ERDC/CRREL MP-21-27

Cold Regions Research and Engineering Laboratory

US Army Corps of Engineers®
Engineer Research and Development Center

ERDC
ENGINEER RESEARCH & DEVELOPMENT CENTER

Urban Noise Distributions and the Influence of Geometric Spreading on Skewness

D. Keith Wilson, Matthew J. Kamrath, Caitlin E. Haedrich, Daniel J. Breton, and Carl R. Hart

November 2021

ERDC/CRREL MP-21-27
November 2021

Urban Noise Distributions and the Influence of Geometric Spreading on Skewness

D. Keith Wilson, Matthew J. Kamrath, Caitlin E. Haedrich, Daniel J. Breton, and Carl R. Hart

Cold Regions Research and Engineering Laboratory
U.S. Army Engineer Research and Development Center
72 Lyme Road
Hanover, NH 03775

Final report

Prepared for U.S. Army Corps of Engineers
Washington, DC 201314

Under Program Element SAB20, Project Number AB2, Task 01

Preface

This study was conducted for the U.S. Army Corps of Engineers under the U.S. Army ERDC Geospatial Research and Engineering basic research, Program Element SAB201, Project AB2, and Task 01.

The work was performed by U.S. Army Engineer Research and Development Center, Cold Regions Research Engineering Laboratory (ERDC-CRREL). At the time of publication of this paper, the Deputy Director for ERDC-CRREL was Mr. David Ringelberg and the Director was Dr. Joseph Corriveau.

This article was originally published online in the *Journal of the Acoustic Society of America* on 3 August 2021.

The Commander of ERDC was COL Teresa A. Schlosser and the Director was Dr. David W. Pittman.

Urban noise distributions and the influence of geometric spreading on skewness

ABSTRACT:
Statistical distributions of urban noise levels are influenced by many complex phenomena, including spatial and temporal variations in the source level, multisource mixtures, propagation losses, and random fading from multipath reflections. This article provides a broad perspective on the varying impacts of these phenomena. Distributions incorporating random fading and averaging (e.g., gamma and noncentral Erlang) tend to be negatively skewed on logarithmic (decibel) axes but can be positively skewed if the fading process is strongly modulated by source power variations (e.g., compound gamma). In contrast, distributions incorporating randomly positioned sources and explicit geometric spreading [e.g., exponentially modified Gaussian (EMG)] tend to be positively skewed with exponential tails on logarithmic axes. To evaluate the suitability of the various distributions, one-third octave band sound-level data were measured at 37 locations in the North End of Boston, MA. Based on the Kullback-Leibler divergence as calculated across all of the locations and frequencies, the EMG provides the most consistently good agreement with the data, which were generally positively skewed. The compound gamma also fits the data well and even outper-forms the EMG for the small minority of cases exhibiting negative skew. The lognormal provides a suitable fit in cases in which particular non-traffic noise sources dominate.

I. INTRODUCTION

Probability distributions can capture more information on the variability of noise in space and time than is possible with conventional integrated or averaged sound levels (e.g., L_{eq}) or a small set of exceedance levels (e.g., L_{10}, L_{50}, and L_{90}). Knowledge of the full distribution is potentially useful for applications such as modeling noise disturbance (García and Faus, 1991; Zuo *et al.*, 2014; De Coensel *et al.*, 2016) and signal detection probabilities (Wilson *et al.*, 2017a). Furthermore, with a suitable understanding of the environmental phenomena responsible for the distribution, observations of the distribution's statistical properties, such as its moments, asymmetry, and asymptotic tails, can possibly be used to infer properties of the environment (Stanton *et al.*, 2018).

This article examines the probability distributions for acoustic ambient noise with a focus on urban environments. In such environments, noise distributions can be impacted by many phenomena, including

- *Geometric spreading*, which leads to noise variations when the distance between the dominant source(s) and receiver is varied. This may occur, for example, when traffic on a roadway moves past a receiver.
- *Wave scattering* from small-scale details of building facades (diffusion), vegetation, and turbulence in the air and *multipath* due to reflections off buildings and diffractions around buildings. We group these phenomena together under the description of *random fading*.
- *Absorption losses* occurring during transmission through air and surface interactions.
- *Spatial and temporal variations in strength and density* of noise sources. The spatial variations may occur at the street or neighborhood levels or due to land usage; temporal variations occur due to traffic movement, the daily work schedule, and other changes in activity.
- *Mixtures of source types*, such as motor vehicles of different classes, aircraft, machinery, and human speech.

It is unclear *a priori* (and, likely, very situationally dependent) which of the preceding phenomena are typically most important in urban environments and should, therefore, be emphasized when devising suitable distribution models. The shape of the distributions is also influenced by temporal, spatial, and frequency-band averaging. Data can be averaged over time scales ranging from a fraction of a

Portions of this work were presented in "Evaluating parametric probability density functions for urban acoustic noise," Internoise 2019, Madrid, 16–19 Jun. 2019 and "Selecting a parametric probability density function for urban sound," 178th Acoustical Society of America Meeting, San Diego, CA, 2–7 Dec. 2019.

second to days or even years. They can be averaged over neighborhoods and cities and analyzed as one-third octave band levels, *A*-weighted sound levels, or with other types of band filtering.

Most of the previous literature on urban noise distributions has involved the analysis of sound-level (decibel) data, that is, logarithms of sound pressure. Wesler (1973; see also Transportation Systems Center, 1971, Appendix B) performed an extensive noise survey of Medford, MA, including a tabulation of histograms of *A*-weighted levels. Various distribution shapes are apparent, including normal (Gaussian), multi-modal, and positively skewed (asymmetric with a longer tail for relatively large sound levels). Don and Rees (1985), who examined distributions arising from traffic noise, concluded that sound-level distributions exhibit significant departures from a normal distribution due to traffic flow and mixtures of the vehicle type. García and Faus (1991) collected data at a number of locations in Spain, finding that levels were often approximately normal in very noisy locations but positively skewed in quiet locations. Zuo *et al.* (2014) collected data at 554 locations around Toronto, Canada; their distributions of the daily average data (as evident from their box plots in Fig. 2) appear to span a wide variety of characteristics. De Coensel *et al.* (2016) presented simulations of road traffic noise; although their input distribution for the vehicle sound powers was negatively skewed, the simulated levels at a distance of 15 m from the roadway appear to have a strong positive skew (their Fig. 5). Song and Lenchine (2017), analyzing noise levels from Adelaide, Australia, found a negative skewness in their daytime data and a relatively smaller positive skewness in their nighttime data. The negative skewness may result from the changing traffic flow volume throughout the day. Albert and Decato (2017) examined noise levels in urban and rural locations; although their analysis focused on the frequency dependence of the noise (as opposed to the distributions), their reported standard deviation values notably exhibited very strong dependence on the frequency range and measurement site.

Whereas it is natural to focus on sound-level (logarithmic) distributions from a noise-control perspective, from a broader scientific perspective, it is just as reasonable to analyze distributions of the sound pressure amplitude or power (amplitude squared). If the sound level is normally distributed, the amplitude and power have lognormal distributions.

In contrast to the previously mentioned, largely empirical, noise-control studies, there is a body of research examining the suitability of distributions that are theoretically justified for various situations involving multiple sources, multipath, and various scattering phenomena. Such studies can be found in the acoustical as well as the radio frequency (RF) literature, and the latter is due to their importance in communication. One important special case is a signal that has been completely randomized by the propagation medium (due, for example, to multipath reflections from buildings or scattering from facades), which results in an exponential distribution for a single source and a chi-square distribution for multiple equal sources (Burdic, 1991). The lognormal distribution arises theoretically for signals that are weakly scattered (Turin *et al.*, 1972; Strohbehn *et al.*, 1975). A steady signal embedded in random Gaussian noise has a Rice distribution for the amplitude (Strohbehn *et al.*, 1975; Suzuki, 1977). Some other distributions in use include the gamma (Suzuki, 1977), generalized gamma (Ewart and Percival, 1989), and Weibull (Tzeremes and Christodoulou, 2002). Based on realistic simulations of the sound propagation in urban environments by the finite-difference time-domain (FDTD) method, Wilson *et al.* (2006) and Ketcham *et al.* (2008) showed that a single source in a dense urban environment produces a distribution close to Rician near to the source and exponential at greater distances when the field becomes diffuse and many propagation paths are present.

This article addresses the problem of modeling urban noise distributions from a broad perspective. First, in Sec. II, the phenomena leading to different functional forms of distributions are described. The discussion begins with distributions appropriate for random fading (scattering and multipath effects) in the presence of one or more sources of approximately equal strength (Sec. II A). The motivations for and interrelationships between a number of conventional distributions appearing in the literature (such as exponential, gamma, and Rice) are described, along with an approach for extending these distributions to account for the source strength modulations (variations in space and time). Next, we consider (Sec. II B) the complementary situation of a random transmission geometry but without random fading. The single-source case is found to lead to a Pareto distribution on a linear axis and an offset exponential distribution on a logarithmic axis. The multiple-source case is found to result in distributions with statistical characteristics falling between those of the offset exponential and normal distributions, which is well approximated by an exponentially modified Gaussian (EMG) distribution. In Sec. III, we compare various theoretical models from Sec. II to measurements made in the city of Boston, MA. Section IV provides concluding remarks.

II. THEORY

This section examines the physical bases underlying a number of distributions when environmental noise is impacted by various phenomena. The distributions are formulated for the signal power, where *power* refers generically to a quantity proportional to the squared amplitude of a signal (as is common practice in signal processing). The signal, in this case, is the sound pressure.

Because environmental noise is typically analyzed in decibels (dB), which is a logarithmic scale, the appearance of the distributions on both the linear and logarithmic scales will be considered. For mathematical simplicity, in this section, we employ the natural logarithm of the power. Designating the complex pressure as $\boldsymbol{p} = x + iy$ (where x and y are the real and imaginary parts, respectively) and the

root mean square (rms) sound pressure as p_{rms}, the power is $s = |\boldsymbol{p}|^2 = 2p_{\mathrm{rms}}^2$, and its natural logarithm, termed the *log-power*, is $\zeta = \ln |\boldsymbol{p}|^2$. The familiar sound pressure level (SPL) in dB is $L_p = 20\log(p_{\mathrm{rms}}/p_{\mathrm{ref}}) = 20\log\left(|\boldsymbol{p}|/\sqrt{2}p_{\mathrm{ref}}\right)$, where log is the base-10 logarithm and $p_{\mathrm{ref}} = 20\,\mu\mathrm{Pa}$ is the standard reference pressure in air. From these relationships, it can be shown that the SPL is a linear transformation of the log-power, namely, $L_p = (10\log e)\zeta - 20\log\left(\sqrt{2}p_{\mathrm{ref}}\right) = 4.34\zeta + 90.97$. Hence, the distribution for L_p is a shifted and rescaled version of the distribution for ζ, although it has the same shape.

A. Fixed transmission geometry, random fading

We first consider conventional distributions for random fading effects, which may be used to describe multipath propagation and scattering in urban environments. The random fading may be conceptualized as predominantly a spatial process, caused by complex interferences between signals reflecting from objects such as buildings, and diffractions around such objects. However, the spatial fading process can also create variations in time when the sources are in motion. Following the discussion of conventional fading distributions, we describe their extension to account for variations in source levels in space and time.

1. Full saturation: Exponential, Erlang, and gamma distributions

Let us start with the relatively simple case in which the signal is *fully saturated*, meaning that the signal reaching the receiver consists entirely of a large number of incoherently scattered or multipath contributions. This situation is also known as *deep* or *Rayleigh fading*. The scattering and multipath may result from interactions with buildings, rough surfaces, turbulence (inhomogeneities in the air), vegetation, and other objects. The assumption of full saturation means that the complex pressure $\boldsymbol{p}$ has zero mean, which implies $\langle x\rangle = \langle y\rangle = 0$. Because the signal consists of a large number of independent, randomized contributions, by the central limit theorem, the real and imaginary parts are normally distributed with equal variance; that is, $x = \varepsilon n_x$ and $y = \varepsilon n_y$, where n_x and n_y are independent, zero-mean, unit variance random variables (rvs), and ε^2 is the variance.

One can then show that the received power $s = |\boldsymbol{p}|^2 = x^2 + y^2$ has an exponential probability density function (pdf) (Burdic, 1991; Flatté *et al.*, 1979), which is given by

$$p(s|\lambda) = \lambda e^{-\lambda s}, \tag{1}$$

where $\lambda = 1/2\varepsilon^2$ is the rate parameter. In this article, vertical lines are used to separate the arguments of the pdf for the rv and the parameters on which it depends. We indicate rvs with uppercase letters; a rv S drawn from an exponential distribution is indicated with the notation $S \sim \mathrm{Exp}(\lambda)$. For the exponential pdf, the mean, $m = \langle S\rangle$, is equal to $1/\lambda = 2\varepsilon^2$, and the variance, $\sigma^2 = \langle (S - \langle S\rangle)^2\rangle$, is equal to $1/\lambda^2$.

