1929).

Figure 6 shows several examples of the Wishart distribution for various cases of the scale matrix $\boldsymbol{V}$. We consider a situation with two transmission paths ($N = 2$), for which we may write the scale matrix as

$$\boldsymbol{V} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

where σ_1^2 and σ_2^2 are the variances along each path, and ρ is the correlation coefficient. The mean along one path is taken to be twice that of the other ($\sigma_1^2 = 2\sigma_2^2$); the cases differ regarding the off-diagonal elements (i.e., the value of ρ is varied).

Figure 6. Matrix Wishart distribution for $d = 2$ and several illustrative cases of the scale matrix V. The mean along the second path is twice that of the first. *Solid lines* are theoretical results, and *circles* are numerical simulations.



3.3 Matrix gamma (weak and strong scattering)

Similar to the manner by which the two degree-of-freedom Wishart distribution generalizes the single-variate exponential pdf to matrices, the matrix gamma distribution generalizes the single-variate gamma distribution to matrices. The matrix gamma distribution is given by (Gupta and Nagar 1999; Wikipedia 2017b)

$$p(\boldsymbol{S}|\alpha,\beta,\boldsymbol{V}) = \frac{|\boldsymbol{S}|^{\alpha-(N+1)/2}}{\beta^{\alpha N}\Gamma_N(\alpha)|\boldsymbol{V}|^{\alpha}}\exp[-\mathrm{tr}(\boldsymbol{V}^{-1}\boldsymbol{S})/\beta]. \quad (17)$$

This distribution reduces to the Wishart when $\beta = 2$ and $\alpha = d/2$.

It seems reasonable that the matrix gamma distribution would be appropriate for both weak and strong scattering along multiple paths. However, given that the gamma pdf is largely empirically motivated to begin with, we should not expect that the matrix gamma distribution can be motivated by a theoretical argument as was possible with the multivariate lognormal and matrix Wishart distributions. We leave the hypothesis that the matrix gamma pdf is appropriate for weak or strong scattering to future investigation.

4 Incorporating Parametric Uncertainties

Parametric uncertainties can be conceptualized as falling into two categories: *aleatory* (or *irreducible*) and *epistemic* (or *reducible*). Turbulent intermittency, as will be discussed further in section 4.3, is an appropriate example of the former as turbulence intensity varies spatially and temporally in a manner that is understood in a statistical sense but is impractical to predict deterministically. Conceptually, this is similar to rolling a die; we understand the probability that each roll will yield a particular number but cannot predict the outcome deterministically. Uncertainty due to imperfect weather forecasts, on the other hand, is primarily an epistemic uncertainty. With better knowledge of the initial conditions, we could in principle generate a better forecast and reduce our uncertainty.

From a modeling perspective, aleatory and epistemic uncertainties are handled the same way, namely by specifying a distribution for the uncertain parameters (rather than assuming they are exactly known). Hence, the following discussion conceptually pertains to both types of uncertainty.

Interested readers may refer to chapter 13 in Ostashev and Wilson (2015) for an overview of uncertainty modeling in the context of sound-wave propagation.

4.1 Compound distributions

We wish to incorporate parametric uncertainties into the scattering distributions discussed in chapters 2 and 3. This can be done by introducing a compound probability density function,* which consists of a distribution representing the solution for fixed values of the signal parameters, modulated by a higher-level distribution representing the variability of those parameters. Mathematically, the compound pdf is given by

$$p(s|\boldsymbol{\chi}) = \int p(s|\boldsymbol{\theta}) p(\boldsymbol{\theta}|\boldsymbol{\chi})\, d\boldsymbol{\theta}. \tag{18}$$

* The terminology *compound pdf* has, confusingly, been used in the literature to refer to several different operations on pdfs. Besides the usage here, *compound pdf* can also describe the pdf of a product or of a summation of rvs. Such products and summations are important for modeling communication channels, propagation, and noise background. A compound pdf as defined here is also called a *hyper* or *nested* distribution elsewhere in the literature.

The dimensionality of the integral equals the number of parameters in $\boldsymbol{\theta}$. The first pdf under the integral, $p(s|\boldsymbol{\theta})$, describes the dependence of the wave scattering process on the parameters $\boldsymbol{\theta}$ as before. The second pdf, $p(\boldsymbol{\theta}|\boldsymbol{\chi})$, describes the uncertainties in the scattering parameters and depends on a new set of parameters $\boldsymbol{\chi}$. These are called the *hyperparameters* as they parameterize the variation of parameters. After accounting for the uncertainty by marginalizing (integrating) over the uncertain parameters $\boldsymbol{\theta}$, we have a new pdf, $p(s|\boldsymbol{\chi})$, which depends only on the hyperparameters.

4.2 Connection to Bayes' theorem

The previous discussion of compound distributions has a useful connection to Bayesian inference and adaptation, which is explored in this section. Readers may refer to Gelman et al. (2014) or other suitable textbooks for an introduction to Bayesian statistics.

Suppose we regard $p(\boldsymbol{\theta}|\boldsymbol{\chi})$ in equation (18) as the *prior* distribution for the parameters $\boldsymbol{\theta}$; that is, $p(\boldsymbol{\theta}|\boldsymbol{\chi})$ is the assumed distribution for $\boldsymbol{\theta}$ before any observations of the signal power s are made. The prior may, for example, be predicted from a propagation model. Next, we perform an experiment to collect a sample of s. We would then like to know the *posterior* distribution, $p(\boldsymbol{\theta}|s,\boldsymbol{\chi})$, which indicates the improvement in our knowledge of $\boldsymbol{\theta}$ resulting from the sample. By Bayes' theorem,

$$p(\boldsymbol{\theta}|s,\boldsymbol{\chi}) = \frac{p(s|\boldsymbol{\theta},\boldsymbol{\chi})\,p(\boldsymbol{\theta}|\boldsymbol{\chi})}{p(s|\boldsymbol{\chi})} = \frac{p(s|\boldsymbol{\theta},\boldsymbol{\chi})\,p(\boldsymbol{\theta}|\boldsymbol{\chi})}{\int p(s|\boldsymbol{\theta}')p(\boldsymbol{\theta}'|\boldsymbol{\chi})\,d\boldsymbol{\theta}'}, \tag{19}$$

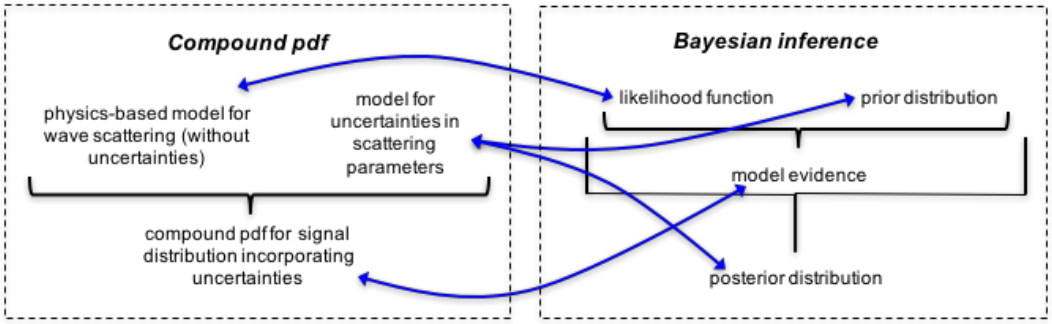
where the second equality follows by using equation (18) to rewrite the denominator. Here, $p(s|\boldsymbol{\theta},\boldsymbol{\chi})$ is called the *likelihood* function for the signal observation; that is, the signal distribution conditioned upon a particular set of parameters. The denominator in equation (19), $p(s|\boldsymbol{\chi})$, is usually called the *model evidence*. Although the denominator is often regarded as not being intrinsically interesting (sometimes simply being a normalizing factor), in the present context it does have an important interpretation. Namely, it is the probability of observing a particular value of the signal after marginalizing over the uncertain parameters.

For some special combinations of priors and likelihoods, the posterior has the same functional form as the prior. In this case, the prior distribution is said to be the *conjugate prior* of the likelihood function. Suppose we then collect a new sample of the signal and use the posterior from the previous

step as the new prior. Since the new prior is still the conjugate prior of the likelihood function, the posterior will again have the same functional form. This process can be repeated indefinitely, which is a great analytical convenience. However, it must be kept in mind that use of a conjugate prior is suitable only if the function is a good representation of the prior distribution.

Figure 7 illustrates the relationships between the distributions occurring in the compound pdf and those in Bayes' theorem. In the next section, we consider a way to exploit these relationships to create new models for the signal pdf.

Figure 7. Relationships between the compound pdf formulation and the various distributions involved in Bayesian inference.



4.3 Scattering by intermittent turbulence (exponential distribution compounded with a lognormal distribution)

A practically important example of parametric uncertainties involves the intermittency of turbulence. By Kolmogorov's (1962) so-called *refined hypothesis*, the structure-function parameters of turbulence vary in time and space according to a lognormal distribution. Since the scattering cross section in the inertial subrange (and hence the scattered signal power) is proportional to the structure-function parameter (e.g., Ostashev and Wilson 2015), Gurvich and Kukharets (1986) proposed modeling strong scattering with an exponential distribution in which the local variations in the scattered signal power are modeled with a lognormal distribution. Wilson et al. (1996) applied this formulation to sound scattering into a shadow zone and found good agreement with experimental data.

The formulation of Gurvich and Kukharets (1986) translates directly to the compound pdf formulation presented in section 4.1. Since the scattered

power is the uncertain parameter, we parameterize the exponential pdf using the scale parameter θ as in equation (4) (as opposed to using the rate parameter). For the lognormal pdf of θ, we have

$$p(\boldsymbol{\theta}|\boldsymbol{\chi}) = p(\theta|\mu,\phi) = \text{Lognorm}(\theta|\mu,\phi) = \frac{1}{\theta\phi\sqrt{2\pi}}\exp\left[-\frac{(\ln\theta-\mu)^2}{2\phi^2}\right]. \quad (20)$$

Thus, we have equations for $p(s|\boldsymbol{\theta})$ and $p(\boldsymbol{\theta}|\boldsymbol{\chi})$ as needed for the compound pdf integral, equation (18). Unfortunately, the compound pdf apparently has no analytical solution in this case. Nonetheless, it can readily be performed numerically, and we can also find an approximate analytical solution that is valid when $\phi^2 \ll 1$. This case was analyzed in section 2.2 and 2.4 and shown to correspond to situations in which the lognormal pdf can be approximated by a normal or gamma pdf. Here, small ϕ^2 would correspond to weak intermittency in the turbulence. Section 2.4 showed that the gamma pdf approximation to the lognormal pdf corresponds to $\alpha = \phi^{-2}$ and $\beta = \phi^2 e^{\mu}$, where α and β are the shape and rate parameters, respectively. This problem will be analyzed in section 4.4 where we show that an exponential pdf compounded with a gamma pdf produces a K-distribution, equation (21).

4.4 K-distributions (gamma compounded with gamma)

The next compound pdf we consider is described in a book on optical wave scattering in random media (Andrews and Phillips 2005). To motivate the formulation, those authors write

> . . . it has been observed that the lognormal PDF [for weakly scattered signals] . . . can underestimate the peak of the probability density function and also underestimate the behavior of the tails as compared with measured data. Underestimating the tails of a PDF has important consequences on radar and communication systems where detection and fade probabilities are calculated over the tails of the PDF.

The compound pdf, or *modulation process* in the terminology of Andrews and Phillips, is introduced to provide distributions with more realistic tails. The result (among others in the Andrews and Phillips book) is the K-distribution, so-called because it involves a modified Bessel function.

