Parameter Estimation and Performance Modeling in Generalized-Pareto-Distributed Clutter

D. A. Abraham

Applied Physics Laboratory
1013 NE 40th Street

University of Washington
Seattle, Washington 98105-6698

Acknowledgements

Abstract

In active sonar systems, a clutter-dominated background is often the limiting factor affecting detection performance. Sources of clutter typically violate the central-limit-theorem conditions that lead to Gaussian distributed bandpass measurements, necessitating more general statistical models to represent their effect on the system. The generalized Pareto distribution (GPD) is a common phenomenological model for clutter, with its shape parameter representing severity through the heaviness of the distribution tail. The focus of this report is on techniques for representing active sonar clutter with the GPD model and assessing the degradation in detection performance as the clutter severity increases. The GPD shape parameter is interpreted through its relationship to the K-distribution shape parameter to understand what values are appropriate in different modeling scenarios (e.g., ranging from mild to extremely heavy-tailed clutter). A comparison of parameter estimators leads to one reliably providing an estimate representative of a physically realizable process. Approximations to the design SNR required to achieve a detector operating-point specification (i.e., the detection threshold term in the sonar equation) for the standard signals in GPD clutter are presented as is the J-divergence detection currency when accounting for thresholding. These simple approximations enable more realistic prediction of active-sonar detection performance by accounting for clutter severity through the GPD model.

Contents

1 Introduction

The performance of active sonar systems is often limited by false alarms arising from reflections of the sensing waveform off physical objects, boundaries, or other discontinuities in the ocean environment. When these reflectors have a myriad of independent scattering elements, the central limit theorem (CLT) dictates that the combined acoustic pressure measurement, termed reverberation, will follow a Gaussian distribution. Subsequent signal processing (basebanding and matched filtering) of the bandpass reverberation measurements converts the Gaussian distribution to a Rayleigh-distributed envelope and an exponentially-distributed instantaneous intensity. This scenario, with its distribution chain,[1] defines the nominal benign background. Although sources of reverberation satisfying the requirements of the CLT are ubiquitous, it is not uncommon to encounter those that do not and these interferences are often the ones driving performance. Termed clutter in active sonar, such objects typically have too few independent elemental scatterers, which leads to a larger number of false alarms than that expected under a benign background. From a modeling perspective, the upper tail of the normalized envelope or intensity probability density function (PDF) is heavier for clutter than for reverberation. For a fixed detector decision threshold, this causes the probability of false alarm (P_f) in clutter-dominated regions to be higher than in the more benign conditions typical of diffuse reverberation or ambient noise. In systems with a variable threshold that adapts to maintain a constant probability of false alarm, the increase required in clutter-dominated areas results in a reduction in the probability of detection. The focus of this report is on using the generalized Pareto distribution (GPD) to represent clutter in modeling the detection performance of active sonar systems.

There are many statistical distributions that can be used to represent clutter [1, Sect. 7.4.3], including ones with physical interpretations (e.g., the K and Poisson-Rayleigh distributions) and those with phenomenological support (e.g., the GPD, Weibull and log-normal distributions). Of these, the K-distribution will be used to interpret the severity of the GPD and to subsequently define different regimes of interest for the GPD parameter controlling tail heaviness. A motivating example illustrating strong support for the GPD model is found in Fig. 1 where P_f is shown as a function of the detector decision threshold for data obtained during the NATO SCARAB 1997 Experiment.[2] The probability of false alarm is also shown for an exponentially-distributed intensity and the heavier-tailed K and GPD models. Although most of the data from the SCARAB 1997 Experiment were fit well by the K-distribution [3], this particular segment, which contained reflections from steep bathymetry, is fit best by the GPD model.

The GPD was shown in [4] to arise from a modulation process between an exponentially-distributed instantaneous intensity (representative of a benign background) and an inverse-gamma-distributed random variable (nominally representing a random multiplicative effect). This model is a scalar form of the more general spherically invariant random vector (SIRV) model [5]. La Cour [6] and Gelb [7] have championed use of the GPD to represent active sonar clutter, with a focus on how it is useful when representing the distribution above the detector decision threshold. Background on the GPD, its SIRV genesis model, properties, and how to interpret the shape parameter (γ) controlling tail heaviness are found in Sect. 2.

Although the K-distribution provides an interpretation of the GPD shape parameter when matching the scintillation index (see Sect. 2.2), estimating γ from measured data allows its use in

---

[1] A distribution chain is the sequence of statistical distributions required to represent a random signal and/or noise as it passes through the operations comprising a signal processing chain [1, pg. 388].

[2] Acknowledgement: [2] with gratitude to Dr. C. Holland (scientist in charge, SCARAB 1997 Experiment).
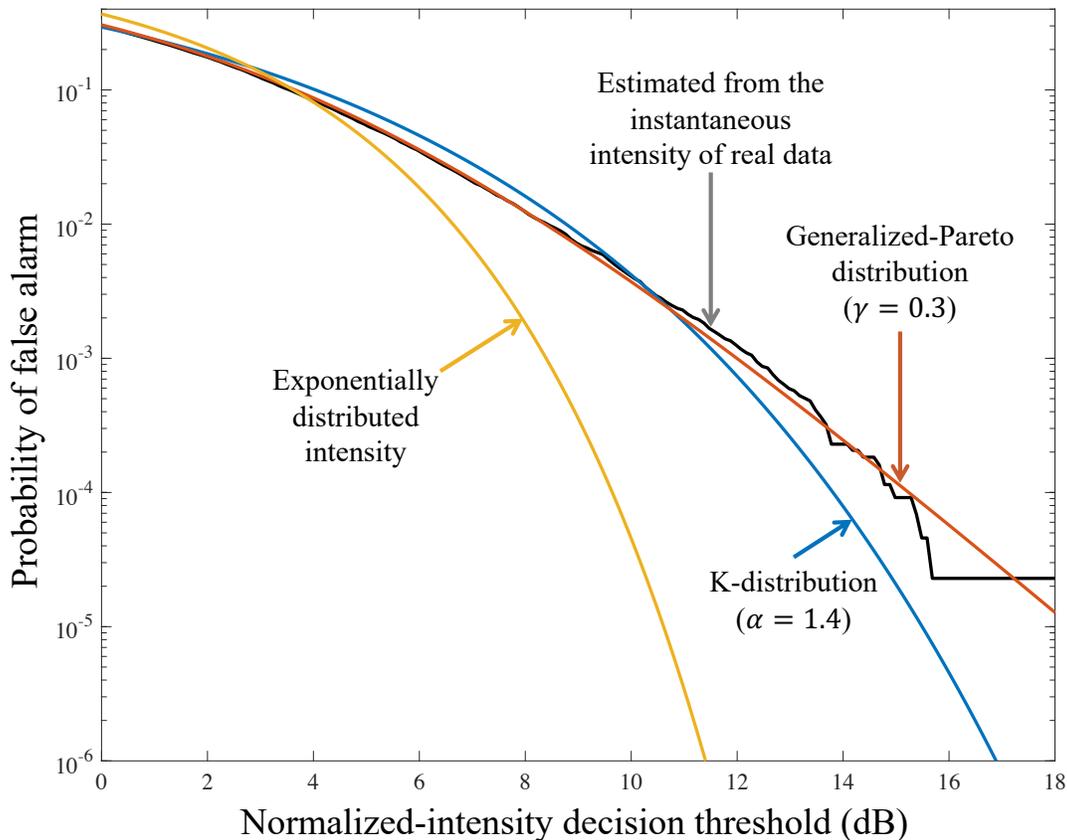
Figure 1: Probability of false alarm as a function of the normalized-intensity detector decision threshold for data from the SCARAB 1997 Experiment [2] along with that obtained for the K and generalized-Pareto distributions.

signal processing algorithms and can provide a more pertinent assessment when modeling detection performance in clutter-dominated backgrounds. Several parameter estimators for γ are presented and compared in Sect. 3. An iterative solution to the maximum likelihood estimator (MLE) for γ was presented in [6]. However, it is unconstrained and therefore can yield values outside of the interval $[0, 0.5)$ for which the GPD has tails at least as heavy as the exponential distribution ($\gamma \geq 0$) and where the intensity has a finite variance ($\gamma < 0.5$). The method-of-moments-based Bayesian approach found in [8] for the K-distribution shape parameter is extended in Sect. 3.4 to produce an estimator of the GPD shape parameter that always lies on the interval $[0, 0.5)$. This enables use of the estimate in modeling and analysis requiring distributions representative of physically realizable processes.

Modeling sonar performance in clutter often starts with forward models of the probabilities of false alarm (P_f) and detection (P_d), given the signal-to-noise power ratio (SNR) and the shape parameter of the clutter distribution. The relationship between the detector decision threshold (h) and P_f, which is defined by the cumulative distribution function (CDF) of the detector's decision statistic, is typically straightforward to evaluate, especially when the detector comprises a single instantaneous intensity. Although evaluating P_d for signals in heavy-tailed noise is more difficult, there exist approximations that can be usefully accurate [1, Sect. 7.5.6]. For example, a Gaussian-noise-background approximation was used in [9] to obtain the design SNR achieving a desired

(P_d, P_f) operating point (i.e., the detection threshold (DT) term of the sonar equation) for signals in K-distributed clutter. This approximation essentially uses the decision threshold dictated by the clutter distribution in the P_d model for a benign background. The approach is applied in Sect. 4 to obtain very simple approximations to DT for deterministic and Gaussian-fluctuating signals in GPD clutter. An interesting result of the analysis is seen in Sect. 4.3 where DT in clutter can be approximated by DT in the benign background plus the decibel change in the intensity decision threshold required to maintain a constant P_f. The case of a detector formed by integrating intensities is also considered and seen to yield useful approximations to DT given the detector decision threshold. However, the relationship between h and P_f for the integrated-intensity detector is quite complicated and requires numerical evaluation (Sect. 4.2.2 & App. C.3) or an empirical approximation (Sect. 4.2.3 & App. C.4). The empirical approximation to the decision threshold provides a simple means for approximating DT in GPD clutter for the integrated-intensity detector, but comes at the expense of a larger error or a smaller viable parameter space than the single-intensity detector.

The final section of this report (Sect. 5) considers the J-divergence detection currency (JDC) [10] for the basic sonar signal models in GPD clutter. JDC is an alternative detection performance measure to the traditional (P_d, P_f) operating point, providing a simple scalar measure of performance that can easily be accrued over multiple measurements (the linear quantities simply add) and can be extended throughout the signal and information processing chain. Its basic application to a Rician signal (which contains both deterministic and Gaussian-fluctuating signals) in a benign background was covered in [10]. Given the complicated form of the intensity PDF for a signal in clutter, a numerical evaluation of J-divergence is mandated (Sect. 5.2). However, when incorporating the effect of thresholding a decision statistic prior to combination across measurements (Sect. 5.3), accurate approximations can be employed to more easily obtain JDC for a Rician signal in GPD clutter.

In addition to the techniques and analysis presented in Sects. 3–5 for parameter estimation and sonar performance modeling in GPD clutter, MATLAB® functions implementing the key results are provided in the appendices.