Suppose, next, that the received signal consists of k independent contributions, each of which is drawn from $\mathrm{Exp}(\lambda)$ and combined at the receiver [that is, we have k independent and identically distributed (iid) samples of S]. This may occur if we sum (or average) k signal power samples in space or time or if there are k independent sources. Increasing the averaging time will also effectively increase k, although the rate of increase depends on the correlation time of the process and can be challenging to quantify. [With multiple sources, the signals arriving at the receiver may actually have different mean powers and, thus, different values of λ. But it may be the case that a few independent contributions dominate, which can be approximated as having equal mean power, as hypothesized, for example, by Dyer (1970).] Because a single sample drawn from $\mathrm{Exp}(\lambda)$ corresponds to the sum of the squares of two independent, zero-mean normal variables with variance ε^2, the sum of k samples corresponds to the sum of $d = 2k$ independent, zero-mean normal variables with variance ε^2. Considering, first, the case of $\varepsilon^2 = 1$, the sum ξ can be shown to have a *chi-squared* distribution with d degrees of freedom, namely, $p_\xi(\xi) = \xi^{d/2-1}/\left(2^{d/2}\Gamma(d/2)\right)e^{-\xi/2}$ (e.g., Burdic, 1991), where $\Gamma(\cdot)$ is the gamma function.

The pdf for non-unit variance can then be calculated by making the transformation $s = \varepsilon^2\xi$ and setting $p(s) = |d\xi/ds|p_\xi(\xi)$, resulting in

$$p(s|k,\lambda) = \frac{s^{k-1}\lambda^k}{\Gamma(k)}e^{-\lambda s}. \tag{2}$$

Here, $\lambda = 1/2\varepsilon^2$, as with the exponential pdf. Equation (2) is called an *Erlang* pdf. The *gamma* pdf is given by the same equation as the Erlang, except that k may be noninteger. A rv with this distribution is indicated with the notation $S \sim \mathrm{Gamma}(k,\lambda)$. For the Erlang and gamma pdfs, $m = k/\lambda = 2k\varepsilon^2$ and $\sigma^2 = k/\lambda^2 = 4k\varepsilon^4$. The variance, normalized by the square of the mean, is $\sigma^2/m^2 = 1/k$. Dyer (1970) derives the means and variances of the log-power when the power is given by a gamma pdf. In particular, $\sigma_\zeta^2 = \pi^2/6 - S_2(k)$, where the function $S_2(k)$ equals one for $k = 1$ and $\sum_{w=1}^{k-1} w^{-2}$ for $k > 1$. Hence, $\sigma_{L_p} = 4.34\sigma_\zeta$ is 5.57 dB for $k = 1$ (the exponential pdf) and decreases to 2.31 dB for $k = 4$ and 1.10 dB for $k = 16$.

The previous results for the exponential and gamma pdfs can be readily transformed to pdfs for log-power by setting $p_\zeta(\zeta) = \left[|ds/d\zeta|p(s)\right]_\zeta = [sp(s)]_\zeta$. Furthermore, they can be transformed to pdfs for the amplitude $a = \sqrt{s}$ by setting $p_a(a) = \left[|ds/da|p(s)\right]_a = 2\left[\sqrt{s}p(s)\right]_a$. For the exponential pdf, the result for $p_a(a)$ is a Rayleigh pdf, whereas for the gamma pdf, the result is a Nakagami pdf. The Rayleigh and Nakagami pdfs are frequently employed for electromagnetic noise and propagation modeling, respectively (Strohbehn *et al.*, 1975; Suzuki, 1977).

Figure 1 shows the gamma pdf on both the linear and log axes. (Also, shown in Fig. 1 is the noncentral Erlang pdf, which will be discussed in Sec. II A 2.) Each curve

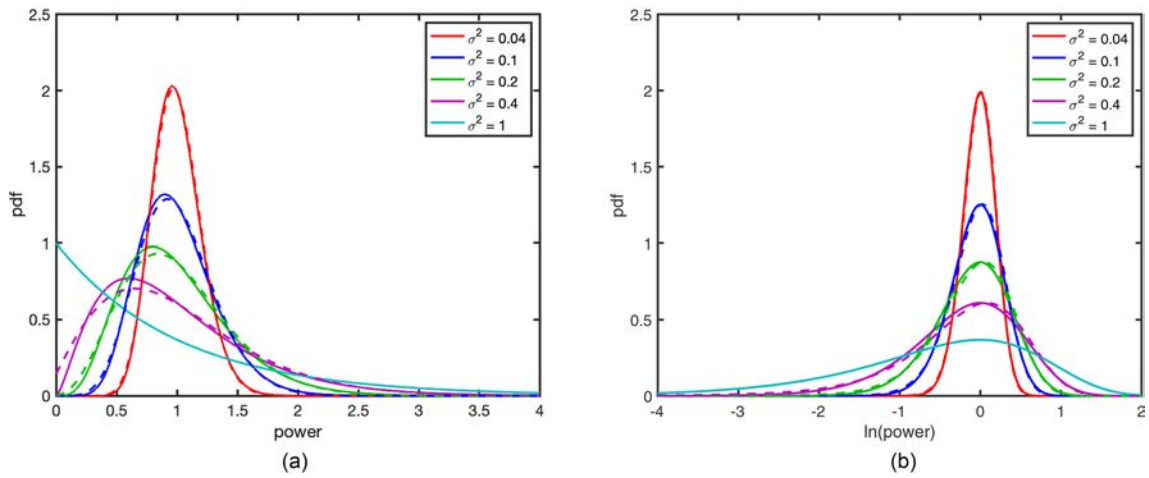

FIG. 1. (Color online) Plots of the gamma (solid lines) and noncentral Erlang (dashed lines) pdfs for various values of the variance σ^2. The mean power m is set to one for all curves. For the gamma pdf, k is set to $1/\sigma^2$, whereas for the noncentral Erlang, $d = 2$ and $\lambda = \sqrt{1-\sigma^2}$ as explained in the text. (a) Linear axis and (b) logarithmic axis.

shows a different value of the variance, namely, $\sigma^2 = 0.04$, 0.1, 0.2, 0.4, and 1. For all of the curves, $m = 1$. Because $\sigma^2 = 1/k$ when $m = 1$, the chosen variance values correspond to $k = 25$, 10, 5, 2.5, and 1, respectively. For $k = 1$, the gamma pdf reduces to the exponential. As $k \to \infty$, the central limit theorem implies that the gamma pdf takes on the appearance of a normal pdf as given by

$$p\left(s|\mu, \phi^2\right) = \frac{1}{\sqrt{2\pi}\phi} e^{-(s-\mu)^2/2\phi^2}, \tag{3}$$

where μ and ϕ^2 are the mean and variance parameters, respectively, which are approximated with k/λ and k/λ^2, respectively. The gamma distribution also has a normal-like appearance for $k \to \infty$ on the logarithmic axis in Fig. 1; that is, the gamma pdf is approximately *lognormal* in this limit as well as being normal. The lognormal pdf, which is derived from the assumption that $\zeta = \ln s$ is normal [i.e., Eq. (3) with ζ replacing s, and then applying the transformation $p(s) = [p_\zeta(\zeta)/s]_s$], is given by

$$p\left(s|\mu_\zeta, \phi_\zeta^2\right) = \frac{1}{\sqrt{2\pi}\phi_\zeta s} e^{-\left(\ln s - \mu_\zeta\right)^2/2\phi_\zeta^2}. \tag{4}$$

The mean and variance are $m = e^{\mu_\zeta + \phi_\zeta^2/2}$ and $\sigma^2 = \left(e^{\phi_\zeta^2} - 1\right)m^2$, respectively. Thus, in the limit $k \to \infty$, for which the normalized variance σ^2/m^2 is small, the gamma, normal, and lognormal pdfs approximately coincide. This explains why a normal pdf might be observed on a logarithmic axis, as with many of the observational studies cited in the Introduction, when sufficient averaging or multiple sources are present [such that by the previously mentioned calculations of Dyer (1970), σ_{L_p} is less than a few dB].

2. Unsaturated signals: Noncentral Erlang and Rice distributions

Next, let us consider the more complicated case in which the complex signal does not have zero mean (i.e., is not fully saturated). Such a situation may occur when a line-of-sight path has power comparable to the randomly scattered paths. In this case, we will need to calculate the distribution of the sum of the squares of non-zero mean normal variables. The noncentral chi-squared distribution provides the desired generalization of the chi-squared; specifically, it describes the distribution of $\xi = \sum_{i=1}^d \eta_i^2$, where $\eta_i = \mu_i + n_i$, μ_i is the mean, and n_i is a normal random variate with zero mean and unit variance. (As previously, $d = 2k$, where k could represent independent spatial or temporal samples or different sources. Conventionally, the odd indices for i represent the real parts of the signals and the even indices are the imaginary parts.) The noncentral chi-squared distribution is given by (Johnson *et al.*, 1995)

$$p_\xi(\xi|d, \lambda) = \frac{1}{2} e^{-(\xi+\lambda)/2} \left(\frac{\xi}{\lambda}\right)^{d/4-1/2} I_{d/2-1}\left(\sqrt{\lambda\xi}\right), \tag{5}$$

where $\lambda = \sum_{i=1}^d \mu_i^2$ and $I_\nu(\cdot)$ is the modified Bessel function of the first kind of order ν.

To generalize this distribution to non-unit variances, we set $\eta_i = \mu_i + \varepsilon n_i$ from which $s = \sum_{i=1}^d \eta_i^2 = \varepsilon^2 \sum_{i=1}^d \left[(\mu_i/\varepsilon) + n_i\right]^2$. Hence, $s/\varepsilon^2 = \sum_{i=1}^d \left[(\mu_i/\varepsilon) + n_i\right]^2$ has a noncentral chi-squared as given by Eq. (5) but with $\lambda' = \left(1/\varepsilon^2\right) \sum_{i=1}^d \mu_i^2$ replacing λ. The distribution for s can then be found from the transformation $p(s) = |d\xi/ds| p_\xi(\xi)$, where $\xi = s/\varepsilon^2$. Making these substitutions and then defining $\lambda = \sum_{i=1}^d \mu_i^2$ as before to eliminate λ', the result is

$$p\left(s|d, \lambda, \varepsilon^2\right) = \frac{1}{2\varepsilon^2} e^{-(s+\lambda)/2\varepsilon^2} \left(\frac{s}{\lambda}\right)^{d/4-1/2} I_{d/2-1}\left(\frac{\sqrt{\lambda s}}{\varepsilon^2}\right). \tag{6}$$

For consistency with our earlier terminology, we will call this result the *noncentral Erlang* distribution. It can be shown that $m = d\varepsilon^2 + \lambda$ and $\sigma^2 = 2\varepsilon^2(d\varepsilon^2 + 2\lambda)$. Figure 2 shows the noncentral Erlang pdf for various combinations of

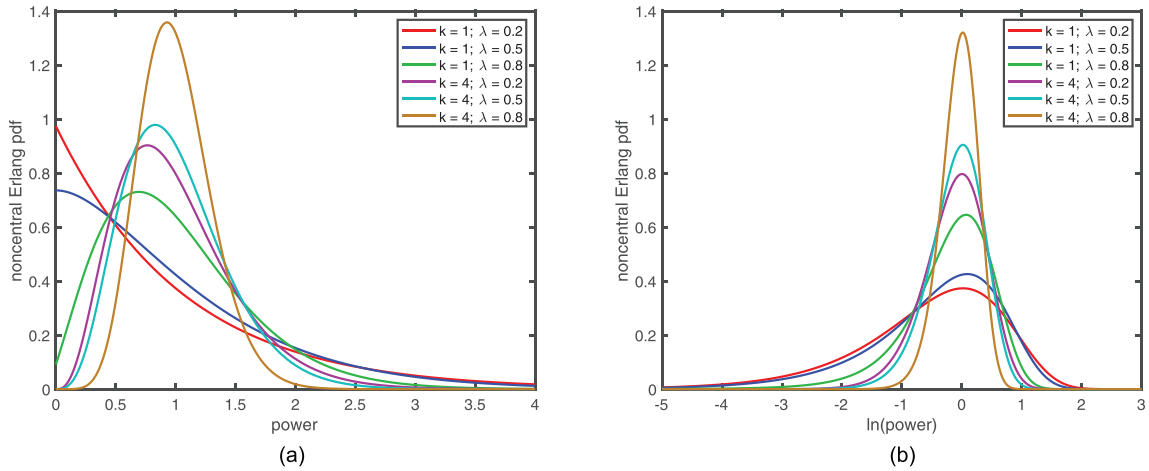

FIG. 2. (Color online) The noncentral Erlang distribution for various combinations of the parameters λ and $k = d/2$. The mean power is set to one. (a) Distributions on a linear axis and (b) logarithmic axis.

λ and $k = d/2$. For all of the curves, $m = 1$ and, hence, $\varepsilon^2 = (1-\lambda)/d$. (Note that $\lambda \leq 1$.) As k and λ increase, the pdf becomes nearly normal and eventually delta-like.

The case of a noncentral Erlang distribution with $d = 2$ corresponds to a single sample of the power, i.e., $x = \eta_1 = \mu_x + \varepsilon n_x$ and $y = \eta_2 = \mu_y + \varepsilon n_y$. We have

$$p(s|\lambda, \varepsilon^2) = \frac{1}{2\varepsilon^2} e^{-(s+\lambda)/2\varepsilon^2} I_0\left(\frac{\sqrt{\lambda s}}{\varepsilon^2}\right), \tag{7}$$

where $\lambda = \mu_x^2 + \mu_y^2$. Equation (7), when transformed from power to amplitude, is known as the Rice distribution and used widely in electromagnetics (e.g., Suzuki, 1977). Figure 1 includes the noncentral Erlang pdf for the case $k = d/2 = 1$. In these plots, m is set to one and the same values for the variance are used as in the gamma pdf. The two distributions appear very similar when the means and variances are matched. These comparisons demonstrate how pdfs for unsaturated signals are very similar to pdfs for fully saturated signals, after applying averaging over multiple samples or sources. Therefore, it is difficult to distinguish experimentally between the impacts of saturation and averaging.