The ordinary K-distribution involves compounding an exponential pdf for s, as appropriate for strong scattering, with a gamma pdf based on the scale parameter, θ. Physically, θ represents the *local* mean of the scattered power, as would be observed if the scattering process were not undergoing modulations. Denoting the parameters of the gamma pdf for θ as α and β (i.e., the hyperparameters $\boldsymbol{\chi}$),* we have, by substituting into equation (18),

$$p(s|\boldsymbol{\chi}) = p(s|\alpha,\beta) = \frac{1}{\theta}\int e^{-s/\theta}\frac{\beta^{\alpha}\theta^{\alpha-1}}{\Gamma(\alpha)}e^{-\beta\theta}\,d\theta\ .$$

The integral can be found in standard integral tables, with result

$$p(s|\boldsymbol{\chi}) = p(s|\alpha,\beta) = \frac{2\beta}{\Gamma(\alpha)}(\beta s)^{(\alpha-1)/2}K_{\alpha-1}\left(2\sqrt{\beta s}\right). \tag{21}$$

The *generalized* K-distribution results from compounding a gamma pdf for s (with parameters k and $\lambda = \theta^{-1}$) with a gamma pdf (hyperparameters α and β) for the local mean scattered power θ. The result is

$$p(s|\alpha,\beta,k) = \mathrm{Kgen}(s|\alpha,\beta,k) = \frac{2\beta}{\Gamma(k)\Gamma(\alpha)}(\beta s)^{(k+\alpha-2)/2}K_{\alpha-k}\left(2\sqrt{\beta s}\right). \tag{22}$$

The generalized K-distribution should be reasonable for either weak or strong scattering. For $k = 1$, the generalized K-distribution reduces to the ordinary K-distribution. Note that the generalized K-distribution, like the generalized gamma distribution, has three parameters.

Figure 8 compares the K-distribution to the exponential distribution for various values of the parameter α. Clearly, the effect of decreasing α is to elevate the tail of the pdf. The K-distribution is thus helpful for explaining overdispersion in the data. Figure 9 provides a similar comparison between the generalized K-distribution and the gamma pdf for weak scattering, specifically $k = 8$.

* We use α and β for the shape and rate parameters of a gamma pdf for the modulating process so as to avoid confusion with the shape and rate parameters k and $\lambda = \theta^{-1}$ when a gamma pdf is used for the scattering process.

Figure 8. K-distribution for various values of the parameter α (*solid lines*). The *dashed line* is the prediction for the exponential distribution.

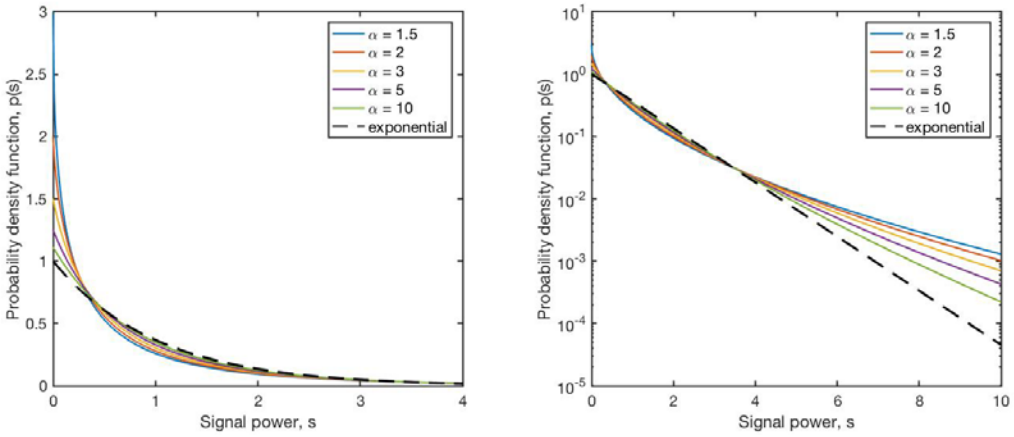
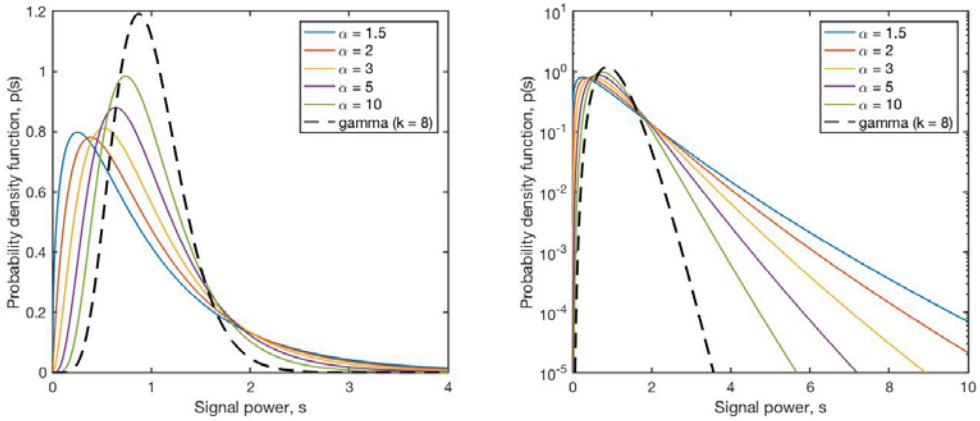


Figure 9. Same as Fig. 8 except that the generalized K-distribution is compared to the gamma distribution for weak scattering ($k = 8$).



4.5 Lomax distribution (exponential distribution compounded with a gamma distribution)

As discussed in section 2.1, the signal power has an exponential pdf for strong scattering. The parameterization of this pdf based on the rate parameter λ was given by equation (3). In the Bayesian context, the exponential pdf corresponds to the likelihood function. The conjugate prior to an exponential likelihood function is known to be the gamma distribution (e.g., Gelman et al. 2014). Hence, we set

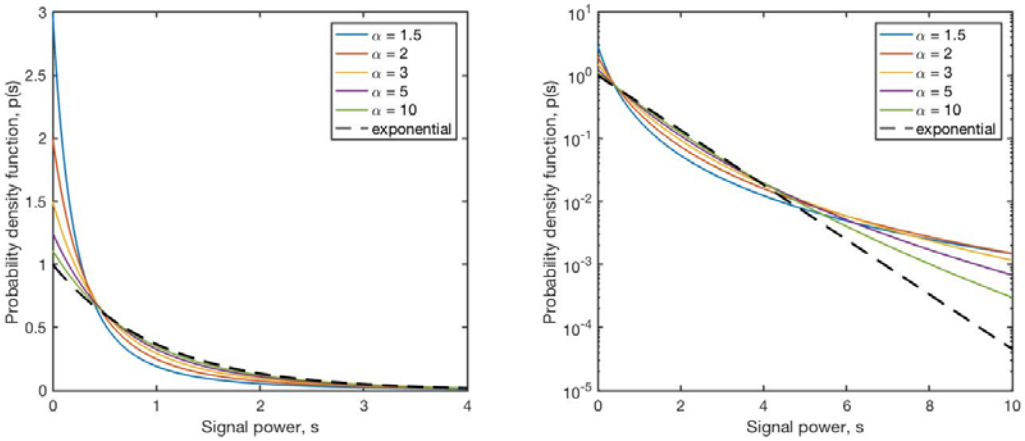
$$p(\lambda|\alpha,\beta) = \text{Gamma}(\lambda|\alpha,\beta) = \frac{\beta^{\alpha}\lambda^{\alpha-1}}{\Gamma(\alpha)}e^{-\beta\lambda}, \tag{23}$$

where α and β are the shape and rate parameters of the gamma distribution describing the mean scattered power, which is assumed to be uncertain. Substituting into equation (18) and integrating, we find for the model evidence (i.e., the compound pdf)

$$p(s|\alpha,\beta) = \text{Lomax}(s|\alpha,\beta) = \frac{\alpha\beta^{\alpha}}{(s+\beta)^{\alpha+1}} = \frac{\alpha}{\beta(1+s/\beta)^{\alpha+1}}. \qquad (24)$$

This pdf is called a *Lomax* (Lomax 1954) or *Pareto Type II* distribution. The mean of the Lomax distribution is $\langle s\rangle = m = \beta/(\alpha-1)$. The derivation of the Lomax distribution is essentially the same as the K-distribution (section 4.4) except that it is based upon a gamma pdf for λ rather than for θ. Figure 10 shows the Lomax pdf for various values of α. As with the K-distribution, decreasing α elevates the tail of the pdf. Note, however, that the Lomax pdf achieves this behavior with a much simpler equation.

Figure 10. Lomax distribution for various values of the parameter α (*solid lines*). The *dashed line* is for the exponential distribution.



It happens that the Lomax pdf can also be derived by compounding the θ parameterization of the exponential pdf, as given by equation (4), with an *inverse* gamma pdf. The inverse gamma pdf is given by

$$\text{InvGamma}(x|\alpha,\beta) = \frac{\beta^{\alpha}x^{-\alpha-1}}{\Gamma(\alpha)}e^{-\frac{\beta}{x}}.$$

Proceeding as before, it can be shown that the compound pdf again evaluates to equation (24). This exercise demonstrates that we obtain equivalent results by formulating the problem using λ and the gamma pdf or by using $\theta = 1/\lambda$ and the inverse gamma pdf.

Interestingly, the Lomax pdf can also be derived by an entirely different approach based on the Student's t-distribution. The t-distribution behaves similarly to the normal distribution except that the tails are elevated. The standardized (zero-mean, unit-variance) form of this pdf is given by

$$p(t|\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

This pdf can be extended to two uncorrelated variables as follows (Wikipedia 2017c):

$$p(t_1, t_2|\nu) = \frac{1}{2\pi}\left(1+\frac{t_1^2+t_2^2}{\nu}\right)^{-\frac{\nu+2}{2}}.$$

Setting $x = \sqrt{\beta/\nu}\,t_1$ and $y = \sqrt{\beta/\nu}\,t_2$, where β is a scaling factor, we have

$$p(x, y|\nu, \beta) = \frac{\nu}{2\pi\beta}\left(1+\frac{x^2+y^2}{\beta}\right)^{-\frac{\nu+2}{2}}.$$

Following the same procedure we used in section 2.1 to derive the exponential pdf for the power *s*, we arrive at

$$p(s|\nu, \beta) = \frac{\nu}{2\beta}\left(1+\frac{s}{\beta}\right)^{-\frac{\nu+2}{2}}.$$

By setting $\alpha = \nu/2$, we can show that this result is equivalent to equation (24). Hence, the Lomax pdf corresponds to uncorrelated real and imaginary parts of a signal given by a joint t-distribution, much as the exponential pdf corresponds to uncorrelated real and imaginary parts given by a joint normal distribution. (We are unaware of any previous publications where this connection has been described.)

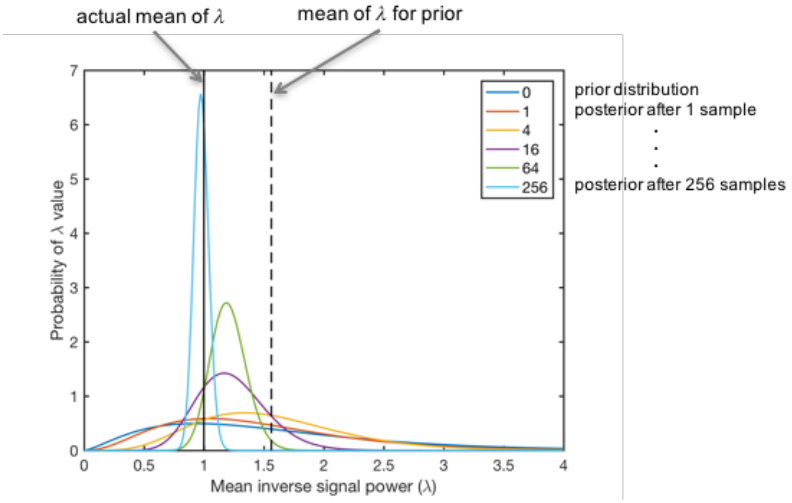
Let us return to exploring the relationship between the compound pdf and Bayes' theorem. Using equation (19), we find for the posterior pdf

$$p(\lambda|s, \alpha, \beta) = \frac{(\beta+s)^{\alpha+1}\lambda^{\alpha}}{\Gamma(\alpha+1)}e^{-\lambda(\beta+s)}. \quad (25)$$

Note that the posterior is a gamma pdf for λ (thus proving that the gamma pdf is indeed the conjugate prior of an exponential likelihood function); however, α updates to $\alpha + 1$, and β updates to $\beta + s$, in comparison to the original prior. Hence, given a new signal observation s, we would update the distribution for λ from Gamma$(\lambda|\alpha, \beta)$ to Gamma$(\lambda|\alpha + 1, \beta + s)$. This process can be applied iteratively to a set of multiple observations of the process s. Namely, suppose we have the observations $s_1, \ldots, s_N$, then α updates to $\alpha + N$, and β updates to $\beta + s_1 + \cdots + s_N$.