2 Genesis and properties of the generalized Pareto distribution (GPD)

The generalized Pareto distribution (GPD) [11, Ch. 20] is a member of the class of spherically invariant random vector (SIRV) models [5, 12], which are formed by modulating a multi-variate, zero-mean, Gaussian-distributed complex envelope by an independent, non-negative random scalar. The GPD is obtained in this formulation [4] when the square of the modulating random scalar follows an inverse gamma distribution (i.e., one over a gamma-distributed random variable). This is more easily described by characterizing the instantaneous intensity of a single sample as the quotient

$$Y = \frac{Y_o}{W} \quad \text{[units: power]} \tag{1}$$

where Y_o follows an exponential distribution with mean λ and W is an independent, gamma-distributed random variable with shape parameter $1/\gamma$ and scale parameter γ. In this formulation, the exponentially-distributed numerator is representative of a Gaussian-distributed complex envelope and the randomness of the gamma-distributed W increases the tails of the PDF of Y, which allows representing active-sonar clutter exhibiting higher false alarm rates than those observed in

a benign background. Although the GPD does not have a genesis as a physical-statistical model (as does the K-distribution), it can be quite effective at representing heavy-tailed sonar clutter, as shown in [6, 7] and the motivating example seen in Fig. 1. As described in Sect. 2.4, it is also the limiting distribution describing the upper tail of many statistical models representing heavy-tailed data.

In the statistical analysis of clutter in active sonar systems, models can be applied to the modulus (i.e., the envelope) or squared modulus (i.e., the instantaneous intensity) of the complex-matched-filter response. These are most easily identified through their units. In this report, the complex-matched-filter response and its modulus will be defined as having "field" units, which could be pressure or a quantity proportional to pressure. The squared modulus will be defined as having "power" units in representation of a generic power quantity obtained by squaring a field quantity. Common power quantities include squared pressure and acoustic intensity. The theoretical analysis often exploits perfect normalization, which produces a unitless quantity. When there might be confusion as to how these quantities are formed, a description is provided (e.g., [unitless: normalized power] when formed from a ratio of two power quantities).

2.1 Distribution, moments, and scintillation index

The characteristics of the generalized Pareto distribution can be found in [11, Ch. 20] or [1, pgs. 302 & 417]. The following definitions use the notation found in the latter reference. The probability density function of the matched-filter intensity is

$$f_Y(y) = \frac{1}{\lambda\left(1 + \frac{\gamma}{\lambda}y\right)^{\gamma^{-1}+1}} \quad \text{for } y \geq 0, \tag{2}$$

where γ is the shape parameter and $\lambda > 0$ is the scale parameter. By letting $\gamma \to 0$ in (2), it can be shown[3] that $f_Y(y) \to e^{-y/\lambda}/\lambda$, which is an exponential distribution with mean λ. Increasing γ from zero clearly slows the decay of the PDF at large arguments, implying it represents a worsening of the clutter.

The GPD cumulative distribution function,

$$F_Y(y) = 1 - \frac{1}{\left(1 + \frac{\gamma}{\lambda}y\right)^{\gamma^{-1}}} \quad \text{for } y \geq 0, \tag{3}$$

is easily obtained by integrating (2). When setting a decision threshold as a function of the probability of false alarm (P_f) and the GPD shape and scale parameters, (3) can easily be solved to yield

$$h = \frac{\lambda}{\gamma}\left(\frac{1}{P_f^\gamma} - 1\right) \quad \text{[units: power]}. \tag{4}$$

The mean of the distribution is

$$E[Y] = \frac{\lambda}{1 - \gamma} \quad \text{for } \gamma < 1 \quad \text{[units: power]} \tag{5}$$

---

[3]Apply L'Hôpital's rule to $\log f_Y(y)$.

and the variance is

$$\text{Var}\{Y\} = \frac{\lambda^2}{(1-\gamma)^2(1-2\gamma)} \quad \text{for } \gamma < 0.5 \quad [\text{units: power}^2]. \tag{6}$$

These are easily combined to produce the scintillation index (SI), which is the ratio of the variance of the instantaneous intensity to the square of its mean,

$$\text{SI} = \frac{\text{Var}\{Y\}}{(E[Y])^2} = \frac{1}{1-2\gamma} \quad [\text{unitless: normalized power}^2]. \tag{7}$$

The scintillation index is a useful measure of PDF tail heaviness relative to the exponential distribution, for which $\text{SI} = 1$. Heavier-tailed distributions have a larger SI and lighter-tailed models a smaller one (e.g., a deterministic signal has $\text{SI} = 0$).

From [1, pg. 302], the kth moment of the GPD is

$$E[Y^k] = \frac{\lambda^k \Gamma(k+1)\Gamma(\gamma^{-1} - k)}{\gamma^k \Gamma(\gamma^{-1})} \quad \text{for } \gamma < 1/k \quad [\text{units: power}^k]. \tag{8}$$

This illustrates that for a given value of γ, only the moments up to $k < 1/\gamma$ exist.

Less common forms of the distribution:

For many heavy-tailed clutter models, it is more convenient to work with the instantaneous intensity as opposed to the envelope. For example, it can be seen from (8) that odd moments of the envelope (for which $k = 0.5, 1.5, 2.5, \dots$) do not simplify to rational functions of γ. However, as will be seen in Sect. 3, it can sometimes be advantageous to work with envelope data. The envelope is formed from the instantaneous intensity by the simple transformation $X = \sqrt{Y}$ [units: field], which in turn transforms the CDF in (3) to

$$F_X(x) = F_Y(x^2) = 1 - \frac{1}{\left(1 + \frac{\gamma}{\lambda}x^2\right)^{\gamma^{-1}}} \quad \text{for } x \geq 0. \tag{9}$$

Differentiating this then produces the PDF of the envelope,

$$f_X(x) = \frac{2x}{\lambda\left(1 + \frac{\gamma}{\lambda}x^2\right)^{\gamma^{-1}+1}} \quad \text{for } x \geq 0. \tag{10}$$

The moments of the envelope are straightforward to construct from (8).

The final step backwards in the distribution chain to the complex envelope (Z) producing a GPD instantaneous intensity (e.g., $Y = |Z|^2$) can be made assuming the phase is uniformly random on $[0, 2\pi)$. This then results in the PDF

$$f_Z(z) = \frac{1}{\pi\lambda\left(1 + \frac{\gamma}{\lambda}|z|^2\right)^{\gamma^{-1}+1}} \tag{11}$$

for z in the complex plane.

2.2 Interpreting the shape parameter

As with other heavy-tailed models, the shape parameter (γ) of the GPD dictates the severity of the clutter it represents. As noted above, $\gamma = 0$ represents a benign background and increasing γ from zero increases tail heaviness. To aid in the interpretation of the GPD shape parameter, it is related here to the physical model of the K-distribution by equating scintillation index. First, however, it is interesting to note that the ratio in (1) can be interpreted as an instantaneous intensity (Y_o) normalized by an estimate of the background power (W). In particular, a cell-averaging constant-false-alarm-rate (CA-CFAR) normalizer [1, Sect. 8.6.1] operating in a benign background produces data following the GPD with $\gamma = 1/L$ when the background power estimate is the average of L independent intensity samples.[4] L is approximately the product of the waveform bandwidth [units: Hz] and the temporal extent [units: s] of the background estimation window. The large quantities of stationary data required to estimate what can be small probabilities of false alarm typically can only be obtained after normalization. Although this is appropriate when assessing system performance, low quality normalizers (e.g., those with small values of L) can obscure the underlying clutter statistics and should therefore be taken into account when analyzing the clutter source is the primary objective.

Under the SIRV representation, the shape parameter of the GPD must be non-negative ($\gamma \geq 0$) because the shape parameter of the gamma distribution describing W must be positive. Although negative values of γ may be permitted by the functions shown in Sect. 2.1, they represent distributions with lighter upper tails than the nominal exponential distribution and impart an upper limit on their argument. As such, they are generally deemed inappropriate when representing active sonar clutter. Letting γ tend to zero causes the distribution of W to tend to an impulse function at $W = 1$ (note that the mean of W is one and its variance is γ), which implies that Y is exponentially-distributed when $\gamma = 0$. As γ increases from zero, the tails of the distribution increase, as can be seen in Fig. 2, which displays the probability of false alarm when a single (perfectly) normalized intensity is compared to the detector decision threshold.

A phenomenological application of the GPD might require a large value of γ to accurately represent extreme observations of active-sonar clutter. However, the variance of the instantaneous intensity under the GPD model is infinite when $\gamma \geq 0.5$. This suggests limiting GPD shape parameters to be on the interval $\gamma \in [0, 0.5)$ to represent physically realizable random processes. Within this regime, a lexicon for describing clutter can be obtained by mapping the first two intensity moments of the GPD to those of the K-distribution to obtain an equivalent K-distribution shape parameter,

$$\alpha = \frac{1}{\gamma} - 2 \tag{12}$$

or $\gamma = 1/(\alpha + 2)$ when starting with α. The physical interpretation of α as being proportional to the number of independent scatters contributing to the clutter [13] can then be employed to define the set of clutter regimes shown in Table 1. Because the first two intensity moments for the GPD and K-distributions are the same under (12), they have the same scintillation index,

$$\text{SI} = \frac{1}{1 - 2\gamma} = 1 + \frac{2}{\alpha} \tag{13}$$

for $\gamma \in [0, 0.5)$ and $\alpha > 0$. This relationship illustrates that the benign-background case of an exponentially-distributed instantaneous intensity, for which SI = 1, occurs when $\gamma \to 0$ and $\alpha \to \infty$.