3. Variable source strength: Compound gamma distribution

Next, we consider an approach for extending the previously described distributions for random fading so as to introduce a variable (unsteady or uncertain) source strength. This could be important, for example, when modeling traffic noise variations caused by changing traffic volume or vehicle types, variations in noise levels between neighborhoods, or sources, such as machines and birds, that produce sound intermittently. The approach involves *compound* probability distributions, meaning that a parameter in the original pdf for the received power is itself considered to be a rv. This rv is described by a second pdf, called the *mixing* (or *modulating*) distribution. The original pdf is termed the *conditional* distribution; the *unconditional* distribution results from integrating (marginalizing) over the randomized parameter. Although, in principle, most any distribution can be can be compounded with another, the modulating distribution should satisfactorily describe the actual source variability; for example, different distributions would be appropriate for relatively continuously varying sources such as traffic noise, whereas a binary modulating distribution would likely be appropriate for machinery that turns on and off. It is also advantageous to pair the conditional and mixing distributions so as to enable analytical results.

A particularly simple case is the compounding of a normal pdf with another. Let us suppose that the conditional pdf for the log-power ζ is normal (i.e., the power is lognormal) with mean $m_\zeta = \langle\zeta\rangle$ and variance $\sigma_\zeta^2 = \langle(\zeta - m_\zeta)^2\rangle$. As discussed in Sec. II A 1, this might occur with fully saturated scattering when a large number of signal samples are incoherently averaged. Furthermore, let us take m_ζ to be random. This might occur, for example, if the strength of the scattering or the power of the source(s) varies in time or space. Suppose that the pdf for m_ζ (the mixing distribution) is normal with mean μ and variance ϕ^2. The unconditional pdf for ζ is then calculated by integrating over all possible values of m_ζ; that is,

$$p\left(\zeta|\sigma_\zeta^2, \mu, \phi^2\right) = \int_{-\infty}^{\infty} p\left(\zeta|m_\zeta, \sigma_\zeta^2\right) p\left(m_\zeta|\mu, \phi^2\right) dm_\zeta. \tag{8}$$

Performing the integration, the unconditional pdf is found to be normal where the unconditional mean for ζ is $m_\zeta + \mu$ and the variance is $\sigma_\zeta^2 + \phi^2$. Hence, the means and variances are simply additive.

Another case involving an uncertain source power or scattering strength that permits analytical results is a gamma distribution in which the rate parameter is modulated by another gamma distribution. Specifically, the pdf for λ is given by Eq. (2) but with λ replacing s and α and β replacing k and λ, respectively. Here, α and β are the shape and rate parameters, respectively, of the mixing distribution.

A detailed discussion of this case can be found in Wilson *et al.* (2017b). The integral to be solved is

$$p(s|k,\alpha,\beta) = \int_0^\infty p(s|k,\lambda)p(\lambda|\alpha,\beta)d\lambda, \tag{9}$$

which can be shown to result in the compound gamma pdf as given by (Dubey, 1970; Wilson *et al.*, 2017b)

$$p(s|k,\alpha,\beta) = \frac{1}{B(k,\alpha)}\frac{(s/\beta)^{k-1}}{\beta(1+s/\beta)^{k+\alpha}}, \tag{10}$$

for $s > 0$. Here, k, α, and β are pdf parameters, and $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. From Jacob (2013), the mean for the compound gamma is $m = \beta k/(\alpha-1)$ and the variance is $\sigma^2 = \beta^2 k(k+\alpha-1)/[(\alpha-1)^2(\alpha-2)]$. Figure 3 shows the compound gamma pdf for various values of k and α. The value of β is set to $(\alpha-1)/k$ so that $m=1$. A large value of α corresponds to a delta-like modulating function; the modulating process is, thus, weak and causes the compound gamma pdf to be very similar to the original conditional pdf. Smaller values of α lead to a more elevated tail on the linear axis. On the logarithmic axis, the compounding creates a negatively skewed pdf when $k<\alpha$ and a positively skewed pdf when $k>\alpha$. The former situation occurs if the variations resulting from the modulating process are relatively weak compared to those of the fading process and vice versa.

B. Random transmission geometry without fading

This subsection analyzes the situation where one or more sources are randomly placed relative to the receiver, resulting in random geometric spreading losses between the source(s) and receiver. Unlike in Sec. II A, random signal fading is not considered here. For simplicity, the model assumes that the sources are randomly placed in a circular region. Depending on the application, other shapes for the source distribution region may be more appropriate. Presumably, qualitatively similar results would be obtained with other non-pathological two-dimensional (2D) shapes. However, the results might differ significantly for one-dimensional (1D) source regions such as a random placement on a line or on a circle. These situations, although not considered here, merit further investigation.

1. Single source: Pareto and offset exponential distributions

Consider a single source randomly placed anywhere in the circle $r \le r_{\max}$ with the receiver at the origin. (By the principle of reciprocity, we could equivalently consider the problem of a source at the origin and a receiver randomly placed in the circle $r \le r_{\max}$.) The probability of the radius taking on a particular value r is proportional to the circumference of the circle, $2\pi r$, which implies that the pdf is $p_r(r) = Ar$, where A is a constant. By integrating the pdf from 0 to $r_{\max}$ and setting the result to one, it can be determined that $A = 2/r_{\max}^2$. Hence, $p_r(r) = 2r/r_{\max}^2$ for $r \le r_{\max}$ and zero, otherwise. Suppose, next, that the signal power decays according to the power law,

$$s = \frac{s_0}{(r/r_0)^\beta}, \tag{11}$$

where s_0 is the source power observed at the reference distance $r = r_0$, and β is the power-law exponent for the geometric spreading. For cylindrical spreading, $\beta = 1$, whereas for spherical spreading, $\beta = 2$. Because $p(s) = |dr/ds|\ p_r(r)$, we have

$$p(s) = \frac{2}{\beta s_0}\left(\frac{r_0}{r_{\max}}\right)^2\left(\frac{s_0}{s}\right)^{1+2/\beta} \tag{12}$$

for $r \le r_{\max}$ and zero, otherwise. Compare this result to the Pareto distribution, for which the pdf is

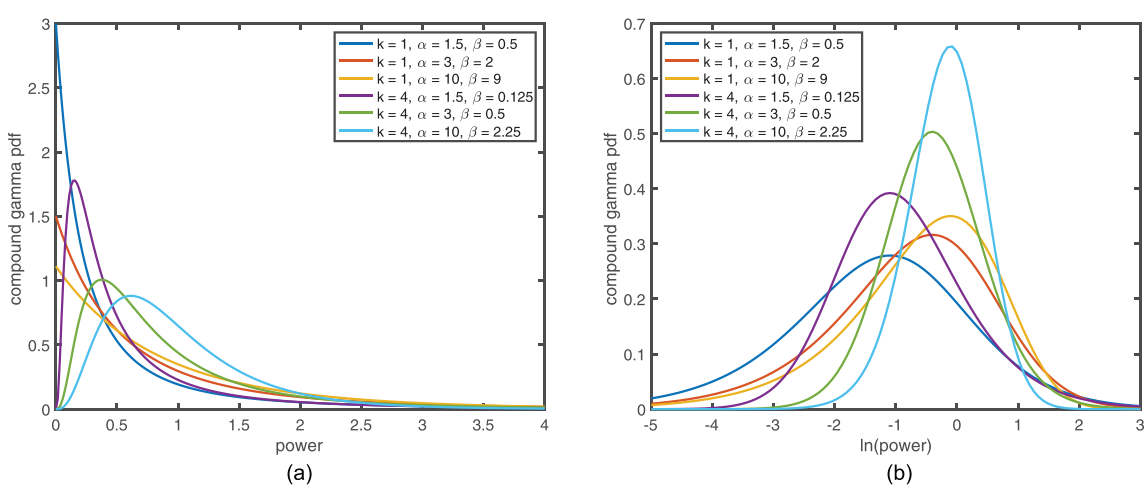

FIG. 3. (Color online) The compound gamma pdf for various values of the parameters k and α. The value of β is chosen such that the mean power is one. (a) Linear axis and (b) logarithmic axis.

$$p(s) = \begin{cases} 0, & s < s_{\min}, \\ \alpha s_{\min}^{\alpha}/s^{\alpha+1}, & s \geq s_{\min}, \end{cases} \tag{13}$$

where $s_{\min} = s_0/(r_{\max}/r_0)^{\beta}$ is the minimum value of s (namely, its value when the source is at $r = r_{\max}$). By equating Eqs. (12) and (13) to the Pareto pdf, we identify $\alpha = 2/\beta$ and $s_{\min} = s_0(r_0/r_{\max})^{2/\alpha}$. We indicate a rv drawn from this pdf as $S \sim \text{Pareto}(s_{\min}, \alpha)$. For spherical spreading ($\beta = 2$), $\alpha = 1$ and $s_{\min} = s_0(r_0/r_{\max})^2$ and, hence, $p(s) = (s_0/s^2)(r_0/r_{\max})^2$. The mean of the Pareto distribution is $\alpha s_{\min}/(\alpha - 1)$ for $\alpha > 1$. For $\alpha \leq 1$, the mean diverges. This behavior is important in the present context because $\alpha = 1$ corresponds to spherical spreading.

To transform the Pareto distribution to a logarithmic axis, we set $\zeta = \ln s$ in Eq. (13), resulting in

$$p_{\zeta}(\zeta) = \begin{cases} 0, & \zeta < \zeta_{\min}, \\ \alpha e^{-\alpha(\zeta - \zeta_{\min})}, & \zeta \geq \zeta_{\min}, \end{cases} \tag{14}$$

where $\zeta_{\min} = \ln s_{\min} = \ln\left[s_0(r_0/r_{\max})^{2/\alpha}\right]$. Hence, a Pareto distribution for s implies an exponential distribution for ζ with an offset equal to $\zeta_{\min}$. The decay rate for this distribution is α, whereas the mean is $\zeta_{\min} + 1/\alpha$ and the variance is $1/\alpha^2$. Thus, the *logarithm* of the power (unlike the power itself) has a well-behaved mean and variance when $\alpha = 1$ (spherical spreading), which are equal to $\zeta_{\min} + 1$ and 1, respectively. Figure 4 shows the Pareto pdf for various combinations of α and $s_{\min}$. When viewed on a logarithmic axis, an important qualitative difference between the Pareto pdf and the gamma and noncentral Erlang pdfs (Figs. 1 and 2) is that the Pareto has no tail to the left (small values of ζ) and an exponential tail to the right (large values of ζ), whereas the latter both have higher tails to the left than to the right. This behavior provides an important indicator as to whether random fading or random geometric spreading dominates the received signal.

Consider next the random placement of a source within the ring $r_{\min} \leq r \leq r_{\max}$. A lower bound on r might be appropriate, for example, if a sound-level meter is placed at a fixed distance from a roadway. As before, $p(r) = Ar$, where A is a constant. Integrating this function from $r_{\min}$ to $r_{\max}$ and setting the result to one, we have $A = 2/(r_{\max}^2 - r_{\min}^2)$ for $r_{\min} \leq r \leq r_{\max}$ and zero, otherwise. Substituting with Eq. (11), we find

$$p(s) = \begin{cases} 0, & s < s_{\min}, \\ \alpha s_{\min}^{\alpha} s^{-\alpha-1}/\left[1 - (s_{\min}/s_{\max})^{\alpha}\right], & s_{\min} \leq s \leq s_{\max}, \\ 0, & s_{\max} < s, \end{cases} \tag{15}$$

where $s_{\min} = s_0(r_0/r_{\max})^{2/\alpha}$ and $s_{\max} = s_0(r_0/r_{\min})^{2/\alpha}$ are the minimum and maximum values of s, respectively. This pdf is called a *bounded* Pareto distribution. Transforming to the logarithmic axis results in

$$p_{\zeta}(\zeta) = \begin{cases} 0, & \zeta < \zeta_{\min}, \\ \alpha e^{-\alpha(\zeta - \zeta_{\min})}/\left[1 - e^{\alpha(\zeta_{\min} - \zeta_{\max})}\right], & \zeta_{\min} \leq \zeta \leq \zeta_{\max}, \\ 0, & \zeta_{\max} < \zeta, \end{cases} \tag{16}$$

where $\zeta_{\min} = \ln s_{\min}$ and $\zeta_{\max} = \ln s_{\max}$. As before, this is an exponential pdf with offset $\zeta_{\min}$, although it is truncated at $\zeta_{\max}$.