Let us consider a numerical example illustrating Bayesian sequential updating of the pdf for λ as signal samples are collected. Without loss of generality, we will assume that the actual value of λ is 1. Initially, we start with a gamma pdf with $\alpha = 2.5$ and $\beta = 1.6$ to represent the uncertainty in our knowledge of λ (the prior). These values imply an initial distribution for λ with a mean of 1.5625 and produce a very broad distribution, as indicative of a situation where our knowledge of the actual value of λ is limited. (In Bayesian terminology, the prior is rather *uninformative*.) We then begin to collect signal samples, s_i, each of which is a sample from the exponential pdf. Because of the random scattering, each signal sample provides additional information on the true value of λ, but we do not know this value exactly. Figure 11 shows an example simulation that involved sequentially updating the gamma pdf for λ after 1, 4, 16, 64, and 256 random trials (samples of s_i). After a single update, the distribution remains close to the prior. After 256 updates, the pdf has a sharp peak near the true value of λ.

Figure 11. Simulation of the refinement of the pdf for the rate parameter λ (inverse of the mean scattered power) as signal samples are collected. The curve labelled 0 is the initial assumed distribution (prior). Subsequent curves show updated (posterior) distributions after collecting the indicated number of samples.



4.6 Compound gamma distribution (gamma distribution compounded with a gamma distribution)

In this section, we generalize the problem in section 4.5 to the situation where the scattering (likelihood function) is described by a gamma pdf rather than an exponential pdf. Recall that the gamma pdf applies to weak or strong scattering whereas the exponential pdf applies only to strong scattering. The treatment here still focuses on the situation where λ is the uncertain parameter; the parameter k in the gamma pdf for the scattered signal is regarded as known (fixed). This assumption simplifies the model by allowing us to use a single-variate pdf for $p(\boldsymbol{\theta}|\boldsymbol{\chi})$ rather than a multivariate (joint) pdf. Physically, the implication is that the relative contributions of the deterministic and random scattered paths to the signal (and hence the shape of the distribution) remain fixed. A simple case where this sort of model is applicable is when the power of the source (or, similarly, the gain of the receiver) is imperfectly known. It would also be reasonable for a scattering process when the amount of energy reaching the receiver varies (or is uncertain) although the relative amount of energy propagating along direct (deterministic) and scattered (random) paths is comparatively stable.

It can be shown that the gamma distribution is the conjugate prior for the gamma likelihood function, just as it was for the exponential likelihood function. Hence, we still use equation (23). Performing the integration for the compound pdf, equation (18), we find

$$p(s|\alpha,\beta,k) = \frac{1}{B(k,\alpha)}\frac{(\beta/s)^{\alpha}}{s(1+\beta/s)^{k+\alpha}} = \frac{1}{B(k,\alpha)}\frac{(s/\beta)^{k-1}}{\beta(1+s/\beta)^{k+\alpha}}, \quad (26)$$

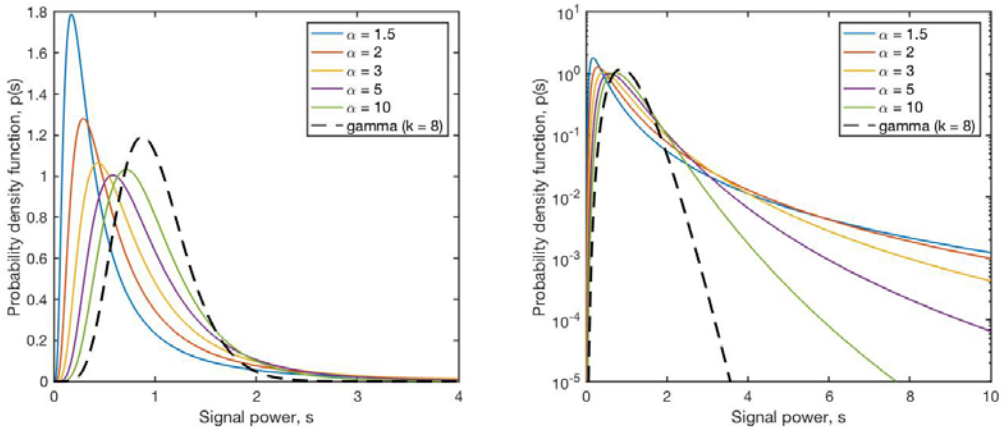
where $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ is the beta function. This result is called a *compound gamma distribution* (Dubey 1970), which is a special case of the generalized beta-prime distribution with the shape parameter set to 1 (Wikipedia 2017a). The compound gamma pdf reduces to the Lomax when $k = 1$ and has mean $\langle s\rangle = m = \beta k/(\alpha - 1)$.

Figure 12 shows the compound gamma distribution for various values of α. The curves are qualitatively similar to the K-distribution.

Although the compound gamma distribution and the generalized K-distribution (section 4.4) are derived in a very similar manner, the resulting equation for a compound gamma pdf is much simpler and enables simple Bayesian sequential updating. Applying Bayes' theorem as before, we find

that the prior Gamma$(\lambda|\alpha,\beta)$ updates to the posterior Gamma$(\lambda|\alpha+k,\beta+s)$ after a new signal observation s. The sequential updating equations are thus essentially the same as the strong scattering case considered in the previous section.

Figure 12. Compound gamma distribution for $k = 8$ for various values of the parameter α (*solid lines*). The *dashed line* is for the gamma distribution for $k = 8$.



4.7 Lognormal compounded with normal (single variate)

Compound models for the lognormal distribution are most naturally formulated using the rv $\eta = \ln s$, in which case the baseline distribution is normal. For the compound pdf, we have the integral

$$p(\eta|\boldsymbol{\chi}_\eta) = \int p(\eta|\boldsymbol{\theta}_\eta)p(\boldsymbol{\theta}_\eta|\boldsymbol{\chi}_\eta)d\boldsymbol{\theta}_\eta,$$

where $\boldsymbol{\theta}_\eta$ are the parameters of the distribution for η and $\boldsymbol{\chi}_\eta$ are the hyperparameters for the distribution of $\boldsymbol{\theta}_\eta$. For the lognormal distribution, $\boldsymbol{\theta}_\eta$ may consist of μ or ϕ, or both.

As discussed in section 4.6, the modeling is much simplified if we consider just one uncertain parameter. Let us for now take this parameter to be μ and assume that μ is normally distributed with mean m_μ and variance σ_μ^2. Hence the preceding equation becomes

$$p\left(\eta\middle|\phi, m_\mu, \sigma_\mu^2\right) = \int p(\eta|\mu,\phi)p\left(\mu\middle|m_\mu,\sigma_\mu^2\right)d\mu,$$

where

$$p(\mu|m_\mu,\sigma_\mu^2) = \mathcal{N}(\mu|m_\mu,\sigma_\mu^2) = \frac{1}{\sigma_\mu\sqrt{2\pi}}\exp\left[-\frac{(\mu-m_\mu)^2}{2\sigma_\mu^2}\right].$$

We then have for the integral

$$p\left(\eta\middle|\phi, m_\mu, \sigma_\mu^2\right) = \frac{1}{2\pi\phi\sigma_\mu}\int_{-\infty}^{\infty}\exp\left[-\frac{(\eta-\mu)^2}{2\phi^2} - \frac{(\mu-m_\mu)^2}{2\sigma_\mu^2}\right]d\mu\,, \quad (27)$$

which can be written in the general form

$$p\left(\eta\middle|\phi, m_\mu, \sigma_\mu^2\right) = \frac{1}{2\pi\phi\sigma_\mu}\int_{-\infty}^{\infty}\exp\left[-(ax^2+bx+c)\right]dx,$$

where

$$a = \frac{\sigma_\mu^2+\phi^2}{2\sigma_\mu^2\phi^2},$$

$$b = -\frac{\sigma_\mu^2\eta + \phi^2 m_\mu}{\sigma_\mu^2\phi^2},$$

and

$$c = \frac{\sigma_\mu^2\eta^2 + \phi^2 m_\mu^2}{2\sigma_\mu^2\phi^2}.$$

This integral can be found in standard tables, with result

$$p\left(\eta\middle|\phi, m_\mu, \sigma_\mu^2\right) = \frac{1}{2\pi\phi\sigma_\mu}\sqrt{\frac{\pi}{a}}\,e^{(b^2-4ac)/4a} = \frac{1}{\sqrt{2\pi(\sigma_\mu^2+\phi^2)}}\exp\left[-\frac{(\eta-m_\mu)^2}{2(\sigma_\mu^2+\phi^2)}\right]. \quad (28)$$

Thus, η is normally distributed with mean m_μ and variance $\sigma_\mu^2+\phi^2$.

We could have inferred the previous result more simply by noticing that equation (27) is the convolution between two normal distributions, namely a distribution with mean m_μ and variance σ_μ^2, and a distribution with zero mean and variance ϕ^2. A convolution between two independent rvs equals the distribution of the sum of those variables. Using the well-known result that the sum of two normally distributed variables is itself normally distributed, with a mean equal to the sum of the means and variance equal to the sum of the variances, it follows that η is normally distributed with mean m_μ and variance $\sigma_\mu^2+\phi^2$.

Using Bayes' theorem, we have the following equation for updating the estimate of μ after a new observation of the log-signal, η:

$$p\left(\mu\middle|\eta,\phi,m_\mu,\sigma_\mu^2\right)=\frac{p(\eta|\mu,\phi)p\left(\mu\middle|m_\mu,\sigma_\mu^2\right)}{p\left(\eta\middle|\phi,m_\mu,\sigma_\mu^2\right)}.$$

At this point, we have worked out equations for all of the distributions on the right. Substituting, we find after some algebra

$$p\left(\mu\middle|\eta,\phi,m_\mu,\sigma_\mu^2\right)=\mathcal{N}\left(\mu\middle|m'_\mu,\sigma'^2_\mu\right)=\frac{1}{\sigma'_\mu\sqrt{2\pi}}\exp\left[\frac{(\mu-m'_\mu)^2}{2\sigma'^2_\mu}\right], \quad (29)$$

where

$$m'_\mu=\frac{\phi^2 m_\mu+\sigma_\mu^2\eta}{\phi^2+\sigma_\mu^2}=\left(\sigma_\mu^{-2}+\phi^{-2}\right)^{-1}\left(\sigma_\mu^{-2}m_\mu+\phi^{-2}\eta\right)$$

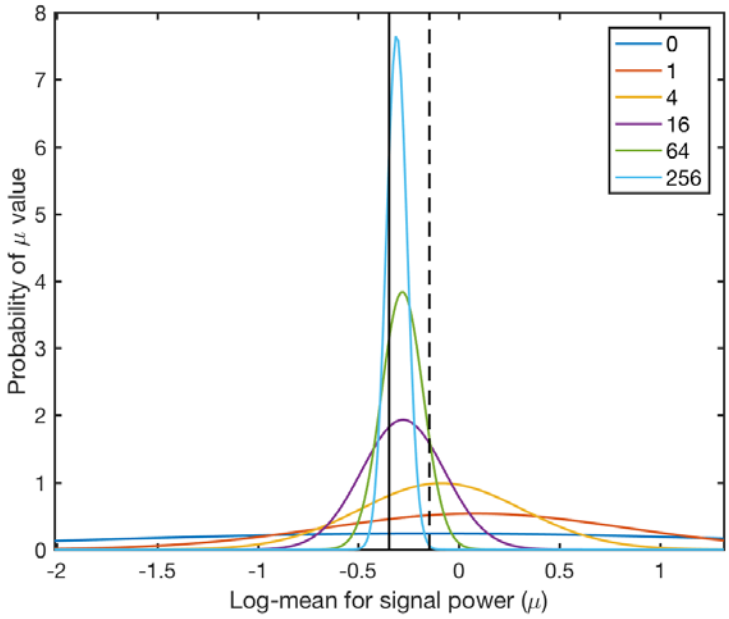
and

$$\sigma'^2_\mu=\frac{\phi^2\sigma_\mu^2}{\phi^2+\sigma_\mu^2}=\left(\sigma_\mu^{-2}+\phi^{-2}\right)^{-1}.$$

Equation (29) is the desired equation for sequentially updating estimates of μ as observations of the log-signal η become available.