---

[4]The GPD is proportional to an F-distributed random variable with $\nu_1 = 2$ and $\nu_2 = 2/\gamma$ degrees of freedom.
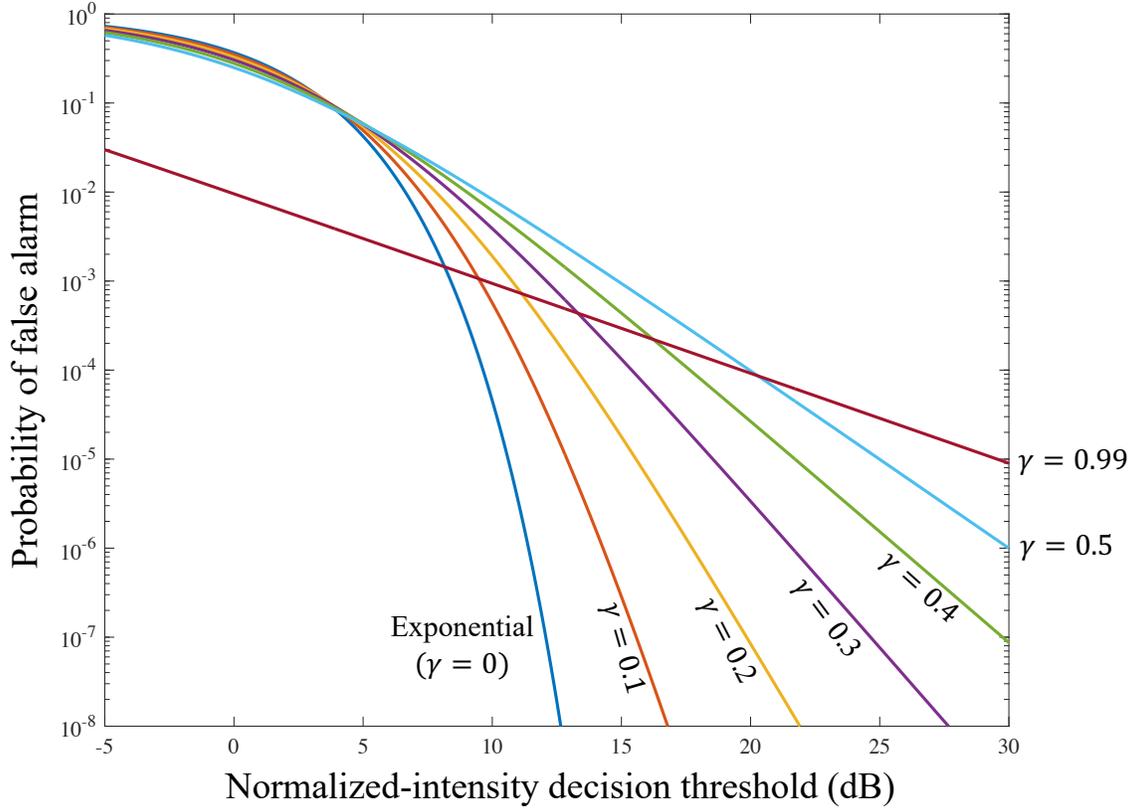
Figure 2: Probability of false alarm as a function of the normalized-intensity detector decision threshold for generalized-Pareto-distributed clutter with different shape parameters. The variance of the instantaneous intensity is only finite when $\gamma < 0.5$.

Table 1: Lexicon for describing heavy-tailed noise.

Regime in scintillation index (SI)	Regime in K shape	Regime in GPD shape	How heavy tailed?	Characterization
$1 \leq \text{SI} < 1.2$	$10 < \alpha \leq \infty$	$0 \leq \gamma < 0.08$	scarcely	diffuse noise or reverberation (very close to benign bkgnd.)
$1.2 \leq \text{SI} < 1.4$	$5 < \alpha \leq 10$	$0.08 \leq \gamma < 0.14$	mildly	diffuse clutter (e.g., rocky ridge, seaweed, fish schools)
$1.4 \leq \text{SI} < 2$	$2 < \alpha \leq 5$	$0.14 \leq \gamma < 0.25$	moderately	discrete, multi-faceted clutter (e.g., oil rig or fragmented shipwreck)
$2 \leq \text{SI} < 9$	$0.25 < \alpha \leq 2$	$0.25 \leq \gamma < 0.44$	very	single dominant scatterer (e.g., with very compact structure)
$\text{SI} \geq 9$	$0 < \alpha \leq 0.25$	$0.44 \leq \gamma < 0.5$	extremely	not explained by a single process (via K-distribution)

As α decreases to the point where there are too few independent scatterers for the central limit theorem to yield an adequate approximation to a Gaussian-distributed complex envelope, the upper tail of the intensity distribution becomes heavier and γ increases.

The first four regimes seen in Table 1 describe increasing severity in terms of tail heaviness along with potential examples of clutter sources. The examples shown in Fig. 2 illustrate how the heavier tails impact the probability of false alarm. When α equals 0.5, which equates to $\gamma = 0.4$, the K-distribution interpretation for scatterers having an exponentially-distributed size is that there is a single dominant scatterer [13, eq. 15]. Accounting for some variation in the size distribution, the regime for which $\alpha \in (0, 0.25]$ or $\gamma \in [0.44, 0.5)$ is not easily explained by a single process.

When data are observed with *extremely* heavy tails ($\gamma \geq 0.44$ or $\alpha < 0.25$) or in the upper end of *very* heavy tails, it may be more appropriate to use mixture models [14] or the clutter sources should be viewed as discrete interferences with average intensities exceeding the noise background. Raising the detector decision threshold is not likely to control these types of sources—their echoes are generally large enough relative to the local background for them to be detected with ease. As such they are typically handled by (or cause problems in) subsequent information processing such as tracking and classification.

2.3 The GPD is a closed distribution family under thresholding

A useful property of the GPD is that it is closed under the process of thresholding. That is, the distribution of the excess over a threshold (given it is exceeded) is another GPD. Suppose

$$Z = Y - h \text{ given } Y > h, \tag{14}$$

where Y is GPD with shape γ and scale λ. The PDF of Z is

$$f_Z(z) = \frac{f_Y(z + h)}{1 - F_Y(h)} = \frac{1}{\tilde{\lambda}\left(1 + \gamma z/\tilde{\lambda}\right)^{\gamma^{-1}+1}} \text{ for } z \geq 0, \tag{15}$$

which is GPD with the same shape parameter and a scale parameter $\tilde{\lambda} = \lambda + \gamma h$. The self-replicating memoryless property of the exponential distribution can be seen by setting $\gamma = 0$, so the excess over the threshold is exponentially distributed with mean λ just like the unthresholded intensity.

The importance of this result lies in the ability to estimate the shape parameter from thresholded data, which can be helpful when analyzing active-sonar clutter.

2.4 The GPD is a limit distribution for the excess over the threshold

When $\gamma = 0$, the GPD simplifies to an exponential distribution with mean λ, which has PDF

$$f_Y(y) = \frac{e^{-y/\lambda}}{\lambda} \text{ for } y \geq 0. \tag{16}$$

The decay of the PDF is clearly exponential with y. For $\gamma > 0$, it can be seen from the PDF in (2) that the decay tends to an inverse power-law as y increases. This tail behavior can be characterized by considering the distribution of the excess over the threshold (EOT). In Sect. 2.3, this was shown

to be another GPD with the same shape parameter and a different scale, so the EOT for the GPD always has an inverse power-law tail behavior when $\gamma > 0$.

One of the arguments made in [6] for use of the GPD to represent sonar clutter was that it is one of the limiting distributions of the EOT. The limit theorem [15, 16] dictates the EOT distribution for random variables with a domain $y \geq 0$ tends to either an exponential or generalized Pareto distribution as the threshold increases to infinity. This implies that the GPD could be a good fit above some finite threshold for distributions having the inverse power-law tail behavior. However, the argument weakens as the threshold decreases and does not apply when the tail behavior has exponential decay, as is the case for the K-distribution. Thus, whether the GPD is appropriate can depend on how far out into the tails is of interest and whether the underlying distribution has inverse power-law tail behavior or not.

Recall that the GPD is a scaled F distribution, which is also representative of an exponentially-distributed instantaneous intensity normalized by a cell-averaging estimate of the power in a benign background. In this example the normalization process transforms a distribution with exponential decay in the tails to one having an inverse power-law behavior. With a high-quality normalizer, however, this might not be apparent except at the highest threshold levels. Conversely, a low-quality normalizer may require the GPD model even at low thresholds. When the clutter statistics are used to infer the state of nature (e.g., a characteristic or condition of the clutter source), it is clear that the normalizer needs to be designed so it does not adversely affect the inference or that the inference takes into account the normalization.

In the case of a normalizer, it is the division by an estimate of the background power ($\hat{\lambda}$) that causes the change in tail behavior. The background power estimate can be assumed to have exponential tail behavior by noting that averaging lightens heavier tails toward the Gaussian distribution (via the central limit theorem). It is then straightforward to show that inverting a random variable with exponential tail behavior produces one following an inverse power-law. That the power-law tail behavior dominates in the product between $1/\hat{\lambda}$ and the test cell in a normalizer can be seen using Mellin transforms, where the transform of the product of two independent random variables is the product of their individual transforms. The Mellin transform essentially provides a spectrum with a power series kernel [17, pg. 256] and some distributions with power-law tail behavior (e.g., the F distribution and the inverse of many exponential-tail-behavior PDFs) exhibit a discontinuity to infinity at their power, which dominates the transform of the lighter-tailed term.

3   Estimating the GPD shape and scale parameters

The GPD shape parameter (γ) provides a measure of the heaviness of the clutter distribution tail, which impacts false-alarm performance in active sonar systems. Several techniques for estimating γ are presented and evaluated in this section. When shape-parameter estimates are utilized in performance prediction (e.g., when evaluating the design SNR in Sect. 4 or the J-divergence detection currency in Sect. 5) or incorporated into signal and information processing algorithms, the estimates ideally represent viable clutter distributions in the sense that they should be no lighter-tailed than a benign background (i.e., $\gamma \geq 0$) and should have a finite variance (i.e., $\gamma < 0.5$). When the first condition is violated, the clutter distribution is referred to as being *sub-Rayleigh* ($\gamma < 0$), referencing the Rayleigh-distributed envelope found in a benign background. The latter condition ($\gamma \geq 0.5$) will be referred to as an *infinite-variance* clutter distribution. As seen in Sects. 3.1–3.3 and illustrated in Fig. 3, the most common approaches to parameter estimation (maximum likelihood

and method of moments) do not necessarily constrain the estimates to lie on the interval $[0, 0.5)$. To satisfy this requirement, the Bayesian approach described in [8] for estimating the K-distribution shape parameter is extended in Sect. 3.4 to the GPD.