2. Variable source strength revisited: EMG distribution

Let us employ the compound distribution approach from Sec. II A 3 to again derive a pdf for a source of variable strength, except this time for a single source randomly placed in a circle. These variations might result, for example, from different types of vehicles or machinery. In this case, the starting point is the offset exponential, Eq. (14), for the log-power, ζ, namely, $p_{\zeta}(\zeta|\alpha, \zeta_{\min}) = \alpha e^{-\alpha(\zeta - \zeta_{\min})}$ for $\zeta \geq \zeta_{\min}$ and 0, otherwise. Further, suppose that the logarithm of the minimum log-power, $\zeta_{\min}$, is normally

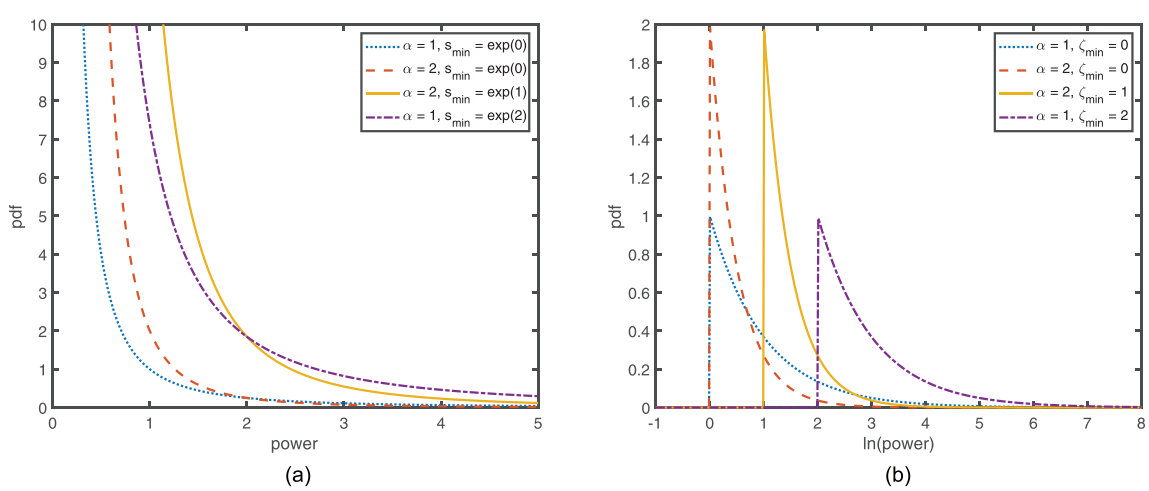

FIG. 4. (Color online) The Pareto distribution for various combinations of the parameters α and $s_{\min}$. (a) Distributions on a linear axis and (b) logarithmic axis (where $\zeta_{\min} = \ln s_{\min}$).

distributed; that is, the minimum power $s_{\min}$ and by implication, the source strength s_0 are lognormally distributed. We, thus, employ Eq. (3) as the mixing distribution, although with $\zeta_{\min}$ replacing s. Substituting the pdfs and keeping in mind that $p_\zeta(\zeta|\alpha,\zeta_{\min})$ is zero when $\zeta_{\min} < \zeta$, we have

$$p_\zeta(\zeta|\alpha,\mu,\phi^2) = \frac{\alpha}{\sqrt{2\pi}\phi}\int_{-\infty}^{\zeta} e^{-\alpha(\zeta-\zeta_{\min})-(\zeta_{\min}-\mu)^2/2\phi^2} d\zeta_{\min}. \quad (17)$$

After substantial algebra, and introducing the error function, $\mathrm{erf}(\cdot)$, we find

$$p_\zeta(\zeta|\alpha,\mu,\phi^2) = \frac{\alpha}{2} e^{-\alpha(\zeta-\mu)+\alpha^2\phi^2/2}\left[1+\mathrm{erf}\ \frac{\zeta-\mu-\alpha\phi^2}{\sqrt{2}\phi}\right]. \quad (18)$$

Because $1+\mathrm{erf}(z) = 1-\mathrm{erf}(-z) = \mathrm{erfc}(-z)$, where $\mathrm{erfc}(\cdot)$ is the complimentary error function, this result can be rewritten as

$$p_\zeta(\zeta|\alpha,\mu,\phi^2) = \frac{\alpha}{2} e^{-\alpha(\zeta-\mu)+\alpha^2\phi^2/2}\mathrm{erfc}\ \frac{\alpha\phi^2-\zeta+\mu}{\sqrt{2}\phi}\Big). \quad (19)$$

The result is known as the EMG pdf (Grushka, 1972). The EMG is ordinarily derived as the sum of a normal rv with mean μ and variance ϕ^2 to an exponential rv with rate parameter α. (In Sec. II B 3, we will discuss how the ordinary derivation relates to the present one.) The mean of the EMG is $m = \mu + 1/\alpha$, whereas the variance is $\sigma^2 = \phi^2 + 1/\alpha^2$. Figure 5 shows the EMG pdf for various values of the variance with $\mu = 0$, $\alpha = 1$, and $\phi^2 = \sigma^2 - 1/\alpha^2$. With these values, $m = 1$. For $\sigma^2 = 1$, the pdf is exponential; as σ^2 is increased, the pdf becomes increasingly normal.

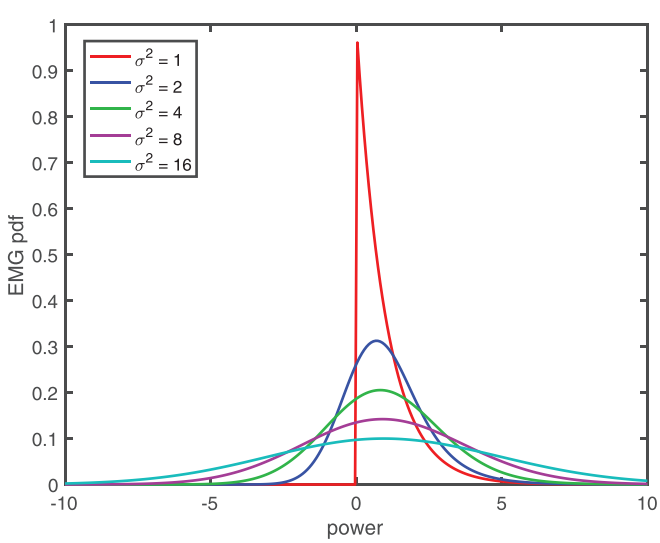

FIG. 5. (Color online) EMG distributions for various values of the variance σ^2. The EMG parameters are set to $\mu = 0$, $\alpha = 1$, and $\phi^2 = \sigma^2 - 1/\alpha^2$. (With these values, the mean is always one.)

3. Multiple sources: EMG distribution (approximate)

Let us next consider the presence of multiple randomly distributed sources. Specifically, suppose N sources are placed within the circle $r \le r_{\max}$. The total field then consists of a sum N iid Pareto rvs given by Eq. (13). Without loss of generality, we can make a change of variables to the normalized power $\overline{s} = s/s_{\min}$ for which the pdf of a single source is $p_{\overline{s}}(\overline{s}) = \alpha/\overline{s}^{\alpha+1}$ for $\overline{s} \ge 1$ and 0, otherwise). For the logarithm of the normalized power of a single source, $\overline{\zeta} = \ln\overline{s} = \ln s - \ln s_{\min} = \zeta - \zeta_{\min}$, we have $p_{\overline{\zeta}}(\overline{\zeta}) = \alpha e^{-\alpha\overline{\zeta}}$ for $\overline{\zeta} \ge 0$ (and zero otherwise); that is, $\overline{Z} \sim \mathrm{Exp}\,(\alpha)$, where $\overline{Z}$ is a rv drawn from the exponential pdf $p_{\overline{\zeta}}(\overline{\zeta}|\alpha)$.

With the problem thus formulated, we seek the distribution of the sum $\overline{S} = \sum_{i=1}^{N}\overline{S}_i$, where each $\overline{S}_i$ is an iid Pareto(1, α) rv. Unfortunately, there does not appear to be a general solution available for the distribution of a sum of iid Pareto rvs because statistics, such as means, are divergent. Let us consider, instead, the pdf of the log-power, namely, $\overline{Z} = \ln\left(\sum_{i=1}^{N} e^{\overline{Z}_i}\right)$, where $\overline{Z}_i = \ln\overline{S}_i$ and $\overline{Z}_i \sim \mathrm{Exp}(\alpha)$. The distribution of $\overline{Z}$ depends on N and α. From here forward, it will be assumed that $\alpha = 1$ (spherical spreading). Although we do not know how to determine the dependence of the pdf on N analytically, it can be readily simulated numerically by generating a large number of random samples of $\overline{Z}_i$ from exponential pdfs, applying the formula $\overline{Z} = \ln\left(\sum_{i=1}^{N} e^{\overline{Z}_i}\right)$, finding the histogram, and normalizing the histogram to determine the pdf. The results are shown in Fig. 6 for histograms based on 2^{21} random samples. We see that the pdfs transition gradually from exponential when $N = 1$ to

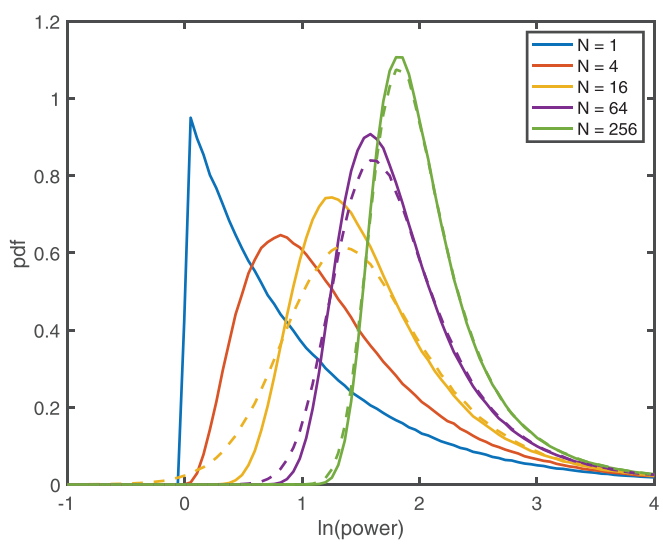

FIG. 6. (Color online) Distributions of the log-power $\overline{\zeta}$ as produced by N sources randomly distributed in a circle with the receiver at the origin. Spherical spreading ($\alpha = 1$) is assumed. All curves are for the power normalized by $Ns_{\min}$, i.e., $\overline{s} = s/(Ns_{\min})$. The solid lines are numerical simulations based on drawing random samples from exponential distributions. The dashed lines (shown for the cases $\langle N\rangle = 16$, 64, and 256) are simulations that further randomize the number of sources by drawing N from a Poisson distribution with the specified $\langle N\rangle$.

normal-like with an exponential tail for large N. For Fig. 6, we have normalized the power emitted by each source, s_0, by N, so that the *total power emitted by all sources* is independent of N. Even with this adjustment, the mean power at the receivers increases with N.

As we do not have analytical results for the distributions shown in Fig. 6, it is of interest to find distributions providing suitable empirical fits. Many of the distributions discussed up to this point are possible candidates; in particular, distributions that reduce to the exponential will, at least, fit the simulations well for the case $N = 1$ when applied on a logarithmic axis. Figure 7(a) shows fits to the simulated distributions for $p_{\overline{\zeta}}(\overline{\zeta})$ for several such candidates: the gamma [Eq. (2), but with $\overline{\zeta}$ replacing s], compound gamma [Eq. (10), but with $\overline{\zeta}$ replacing s], and EMG [Eq. (19)]. The parameters for these distributions were determined by maximum-likelihood estimation (MLE). The fits were applied on a *logarithmic* axis because (as previously discussed) the statistics for this model are unstable on a linear axis. In fitting the gamma and compound gamma to the simulation data for $\overline{\zeta}$, the origin is shifted to $\ln N$ because that is the minimum possible value for $\overline{\zeta}$. (It should be stressed that the gamma and compound gamma pdfs are being employed here simply to examine their suitability for approximating the multisource simulations with random transmission geometry; no random fading is involved in this problem.)

As indicated by Fig. 7(a), because the gamma, compound gamma, and EMG pdfs all reproduce the exponential when $N = 1$, they provide exact fits for that case as expected. For the compound gamma, the fit is essentially indistinguishable from the simulation when $N = 4$ and barely distinguishable when $N = 16$. However, we do not show fits for the compound gamma when $N = 64$ and $N = 256$ because the MLE did not converge to a valid solution in those cases. The MLE did, however, converge for all values of N with the gamma and EMG and, hence, fits are shown for these pdfs for all cases. Overall, we see from Fig. 7(a) that the compound gamma provides better fits than the gamma (when $N \leq 16$), which is not surprising as the former has an additional parameter. Both of these distributions provide *worse* fits as N increases. In contrast, the EMG provides *better* fits as N increases. The EMG is also quite satisfactory for small N (e.g., $N = 4$ or $N = 16$).

Some additional results shown in Fig. 7(b) help to explain why the EMG provides a consistently good empirical fit. In this plot, the full distributions for $N = 1$, $N = 16$, and $N = 256$ are shown as solid lines. These are the same curves that are shown as solid lines in Fig. 7(a). Shown as dashed lines are the distributions for the single strongest among the N sources; that is, we simulate all N sources, find the strongest one, and determine the pdf for that source only. The pdfs for the remaining $N - 1$ weaker sources are shown as dashed-dotted lines. Finally, the dotted curves are exponential curves for $\alpha e^{-\alpha\left(\overline{\zeta} - \log N\right)}$. This curve corresponds to the pdf for a single source distributed within a circle of area $\pi r_{\max}^2/N$, where $r_{\max}$ is the radius of the circle in which all N sources are distributed. Hence, this is the circle within which, on average, the closest source is located. For relatively large values of $\overline{\zeta}$, the dotted line coincides with the pdfs of the strongest source and the overall pdf. We observe that *the exponential tail of the overall pdf is determined by the single source closest to the receiver* (the strongest one). The remaining part of the pdf (for the more distant $N - 1$ sources) becomes increasingly normal as N increases. This result can be explained as follows. The more distant sources are distributed within a ring (the circle $\pi r_{\max}^2/N$ having been removed from the larger circle with area $\pi r_{\max}^2$) and as we showed earlier, this removes the exponential tail [Eq. (16)]. Because the central limit theorem should apply for large N, the overall pdf consists of an exponential pdf corresponding to the closest source and an approximately normal pdf for the remaining $N - 1$ more distant sources.