Figure 13 is similar to Figure 11 except that a simulation of Bayesian updating for the log-mean (as based on the just-described lognormal signal model) is shown. The parameter updates are performed with equation (29). In this simulation, the mean m and variance σ^2 were both set to 1, as expected in strong scattering. The resulting values of the log-signal parameters are $\mu=-0.3466$ and $\phi^2=0.6931$. For the prior distribution of μ, we somewhat arbitrarily set $m_\mu=-0.1466$ and $\sigma_\mu^2=2.773$ to represent an uninformative prior. The figure clearly demonstrates the convergence to a peaked distribution around the correct value of μ as more data samples become available.

Figure 13. Refinement of the pdf for the log-mean signal distribution parameter, μ, as more log-signal samples are collected. The curve labelled 0 is the initial assumed distribution (prior). The subsequent curves show updated (posterior) distributions after 1, 4, 16, 16 and 256 random trials. The *vertical dashed line* indicates the initial mean of the prior; the *solid line* indicates the correct value of μ.



Besides the case considered here where there is parametric uncertainty in the mean μ of the distribution for the log-signal, solutions are available for the cases when the variance ϕ^2 is uncertain and when both the mean and variance are uncertain. In these cases, the conjugate priors consist of an inverse gamma pdf and a normal-inverse gamma pdf, respectively.

4.8 Lognormal compounded with normal (multivariate)

In the multivariate case, the compound pdf for the log-signal vector $\boldsymbol{\eta}$ (in the case of unknown $\boldsymbol{\mu}$ and known $\boldsymbol{\Phi}$) is

$$p\left(\boldsymbol{\eta}\middle|\boldsymbol{\Phi},\boldsymbol{m}_{\mu},\boldsymbol{\Sigma}_{\mu}\right)=\int p(\boldsymbol{\eta}|\boldsymbol{\mu},\boldsymbol{\Phi})p\left(\boldsymbol{\mu}\middle|\boldsymbol{m}_{\mu},\boldsymbol{\Sigma}_{\mu}\right)d\boldsymbol{\mu}. \tag{30}$$

The multivariate pdf describing the uncertainty in $\boldsymbol{\mu}$ is

$$p\left(\boldsymbol{\mu}\middle|\boldsymbol{m}_{\mu},\boldsymbol{\Sigma}_{\mu}\right)=\frac{1}{\sqrt{(2\pi)^{K}|\boldsymbol{\Sigma}_{\mu}|}}\exp\left[-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{m}_{\mu})^{T}\boldsymbol{\Sigma}_{\mu}^{-1}(\boldsymbol{\mu}-\boldsymbol{m}_{\mu})\right], \tag{31}$$

where $\boldsymbol{m}_\mu$ is the mean vector and $\boldsymbol{\Sigma}_\mu$ the mean covariance matrix for $\boldsymbol{\mu}$. Substituting equation (31) for $p(\boldsymbol{\mu}|\boldsymbol{m}_\mu, \boldsymbol{\Sigma}_\mu)$ and equation (13) for $p(\boldsymbol{\eta}|\boldsymbol{\mu}, \boldsymbol{\Phi})$ into equation (30), we find

$$p(\boldsymbol{\eta}|\boldsymbol{\Phi}, \boldsymbol{m}_\mu, \boldsymbol{\Sigma}_\mu) = \frac{1}{(2\pi)^K \sqrt{|\boldsymbol{\Phi}||\boldsymbol{\Sigma}_\mu|}} \int \exp\left\{-\frac{1}{2}\left[(\boldsymbol{\eta}-\boldsymbol{\mu})^T \boldsymbol{\Phi}^{-1}(\boldsymbol{\eta}-\boldsymbol{\mu}) + (\boldsymbol{\mu}-\boldsymbol{m}_\mu)^T \boldsymbol{\Sigma}_\mu^{-1}(\boldsymbol{\mu}-\boldsymbol{m}_\mu)\right]\right\} d\boldsymbol{\mu}.$$

Direct solution of this integral is rather involved. However, we can save some work by observing that the sum of two normally distributed multivariates works essentially the same way as the sum of two normally distributed single rvs except that we sum the mean vectors and covariance matrices instead of the scalar means and variances. Hence the result of integrating the preceding equation is

$$p(\boldsymbol{\eta}|\boldsymbol{\Phi}, \boldsymbol{m}_\mu, \boldsymbol{\Sigma}_\mu) = \frac{1}{\sqrt{(2\pi)^K |\boldsymbol{\Phi} + \boldsymbol{\Sigma}_\mu|}} \exp\left[-\frac{1}{2}(\boldsymbol{\eta}-\boldsymbol{m}_\mu)^T (\boldsymbol{\Phi}+\boldsymbol{\Sigma}_\mu)^{-1}(\boldsymbol{\eta}-\boldsymbol{m}_\mu)\right]. \quad (32)$$

Let us next consider application of the multivariate distribution to Bayesian inference. The multivariate form of Bayes' theorem for this problem is

$$p(\boldsymbol{\mu}|\boldsymbol{\eta}, \boldsymbol{\Phi}, \boldsymbol{m}_\mu, \sigma_\mu^2) = \frac{p(\boldsymbol{\eta}|\boldsymbol{\mu}, \boldsymbol{\Phi}) p(\boldsymbol{\mu}|\boldsymbol{m}_\mu, \boldsymbol{\Sigma}_\mu)}{p(\boldsymbol{\eta}|\boldsymbol{\Phi}, \boldsymbol{m}_\mu, \boldsymbol{\Sigma}_\mu)}.$$

Although we will not provide derivations here, it can be shown that the equations for the priors and parameter updates have the same forms for the multivariate distribution and for the single variate distribution. Specifically, with equation (13) as the likelihood function, the prior distribution is given by equation (31), and the posterior is found to be

$$p(\boldsymbol{\mu}|\boldsymbol{\eta}, \boldsymbol{\Phi}, \boldsymbol{m}_\mu, \boldsymbol{\Sigma}_\mu) = \frac{1}{\sqrt{(2\pi)^K |\boldsymbol{\Sigma}'_\mu|}} \exp\left[-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{m}'_\mu)^T (\boldsymbol{\Sigma}'_\mu)^{-1}(\boldsymbol{\mu}-\boldsymbol{m}'_\mu)\right], \quad (33)$$

where

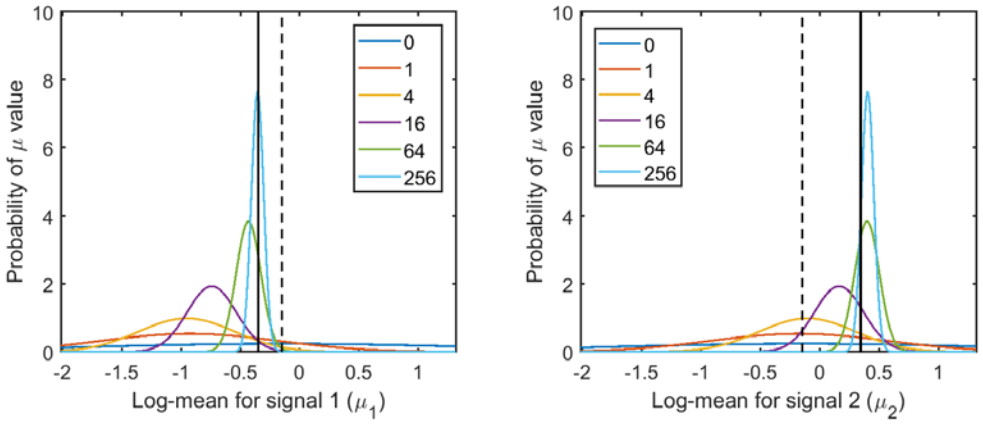
$$\boldsymbol{m}'_\mu = \left(\boldsymbol{\Sigma}_\mu^{-1} + \boldsymbol{\Phi}^{-1}\right)^{-1}\left(\boldsymbol{\Sigma}_\mu^{-1}\boldsymbol{m}_\mu + \boldsymbol{\Phi}^{-1}\boldsymbol{\eta}\right)$$

and

$$\boldsymbol{\Sigma}'_\mu = \left(\boldsymbol{\Sigma}_\mu^{-1} + \boldsymbol{\Phi}^{-1}\right)^{-1}.$$

Figure 14 is like Figure 13 except that a bivariate case ($N = 2$) is shown, and equation (33) is used for the parameter updates. In this simulation, the mean $\boldsymbol{m}$ was set to $[1; 2]$ and the variances $\boldsymbol{\sigma}^2$ to $[1; 4]$ in accordance with strong scattering (for which the ratio of the variance to the squared mean is one). The resulting mean for the log-signal parameters is $\boldsymbol{\mu} = 0.3466[-1; 1]$. (Note that the means of the two signals differ to add realism to the problem.) The matrix $\boldsymbol{\Phi}$ is $\phi^2[1\ \rho; \rho\ 1]$, where $\phi^2 = 0.6931$ and $\rho = 0.5$. For the prior distribution of $\boldsymbol{\mu}$, we set $\boldsymbol{\mu}_\mu = -0.1466[1; 1]$ and $\Sigma_\mu = 2.773[1\ \rho; \rho\ 1]$ to represent an uninformative prior. As with the previous single variate example, the bivariate update converges well to distributions that are sharply peaked around the correct values of the means.

Figure 14. Same as Fig. 13 except for a bivariate normal distribution. *Left* is the prior distribution and its updates for the first transmission path; *right* is for the second transmission path.

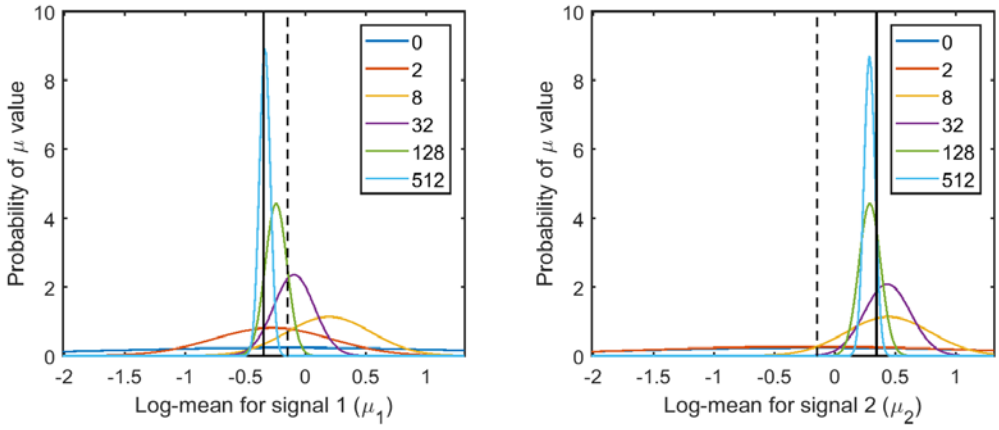


In a typical experiment, it may happen that just a subset of the paths is sampled at each iteration. For example, if we have multiple source and receiver locations, typically each "event" may involve just one source location so that the sample consists of only those paths corresponding to that one source location and each receiver location able to detect the source. Only those transmission paths should be used to update the prior.