To illustrate the issues surrounding estimation of the GPD shape parameter, consider the histograms of parameter estimates shown in Fig. 3. Although the true value of $\gamma = 0.25$ is in the middle of the viable range, it represents moderate to very heavy-tailed clutter (e.g., it has SI = 2 and is equivalent to a K-distribution shape parameter of $\alpha = 0.25$). In this example, only $N = 100$ independent intensity samples are used in the parameter estimation so the issues are apparent in the histograms. Most applications will utilize more intensity samples in the estimation and encounter failures less frequently. Examples using $N = 1000$ intensity samples are shown in Fig. 4 for $\gamma = 0.44$ (extremely heavy-tailed clutter) and in Fig. 5 for $\gamma = 0.01$ (which is essentially a Rayleigh-distributed envelope).
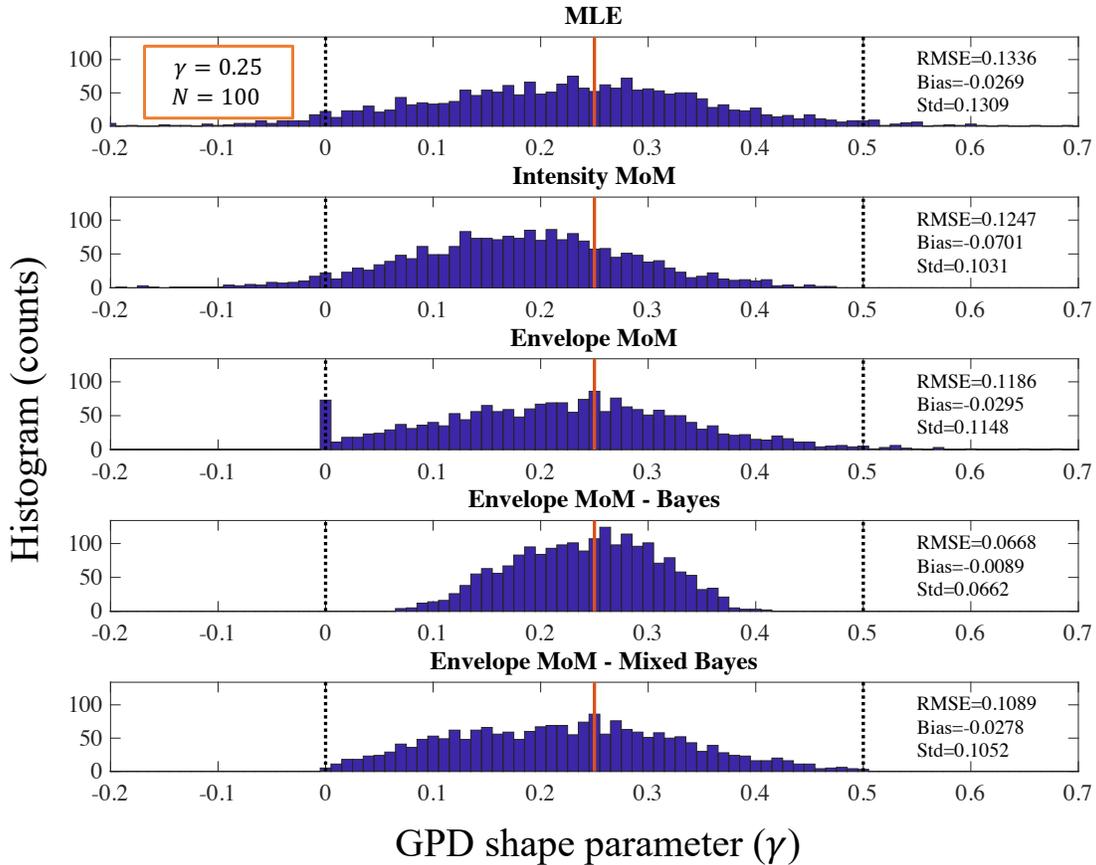
Figure 3: Histograms of GPD shape parameter estimates using $N = 100$ independent intensity samples when $\gamma = 0.25$.

In discussing these examples, the following comments refer to Fig. 3; however, many also apply to the other scenarios. The maximum likelihood estimator (Sect. 3.1),[5] exhibits both sub-Rayleigh and infinite-variance estimates and (in this example) the highest root mean-squared-error (RMSE). Although the intensity method-of-moments (MoM) estimator (Sect. 3.2) inherently does not produce

---

[5]The MLE iteration in this example is initialized using the intensity MoM estimate because it is always available. It is not necessarily a global maximum in every example as a better initialization might improve MLE performance.

infinite-variance estimates because it utilizes the moment equation for a squared intensity, it can produce sub-Rayleigh ones. When envelope moments are used in an MoM estimator (Sect. 3.3), the converse occurs: it produces no sub-Rayleigh estimates, but can produce infinite-variance estimates. The envelope MoM estimator can also fail to produce invertible moment equations, in which case a Rayleigh-envelope distribution ($\gamma = 0$) is assumed. This can be seen by the large number of occurrences of $\gamma = 0$ in the middle plot of Fig. 3. The Bayesian adaptation of the envelope-MoM estimator (Sect. 3.4) illustrates the desired restriction to viable clutter distributions and the smallest RMSE in this example. An alternative estimator that uses the basic envelope-MoM estimator when it produces a viable result and otherwise resorts to the Bayesian adaptation is shown in the bottom plot of Fig. 3. In this example, the mixed approach performs poorly compared to the straight Bayesian envelope-MoM estimator. However, it can be slightly better when the shape parameter is closer to the boundaries of the viable-clutter regime and has compelling performance when a larger number of intensity samples are utilized in the estimators (e.g., see Figs. 4 & 5).

With respect to computational effort, the MLE requires the most owing to its implementation as an iteration. The intensity-MoM estimator requires the least as it is a linear function of a ratio of sample moments. The envelope-MoM estimators require inversion of a one-dimensional non-linear function and so require more effort than the intensity-MoM estimator but less than the MLE. The functional inversion can be avoided using the simple approximation presented in Sect. 3.3.1 at the cost of a small loss in performance. The Bayesian envelope-MoM estimator requires a step to adjust the sample-moment-ratio into the viable-clutter regime. This involves incomplete gamma functions, so the additional computational requirement is not insignificant. However, it is still significantly quicker than the MLE.

The different estimators are described or derived in Sect. 3.1–Sect. 3.4. An analysis of the errors, including a comparison to the Cramér-Rao lower bound, is presented as a function of γ in Sect. 3.5.

3.1 Maximum likelihood estimator

Although no closed-form solution exists to the maximum likelihood estimator (MLE) of the GPD parameters, it is amenable to an iterative solution. From [6, eqs. 13a & 13b] the updates for the shape and scale parameters are

$$\hat{\gamma} := \frac{1}{N} \sum_{n=1}^{N} \log\left(1 + \frac{\hat{\gamma} Y_n}{\hat{\lambda}}\right) \tag{17}$$

and

$$\hat{\lambda} := \frac{1 + \hat{\gamma}}{N} \sum_{n=1}^{N} \frac{Y_n}{1 + \hat{\gamma} Y_n / \hat{\lambda}}, \tag{18}$$

where the operator := represents replacement in an iterative numerical evaluation. The initial parameter estimates can be obtained from one of the method-of-moment procedures described in the following sections. Starting it at arbitrary values generally requires more iterations for convergence. Examining (17), it can be seen that the iteration does not alter the sign of the shape parameter and if $\hat{\gamma}$ starts out at zero it stays there. The primary issues with the MLE are (i) it can produce shape parameter estimates outside of $[0, 0.5)$ and (ii) convergence can take many iterations.
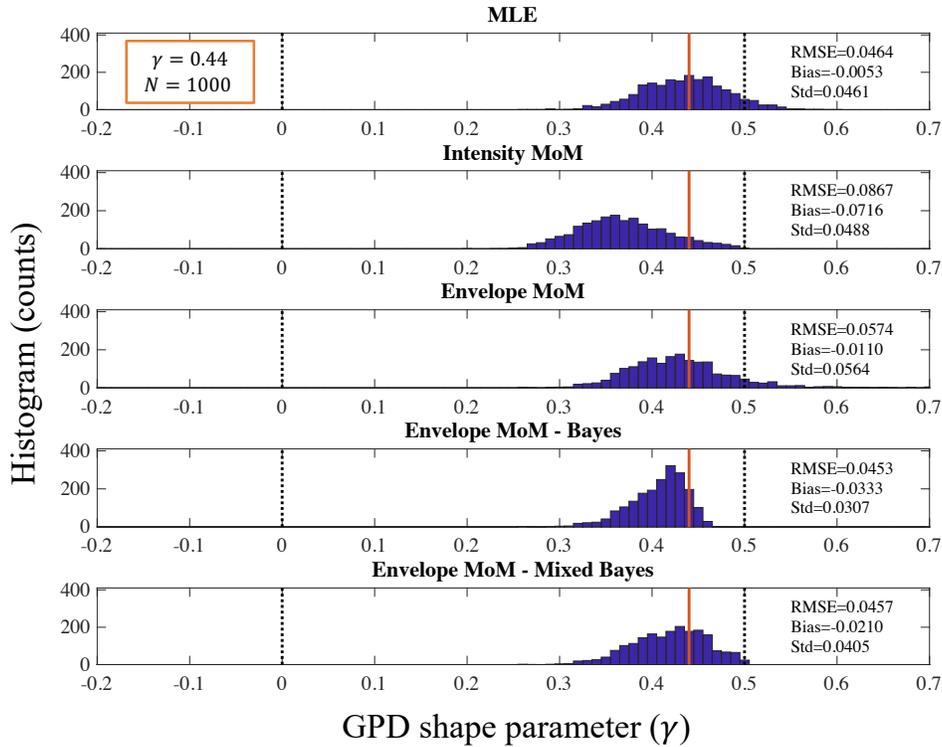
Figure 4: Histograms of GPD shape parameter estimates using $N = 1000$ independent intensity samples when $\gamma = 0.44$.
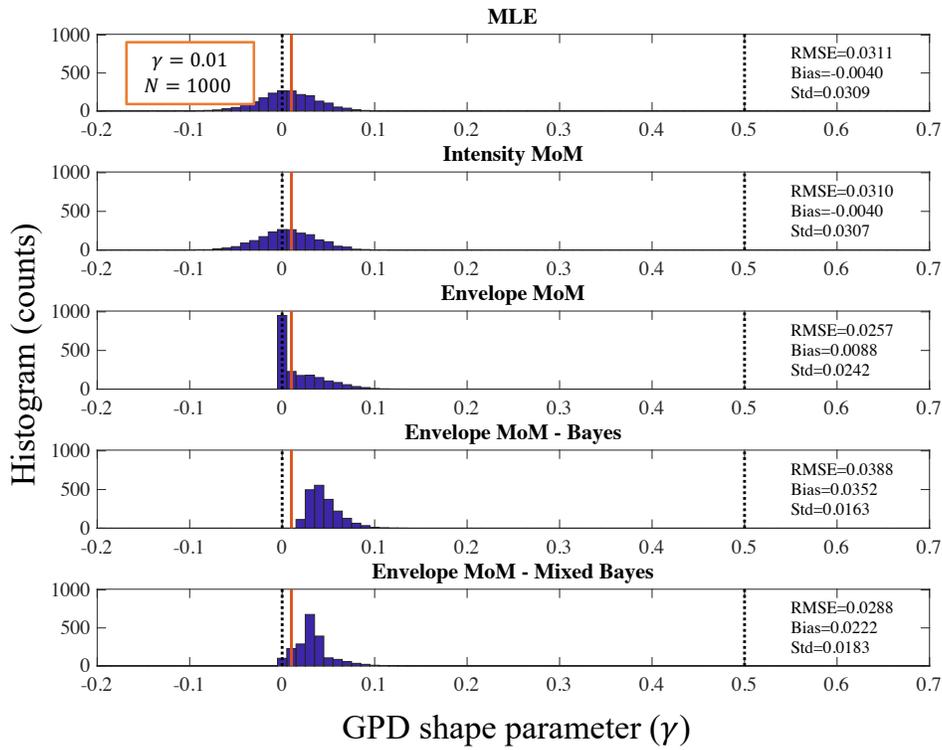
Figure 5: Histograms of GPD shape parameter estimates using $N = 1000$ independent intensity samples when $\gamma = 0.01$.

When implementing the iteration with an initial value of $\hat{\gamma} < 0$, the argument of the logarithm in (17) must be tested to ensure it exceeds zero (i.e., that $\hat{\lambda} > -\hat{\gamma}\max_n\{Y_n\}$). If an iteration violates this, it should be restarted using a positive value of $\hat{\gamma}$.