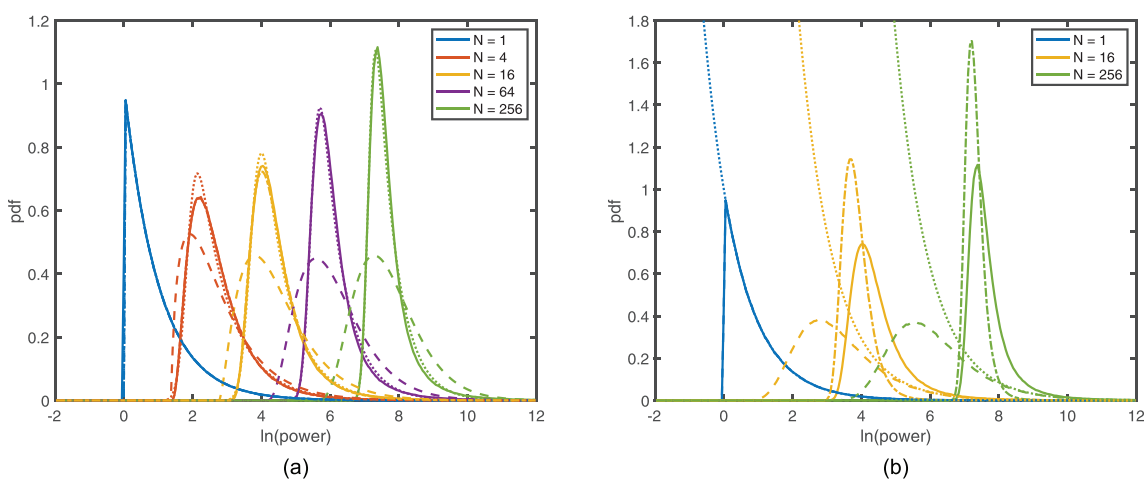

FIG. 7. (Color online) Distributions of the log-power $\overline{\zeta}$ produced by N sources randomly distributed in a circle. The solid lines are the same simulated pdfs shown as solid lines in Fig. 6 (but without the normalization by N). The other lines show various other model pdfs. (a) The dashed lines are gamma pdfs, dashed-dotted lines are compound gamma pdfs (for $N = 1$, 4, and 16), and dotted lines are EMG pdfs. (b) The dashed lines are pdfs for the single strongest among the N sources, dashed-dotted lines are pdfs for the sum of the remaining weaker $N - 1$ sources, and dotted lines are exponential pdfs given by $\alpha e^{-\alpha\left(\overline{\zeta} - \log N\right)}$.

Interestingly, we have arrived at the EMG distribution by two distinct approaches. In Sec. II B 2, it was derived for the log-power of a single source randomly placed in a circle (with the receiver at the origin) under the assumption that the source power is lognormally distributed. Here, it resulted from approximating the multisource simulations with the sum of an exponential rv representing the nearest source and a normal rv representing the remaining sources. The connection between these two interpretations can be understood from the definition of the compound pdf as an integral of the product of two pdfs. The pdf of the sum of two independent rvs is given by the convolution of the pdfs of those rvs, which is also an integral involving a product of pdfs. In fact, the convolution is a special case of the compound pdf in which a location parameter (offset), such as $\zeta_{\min}$, is the variable of integration. Therefore, either of the two interpretations is mathematically valid. They arise, however, from differing physical perspectives.

Figure 8 plots the mean, variance, and skewness of $\overline{\zeta}$ as a function of N. The mean is one for $N = 1$ and thereafter appears to increase linearly with N. The variance is also 1 for $N = 1$ and but then appears to *decrease* linearly with N to a value of about 0.5 when $N = 256$. The skewness takes on a value of two (as expected for an exponential pdf) when $N = 1$ and gradually increases to about three when $N = 256$. The persistent positive skewness is attributable primarily to the exponential tail, which is a consequence of geometric spreading.

The logarithm of the *unnormalized* power is $Z = \ln\left(\sum_{i=1}^{N} e^{Z_i}\right) = \overline{Z} + \zeta_{\min} = \overline{Z} + \ln\left[s_0 (r_0/r_{\max})^{2/\alpha}\right]$. Hence, assuming s_0 and r_0 are fixed, the unnormalized power introduces a dependence on $r_{\max}$. Thus, distributions for Z effectively depend on two adjustable parameters, N and $r_{\max}$. The dependence on $r_{\max}$ is relatively straightforward: it simply offsets the distribution to the left (for larger $r_{\max}$) or right (for smaller $r_{\max}$). In fact, changing $r_{\max}$ has the same impact as changing the source strength by $(1/r_{\max})^{2/\alpha}$.

Figure 9 shows distributions of Z for various combinations of $r_{\max}$ and the spatial density of the sources (number of sources per unit area), $\eta = N/\left(\pi r_{\max}^2\right)$. These curves were calculated from the histograms of 2^{23} random samples. Note that the pdf for $r_{\max} = 2$, $\eta = 4/\pi$ is a shifted version of the pdf for $r_{\max} = 4$, $\eta = 1/\pi$, because both of these cases correspond to $N = 16$. Similarly, $r_{\max} = 4$, $\eta = 4/\pi$ is a shifted version of the pdf for $r_{\max} = 8$, $\eta = 1/\pi$, because both of these cases correspond to $N = 64$. Another notable property of the distributions is that the tails (for large values of the log-power) appear to coincide for cases with the same value of η; that is, the *source density controls the tail*. This can be understood in light of the earlier discussion of Fig. 7(b). The exponential tail is controlled by the average area of the circle in which the closest source is distributed; this area equals $\left(\pi r_{\max}^2\right)/N = 1/\eta$. An interesting consequence of the model is that the pdf for log-power is always impacted to some degree by distant sources and the distance to the boundary even when the source region is very large ($r_{\max} \to \infty$).

Up to this point, we have considered the number of sources N to be fixed, whereas their spatial placement is random. A better assumption would be that the sources occur randomly in space at a fixed rate η. Therefore, the actual number of sources in a particular region is random. The constant-rate assumption corresponds to a Poisson spatial process. In particular, ηA is the probability that a source will occur in a subregion with area A provided that A is small enough that the probability of multiple sources appearing in the subregion is negligible. Therefore, we can think of the Poisson multisource model as having two adjustable parameters, η and the radius of the circle $r_{\max}$. From the two parameters η and $r_{\max}$, we can calculate the expected values

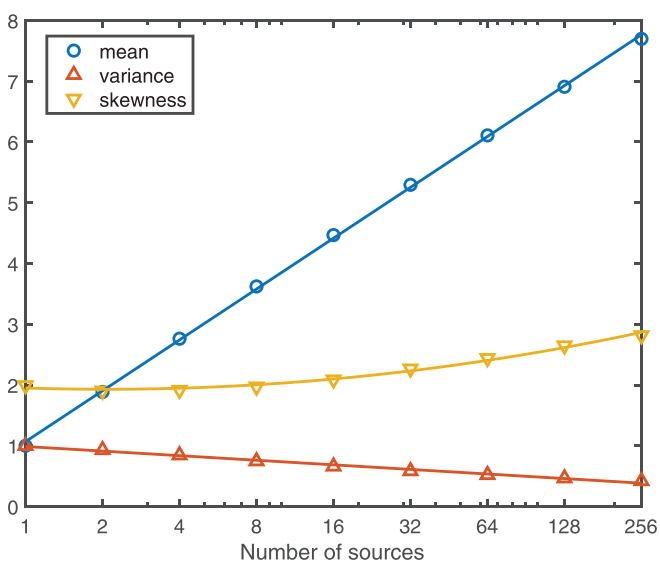

FIG. 8. (Color online) The mean, variance, and skewness of the log-power as a function of the number of sources N as deduced from the simulations shown in shown in Fig. 6.

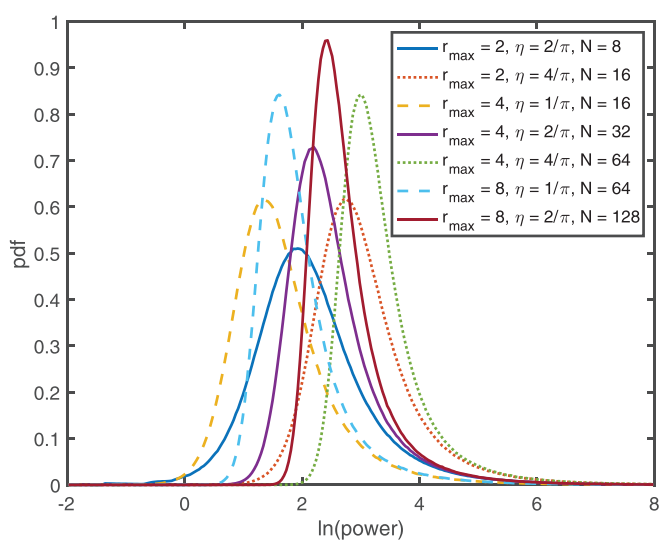

FIG. 9. (Color online) Distributions for the log-power for various values of the radius of the source region ($r_{\max}$) and areal source density (η). The cases with $\eta = 1/\pi$ are shown as dashed lines, $\eta = 2/\pi$ are shown as solid lines, and $\eta = 4/\pi$ are shown as dotted lines.

of the number of sources in the circle, $\langle N \rangle = \pi r_{\max}^2 \eta$. The number of sources in $\pi r_{\max}^2$ follows a Poisson distribution, which is given by $p_N(N) = \langle N \rangle^N e^{-\langle N \rangle}/N!$. The Poisson process can be efficiently simulated by drawing a random sample N from the Poisson distribution and then randomly placing this number of sources in the circle $\pi r_{\max}^2$. Distributions for ζ can then be calculated by the same simulation approach as before but with N changing with each realization of the source positions. The resulting pdfs are shown as dashed curves in Fig. 6 for the cases $\langle N \rangle = 16$, $\langle N \rangle = 64$, and $\langle N \rangle = 256$. (The cases $\langle N \rangle = 1$ and $\langle N \rangle = 4$ result in a large number of random realizations with zero sources and, thus, are not considered.) For $\langle N \rangle = 16$ and $\langle N \rangle = 64$, the pdfs differ somewhat when the Poisson process is employed. When $\langle N \rangle$ is increased to 256, however, there is very little difference between the Poisson process and setting N to a constant.

III. EXPERIMENT

A. Procedure

To assess the suitability of the various noise models discussed in Sec. II, noise data were collected at 37 locations in the Boston North End (Boston, MA). The experiment was conducted on 7 June 2018 and involved measuring the one-third octave SPL for 5 min at each site. These measurements were conducted consecutively with 10 min, usually, between the start of one measurement and the next. The dominant noise source was road traffic noise with secondary contributions from construction sites, pedestrians, air conditioners, and Boston Logan International Airport, which is located approximately 4 km from the North End. Although 5 min is rather short for characterizing statistics at a particular site (Can *et al.*, 2011; Romeu *et al.*, 2011), it should be kept in mind that our priority was to sample the spatial variations of the noise and, hence, the experimental protocol emphasized sampling a large number of sites.

Figure 10 illustrates the North End neighborhood and indicates the locations of each of the 37 locations with blue circles (waypoints 104–140). The North Washington Street Bridge is at the northwest corner of the neighborhood. Wharves dominate the north and east sides of the North End along Commercial Street and Atlantic Avenue. Cross Street runs along the southwest edge of the North End and is the busiest street in this neighborhood. The interstate (I-93) lies in a tunnel below Cross Street. Hanover Street, which runs through the middle of the neighborhood, has numerous restaurants and shops and a mixture of pedestrian and road noise.