A simple approach to handling such situations where data are missing is to assign the variances associated with the missing data (in the covariance matrix $\boldsymbol{\Phi}$) to infinity. In effect, this means that whatever values are used for the missing data in $\boldsymbol{\eta}$, they will be viewed as entirely uninformative and thus have no impact on the adaptation. Numerically, of course, we actually set the variances to a very large value, which is large enough to effectively remove the missing data entries while still enabling the $\boldsymbol{\Phi}^{-1}$ to be calculated without errors.

Figure 15 shows a simulation based on this approach. It is essentially the same as the simulation shown in Figure 14 except that we randomly select one of the transmission paths as missing at each iteration. If the first path is missing, we set $\boldsymbol{\Phi} = [\varsigma^2 \ \phi\varsigma\rho; \phi\varsigma\rho \ \phi^2]$, where ς is a very large value. If the second path is missing, we set $\boldsymbol{\Phi} = [\phi^2 \ \phi\varsigma\rho; \phi\varsigma\rho \ \varsigma^2]$. For present purposes, a value of 10^6 is used for ς. Since only half as much data is available at each iteration for the simulation shown in Figure 15 in comparison to the simulation shown in Figure 14, we would expect the results of the former to be qualitatively similar to the latter when the number of iterations is doubled. Hence, in Figure 15, results have been plotted after twice the number of iterations in comparison to Figure 14. Indeed, there is a strong qualitative similarity between the two figures.

Figure 15. Same as Fig. 14 except that one of the two transmission paths is randomly unavailable at each trial. Note that the number of trials in the legend is doubled in comparison to Fig. 14.



Although we make no attempt here to derive equations for parametric uncertainties and Bayesian sequential updating with the matrix Wishart and gamma distributions, this can probably be done by similar procedures to the single variate cases. We leave that effort to future research.

5 Influence on Receiver Operating Characteristics

Classically, the trade-off between the detections and false alarms for a sensor system is characterized by the receiver operating characteristic (ROC) curve. A high probability of detection must usually be maintained at a low false-alarm rate for a system to be practical. The ROC curves depend on the pdfs of the signal and noise, which are often calculated from model distributions (e.g., exponential, gamma, lognormal, and Rice) which appropriately describe the random processes underlying the signal variations. In practice, however, these model distributions are often based on greatly simplified characterizations of the underlying processes, including the signal production mechanism (the signature of the target), propagation of the signal through the environment, and the processing performed by the sensor. When we fail to reckon with the parametric uncertainties, the ROC curves may not be meaningfully characterized, and we may be overconfident in our predictions of the system performance. We would thus like to ascertain how parametric uncertainties impact the ROC curves.

Many textbooks discuss the basic problem of detecting a signal in noise (e.g., Burdic 1991). The general starting point is the joint distribution (pdf) for the signal s and noise n, $p(s,n|\boldsymbol{\theta})$, which depends on the combined signal/noise parameter set $\boldsymbol{\theta}$. If the signal and noise are independent, $p(s,n|\boldsymbol{\theta}) = p(s|\boldsymbol{\theta}_s)p(n|\boldsymbol{\theta}_n)$, where $\boldsymbol{\theta}_s$ represents the subset of parameters from θ impacting the signal pdf and $\boldsymbol{\theta}_n$ is likewise the subset of parameters impacting the noise pdf. (Note that the symbol $\boldsymbol{\theta}$ in previous sections corresponds to $\boldsymbol{\theta}_s$ in this section.)

Formulation of the detection problem requires the pdf for two cases: (1) when the observed signal x consists only of noise ($x = n$) and (2) when x consists of signal plus noise ($x = s + n$). The pdf for the first case, which we designate as $p_0(x)$, can be found by integrating (marginalizing) the joint pdf over the signal:

$$p_0(x|\boldsymbol{\theta}_n) = p(n|\boldsymbol{\theta}_n) = \int p(s,n|\boldsymbol{\theta})\, ds. \tag{34}$$

The pdf for the second case, $p_1(x)$, is

$$p_1(x|\boldsymbol{\theta}) = \int p(s, x-s|\boldsymbol{\theta})\, ds. \tag{35}$$

Given the pdfs $p_0(x)$ and $p_1(x)$, we can determine the probabilities of false alarm and detection from the integrals

$$P_{\mathrm{fa}}(\boldsymbol{\theta}_n) = \int_{\gamma}^{\infty} p_0\,(x|\boldsymbol{\theta}_n)\, dx \tag{36}$$

and

$$P_{\mathrm{d}}(\boldsymbol{\theta}) = \int_{\gamma}^{\infty} p_1\,(x|\boldsymbol{\theta})\, dx \tag{37}$$

in which γ is the detector threshold, a quantity set by the detection algorithm. Of course, we would normally desire a high probability of detection and low probability of false alarm. The constant false-alarm rate detector provides a simple and useful approach to setting the threshold. Specifically, the threshold is set to achieve a specified probability of false alarm; in effect, equation (36) is inverted to calculate γ from P_{fa}. Although this inversion is nontrivial for most noise pdfs, in practice it can be done readily by numerical methods.

Let us consider application of the compound gamma formulation from section 4.6 (which incorporates parametric uncertainties) in the context of detecting a signal in noise. We assume that the signal and noise both obey gamma distributions, $p(s|\boldsymbol{\theta}_s) = \mathrm{Gamma}(s|k_s, \lambda_s)$ and $p(n|\boldsymbol{\theta}_n) = \mathrm{Gamma}(n|k_n, \lambda_n)$, respectively. The parameters λ_s and λ_n are uncertain and modeled by gamma distributions, with parameters (α_s, β_s) and (α_n, β_n), respectively, whereas the shape parameters k_s and k_n are considered fixed, as discussed previously.

Since *s* and *n* are given by compound gamma pdfs, the mean values are $\langle s\rangle = \beta_s k_s/(\alpha_s - 1)$ and $\langle n\rangle = \beta_n k_n/(\alpha_n - 1)$. In the following calculations, we will set $\langle s\rangle = 1$ without loss of generality. Defining the mean signal-to-noise ratio as $\mathrm{SNR} = \langle s\rangle/\langle n\rangle$, we have $\langle n\rangle = 1/\mathrm{SNR}$. The values $\alpha_s = \beta_s k_s + 1$ and $\alpha_n = \beta_n k_n(\mathrm{SNR}) + 1$ then follow. Hence there are four free parameters in the problem as defined, namely β_s, k_s, β_n, and k_n.

Analytical solutions are not available for all the distributions necessary to calculate the ROC curves, particularly $p_1(x)$ in equation (35). However, Monte Carlo methods provide a viable and relatively simple way to study the behavior of these curves. The basic idea is to draw a large number M of random samples for λ_s from a gamma distribution with the specified values of α_s and β_s. Using these values for λ_s and the specified value for k_s as parameters in a gamma distribution, we then draw M random values for s. A set of M random values for n are generated similarly.

Figure 16 and Figure 17 show ROC curves based on such Monte Carlo simulations. A constant false-alarm rate detector is employed. For both figures, the SNR is set to 2. The signal and noise have gamma distributions with shape factors $k_s = 1$ and $k_n = 4$ for Figure 16 and $k_s = k_n = 4$ for Figure 17. The scaling parameters λ_s and λ_n also follow gamma distributions with β_s and β_n being varied together from $\beta = 0.5$, 1, 2, 4, and 8, as shown in the legends of the figures. The parameters α_s and α_n are determined as described above.

Figure 16. Receiver operating characteristic (ROC) curves corresponding to a gamma-distributed signal with shape factor $k_s = 1$ and gamma-distributed noise with $k_n = 4$. The signal-to-noise ratio is 2, and the hyperparameter β is varied as shown in the legend.

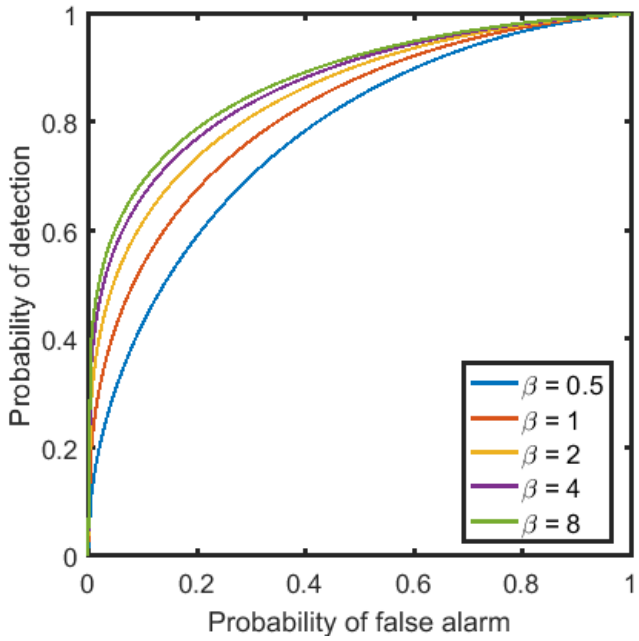
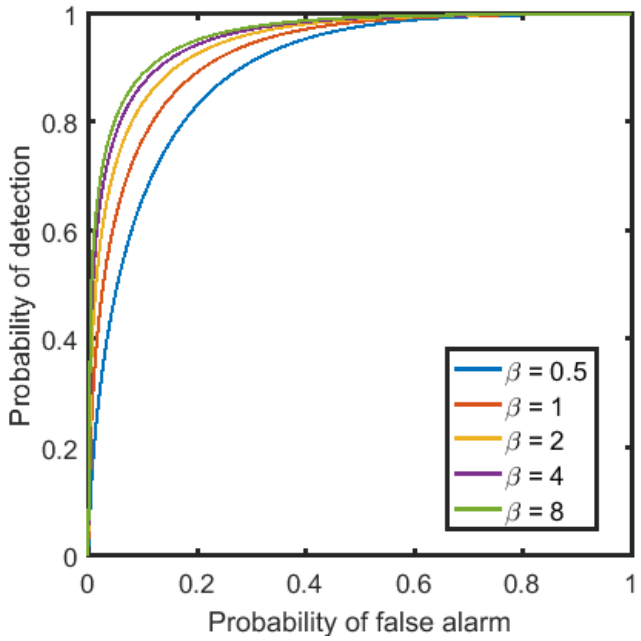


Figure 17. Same as Fig. 16 except that $k_s = 4$ and $k_n = 4$.



The value of β is seen to have a very substantial impact on the ROC curves. For large β (cases where the mean is much larger than the variance), the calculations converge on the conventional prediction, which does not include variations in the value of λ (i.e., there is no intermittency in the scattering). Smaller values of β lead to the occurrence of more extreme values for λ, which unfavorably impacts detection, particularly by increasing the occurrence of false alarms.

6 Practical Modeling with Multiple Sources of Uncertainty

Many different pdfs for signals and parametric uncertainties were discussed in chapters 2–4. In this section, we consider practical application of those pdfs to real situations involving multiple sources of uncertainty. Such uncertainties may pertain, for example, to the source strength, the sensitivity of the receiver, the terrain properties, and the weather.