3.2 Method of moments: Intensity

The method-of-moments (MoM) estimator using the first two intensity moments is formed by solving (5) and (6) for γ and replacing the mean and variance by their estimators,

$$\hat{\gamma} = \frac{1}{2}\left(1 - \frac{\hat{\mu}^2}{\hat{\sigma}^2}\right) \tag{19}$$

where

$$\hat{\mu} = \frac{1}{N}\sum_{n=1}^{N} Y_n \tag{20}$$

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

Noting that the sample mean and sample variance of intensity measurements are always non-negative, it can be seen that $\hat{\gamma}$ in (19) must be below 0.5. This follows from the requirement on the variance. However, it can produce estimates that are negative. From the form of (19), it can be seen that this is more likely to occur when the scintillation index SI $= \sigma^2/\mu^2$ equals one (i.e., when the envelope is Rayleigh distributed), which is when $\gamma = 0$. Although this suggests limiting the estimate to zero when a sub-Rayleigh value is observed, this can occur for larger values of γ with some frequency when N is not large (e.g., see Fig. 3).

Given an estimate of the shape parameter, the MoM estimate of the scale parameter is simply

$$\hat{\lambda} = \hat{\mu}(1 - \hat{\gamma}). \tag{22}$$

The intensity-MoM estimator is the simplest to implement and always produces a solution. However, the possibility of observing negative shape-parameter estimates and poor performance for very heavy-tailed distributions make it less desirable.

3.3 Method of moments: Envelope

In [18], it was shown for the K-distribution that a method of moments estimator using the first two moments of the envelope (which is the square root of the intensity) performed better than one based on the first two intensity moments. The squaring of intensity measurements is likely exacerbated by heavy-tailed distributions, which suggests the approach may be similarly beneficial when estimating the GPD shape parameter.

Let $X = \sqrt{Y}$ be the envelope measurement. Using the $k = 0.5$ moment from (8) produces the envelope mean,

$$E[X] = E[\sqrt{Y}] = \frac{\sqrt{\lambda}\Gamma(1.5)\Gamma\left(\gamma^{-1} - 0.5\right)}{\sqrt{\gamma}\Gamma(\gamma^{-1})}. \quad \text{[units: field]} \tag{23}$$

The variance of the envelope can then be obtained from the intensity mean, $E[Y] = E[X^2] = \lambda/(1-\gamma)$ and (23), via $\mathrm{Var}\{X\} = E[X^2] - (E[X])^2$. These combine to produce the ratio of the squared average envelope to its variance,

$$d = \frac{E[X]^2}{\mathrm{Var}\{X\}} = \left[\frac{4\gamma\Gamma^2(\gamma^{-1})}{\pi(1-\gamma)\Gamma^2(\gamma^{-1}-1/2)} - 1 \right]^{-1} = g(\gamma) \quad \text{[unitless: normalized power]}. \qquad (24)$$

The envelope MoM estimator is then obtained by replacing the mean and variance of the envelope by the corresponding sample moments and solving (24) for γ.

The moment equations used to form (24) only require $\gamma < 1$, which implies errors in the sample moments can cause shape-parameter estimates exceeding 0.5 even when the true value is less. This approach also suffers from the same problem as the K-distribution where any moment-ratio estimate

$$d \geq d_{\max} = \frac{\pi}{4-\pi} \quad \text{[unitless: normalized power]} \qquad (25)$$

results in non-invertible moment equations. Similar to the intensity MoM estimator, this is more likely to occur when γ is near zero, but is not unlikely for higher values of γ when N is small. The estimate of the scale parameter follows (22).

The deficiencies of this estimate include (i) it can produce shape parameter estimates in the infinite-variance regime ($\gamma \geq 0.5$) and (ii) it can fail to produce an estimate.

3.3.1 Solving the envelope moment-ratio equation

In most scenarios, the following approximation can be used to solve the envelope moment-ratio equation $d = g(\gamma)$ in (24) for γ,

$$\gamma = g^{-1}(d) \approx \left(1 - \frac{d}{d_{\max}}\right)\left[1 - \frac{8}{19}\left(\frac{d}{d_{\max}}\right) + \frac{26}{53}\left(\frac{d}{d_{\max}}\right)^2 - \frac{3}{14}\left(\frac{d}{d_{\max}}\right)^3\right]. \qquad (26)$$

The approximation was developed as a least-squared-error polynomial fit to $g^{-1}(d)$ subject to an exact fit at $d = 0$ where $\gamma = 1$ and $d = d_{\max}$ where $\gamma = 0$. The relative absolute error is less than 0.4%, which will generally be much smaller than the standard deviation of the parameter estimator. The analysis presented in this report uses (26) without any refinement.

When greater accuracy is required, the following Newton-Raphson iteration can be applied,

$$\hat{\gamma} := \hat{\gamma} - \frac{g(\hat{\gamma}) - d}{g'(\hat{\gamma})}. \qquad (27)$$

This requires the derivative of (24) with respect to γ, which is

$$g'(\gamma) = \frac{-4\pi\Gamma^2(\gamma^{-1})\,\Gamma^2(\gamma^{-1}-1/2)\{1 + 2(1-\gamma^{-1})[\psi(\gamma^{-1}) - \psi(\gamma^{-1}-1/2)]\}}{[4\gamma\Gamma^2(\gamma^{-1}) - \pi(1-\gamma)\Gamma^2(\gamma^{-1}-1/2)]^2}, \qquad (28)$$

where $\psi(\cdot)$ is the digamma function. To account for values of γ near zero, it is prudent to evaluate the gamma functions indirectly through routines providing the logarithm of the gamma function.

Note that (24) can be written as

$$g(\gamma) = \left[\frac{4(\gamma^{-1} - 1)\Gamma^2(\gamma^{-1} - 1)}{\pi\Gamma^2(\gamma^{-1} - 1 + 1/2)} - 1 \right]^{-1}, \tag{29}$$

which is the same as [8, eq. 2] when $\alpha = \gamma^{-1} - 1$. This implies that the K-distribution shape parameter α can be approximated by

$$\alpha \approx \left(1 - \frac{d}{d_{\max}} \right)^{-1} \left[1 - \frac{8}{19}\left(\frac{d}{d_{\max}} \right) + \frac{26}{53}\left(\frac{d}{d_{\max}} \right)^2 - \frac{3}{14}\left(\frac{d}{d_{\max}} \right)^3 \right]^{-1} - 1 \tag{30}$$

when using envelope-based MoM estimators.

3.4   Method of moments: Bayesian posterior mean

The approach presented in [8] for the development of a Bayesian estimator of the K-distribution shape parameter is applied here to estimate the GPD shape parameter. Taking the moment ratio d in (24) as the parameter to be estimated, the Bayesian construct is based on the posterior distribution of d given the data,

$$f(d|\text{data}) = \frac{f(\text{data}|d)}{f(\text{data})} f_{\text{pri}}(d), \tag{31}$$

which is formed by Bayes' rule from the distribution of the data given d and a prior distribution, $f_{\text{pri}}(d)$. In [8], the shape of $f(\text{data}|d)$ when taken as a function of d was approximated as being gamma distributed with shape parameter a_t and scale parameter b_t. As described in [8, Sect. II.C-1], the gamma-distribution parameters are obtained through moment matching to an error-based characterization of d given the data. The posterior mean estimate of d was then restricted to the interval allowing inversion of the envelope moment equation with a uniform prior distribution.

Suppose the GDP shape parameter is required to be on the interval (γ_0, γ_1), which is nominally set to $[0, 0.5)$.[6] Noting that $d = g(\gamma)$ is a monotonically decreasing function of γ, this results in the interval $(d_{\min}, d_{\max})$ for the moment ratio where

$$d_{\min} = g(\gamma_1) = \frac{\pi^2}{16 - \pi^2} \quad [\text{unitless: normalized power}] \tag{32}$$

when $\gamma_1 = 0.5$ and $d_{\max} = g(\gamma_0) = \pi/(4 - \pi)$ as described in (25) when $\gamma_0 = 0$. In scenarios where γ_1 is set to one, $d_{\min} = 0$.

Following [8], the prior distribution on the moment ratio is assumed to be uniform on the interval $(d_{\min}, d_{\max})$. This has the effect of altering [8, eq. 12] to

$$f_D(d|t) = \frac{d^{a_t - 1} e^{-d/b_t}}{\Gamma(a_t) b_t^{a_t} [F_G(d_{\max}|a_t, b_t) - F_G(d_{\min}|a_t, b_t)]}, \tag{33}$$

---

[6]Whether or not the boundary of the interval is included depends on the prior PDF and the type of estimator. Although $\hat{\gamma} = 0$ is a viable estimate for the present application, the posterior-mean estimator shown in (34) is very unlikely to produce an estimate on either boundary.

where $F_G(x|\alpha, \beta)$ is the gamma cumulative distribution function (CDF) with shape α and scale β and a_t and b_t are the gamma-distribution parameters approximating the posterior distribution on the moment ratio.

The Bayesian posterior mean estimate of the moment ratio, when restricted by the prior PDF on γ, is the mean of the PDF in (33), which is

$$\hat{d} = a_t b_t \frac{F_G(d_{\max}|a_t + 1, b_t) - F_G(d_{\min}|a_t + 1, b_t)}{F_G(d_{\max}|a_t, b_t) - F_G(d_{\min}|a_t, b_t)} \qquad \text{[unitless: normalized power].} \qquad (34)$$

An estimate of the GPD shape parameter is then obtained from $\hat{d}$ as described in Sect. 3.3.1. MATLAB® code implementing this technique can be found in App. B.5.