At each location, a Norsonic Sound Analyzer Nor 140 (Tranby, Norway) measured the SPL for the 6.3 Hz–20 kHz one-third-octave bands using the fast response (decay) time setting with a 125 ms time constant. Data were recorded

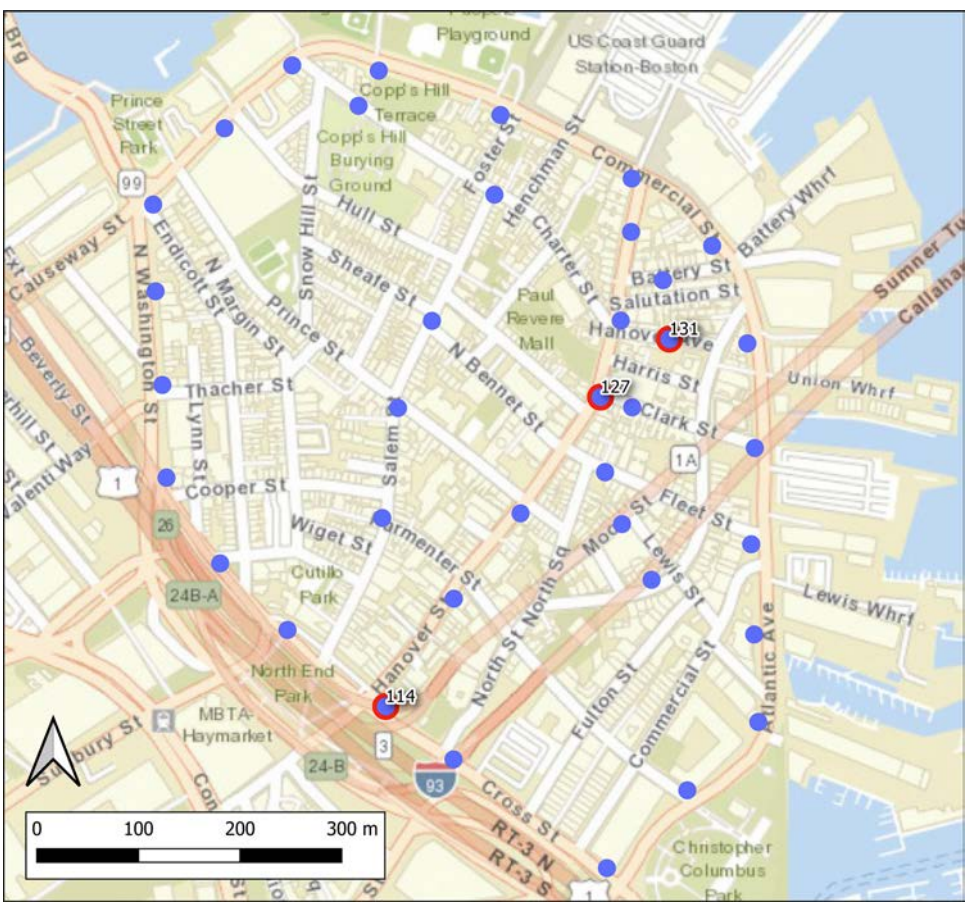

FIG. 10. (Color online) The measurement locations are in the North End (Boston, MA, USA). A sound level meter (SLM) recorded data for 5 min at each blue circle. The three numbered locations with red circles correspond to those shown in Fig. 11. The basemap is courtesy of ESRI® (Redlands, CA).

FIG. 11. (Color online) Pictures of three measurement sites representing loud, medium, and quiet areas.

every 0.5 s during 5-min intervals. The sound-level meter (SLM) was mounted horizontally on a stationary tripod (Fig. 11), and the microphone was 1.25 ± 0.05 m above the ground. Each site was documented using a picture, the time, GPS coordinates, and a subjective assessment of the dominant sound sources.

For simplicity, 3 of the 37 locations are primarily discussed here. These three locations (Fig. 11) give a snapshot of both typical levels (127) for this neighborhood and the great variance from loud (114) to quiet (131) within this neighborhood. Later, we evaluate models using all of the sites. Table I qualitatively describes each of these three sites.

Figure 12 illustrates the corresponding SLM data. The first location (waypoint 114) had very high sound levels, especially at 63 Hz, due to nearby road traffic and construction noise. A police car with its sirens on (0.8–1.6 kHz) passed at about 1 min into the measurement. At 185–240 s, a large truck stopped at the stoplight just meters from the SLM. The second location (waypoint 127) had moderate sound levels from predominately road traffic and pedestrians. The middle row of Fig. 12 demonstrates the intermittent nature of the traffic noise. The third location (waypoint 131) had low overall sound levels, mainly from an air conditioner (likely in the frequency range 250–315 Hz) and chirping birds (3–5 kHz).

B. Results

The SLM data contain about 600 points for each location and one-third octave band. Figure 13 presents the data as probability density versus SPL and frequency by calculating the pdf for each frequency band. Each column shows a different approximation of the data: left (histogram), center (normal distribution), and right (EMG distribution), again calculated for each frequency. Each row represents a different location: top (loud: 114), middle (medium: 127), and bottom (quiet: 131). Maximizing the log-likelihood yielded the distribution parameters for each location and frequency combination. This analysis used Python's scipy.stats distributions (SciPy Community, 2020) fit functions for the EMG (exponnorm) and normal (norm) distributions.

The histograms reveal patterns across all three sites. The sound field is generally broadband and unimodal with a maximum one-third octave band SPL around 40 Hz, which creates a sideways V-shape. The sound levels decrease for the lowest and highest frequencies. This result is expected for the slow-moving road traffic noise (mostly engine noise) that typically dominates.

There are a few discontinuities for adjacent frequency bands that can be readily explained by a narrowband sound source dominating at that location and frequency. For example, the quiet area at 315 Hz has a discontinuity because an air conditioner was near the SLM, and the loud area at 63 Hz has a discontinuity because an electrical generator (and perhaps other equipment) was nearby. The quiet location at 4 kHz had birds chirping as the primary sound source. We speculate that the bimodal behavior for the loud location between about 250 Hz and 1 kHz represents two different vehicle types stopped next to the SLM (i.e., cars and large trucks). Table I indicates that a truck idled next to the SLM,

TABLE I. Information on three of the measurement sites selected for detailed analysis.

Site	Time (EDT)	Cross streets	GPS coordinates (latitude, longitude)	Noise level	Description and dominant sources
114	10:30	Cross St. Hanover St.	42.362492°, −71.055568°	High	Very busy street—sirens (10:31); truck idling; distant generator at 63 Hz
127	13:30	Hanover St. Clark St.	42.365222°, −71.053003°	Medium	Semi-busy street—road/engine noise; pedestrian noise
131	14:10	Hanover Ave.	42.365734°, −71.052181°	Low	Quiet side street—birds; air conditioner; distant pedestrians

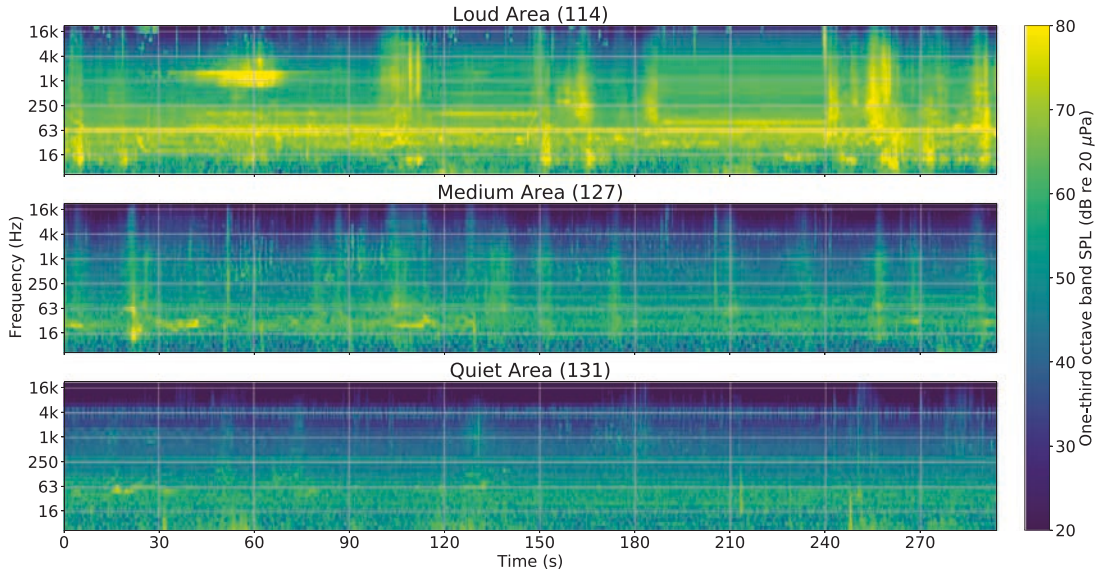

FIG. 12. (Color online) The one-third octave SPL versus time at three locations: top row (loud), middle row (medium), and bottom row (quiet).

which possibly occurred at 185–240 s in Fig. 12. If that is the case, the 57 and 63 dB peaks in Fig. 13 correspond to a car or truck stopped near the SLM.

Comparing the normal and EMG models to the histogram (Fig. 13), the EMG model qualitatively better approximates the histogram data primarily because the EMG model allows for positive skew, which is present at most frequencies and locations. The theoretical analysis in Sec. II suggests that this positive skew comes from geometric spreading losses for sources randomly distributed in space. When a single, immobile, and continuous source is near the receiver (e.g., the construction generator at the loud location and 63 Hz and the air conditioner at the quiet location and 315 Hz), the distribution has smaller variance with less skew (similar to the normal distribution).

Figure 14 plots frequency slices from Fig. 13 to facilitate the quantitative model comparisons. Each column shows a different location: left (loud), center (medium), and right (quiet). Each row shows a different frequency: from top to bottom, 40, 160, 630 Hz, and L_{Aeq}. These frequencies avoid the construction generator and police siren at the loud site and the air conditioner and the chirping birds at the quiet site. The idling truck, a broadband source, could not be avoided at the loud location. Fits are shown for the normal, EMG, and compound gamma distributions. The normal and EMG distributions were fit to the SPL data, whereas the compound gamma was fit to the linear power data after applying the transformation $y = 0.01 \times 10^{L_p/10}$, which was found by trial and error to result in good numerical convergence. The normal model for SPL is, in essence, a lognormal model for the linear power.

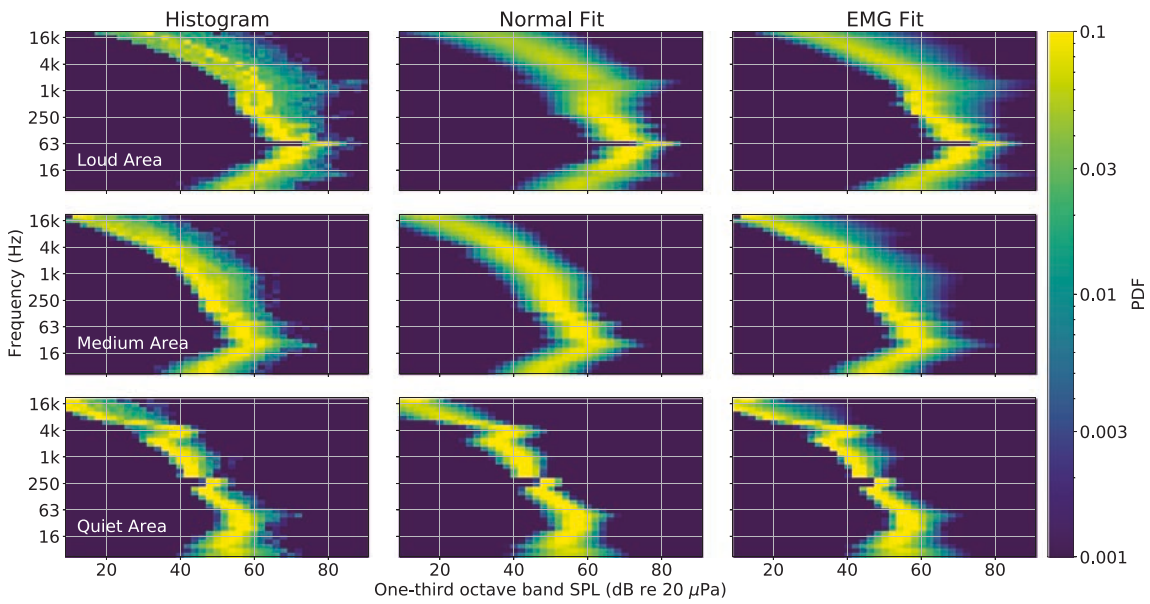

FIG. 13. (Color online) PDFs versus both frequency and SPL. The left column shows the measured histogram. The center (normal fit) and right (EMG fit) columns give two maximum-likelihood estimates of the data. The top (loud), middle (medium), and bottom (quiet) rows show three sites.