First, in section 6.1, we describe a practical analytic approach based on the assumption that the parametric uncertainties are weak and describable by lognormal pdfs. Next, in section 6.2, we discuss how multiple sources of uncertainty can be addressed through a multilevel probabilistic framework. Lastly, in section 6.3, we consider a general approach based on Monte Carlo sampling of the parametric uncertainties.

6.1 Weak parametric uncertainties: Lognormal approach

The lognormal pdf is useful for dealing with products of uncertain quantities. For example, the received signal generally has the form $s = a\psi g$, where a is the source power, ψ is the transmission gain (the gain over the propagation path from the source to the receiver), and g is the gain (sensitivity) of the receiver. Taking logarithms, we have $\eta = \ln s = \ln a + \ln \psi + \ln g$. Let us suppose that a, ψ, and g are all lognormally distributed with parameters (μ_a, ϕ_a^2), etc. Then, since the sum of independent normally distributed rvs is itself normally distributed, η must be normally distributed with log-mean $\mu = \mu_a + \mu_\psi + \mu_g$ and log-variance $\phi^2 = \phi_a^2 + \phi_\psi^2 + \phi_g^2$.

Next, let us suppose that there is a signal scattering process, which can be described by a gamma pdf for the scattered power. Furthermore, there are weak uncertainties in the source power, mean scattered signal power, and receiver gain. The latter can be described in accordance with a lognormal model as described in the preceding paragraph. In essence, we have a gamma wave scattering process, the mean of which is conditioned upon the lognormal variate $\theta = a\psi g$. Mathematically, this is just like the situation analyzed in section 4.3 (strong scattering by intermittent turbulence) except that we have decomposed θ into several independent contributions and do not assume that the scattering is strong. The compound pdf, equation (18), becomes

$$p(s|\mu,\phi,k) = \int p(s|k,\theta)p(\theta|\mu,\phi)\,d\theta, \tag{38}$$

where $p(s|k,\theta) = \text{Gamma}(s|k,\theta^{-1})$ (k being the shape factor for the gamma pdf), $p(\theta|\mu,\phi) = \text{Lognorm}(\theta|\mu,\phi)$, $\mu = \mu_a + \mu_\psi + \mu_g$, and $\phi^2 = \phi_a^2 + \phi_\psi^2 + \phi_g^2$.

Let us furthermore assume that the log-variances for a, g, and r are all small; that is, $\phi_a^2 \ll 1$, $\phi_\psi^2 \ll 1$, and $\phi_g^2 \ll 1$. The log-variance for θ must then also be small. For this reason, we call the analysis here the case of weak parametric uncertainty. As discussed in sections 2.2 and 2.4, we can approximate the lognormal pdf for θ with a gamma pdf in which $\alpha = \phi^{-2}$ and $\beta = \phi^2 e^{\mu}$. That is, $p(\theta|\mu,\phi) = \text{Lognorm}(\theta|\mu,\phi) \cong \text{Gamma}(\theta|\phi^{-2},\phi^2 e^{\mu})$. Substituting into equation (38), the integral becomes a gamma pdf for θ compounded by another gamma pdf. This was the case analyzed in section 4.4, which led to a generalized K-distribution, equation (22). Thus, we have the following solution for scattering with a gamma pdf with weak parametric uncertainties:

$$p(s|\mu,\phi,k) = \text{Kgen}(s|\phi^{-2},\phi^2 e^{\mu},k). \tag{39}$$

6.2 Multilevel modeling

Following the basic pattern of the compound pdf, equation (18), we can extend the compound modeling approach to the case where the hyperparameters are uncertain and depend on an additional, higher-level parameter set, $\boldsymbol{\psi}$. Integrating both sides of equation (18) over $\boldsymbol{\chi}$, we have

$$p(s|\boldsymbol{\psi}) = \int p(s|\boldsymbol{\chi})p(\boldsymbol{\chi}|\boldsymbol{\psi})d\boldsymbol{\chi} = \int \left[\int p(s|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\chi})d\boldsymbol{\theta}\right] p(\boldsymbol{\chi}|\boldsymbol{\psi})\,d\boldsymbol{\chi}. \tag{40}$$

Note that the first of these equalities (the one with the single integral over $\boldsymbol{\chi}$) is functionally the same as equation (18); $\boldsymbol{\chi}$ appears in the place of $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ in the place of $\boldsymbol{\chi}$. Equations (18) and (40) are examples of hierarchical stochastic models: randomness at a lower level in the model is linked to a higher-level random process. These equations have two and three levels in the hierarchy, respectively.

Wilson et al. (2008) formulated a hierarchical model for predicting the signal and noise distributions in the presence of uncertainties in the environment and source/receiver characteristics. In their formulation, the joint pdf $p(s,n)$ (where n is the noise power) was written

$$p(s,n) = \int\int\int p(s,n|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\chi})p(\boldsymbol{\chi}|\boldsymbol{\psi})p(\boldsymbol{\psi})\, d\boldsymbol{\theta}\, d\boldsymbol{\chi}\, d\boldsymbol{\psi},$$

where $\boldsymbol{\theta}$ is a set of parameters for the joint pdf of the signal and noise, $\boldsymbol{\chi}$ is a set of parameters needed by the propagation model, and $\boldsymbol{\psi}$ is a set of environmental parameters needed by the propagation model. This equation is equivalent to equation (40) after marginalization over $\boldsymbol{\psi}$. Wilson et al. (2008) associated $\boldsymbol{\psi}$ with uncertain atmospheric and terrain parameters and $\boldsymbol{\chi}$ with uncertainties in the parameters of a signal propagation (transmission) model.

Although models with three or more levels may be useful in certain applications, analytical solutions are usually unavailable except for very simple cases, such as normal distributions with uncertain means. One other case where we have found an analytical solution involves extending the compound gamma pdf from section 4.6 with an additional level of parametric uncertainty in which the scaling parameter β is given by a gamma pdf. That is, we solve equation (40) with $p(s|\boldsymbol{\theta}) = p(s|\lambda) = \text{Gamma}(s|k,\lambda)$, $p(\boldsymbol{\theta}|\boldsymbol{\chi}) = p(\lambda|\beta) = \text{Gamma}(\lambda|\alpha,\beta)$, and $p(\boldsymbol{\chi}|\boldsymbol{\psi}) = p(\beta|a,z) = \text{Gamma}(\beta|\kappa,\eta)$, where κ and η are the hyperparameters of the pdf for β. As we have already shown in section 4.6, the result of the inner integral (the integral in square brackets) in equation (40) is a compound gamma pdf, equation (26). We thus have the following integral to solve:

$$p(s|\boldsymbol{\psi}) = p(s|\kappa,\eta,\alpha,k) = \frac{\Gamma(k+\alpha)}{\Gamma(k)\Gamma(\alpha)\Gamma(\kappa)}\int_0^\infty \frac{(\beta/s)^\alpha}{s(1+\beta/s)^{k+\alpha}}\,\eta^\kappa\beta^{\kappa-1}e^{-\eta\beta}d\beta.$$

Changing the variable of integration to $t = \beta/s$, we find

$$p(s|\boldsymbol{\psi}) = p(s|\kappa,\eta,\alpha,k) = \frac{\Gamma(k+\alpha)(\eta s)^\kappa}{\Gamma(k)\Gamma(\alpha)\Gamma(\kappa)s}\int_0^\infty t^{\kappa+\alpha-1}(1+t)^{-k-\alpha}e^{-\eta s t}\,dt.$$

The remaining integral is proportional to the Tricomi confluent hypergeometric function $U(a,b,z)$ (e.g., Abramowitz and Stegun 1984, equation [13.2.5]), which is given by

$$U(a,b,z) = \frac{1}{\Gamma(a)} \int_0^\infty t^{a-1}(1+t)^{b-a-1} e^{-zt}\, dt.$$

Identifying $a = \kappa + \alpha$, $b = \kappa - k + 1$, and $z = \eta s$ leads to

$$p(s|\kappa,\eta,\alpha,k) = \frac{\Gamma(k+\alpha)\Gamma(\kappa+\alpha)(\eta s)^\kappa}{\Gamma(k)\Gamma(\alpha)\Gamma(\kappa)s} U(\kappa+\alpha, \kappa-k+1, \eta s). \quad (41)$$

Although it is interesting to see that this three-level hierarchical model can be written using a known function, this result does not appear to be particularly useful since very few numerical libraries include the confluent hypergeometric function.

Because of the challenges of finding analytical solutions for multilevel models, it is natural to consider numerical methods. However, numerical evaluation of the multidimensional integrals by conventional methods such as trapezoidal integration can also be prohibitive. Hence, stochastic methods, such as Monte Carlo integration, are an attractive approach. That is the topic of the next section.

6.3 Monte Carlo approach to parametric uncertainties

For problems involving multiple uncertain parameters, it often becomes prohibitive to perform the integrations over each parameter explicitly by either analytical or conventional numerical methods. In such situations, Monte Carlo methods, which randomly sample the integrand, are particularly useful. In this section, we describe the basic idea behind Monte Carlo integration and show how it applies to the sampling of parametric uncertainties.

Let us approximate the integral in the compound pdf, equation (18), as a discrete sum over the integration variable, $\boldsymbol{\theta}$. For simplicity, we consider for now the case where $\boldsymbol{\theta}$ consists of a single random variable so that the integral is one dimensional. We have (formally, using the right Riemann sum to approximate the integral)

$$p(s|\boldsymbol{\chi}) \approx \sum_{i=1}^{I} p(s|\theta_i) p(\theta_i|\boldsymbol{\chi})\, \Delta\theta_i,$$

where θ_i, $i = 1, \ldots, I$ are points along the θ-axis at which the integrand is sampled and $\Delta\theta_i = \theta_i - \theta_{i-1}$. Many strategies for selecting the sample points and intervals have been developed. One strategy, which may not seem immediately obvious or useful, is to *randomly* select the θ_i on the interval over which θ is defined. Taking the nominal spacing between the samples to be $\Delta\theta_i = 1/I$, we have

$$p(s|\boldsymbol{\chi}) \approx \frac{1}{I}\sum_{i=1}^{I} p(s|\theta_i)p(\theta_i|\boldsymbol{\chi}).$$

Note that this approximation is simply an average of the integrand as evaluated at the sample points θ_i. This is the simplest (often called the *naïve*) form of Monte Carlo sampling. One advantage of this approach is that we can readily adjust I to obtain an appropriate trade-off between computation time and accuracy. A small value can be used to obtain a fast but rough approximation to the integral whereas a large value can be used for a more accurate approximation. A less obvious advantage is that this method can be readily extended to multidimensional integrals; essentially, one just generates random samples of the vector $\boldsymbol{\theta}$ instead of scalar values. Therein lies the primary attraction of Monte Carlo sampling: it turns out to be a very simple and numerically efficient approach for multidimensional integrals. Practically, the implementation is the same for integrals with one dimension or with many dimensions. Although we do not discuss Monte Carlo methods in detail here, the interested reader may refer to Evans and Swartz (2000), O'Leary (2004), and many other sources for more-detailed discussion.

Note that the preceding approximation can be viewed as a weighted sum of samples of $p(s|\boldsymbol{\theta}_i)$; that is, we could write it in the form

$$p(s|\boldsymbol{\chi}) \approx \sum_{i=1}^{I} w_i(\boldsymbol{\chi})\, p(s|\boldsymbol{\theta}_i), \tag{42}$$

where $w_i(\boldsymbol{\chi}) = p(\boldsymbol{\theta}_i|\boldsymbol{\chi})/I$. The interpretation of this equation is that for each of the random samples of $\boldsymbol{\theta}_i$, we evaluate the function $p(s|\boldsymbol{\theta}_i)$ based on our model for the wave scattering process. We then estimate the overall distribution of the scattered signal by averaging the samples of $p(s|\boldsymbol{\theta}_i)$ in accordance with their probability of occurrence (i.e., the weights $w_i(\boldsymbol{\chi})$).