When computational effort is a limiting factor, the numerator in (34) can be simplified through integration by parts to remove one set of incomplete-gamma-function evaluations,

$$\hat{d} = a_t b_t \left\{ 1 - \frac{d_{\max}^{a_t} e^{-d_{\max}/b_t} - d_{\min}^{a_t} e^{-d_{\min}/b_t}}{a_t \Gamma(a_t) b_t^{a_t} [F_G(d_{\max}|a_t, b_t) - F_G(d_{\min}|a_t, b_t)]} \right\} \qquad \text{[unitless: normalized power].} \quad (35)$$

3.5   Performance analysis

The results of a simulation analysis of the GPD shape-parameter estimators is performed when the estimate is formed from $N = 100$, 200, 500, or 1000 independent intensity samples. The error analysis employed 10^4 trials. To determine the efficiency of the estimators, the square root of the Cramér-Rao lower bound (CRLB) on the variance of unbiased estimators of γ when the scale parameter is also unknown is shown along with the estimator RMSE and standard deviation. The CRLB is obtained from [6, eq. 14],

$$\text{Var}\{\hat{\gamma}\} \geq \frac{(1 + \gamma)^2}{N}. \qquad (36)$$

As seen in Fig. 6, the MLE achieves the expected efficiency asymptotically as N increases. The MoM estimators are biased near one or both boundaries of the viable shape parameter region (i.e., $\gamma \in [0, 0.5)$), which adversely impacts their RMSE. Of the two estimators always providing viable shape-parameter estimates, the Bayesian envelope MoM (purple line) is a compelling choice for small sample sizes. The mixed Bayesian envelope MoM estimator (which only applies the prior when the envelope MoM estimate is not in the valid region) is competitive for large sample sizes and allows trading an increase in standard deviation for a smaller bias near the boundaries.
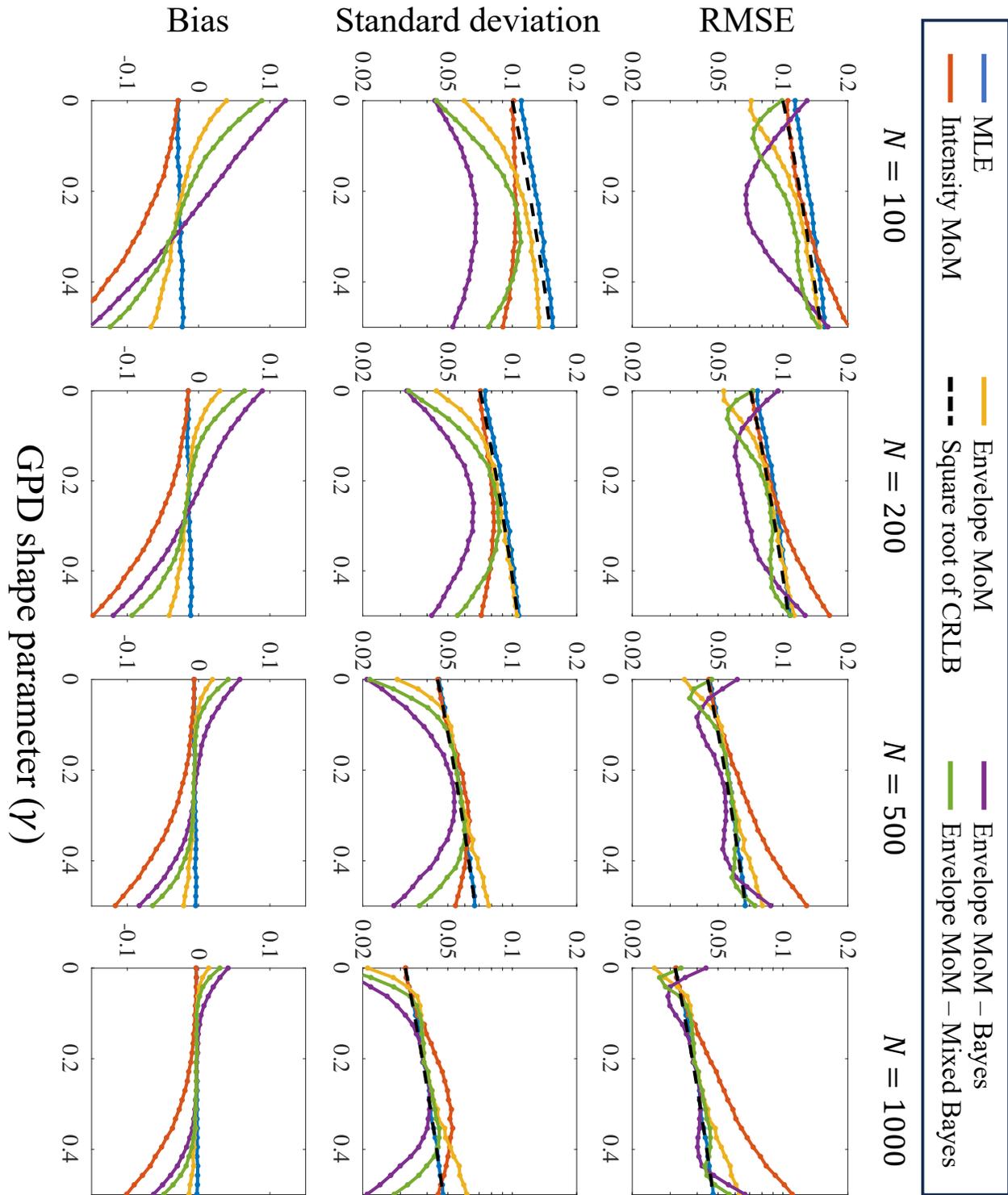
Figure 6: Error analysis for the estimators of the GPD shape parameter from 10^4 simulation trials.

4 Design SNR for a (P_d, P_f) specification in GPD clutter

A common performance measure in the analysis of a sonar system is the SNR at which it achieves a desired detection performance. When the performance is quantified by the probabilities of detection (P_d) and false alarm (P_f) in a single sonar resolution cell, the corresponding *design* SNR is the detection threshold term (DT [units: dB]) in the sonar equation. This design SNR is evaluated by first obtaining the detector decision threshold (h) achieving the design P_f for the given background model and then finding the SNR achieving the design P_d for the detector using h. This process requires definition of the detector and statistical models for the signal and noise. The noise here is assumed to follow the generalized Pareto distribution (GPD) and the standard deterministic and Gaussian-fluctuating signals will be evaluated. The detectors considered include using a single intensity sample (Sect. 4.1) and a detector integrating multiple independent intensity samples (Sect. 4.2).

For these detectors, one or both of the forward models for P_f and P_d involve integral equations, necessitating at least one numerical inversion of the functions to obtain first the decision threshold and then the design SNR. Fortunately, accurate approximations to the design SNR can be obtained in many scenarios through the approach described in [9], which entails approximating P_d by that for a signal in a benign background while using the correct detector decision threshold for the heavier-tailed clutter. For a single intensity in GPD clutter, this results in very simple and yet accurate approximations (Sect. 4.1.3). For the integrated-intensity detector, however, an additional approximation (Sect. 4.2.3) is required to obtain the detector decision threshold as a function of P_f, the number of independent intensities being integrated, and the GPD shape parameter. Although the additional error reduces the region over which the DT approximation is accurate, it will still be useful in many scenarios. MATLAB® code implementing the approximations to DT in GPD clutter can be found in App. C.1.

The simple approximations presented in Sects. 4.1.3 and 4.2.1 provide the design SNR as a function of the (P_d, P_f) specification and the shape parameter of the clutter distribution. In scenarios where the design SNR for a particular system is known in a benign background, an approach for approximating how much it increases in a clutter-dominated background is presented in Sect. 4.3. Using the benign-background approximation to P_d described above, the increase in the design SNR is shown to be approximately the decibel increase in the detector decision threshold required to maintain the P_f specification. The approximation is quite accurate when $P_d = 0.5$ for both the deterministic and Gaussian-fluctuating signals. Given the ubiquity of clutter, this is particularly useful when system data are available to determine the decision thresholds required to meet the P_f specification under different levels of clutter severity.

As might be anticipated, the results presented in this section illustrate that a higher SNR is required to maintain the (P_d, P_f) design operating point when the background becomes heavier tailed. It is important to recall that this analysis only pertains to systems that adjust their decision thresholds to maintain a constant cell-level P_f as the clutter background varies in severity over time and space (e.g., using the statistical normalizer described in [19]). For systems with fixed decision thresholds, P_f increases or decreases with the severity of the clutter and the design SNR (DT) and P_d are predominantly driven by the signal model and the fixed detector decision threshold.

As a final note, the analysis presented in this report assumes perfect normalization of the detector decision statistics. Although approximations exist for the SNR loss incurred by cell-averaging constant-false-alarm-rate (CFAR) normalizers in a benign background [1, Sect. 8.6.1.3], an extension

to heavy-tailed clutter is quite difficult.

4.1 Single-intensity detectors

Consider a detector testing a single intensity sample for the presence of an active-sonar echo from an object of interest (OOI). If Y [unitless: normalized power] is the perfectly normalized intensity and the detector decision threshold is h_c [unitless: normalized power], then the detection-performance probabilities are

$$P_f = \Pr\{Y \geq h_c | H_0\} \ \text{ and } \ P_d = \Pr\{Y \geq h_c | H_1\}, \tag{37}$$

where H_0 is the noise-only hypothesis and H_1 is the signal-present hypothesis. The design SNR achieving a (P_d, P_f) specification is obtained by inverting (i.e., solving) the first equation for h_c as a function of P_f and then using it while inverting the second equation to find the SNR achieving the desired P_d.

In order to avoid confusion, terms related to a clutter-dominated background will be identified using a subscript 'c'. For example, the detector decision threshold in (37) is h_c rather than h. In a benign background, the subscript 'g' (for Gaussian) will be employed when necessary to avoid ambiguity.

4.1.1 Detector decision threshold in GPD clutter

For GPD clutter, the detector decision threshold for testing a single intensity sample,

$$h_c = \frac{1 - \gamma}{\gamma} \left(\frac{1}{P_f^{\gamma}} - 1 \right) \quad \text{[unitless: normalized power]}, \tag{38}$$

is obtained directly from (4) by setting $\lambda = 1 - \gamma$ so the noise intensity has unit mean, as if the normalization were perfect. As $\gamma \to 0$, this decision threshold tends to that for the nominal exponentially-distributed intensity, $h_c \to -\log P_f$.

4.1.2 Integral equations for P_d

The instantaneous intensity when signal is present amid GPD clutter can be described as

$$Y = \left| A e^{j\psi} + \frac{\tilde{V}}{\sqrt{W}} \right|^2 \quad \text{[units: power]}, \tag{39}$$

where A and ψ are, respectively, the amplitude and phase of the signal, $\tilde{V}$ is a zero-mean complex Gaussian random variable with power λ and W is gamma distributed with shape parameter $1/\gamma$ and scale parameter γ. In terms of the SIRV model described in Sect. 2 for the GPD clutter, $Y_o = |\tilde{V}|^2$ is the instantaneous intensity in the numerator of (1). Setting $\lambda = 1 - \gamma$ so the average noise intensity equals one, which represents perfect normalization, implies that $E[A^2] = s$ where s is the linear-quantity SNR.

Conditioning on W implies that $\tilde{V}/\sqrt{W}$ is a zero-mean complex-Gaussian distributed random variable with power $(1 - \gamma)/W$. This allows characterizing Y through the traditional deterministic

and Gaussian-fluctuating signal models. For the Gaussian-fluctuating signal, $Ae^{j\psi}$ is a zero-mean complex Gaussian random variable with power s. Adding this to the noise while conditioning on W causes Y to follow an exponential distribution,

$$Y|W \sim \text{Expon}\left\{s + \frac{1-\gamma}{W}\right\}, \tag{40}$$

with mean $s + (1-\gamma)/W$. The probability of detection can then be obtained by removing the conditioning on W,

$$P_d = E_W[\text{Pr}\{Y \geq h_c|H_1, W\}] \tag{41}$$

$$= \int_0^\infty e^{-h_c/[s+(1-\gamma)/w]} f_W(w)\,\mathrm{d}w, \tag{42}$$

where

$$f_W(w) = \frac{w^{\gamma^{-1}-1}\,e^{-w/\gamma}}{\Gamma(\gamma^{-1})\,\gamma^{\gamma^{-1}}} \quad \text{for} \ \ w \geq 0 \tag{43}$$

is the gamma PDF for W.