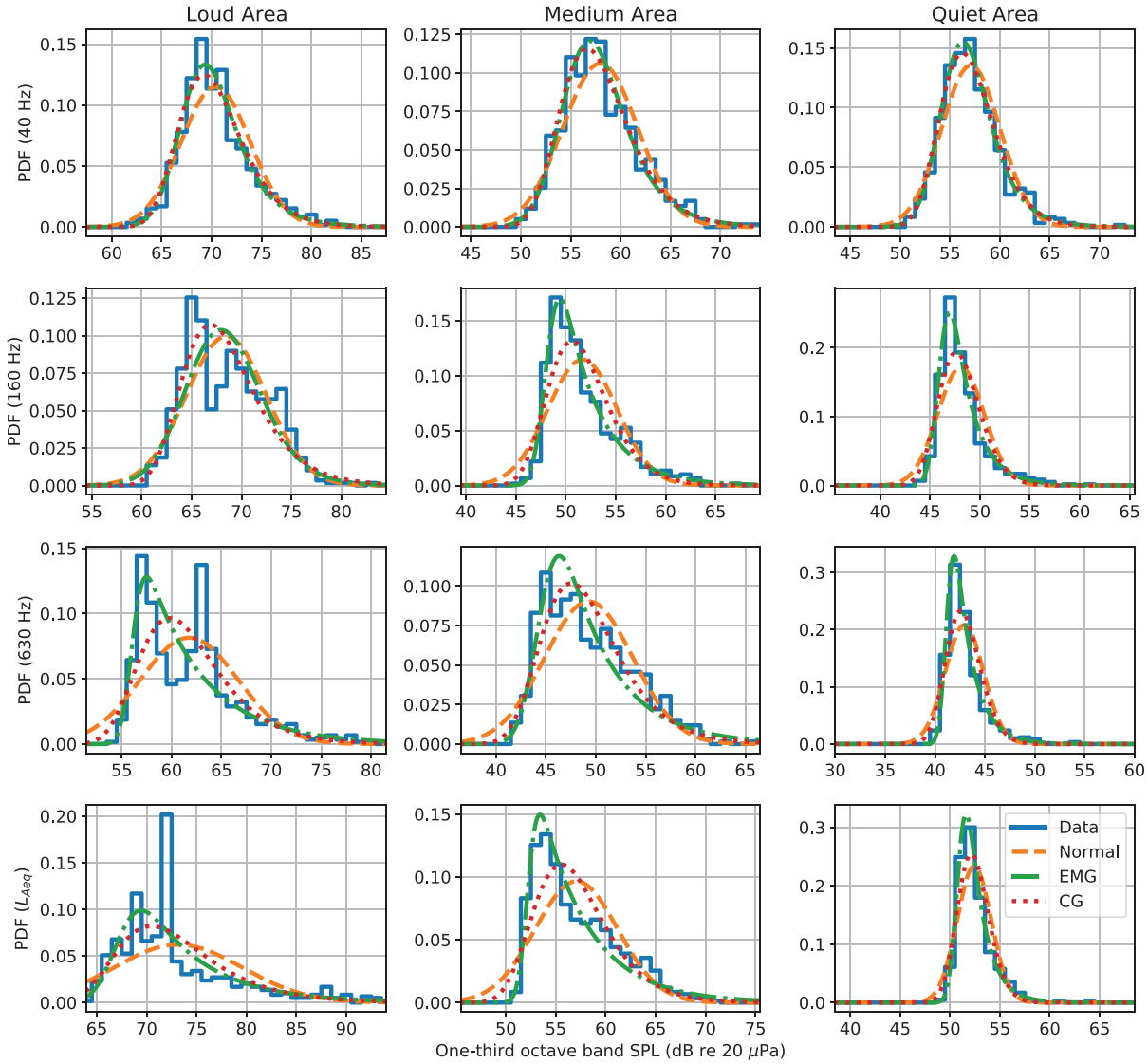

FIG. 14. (Color online) PDFs versus the one-third octave band SPL for primarily road noise. The left (loud), middle (medium), and right (quiet) columns show three different sites. The top three rows give the pdfs for three different frequencies (40, 160, and 630 Hz), whereas the bottom row gives the pdf for L_{Aeq}. The solid blue curves are histograms for the SPL data. The other curves show various fitted distributions: dashed orange is the normal, dashed-dotted green is the EMG, and dotted red is the compound gamma.

At 40 Hz, the EMG and compound gamma distributions are similar and provide only a marginal benefit over the normal distribution, which adequately approximates the lowest frequencies. At the medium and quiet locations for 160 Hz, 630 Hz, and L_{Aeq}, the EMG better approximates the measured data compared to the normal distribution, with the compound gamma falling somewhere in between. At the loud location, especially for 630 Hz but also for 160 Hz and L_{Aeq}, none of the distributions provide a particularly good fit because the histograms are multimodal. Despite summing multiple frequency contributions, the A-weighted levels are still highly skewed and the EMG provides a better fit.

For the loud location, the distribution at 160 Hz appears trimodal with peaks near 65, 69, and 74 dB. Figure 12 reveals that these three peaks could be associated with the idling truck (65 dB, 185–240 s), cars nearly stopped as the police car passes (69 dB, 20–100 s), and idling cars (74 dB, 120–150 s). The loud location at 630 Hz seems bimodal, and the idling truck may have caused the 63 dB peak. The truck idled for 55 s of a 300 s measurement (18%), and the area of the 63 dB peak above the trend is 14%.

Figure 15 is similar to Fig. 14 except the locations and frequencies are chosen to highlight non-road noise sources. The top-left subplot is for the loud area at 63 Hz, which is the construction generator. In this case, there is one continuous, immobile source a medium distance (on the order of 100 m) from the SLM so the SPL pdf is similar to a normal distribution, and all models fit adequately.

The top-center and top-right subplots in Fig. 15 illustrate the contrast between the loud and medium locations at 16 kHz. The top-right plot shows data typical of freely flowing traffic, which can be approximated well with the EMG distribution. However, the EMG does not provide a good fit when a stationary source (such as a parked truck, in this

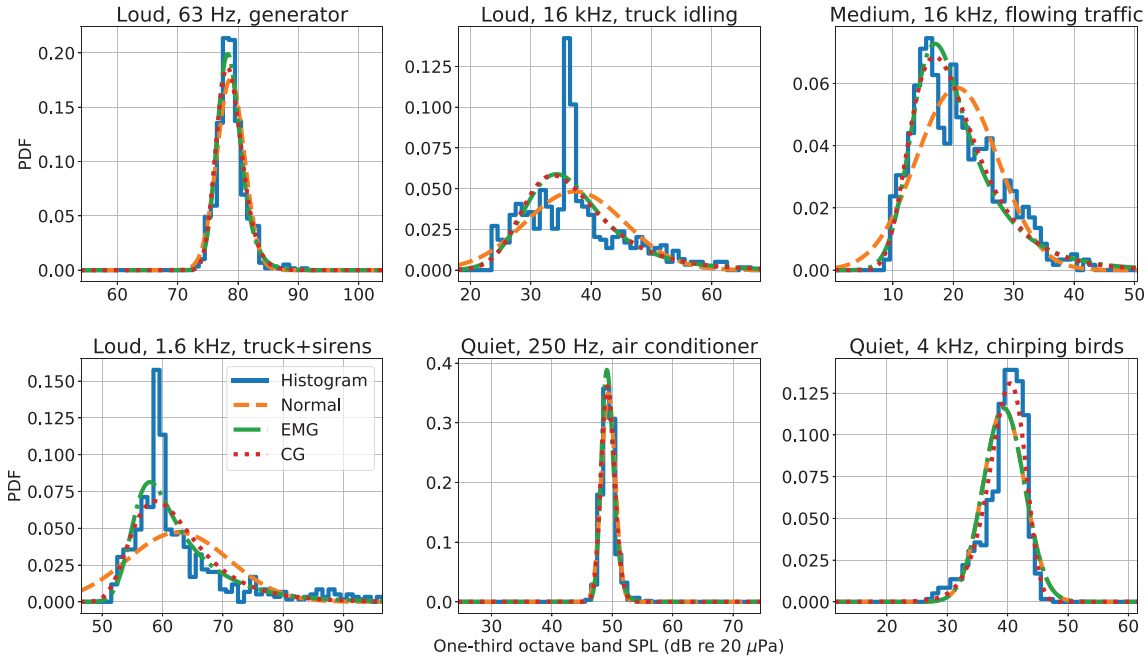

FIG. 15. (Color online) PDFs versus the one-third octave band SPL for primarily non-road noise. The curve colors and styles are the same as those in Fig. 14.

case) is present during significant portions but not for all of a measurement. The area of this peak above the general trend is approximately 16%, which is close to the 18% of the measurement when the idling truck was present. Again, a longer duration measurement or moving the microphone further from the road might mitigate this issue.

The bottom left subplot in Fig. 15 demonstrates the impact of both the idling truck and police sirens. The idling truck introduces that same narrow peak (area about 16% above the trend), and the siren increases the mean and variance of the SPL. The siren likely caused the small peak at 87 dB. Neither model fits this dataset well, but the EMG and compound gamma distributions perform considerably better than the normal distribution.

The bottom-center subplot in Fig. 15 is very similar to the top-left because they both have a single dominant, continuous, and immobile source. Both models work well for this case. The peak for the air conditioner is narrower than that for the construction generator, which might be because the air conditioner was considerably closer to the SLM (about only 30 m away). The bottom right subplot has a negative skew that the compound gamma distribution can model better than either the EMG or normal. In this case, the noise source is chirping birds, which exemplifies a mixture of two distributions: one for when there are chirping sounds and the other for when there is no chirping. Thus, a compound (mixture) distribution, such as the compound gamma, is more appropriate for this situation.

Figure 16 analyzes the frequency dependence of the sample mean, variance, skewness, and kurtosis of the one-third octave band SPL data. The mean is the arithmetic mean,

$$\mu = \frac{1}{N}\sum_{n=1}^{N} L_{p,n}, \tag{20}$$

where $L_{p,n}$ is the SPL of the nth sample. The variance is the unbiased variance,

$$\sigma^2 = \frac{1}{N-1}\sum_{n=1}^{N} (L_{p,n} - \mu)^2. \tag{21}$$

Defining the i th central moment as

$$m_i = \frac{1}{N}\sum_{n=1}^{N} (L_{p,n} - \mu)^i, \tag{22}$$

the skewness and excess kurtosis are, respectively,

$$g_1 = \frac{m_3}{m_2^{3/2}} \quad \text{and} \quad g_2 = \frac{m_4}{m_2^2} - 3. \tag{23}$$

The SPL mean results in Fig. 16 are consistent, predominately, with road noise for city traffic traveling at about 30 mph. The engine noise peak is at about 50 Hz, and the tire noise peak is at about 1 kHz. The large max value at 63 Hz is the construction generator (not road noise). The variance is largest for very high and very low frequencies where the road noise may not be dominant. The skewness is smaller for frequencies below 40 Hz, and the normal distribution works better for these data. The EMG distribution provides a larger benefit for frequencies above 40 Hz because this distribution can model that positive skew. A few cases have negative skew, but the EMG distribution does not provide any benefit in those cases. The kurtosis is also smallest for frequencies below 40 Hz.

Figure 17 presents a quantitative analysis across all locations and frequencies. These fits were evaluated using the Kullback-Leibler (KL)-divergence (Kullback and Leibler, 1951) versus the histogram of the one-third octave band SPL data with 1 dB bin width over the range (–10, 100 dB).

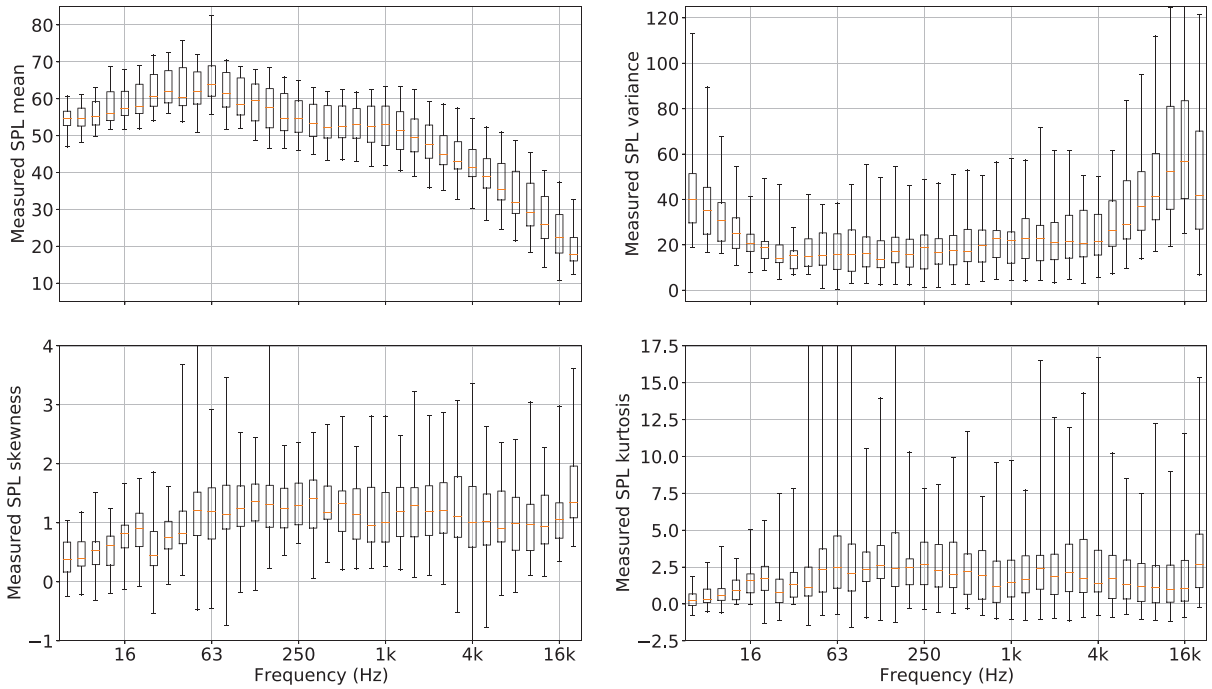

FIG. 16. (Color online) Box plots for the measured mean (top left), variance (top right), skewness (bottom left), and excess kurtosis (bottom right) versus the frequency. Each box represents the data for 37 locations for a one-third octave frequency band. The whiskers give the full range of the data.

The KL-divergence provides a metric for comparing two probability distributions. Figure 17 illustrates the cumulative distribution function (cdf) for the KL-divergence for all 1332 combinations (37 locations and 36 frequency bands) for the EMG, compound gamma, and normal distributions. A rapid increase in the cdf near the origin indicates more small values for the KL-divergence and, hence, a better fitting model.

Figure 17 indicates that the EMG model best approximates the histogram data, followed by the compound gamma, and last by the normal distribution. For example, the median KL-divergences are 0.04 for the EMG, 0.07 for the compound gamma, and 0.15 for the normal. Although the absolute values are not very meaningful because they would all increase if the histogram bin size was smaller, the relative improvement is substantial. This suitability of the EMG model apparently stems from including geometric spreading, which leads to positive skew for many frequencies and locations.