The more samples we take, the better our estimate. Equation (42) exemplifies a *mixture distribution*, meaning that it is a weighted sum of distributions.

The observation that the integral can be interpreted as a set of samples of $p(s|\boldsymbol{\theta}_i)$, weighted by factors proportional to $p(\boldsymbol{\theta}_i|\boldsymbol{\chi})$, suggests an alternative sampling strategy: instead of drawing the $\boldsymbol{\theta}_i$ uniformly from the entire interval and then weighting by $p(\boldsymbol{\theta}_i|\boldsymbol{\chi})/I$, we could alternatively draw random samples of $\boldsymbol{\theta}_i$ from $p(\boldsymbol{\theta}_i|\boldsymbol{\chi})$ so that they are intrinsically weighted in proportion to their probability of occurrence. Our approximation would then become simply

$$p(s|\boldsymbol{\chi}) \approx \frac{1}{I}\sum_{i=1}^{I} p(s|\boldsymbol{\theta}_i). \tag{43}$$

Thus, given a method for sampling from the distribution for $p(\boldsymbol{\theta}_i|\boldsymbol{\chi})$, we can quickly form numerical estimates of $p(s|\boldsymbol{\chi})$.

The Monte Carlo approach can be readily extended to formulations with additional model levels as in equation (40). For example, we could randomly generate a sequence of values $\boldsymbol{\chi}_j$, where $j = 1, \ldots, J$, by sampling from the distribution $p(\boldsymbol{\chi}|\boldsymbol{\psi})$. In the multilevel formulation, however, for each $\boldsymbol{\chi}_j$, there will be a distinct distribution for $\boldsymbol{\theta}$, namely $p(\boldsymbol{\theta}|\boldsymbol{\chi}_j)$. We indicate the samples drawn from $p(\boldsymbol{\theta}|\boldsymbol{\chi}_j)$ as $\boldsymbol{\theta}_{ij}$, $i = 1, \ldots, I$. Hence, the inner integral (in square brackets in equation (40)) is approximated as

$$I_j = \int p(s|\boldsymbol{\theta}) p(\boldsymbol{\theta}|\boldsymbol{\chi}_j) d\boldsymbol{\theta} \approx \frac{1}{I}\sum_{i=1}^{I} p(s|\boldsymbol{\theta}_{ij}) .$$

The multilevel integral then evaluates as

$$p(s|\boldsymbol{\psi}) \approx \frac{1}{J}\sum_{j=1}^{J} I_j \approx \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I} p(s|\boldsymbol{\theta}_{ij}). \tag{44}$$

Although equation (44) retains the double summation resulting from the double integral in equation (40), this structure is unnecessary when we are approximating the integrals by Monte Carlo methods. It is just as effective to sample over all of the rvs at once. Conceptually, we can think of sam-

pling just one value of $\boldsymbol{\theta}_{ij}$ for each j; that is, we set $I = 1$. Hence, equation (44) reduces to the single summation

$$p(s|\boldsymbol{\psi}) \approx \sum_{j=1}^{J} p\left(s\middle|\boldsymbol{\theta}_j\right), \tag{45}$$

where $\boldsymbol{\theta}_j$ is a single sample drawn from the distribution $p\left(\boldsymbol{\theta}\middle|\boldsymbol{\chi}_j\right)$ with $\boldsymbol{\chi}_j$ being drawn from the distribution $p(\boldsymbol{\chi}|\boldsymbol{\psi})$.

Alternatively, we could envision performing the integral over $\boldsymbol{\theta}$ appearing in equation (40) analytically by employing one of the distributions derived in chapter 4 to model $p(s|\boldsymbol{\chi})$. We then have only a single-level problem remaining for the Monte Carlo sampling. The solution to this problem is the same as equation (42) except with $\boldsymbol{\chi}$ replacing $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ replacing $\boldsymbol{\chi}$.

A situation of practical interest where we might use one of the formulations in this section involves ensemble weather forecasting in which multiple forecasts are generated by systematically perturbing the initial conditions in accordance with the ranges of uncertainty in the weather observations. This results in a randomized set of forecasts, each of which is a plausible outcome given the uncertainty and chaos in the modeling process (e.g., Gneiting and Raftery 2005). The forecast ensemble includes several different atmospheric fields that are responsible for wave scattering phenomena. For example, acoustic scattering is caused by wind velocity and temperature fluctuations whereas RF scattering is caused by pressure and water vapor fluctuations. Depending on the implementation details, the ensemble could correspond to either the index i or j above.

7 Conclusion

In this report, we reviewed a number of basic probability distributions (more precisely, probability density functions, or pdfs) for scattered signals and considered the extension of these basic pdfs to incorporate uncertainties in the statistical attributes of the signals.

Among the basic pdfs considered were the exponential pdf, which is a one-parameter pdf intended for strongly scattered signals; the lognormal pdf, which is a two-parameter pdf intended for weakly scattered signals; the Rice and gamma pdfs, which are two-parameter pdfs applicable to either weakly or strongly scattered signals; and the generalized gamma pdf, which extends the gamma pdf so as to address the frequency of extreme fluctuations (the tails of the pdf).

Uncertainties can be conveniently and systematically addressed with a compound pdf, which incorporates separate pdfs for the wave scattering process and for the uncertain parameters in the wave scattering. A number of formulations based on this approach (K-distribution, Lomax distribution, and compound gamma distribution) were described and compared. The consistent theme is that uncertainty raises the tails of the pdfs, which can have important implications for detection and communication-system performance.

We also considered extension of the basic pdfs to transmissions along multiple paths. Mathematically, this corresponds to extending single-variate pdfs to multivariate or matrix pdfs. Such an extension is straightforward for the lognormal pdf. The matrix Wishart distribution, with two degrees of freedom, appears to provide a useful matrix pdf for strong scattering. We also speculated that the matrix gamma pdf might be useful for weak or strong scattering although we did not investigate this possibility in detail.

Table 1 summarizes the various pdfs that were discussed and their attributes. Of the many single-variate pdfs discussed here, the compound gamma pdf would appear to be the most generally useful; it applies to weak or strong scattering, incorporates an elevated tail as is characteristic of intermittent phenomena, and has convenient analytical properties.

Table 1. Summary of signal distribution models and their physical associations. Cases with known Bayesian conjugate priors are indicated in *red*. In the table, "prior" and "posterior" refer to the distributions for the *parameters* describing the signal power (the hyperparameters). Cases where results are not known are indicated by question marks.

Likelihood function (random signal model)	Prior (for mean signal power)	Physical interpretation	Posterior (for mean signal power)	Model evidence (signal model with uncertainty)
Exponential	Gamma (mean)	Strong scattering, single receiver	(non-analytic)	K-distribution
Exponential	Gamma (rate)	Strong scattering, single receiver	Gamma (rate)	Lomax
Exponential	Log-normal	Strong scattering w/turbulent intermittency, single receiver	(non-analytic)	(non-analytic)
Rice	(?)	Weak (Born) or strong scattering, single receiver	(?)	(?)
Gamma	Gamma (mean)	Weak (empirical) or strong, single receiver	(non-analytic)	Generalized K-distribution
Gamma	Gamma (rate)	Weak (empirical) or strong, single receiver	Gamma (rate)	Compound gamma distribution
Log-normal	Normal	Weak (Rytov) scattering, single receiver	Normal	T distribution
Log-normal, multivariate	Multivariate normal	Weak (Rytov) scattering, multiple receivers	Multivariate normal	T distribution, multivariate
Wishart, 2 degree-of-freedom	(matrix inverse gamma?)	Strong scattering, multiple receivers	(matrix inverse gamma?)	(?)
Matrix gamma	(matrix inverse gamma?)	Weak (empirical) or strong, multiple receivers	(matrix inverse gamma?)	(?)

The report also explored how the modeling of parametric uncertainties naturally relates to Bayesian inference of the wave scattering parameters and how this relationship can be exploited. In particular, by interpreting the basic scattering model as a likelihood function and modeling the parametric uncertainties using the Bayesian conjugate prior, convenient analytical solutions, already well known in the literature, can be found. The Bayesian connection also leads to convenient sequential updating algorithms, which refine an initial prediction of the wave scattering parameters as new signal observations (i.e., observations of the random scattering process) become available. In practice, the sequential updating algorithm provides a recipe for scenarios in which the signal parameters are uncertain and repeated signal transmissions are made along the same path to gradually refine the parameter values.

Finally, the report addressed the problem of calculating probabilities of false alarm and detection (i.e., receiver operating characteristics, or ROC curves) when parametric uncertainties are present. The parametric uncertainties were shown to substantially degrade the ROC curves. The primary

issue is that the tails of the pdfs for the signal and noise are significantly elevated when uncertainties are included. Since detector thresholds are normally set to make false alarms very rare and we normally wish to operate with a probability of detection close to 1, the tails of the signal and noise distributions are very important. It is plausible that, in practice, extreme values of the signal and noise are more frequent than expected, due to our limited knowledge of all the processes impacting the signal and noise distributions.

In this technical report, we consolidated and extended many practical statistical treatments for uncertainties in signal properties and linked those treatments to methods for efficiently reducing the uncertainties as additional signal observations become available. Although the only application considered in detail here was calculation of ROC curves, the framework provided can actually be applied to many other problems involving battlefield sensing and communication for which random signal behavior is important. Another application would be to optimally predict signal characteristics along a relatively poorly sampled transmission path based on characteristics along a better-sampled path. This would be useful, for example, for assessing the properties of an unknown signal emitter on the basis of transmission samples between nearby friendly units. Another potentially important application is the development of improved automated target recognition (ATR) algorithms, which account for random wave scattering and while providing realistic self-assessments of the uncertainties in the target classification. Taken together, these advances will help the Army operate more effectively in complex environments where communication and sensing systems are strongly influenced by random signal behavior.

References

Abramowitz, M., and I. A. Stegun, ed. 1984. *Pocketbook of Mathematical Functions: Abridged Edition of Handbook of Mathematical Functions*. Material selected by M. Danos and J. Rafelski. Frankfurt: Verlag Harri Deutsch.

Andrews, L. C., and R. L. Phillips. 2005. *Laser Beam Propagation through Random Media*. Bellingham, WA: SPIE Press.

Burdic, W. S. 1991. *Underwater Acoustic System Analysis*. Englewood Cliffs, NJ: Prentice Hall.

Dubey, S. D. 1970. Compound Gamma, Beta and F Distributions. *Metrika* 16:27–31.

Dyer, I. 1970. Statistics of Sound Propagation in the Ocean. *Journal of the Acoustical Society of America* 48 (2): 337–345.

Evans, M., and T. Swartz. 2000. *Approximating Integrals via Monte Carlo and Deterministic Methods*. Oxford: Oxford University Press.

Ewart, T. E., and D. B. Percival. 1986. Forward Scattered Waves in Random Media—The Probability Distribution of Intensity. *Journal of the Acoustical Society of America* 80 (6):1745–1753.

Flatté , S. M., ed., R. Dashen, W. H. Munk, K. M. Watson, and F. Zachariasen. 1979. *Sound Transmission through a Fluctuating Ocean*. New York: Cambridge University Press.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2014. *Bayesian Data Analysis, Third Edition*. Boca Raton, FL: Chapman and Hall/CRC.

Gneiting, T., and A. E. Raftery. 2005. Weather Forecasting with Ensemble Methods. *Science* 310 (5746): 248–249.

Gupta, A. K., and D. K. Nagar. 1999. *Matrix Variate Distributions*. Boca Raton, FL: Chapman and Hall/CRC.

Gurvich, A. S., and V. P. Kukharets. 1986. The Influence of Intermittence of Atmospheric Turbulence on the Scattering of Radio Waves. *Soviet Journal of Communication Technology and Electron*ics 30:52–58.