A deterministic signal in a benign background leads to an instantaneous intensity proportional to a non-central chi-squared distributed random variable with two degrees of freedom. When conditioned on W, the instantaneous intensity in (40) can therefore be described by

$$T = \tfrac{2W}{1-\gamma}Y \sim \chi^2_{2,\delta} \tag{44}$$

where the non-centrality parameter is

$$\delta = \frac{2sW}{1-\gamma}. \tag{45}$$

The probability of detection is then

$$P_d = 1 - \int_0^\infty F_{\chi^2_{2,\delta}}\left(\frac{2wh_c}{1-\gamma}\right) f_W(w)\,\mathrm{d}w \tag{46}$$

where $F_{\chi^2_{2,\delta}}(t)$ is the CDF of the non-central chi-squared distribution with 2 degrees of freedom and non-centrality parameter δ.

4.1.3 Design SNR using a Gaussian-noise approximation

For moderate to high SNR, a useful approximation to the probability of detection can be obtained by assuming the signal exists in a benign background rather than heavier-tailed clutter. This was the approach taken in [9] to obtain the design SNR for a K-distributed background. Its efficacy was explained in [1, Sect. 7.5.6.1] by the rapid decay of the signal-plus-noise complex-envelope characteristic function compared to that of the noise alone.

The approximation is known to work quite well for Gaussian-fluctuating signals and improves with SNR. It can be accurate near the mean of the signal-plus-noise distribution for a deterministic signal, but has trouble in the upper or lower tails. These expectations are seen in the comparisons shown in Figs. 7–10 between the lines in color and the gray dashed lines. The largest errors occur

for the most extreme clutter ($\gamma = 0.44$) and the highest P_f, which has the lowest detector decision threshold for a given γ. However, the errors are small enough for the approximation to be of use in most practical scenarios.

Also shown in Figs. 7–10 (dots) is an approximation based solely on the change in the decision threshold required to maintain the P_f specification. As described in Sect. 4.3, this can be accurate for the Gaussian-fluctuating signal or, when P_d is near 0.5 for a deterministic signal. Given the simplicity and greater accuracy of the benign-background approximations found in this section, the change-in-threshold approximations are primarily useful when starting with a design SNR estimated from system data in a benign background (see Sect. 4.3).

Approximation for a Gaussian-fluctuating signal: The forward model of P_d for a Gaussian-fluctuating signal in a benign background is simply

$$P_d = e^{-h/(1+s)}, \tag{47}$$

where s is the linear-quantity SNR and h is the normalized-intensity detector decision threshold. The benign-background approximation to P_d simply replaces h with h_c from (38), which increases as the clutter becomes heavier tailed. Solving for SNR and converting to decibels leads to the benign-background approximation to the design SNR for a Gaussian-fluctuating signal in GPD clutter,

$$\mathsf{DT_c} \approx 10 \log_{10} \left[\frac{-h_c}{\log P_d} - 1 \right] \tag{48}$$

$$= 10 \log_{10} \left\{ \frac{1 - \gamma^{-1}}{\log P_d} \left(\frac{1}{P_f^\gamma} - 1 \right) - 1 \right\} \quad \text{[units: dB]}. \tag{49}$$

As seen in Figs. 7 & 8, the approximation is very accurate, with the largest error of 0.03 dB observed when $\gamma = 0.44$, $P_d = 0.5$, and $P_f = 10^{-2}$.

Using the approximation in (49), it can be shown that when P_f is small (so the decision threshold is appropriately large), the increase in the design SNR induced by an order-of-magnitude reduction in P_f is approximately

$$\text{Increase in } \mathsf{DT} \approx 10\gamma + e^{-10\gamma} \cdot 10 \log_{10} \left(1 + \frac{\log(10)}{\gamma_f} \right) \quad \text{[units: dB]}, \tag{50}$$

where γ_f is the decision threshold for an integrated-intensity detector as defined in (58). For a single intensity sample ($M = 1$), $\gamma_f = -\log(P_f)$. When γ is near zero, the latter term dominates and enforces a lower limit that tends to zero as P_f gets smaller. A useful approximation when γ is larger can be obtained from the first term, which implies an order-of-magnitude improvement in P_f has a cost of an increase in DT of 10γ decibels, which can be significant when the clutter is severe. Because this approximation relies on h_c being large, it applies to both deterministic and Gaussian-fluctuating signals and also appears to have reasonable accuracy for an integrated-intensity detector.

Approximation for a deterministic signal: Although the statistical distributions for a deterministic signal in a benign background are not as simple to describe as the Gaussian-fluctuating signal, Albersheim's [20] or Hmam's [21] approximations to the design SNR can be applied to obtain
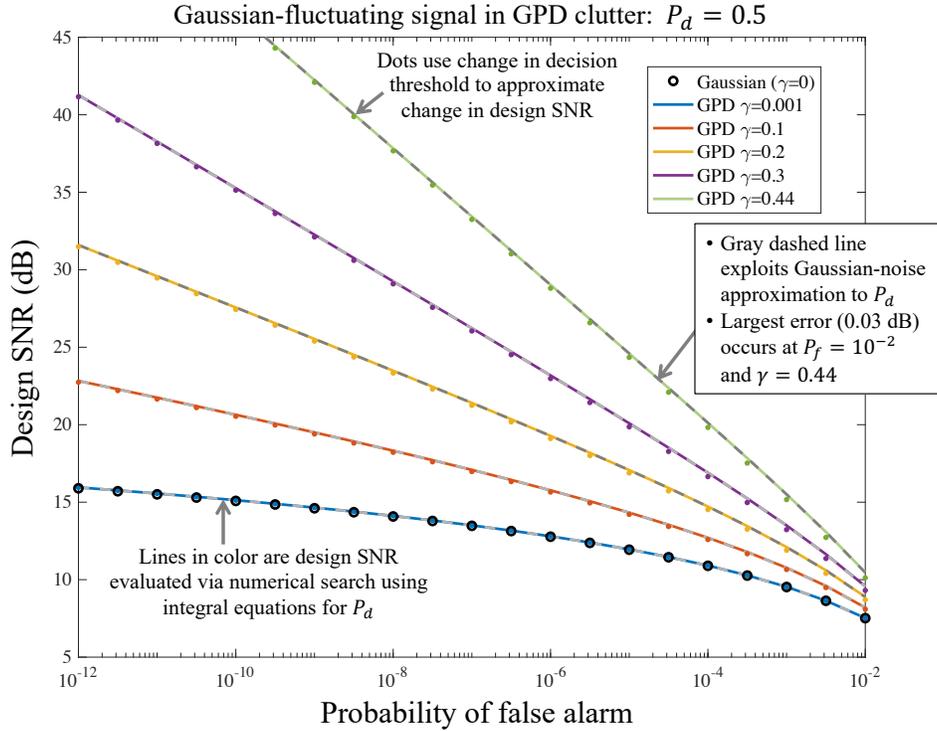
Figure 7: Design SNR and its approximations as a function of the probability of false alarm when $P_d = 0.5$ for a Gaussian-fluctuating signal in GPD clutter with different shape parameters.
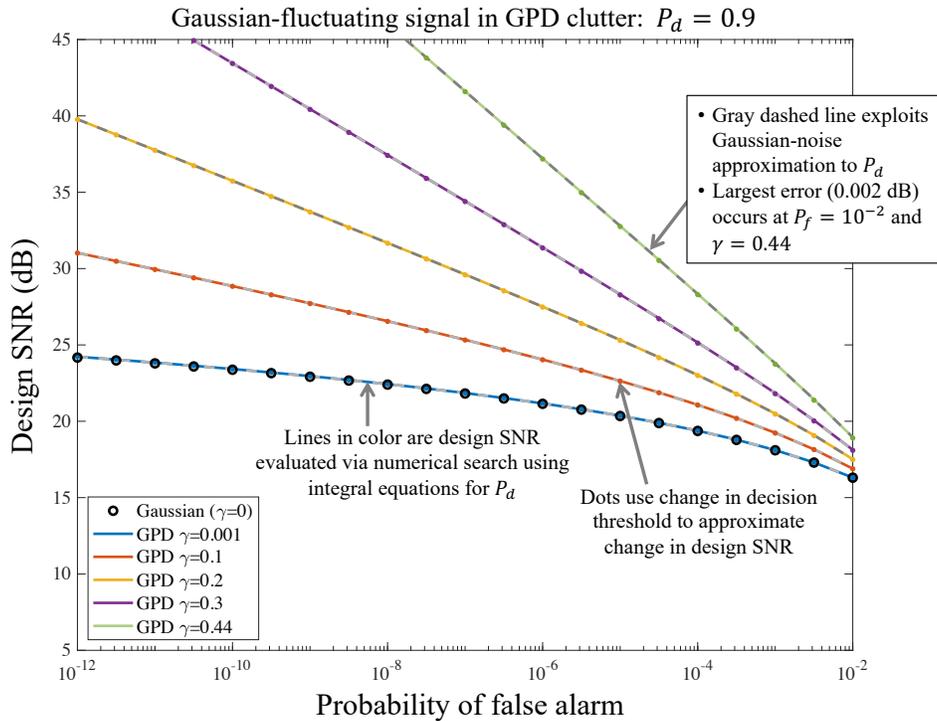
Figure 8: Design SNR and its approximations as a function of the probability of false alarm when $P_d = 0.9$ for a Gaussian-fluctuating signal in GPD clutter with different shape parameters.
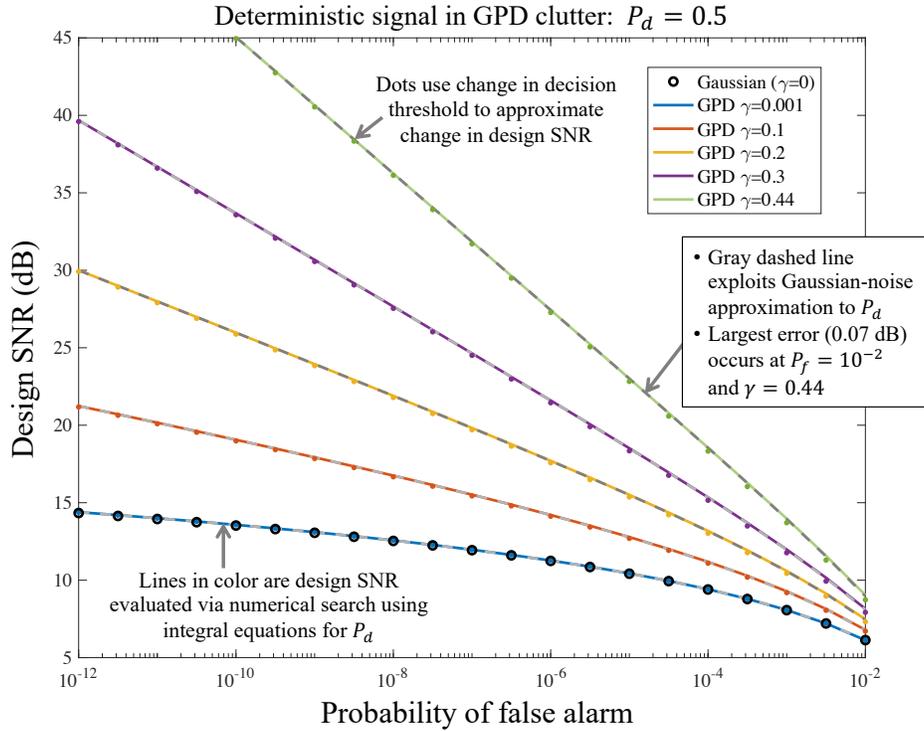
Figure 9: Design SNR and its approximations as a function of the probability of false alarm when $P_d = 0.5$ for a deterministic signal in GPD clutter with different shape parameters.
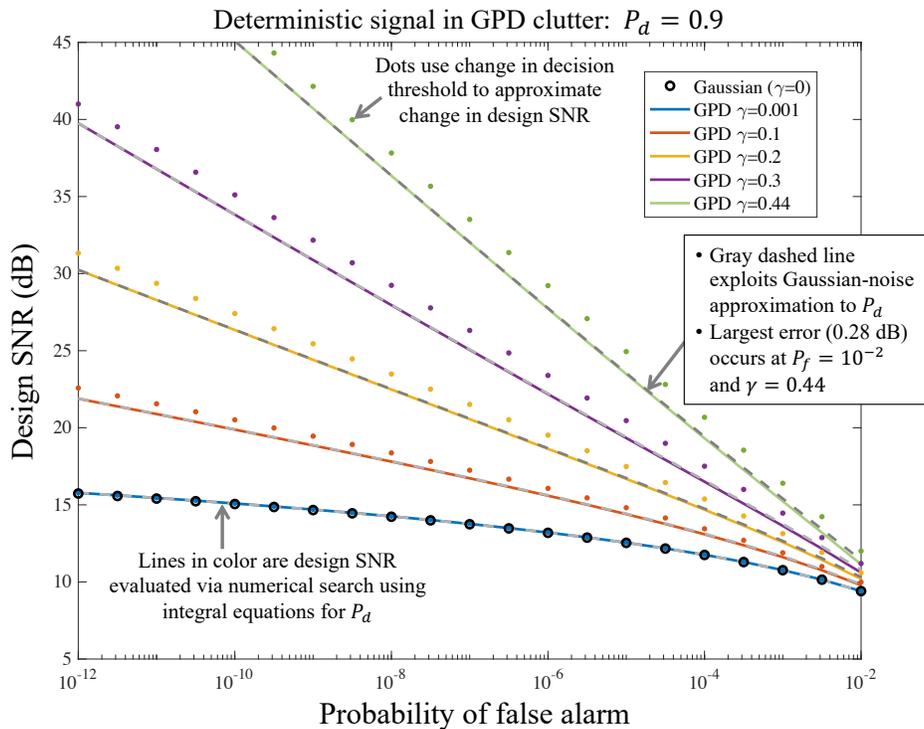
Figure 10: Design SNR and its approximations as a function of the probability of false alarm when $P_d = 0.9$ for a deterministic signal in GPD clutter with different shape parameters.