Finally, Fig. 18 considers the EMG KL-divergence versus the frequency. The KL-divergence is low for frequencies below 40 Hz and then tends to increase with an increase in the frequency. This may indicate that the EMG model is less suitable for some of the noise sources occurring at higher frequencies.

IV. CONCLUSION

The experiment described in this article measured one-third octave levels during a 5-min interval at each of 37

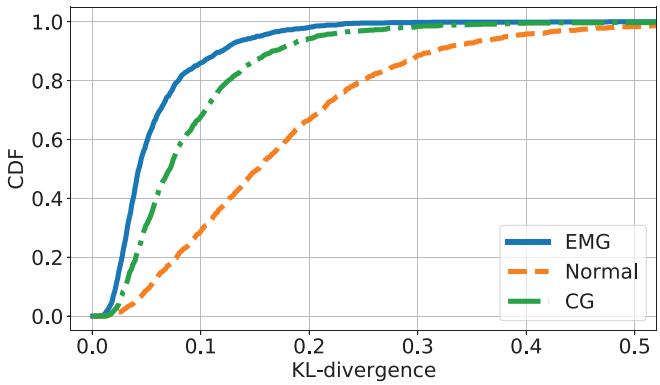

FIG. 17. (Color online) The cumulative distribution function (cdf) for the KL-divergence for all location and frequency combinations. Each curve represents a different parametric model versus the histogram data: solid blue (EMG), dashed orange (normal), and dashed-dotted green (compound gamma).

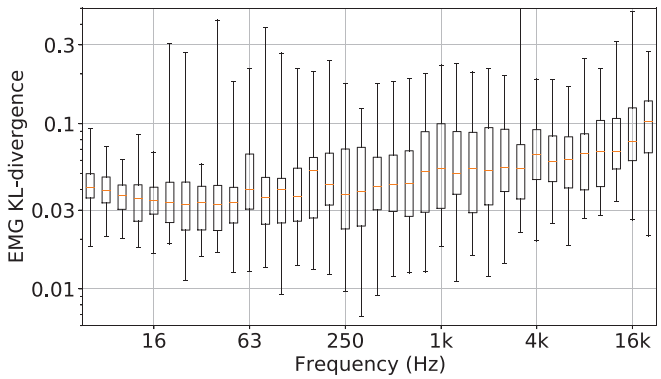

FIG. 18. (Color online) Box plot of the KL-divergence for the EMG model versus the frequency. Each box represents the data for 37 locations for a one-third octave frequency band. The whiskers give the full range of the data.

locations in the Boston North End. The dominant noise source was road traffic noise with secondary contributions from construction sites, pedestrians, birds, air conditioners, and a major airport. The experimental results in Sec. III indicate that urban sound levels are close to normally distributed (Gaussian) for relatively low frequencies but have significant positive skewness at frequencies of about 100 Hz and higher. Exceptions to this trend appear to be attributable to cases in which particular isolated noise sources dominate.

Similar behavior has been reported in previous studies (e.g., Don and Rees, 1985; García and Faus, 1991; De Coensel *et al.*, 2016). Nonetheless, this behavior is surprising from the perspective that conventional theoretical models for signal distributions incorporating random fading effects, averaging, and summation of multiple sources (including typical traffic noise sources; De Coensel *et al.*, 2016) predict that signals will exhibit a variety of behaviors from exponential to approximately normal on a *linear axis* and, subsequently, appear as negatively skewed or approximately normal on a *logarithmic axis* (that is, when sound levels are plotted in decibel units).

A plausible explanation for this discrepancy was described in the theoretical section of the article (specifically, Sec. II B), where it was shown that pdfs with positive skewness arise when sources are randomly positioned relative to a receiver. In particular, we considered N equally strong sources placed randomly within a circle at which the receiver is at the center with the signals decaying away from the sources according to the spherical spreading law. Each individual source was shown to produce a Pareto distribution for the received power. Although the pdf for the sum of the sources could not be expressed in analytical form for this model, we showed that its shape depends only on N, whereas the offset (on a logarithmic axis) depends on the radius of the circle and strength of the sources. Furthermore, the contribution to the pdf from the $N-1$ sources farthest away from the receiver has an approximately normal shape, whereas the contribution from the source closest to the receiver creates an exponential tail on a logarithmic axis. The exponential tail is what leads to the positive skewness of the overall distribution.

Because the theoretical pdf combines characteristics of normal and exponential pdfs, it can be well approximated by the EMG pdf, which describes the sum of normally and exponentially distributed rvs. For most sites and frequency ranges, the EMG was also found to provide a good fit to the urban noise data from Boston. For some cases in which the noise was dominated by particular, identifiable sources at a fixed distance (including a construction generator, an air conditioner, and birds), conventional random fading models, such as the lognormal and gamma distribution, appear suitable for describing the data. We also considered an extension to the gamma distribution, namely, the *compound gamma* distribution, which can incorporate spatially and temporally varying source levels. The compound gamma is able to produce both negatively and positively skewed distributions on a logarithmic axis. The negative skew is consistent with sources, such as chirping birds, and may also be useful for describing traffic noise flow that is unsteady in time (e.g., Don and Rees, 1985; Song and Lenchine, 2017).

This article examined the problem of pdfs for urban noise from a broad perspective with emphasis on how different physical phenomena produce pdfs with differing shapes, and which of these pdfs are consistent with observed distributions. In future work, it would be desirable to obtain a better qualitative understanding of relationships between measurable characteristics of the environment and parameterizations of the noise pdfs such as the EMG. It would also be necessary to obtain a better understanding of the consistency of the pdf shapes and parameterizations from one environment to another.

ACKNOWLEDGMENTS

This work was supported by the U.S. Army Engineer Research and Development Center, Geospatial Research Engineering basic research program. Permission to publish was granted by the Director, Cold Regions Research and Engineering Laboratory. Any opinions expressed in this article are those of the authors and are not to be construed as official positions of the funding agency or the Department of the Army unless so designated by other authorized documents.

Albert, D. G., and Decato, S. N. (**2017**). "Acoustic and seismic ambient noise measurements in urban and rural areas," Appl. Acoust. **119**, 135–143.

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

Can, A., Van Renterghem, T., Rademaker, M., Dauwe, S., Thomas, P., De Baets, B., and Botteldooren, D. (**2011**). "Sampling approaches to predict urban street noise levels using fixed and temporary microphones," J. Environ. Monit. **13**(10), 2710–2719.

De Coensel, B., Brown, A. L., and Tomerini, D. (**2016**). "A road traffic noise pattern simulation model that includes distributions of vehicle sound power levels," Appl. Acoust. **111**, 170–178.

Don, C. G., and Rees, I. G. (**1985**). "Road traffic sound level distributions," J. Sound Vib. **100**(1), 41–53.

Dubey, S. D. (**1970**). "Compound gamma, beta and *F* distributions," Metrika **16**, 27–31.

Dyer, I. (**1970**). "Statistics of sound propagation in the ocean," J. Acoust. Soc. Am. **48**(1), 337–345.

Ewart, T. E., and Percival, D. B. (**1986**). "Forward scattered waves in random media—The probability distribution of intensity," J. Acoust. Soc. Am. **80**(6), 1745–1753.

Flatté, S. M. (**1979**). *Sound Transmission through Fluctuating Ocean* (Cambridge University Press, London).

García, A., and Faus, L. J. (**1991**). "Statistical analysis of noise levels in urban areas," Appl Acoust. **34**(4), 227–247.

Grushka, E. (**1972**). "Characterization of exponentially modified Gaussian peaks in chromatography," Anal. Chem. **44**(11), 1733–1738.

Jacob, O. (**2013**). "From the classical beta distribution to generalized beta distributions," M.S. thesis, University of Nairobi, pp. 229–231.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (**1995**). *Continuous Univariate Distributions*, second ed. (Wiley, New York), Vol. 2.

Ketcham, S. A., Wilson, D. K., Parker, M. W., and Cudney, H. H. (**2008**). "Signal fading curves from computed urban acoustic wave fields," in *SPIE Defense and Security Symposium* (International Society for Optics and Photonics, Bellingham, WA), p. 69630K.

Kullback, S., and Leibler, R. A. (**1951**). "On information and sufficiency," Ann. Math. Statist. **22**, 79–86.

Romeu, J., Genescà, M., Pàmies, T., and Jiménez, S. (**2011**). "Street categorization for the estimation of day levels using short-term measurements," Appl. Acoust. **72**(8), 569–577.

SciPy Community (**2020**). "Statistical functions (scipy.stats)," available at https://docs.scipy.org/doc/scipy/reference/stats.html (Last viewed 02/07/2021).

Song, J., and Lenchine, V. V. (**2017**). "Statistical properties of urban noise—Results of a long term monitoring program," in *Proceedings of Acoustics 2017*, Perth, Australia, available at https://www.acoustics.asn.au/conference_proceedings/AAS2017/papers/p36.pdf (Last viewed 02/07/2021).

Stanton, T. K., Lee, W. J., and Baik, K. (**2018**). "Echo statistics associated with discrete scatterers: A tutorial on physics-based methods," J. Acoust. Soc. Am. **144**(6), 3124–3171.

Strohbehn, J. W., Wang, T., and Speck, J. P. (**1975**). "On the probability distribution of line-of-sight fluctuations of optical signals," Radio Sci. **10**(1), 59–70, https://doi.org/10.1029/RS010i001p00059.

Suzuki, H. (**1977**). "A statistical model for urban radio propagation," IEEE Trans. Commun. **25**(7), 673–680.

Transportation Systems Center. (**1971**). "A community noise survey of Medford, Massachusetts," U.S. Department of Transportation Report DOT-TSC-OST-72-1 (U.S. Department of Transportation, Washington, DC).

Turin, G. L., Clapp, F. D., Johnston, T. L., Fine, S. B., and Lavry, D. (**1972**). "A statistical model of urban multipath propagation," IEEE Trans. Veh. Technol. **21**(1), 1–9.

Tzeremes, G., and Christodoulou, C. G. (**2002**). "Use of Weibull distribution for describing outdoor multipath fading," *Antennas and Propagation Soc. International Symposium* **1**, 232–234.

Wesler, J. E. (**1973**). "Community noise survey of Medford, Massachusetts," J. Acoust. Soc. Am. **54**(4), 985–995.

Wilson, D. K., Bey-Hernandez, R., and Ostashev, V. E. (**2006**). "Statistical models for fading and coherence of sound in urban environments," J. Acoust. Soc. Am. **120**(5), 3336–3337.

Wilson, D. K., Breton, D. J., Hart, C. R., and Pettit, C. L. (**2017a**). "Impact of signal scattering and parametric uncertainties on receiver operating characteristics," in *Signal Processing, Sensor/Information Fusion, and Target Recognition XXVI* (International Society for Optics and Photonics, SPIE), Vol. 10200, p. 1020018.

Wilson, D. K., Hart, C. R., Pettit, C. L., Breton, D. J., Nykaza, E. T., and Ostashev, V. E. (**2017b**). "Scattered signal distributions, parametric uncertainties, and Bayesian sequential updating," Proc. Mtgs. Acoust. **31**, 055002.

Zuo, F., Li, Y., Johnson, S., Johnson, J., Varughese, S., Copes, R., Liu, F., Wu, H. J., Hou, R., and Chen, H. (**2014**). "Temporal and spatial variability of traffic-related noise in the City of Toronto, Canada," Sci. Total Environ. **472**, 1100–1107.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
November 2021

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Urban Noise Distributions and the Influence of Geometric Spreading on Skewness

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
SAB201

5d. PROJECT NUMBER
AB2

5e. TASK NUMBER
01

5f. WORK UNIT NUMBER

6. AUTHOR(S)
D. Keith Wilson, Matthew J. Kamrath, Caitlin E. Haedrich, Daniel J. Breton, and Carl R. Hart

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cold Regions Research and Engineering Laboratory
U.S. Army Engineer Research and Development Center
72 Lyme Road
Hanover, NH 03755

8. PERFORMING ORGANIZATION REPORT NUMBER
ERDC/CRREL MP-21-27

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Corps of Engineers
Washington, DC 20314

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
This article was originally published online in the *Journal of Acoustic Society of America* on 3 August 2021.
This work was supported by the Army ERDC Geospatial Research and Engineering basic research program.

14. ABSTRACT
Statistical distributions of urban noise levels are influenced by many complex phenomena, including spatial and temporal variations in the source level, multisource mixtures, propagation losses, and random fading from multipath reflections. This article provides a broad perspective on the varying impacts of these phenomena. Distributions incorporating random fading and averaging (e.g., gamma and noncentral Erlang) tend to be negatively skewed on logarithmic (decibel) axes but can be positively skewed if the fading process is strongly modulated by source power variations (e.g., compound gamma). In contrast, distributions incorporating randomly positioned sources and explicit geometric spreading [e.g., exponentially modified Gaussian (EMG)] tend to be positively skewed with exponential tails on logarithmic axes. To evaluate the suitability of the various distributions, one-third octave band sound-level data were measured at 37 locations in the North End of Boston, MA. Based on the Kullback-Leibler divergence as calculated across all of the locations and frequencies, the EMG provides the most consistently good agreement with the data, which were generally positively skewed. The compound gamma also fits the data well and even outperforms the EMG for the small minority of cases exhibiting negative skew. The lognormal provides a suitable fit in cases in which particular non-traffic noise sources dominate.

15. SUBJECT TERMS
Noise control, Noise – Measurement, Mathematical statistics – Noise, City noise, Boston (Mass.)

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	UU	23	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. 239.18