Kolmogorov, A. N. 1962. A Refinement of Previous Hypotheses Concerning the Local Structure of Turbulence in a Viscous Incompressible Flow at High Reynolds Number. *Journal of Fluid Mechanics* 13 (1): 82–85.

Lomax, K. 1954. Business Failures: Another Example of the Analysis of Failure Data. *Journal of the American Statistical Association* 49 (268): 847–852.

O'Leary, D. P. 2004. Multidimensional Integration: Partition and Conquer. *Computing in Science and Engineering* 6 (6): 58–66.

Ostashev, V. E., and D. K. Wilson. 2015. *Acoustics in Moving Inhomogeneous Media, Second Edition*. Boca Raton, FL: CRC Press.

———. 2017. Strength and Wave Parameters for Sound Propagation in Random Media. *Journal of the Acoustical Society of America* 141 (3): 2079–2092.

Ostashev, V. E., D. K. Wilson, and S. N. Vecherin. 2011. Effect of Random Impedance on the Interference of the Direct and Ground-Reflected Waves. *Journal of the Acoustical Society of America* 130 (4): 1844–1850.

Pearson, K., G. B. Jeffery, and E. M. Elderton. 1929. On the Distribution of the First Product Moment-Coefficient, in Samples Drawn from an Indefinitely Large Normal Population. *Biometrika* 21:164–201.

Rytov, S. M., Y. A. Kravtsov, and V. I. Tatarskii. 1989. *Principles of Statistical Radio Physics 4: Wave Propagation through Random Media*. Berlin: Springer.

StackExchange. 2017. Multivariate Log-Normal Probabiltiy Density Function (PDF). *StackExchange: Cross Validated*. https://stats.stackexchange.com/questions/214997/multivariate-lognormal-probabiltiy-density-function-pdf.

Suzuki, H. 1977. A Statistical Model for Urban Radio Propagation. *IEEE Transactions on Communications* 25 (7): 673–680.

Wikipedia. 2017a. Beta Prime Distribution. *Wikipedia: The Free Encyclopedia*. https://en.wikipedia.org/w/index.php?title=Beta_prime_distribution&oldid=796297710 (accessed 29 December2017).

———. 2017b. Matrix Gamma Distribution. *Wikipedia: The Free Encyclopedia*. https://en.wikipedia.org/w/index.php?title=Matrix_gamma_distribution&oldid=787262735 (accessed 27 December 2017).

———. 2017c. Multivariate t-Distribution. *Wikipedia: The Free Encyclopedia*. https://en.wikipedia.org/w/index.php?title=Multivariate_t-distribution&oldid=808355795 (accessed 3 January 2018).

Wilson, D. K., J. Wyngaard, and D. Havelock. 1996. The Effect Of Turbulent Intermittency on Scattering into an Acoustic Shadow Zone. *Journal of the Acoustical Society of America* 99 (6): 3393–3400.

Wilson, D. K., C. L. Pettit, M. S. Lewis, S. Mackay, and P. M. Seman. 2008. Probabilistic Framework for Characterizing Uncertainty in the Performance of Networked Battlefield Sensors. In *Proceedings of SPIE, Defense Transformation and Net-Centric Systems 2008*, 18–20 March, Orlando, FL, ed. R. Suresh, 6981: 698104. Bellingham, WA: Society of Photo-Optical Instrumentation Engineers.

Appendix A: Determination of Hyperparameters for the Exponential Family of Distributions

In the body of this report, we did not discuss in detail how to choose the parameters of the distributions. By exploring the general notions of exponential distributions and their implications for conjugate priors, we show in this appendix how it is possible to rigorously choose the parameters.

To start, we note that pdfs for the exponential family have the following general form:

$$p(\boldsymbol{x}|\boldsymbol{\eta}) = h(\boldsymbol{x})g(\boldsymbol{\eta})\exp[\boldsymbol{\eta}^T\boldsymbol{u}(\boldsymbol{x})],$$

where $\boldsymbol{\eta}$ are the natural parameters, $\boldsymbol{u}(\boldsymbol{x})$ are the sufficient statistics, and $g(\boldsymbol{\eta})$ can be interpreted as a normalization constant. In the case of a univariate normal distribution, it can be cast in this exponential form:

$$\mathcal{N}(x|\mu,\tau) = h(x)g(\boldsymbol{\eta})\exp[\boldsymbol{\eta}^T\mathbf{u}(x)],$$

where the natural parameters, sufficient statistics, and normalization coefficient are

$$h(x) = (2\pi)^{-1/2},$$

$$\boldsymbol{\eta} = [\mu\tau, -\tau/2, \ln\tau]^T,$$

$$\boldsymbol{u}(x) = [x, x^2, 0]^T,$$

$$g(\boldsymbol{\eta}) = \exp[\eta_1^2/4\eta_2 + \eta_3/2].$$

In a similar manner, conjugate priors can be formulated as a distribution in the exponential family

$$p(\boldsymbol{\eta}|\boldsymbol{\chi},\nu) = f(\boldsymbol{\chi},\nu)g(\boldsymbol{\eta})^\nu\exp[\nu\boldsymbol{\eta}^T\boldsymbol{\chi}].$$

The normal-Gamma cast in this form is

$$\mathcal{N}(\mu|\mu_0,(\beta\tau)^{-1})\mathrm{Gam}(\tau|a_0,b_0) = f(\boldsymbol{\chi},\nu)g(\boldsymbol{\eta})^\nu\exp[\nu\boldsymbol{\eta}^T\boldsymbol{\chi}],$$

$$\nu = \beta,$$

$$\boldsymbol{\chi} = [\mu_0, \mu_0^2 + 2b_0/\beta, (2a_0 - \beta - 1)/2\beta]^T,$$

$$f(\boldsymbol{\chi}, \nu) = (\beta/2\pi)^{1/2}[b_0^{a_0}/\Gamma(a_0)].$$

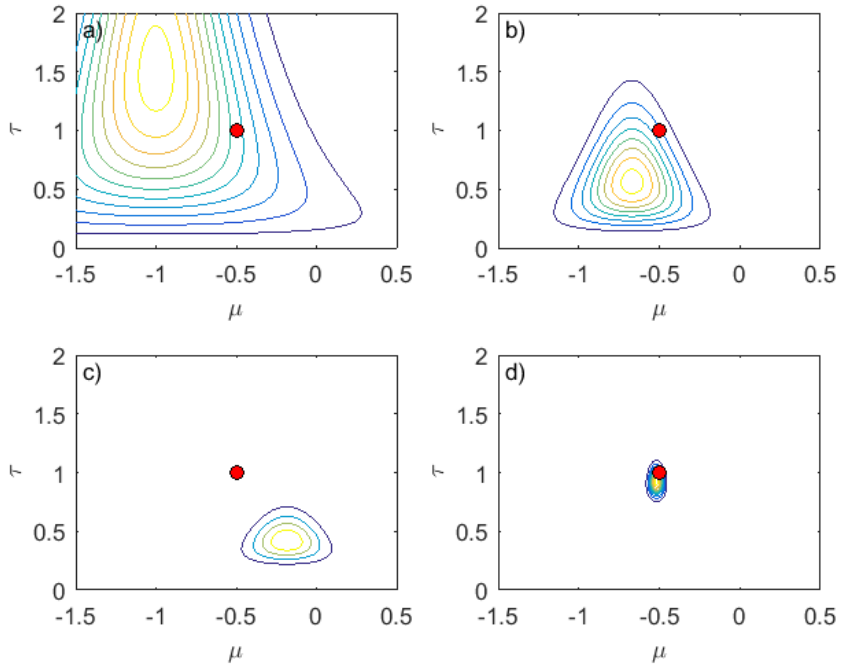
In this way, the posterior for the natural parameters, following multiplication of the conjugate prior with the likelihood function, follows in a straightforward manner:

$$p(\boldsymbol{\eta}|\mathbf{x}, \boldsymbol{\chi}, \nu) \propto g(\boldsymbol{\eta})^{\nu+N} \exp\left[\boldsymbol{\eta}^{\mathrm{T}}\left(\sum_n \mathbf{u}(x) + \nu\boldsymbol{\chi}\right)\right],$$

where we can interpret ν as an effective number of pseudo-observations of the sufficient statistics quantified by $\boldsymbol{\chi}$. Practically, we can generate an ensemble of predictions and relate the statistics of those predictions (pseudo-observations) to the hyperparmeters by the relationship of $\boldsymbol{\chi}$ to the sufficient statistics. Hence, for the normal-Gamma prior, we can interpret β as the number of ensemble predictions generated, μ_0 as the mean of the predictions, b_0 is related to the mean square of the predictions, and a_0 is determined from some defining characteristics of sound scattering. More specifically, we can choose a_0 such that the mode of the gamma distribution corresponds to experimentally observed variations in the logarithm of the signal intensity.

Figure A-1 illustrates an example of Bayesian inference for the mean and precision. The log-mean is chosen to be -0.5, and the log-precision is chosen to be 1.0. The hyperparameters for the conjugate prior are $\mu_0 = -1.0$, $\beta = 5$, $a_0 = 2$, and $b_0 = 1$. As the number of observations increases, uncertainty in the log-mean and log-precision decreases.

Figure A-1. Bayesian inference for the log-mean, μ, and log-precision, τ. Part (*a*) shows the normal-Gamma prior for the log-mean and log-precision. As new observational data is collected, the uncertainty in the log-mean and log-precision is reduced. A total of (*b*) 4, (*c*) 16, and (*d*) 256 random observations are used to update the posterior for the lognormal parameters. The *red circle* indicates the true log-mean and log-precision.



REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
July 2018	Technical Report/Final	

4. TITLE AND SUBTITLE

Impact of Parametric Uncertainties on Scattered Signal Distributions and Receiver Operating Characteristics

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

D. Keith Wilson, Daniel J. Breton, Carl R. Hart, Chris L. Pettit, Edward T. Nykaza, and Vladimir E. Ostashev

5d. PROJECT NUMBER

5e. TASK NUMBER
62272089604 & 611102T2400

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Army Engineer Research and Development Center (ERDC)
Cold Regions Research and Engineering Laboratory (CRREL)
72 Lyme Road
Hanover, NH 03755-1290

8. PERFORMING ORGANIZATION REPORT NUMBER

ERDC TR-18-7

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Headquarters, U.S. Army Corps of Engineers
Washington, DC 20314-1000

10. SPONSOR/MONITOR'S ACRONYM(S)
USACE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Environmental Quality/Installations and Geospatial Research and Engineering

14. ABSTRACT

Many different distributions are used to model statistics of waves that have been randomly scattered in atmospheric and terrain environments. These distributions have varying analytical advantages and ranges of physical applicability. This report reviews several basic distributions and discusses how they can be extended to include spatial and temporal variability in the scattering process. For this purpose, a compound probability density function (pdf) can be introduced in which a basic pdf describing the underlying scattering process is modulated by a second pdf describing parametric uncertainties in the scattering. We describe some useful new formulations based on the compound pdf, including strong and Rytov (lognormal) scattering processes modulated by the environment. These new formulations lead to relatively simple marginalized signal power distributions (Lomax and lognormal, respectively). Furthermore, we show how the conditional scattered signal pdf may be viewed as a likelihood function in which the modulating pdf is the Bayesian conjugate prior. The parameters of the modulating process can thus be refined by simple sequential Bayesian updating. Finally, the impact of the parametric uncertainties on signal detection and receiver operating characteristic curves is discussed and shown to be a very important consideration in practical applications.

15. SUBJECT TERMS
Battlefield signal propagation, Electromagnetic waves, Parametric uncertainties, Receiver operating characteristic, Scattering (Physics), Seismic waves, Sensor performance, Signal processing, Sound-waves, Wave scattering

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
Unclassified	Unclassified	Unclassified	SAR	66	