an accurate approximation for clutter-dominated backgrounds. Hmam's approximation is employed here owing to its formulation in terms of the detector decision threshold. Although application of Albersheim's equation can be performed by setting $P_f = e^{-h_c}$, the greater accuracy of Hmam's equation is preferred.

From [1, pg. 79, eq. 2.87], Hmam's equation for the design SNR (per intensity sample) when integrating M independent intensities is

$$\mathsf{DT} \approx -10\log_{10} M + 10\log_{10}\left[\left(\sqrt{h - \frac{M}{2} + 0.25} - A\right)^2 - \frac{M}{2} + 0.25\right] \quad \text{[units: dB]}, \qquad (51)$$

where $h = \tilde{\gamma}^{-1}(1 - P_f; M)$ for a benign background,

$$A = \text{sign}(0.5 - P_d)\sqrt{0.85616\log_e B} \qquad (52)$$

and

$$B = \frac{0.19}{\sqrt{0.819025 + 1.5206 P_d(0.9998 - P_d)} - 0.905}. \qquad (53)$$

Note that (51) represents the average SNR required in each of the M intensities being integrated.

When $M = 1$, this simplifies to

$$\mathsf{DT} \approx 10\log_{10}\left[\left(\sqrt{h - 0.25} - A\right)^2 - 0.25\right] \quad \text{[units: dB]}, \qquad (54)$$

with $h = -\log P_f$ for a benign background. Using (38) for the detector decision threshold produces

$$\mathsf{DT_c} \approx 10\log_{10}\left[\left(\sqrt{\frac{1-\gamma}{\gamma P_f^\gamma} - \frac{1}{\gamma} + \frac{3}{4}} - A\right)^2 - 0.25\right] \quad \text{[units: dB]} \qquad (55)$$

as an approximation to the design SNR for a deterministic signal in GPD clutter.

With the understanding that the benign-background approximation works best near the mean of the signal-plus-noise distribution for a deterministic signal, it is not surprising that (55) works quite well when $P_d = 0.5$, as can be seen in Fig. 9 where the largest error is 0.07 dB. The accuracy of the approximation degrades as the P_d specification loosens or tightens, as seen in Fig. 10 where $P_d = 0.9$ and the maximum error rises to 0.28 dB. Similar to the Gaussian-fluctuating signal, these errors occur when $\gamma = 0.44$ and $P_f = 10^{-2}$ and the approximation improves as either of these parameters decreases. As such, the approximation in (55) will be useful in most practical scenarios. For situations where the performance specification entails a large value for P_f and a value of P_d not near 0.5, the approximation can be refined numerically using the integral equation in (46).

Examining (51), it can be seen that when h is large, the effect of A and therefore P_d is muted. When $P_d = 0.5$, $A = 0$ and it decreases to -0.9 when $P_d = 0.9$ and -1.64 when $P_d = 0.99$. In contrast, the radical term in (51) will be significantly larger when P_f is small and γ increases away from zero (e.g., it is over 20 when $P_f = 10^{-6}$ and $\gamma = 0.44$). The impact of this on the design SNR can be seen in Figs. 9 & 10 where there is little difference between the $P_d = 0.5$ and 0.9 specifications when P_f is small and γ is large. Conversely, this implies a small change in SNR at

these operating points can lead to a large change in P_d (i.e., the transition curve is very sharp). This sensitivity can be explained by considering the complex envelope in a benign background, $Ae^{j\psi} + \tilde{V}$, which can be obtained from (39) by setting $W = 1$. Given perfect normalization and a deterministic signal, $A = \sqrt{s}$ and $E\left[|\tilde{V}|^2\right] = 1$. The joint PDF of the real and imaginary parts is a symmetric Gaussian bell-shaped curve centered at $\sqrt{s}e^{j\psi}$. Because the deterministic signal only affects the mean, the variance of the real or imaginary components arises solely from the noise, which is $\text{Var}\{\text{Real}\{\tilde{V}\}\} = \text{Var}\{\text{Imag}\{\tilde{V}\}\} = 1/2$ after perfect normalization. When h is large, the SNR required to achieve $P_d = 0.5$ is also large (from Hmam's equation in (54), it is $\mathsf{DT} \approx 10\log_{10}(h - 0.5)$). As such, it takes very little change in SNR to move the complex-envelope PDF from being mostly within the circle representing the decision threshold (i.e., one with radius $\sqrt{h}$) to being mostly outside of it. An example of this transition, where a change of less than one decibel transforms a low P_d into a high one, is shown in Fig. 11 for $\gamma = 0.3$ and $P_f = 10^{-8}$. This effect is not observed in Figs. 7 & 8 for the Gaussian-fluctuating signal because the spread of the complex-envelope PDF increases with SNR. Although scenarios exhibiting the sensitivity seen in Fig. 11 are unlikely (owing to the use of a very large decision threshold), it strengthens the argument for not using the deterministic signal at high SNR.
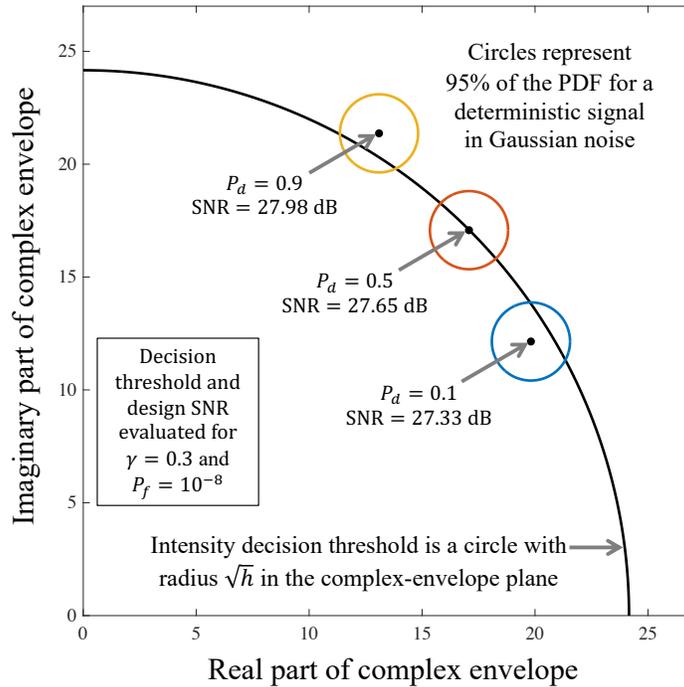
Figure 11: Representation of the complex-envelope PDF (95% coverage) for a deterministic signal in Gaussian noise at the design SNRs required to meet $P_d = 0.1$, 0.5, and 0.9 when $P_f = 10^{-8}$ in GPD clutter with shape parameter $\gamma = 0.3$. The signal phases are chosen arbitrarily to separate the distributions on the figure.

4.2 Integrated-intensity detectors

When the effects of propagation and scattering from the object of interest (OOI) spread an active-sonar echo in time, it is common to apply post-matched-filter integration (PMFI) to recoup some of the spreading losses [1, Sect. 8.8]. The resulting detector decision statistic can be modeled by the sum

$$U = \sum_{m=1}^{M} Y_m \quad \text{[units: power]}, \tag{56}$$

where $Y_1, \ldots, Y_M$ are independent instantaneous intensity samples and M is approximately the product of the integration time and the signal bandwidth. This decision statistic can also be used to represent track-level detection when summing intensities over M observations (although it assumes perfect measurement-to-track assignment).

In a benign background, the standard signals in an integrated-intensity detector are analyzed using chi-squared and non-central-chi-squared distributions where the degrees-of-freedom parameter increases from two to $2M$. Analysis of this detector in clutter-dominated backgrounds, however, is quite difficult. Although some approximations are accessible [1, Sect. 7.5.6.4], they do not provide the accuracy required when evaluating the design SNR from a detection operating point.[7] Numerical approaches for evaluating P_f and P_d requiring the equivalent of a two-dimensional numerical integral are presented in Sects. 4.2.2 & 4.2.4, respectively. They are employed here to assess the accuracy of the benign-background approximations to the design SNR in Sect. 4.2.1, which have a similar efficacy to what was seen in Sect. 4.1.3 for a single intensity sample.

The techniques found in this section are employed in Figs. 12 and 13, which extend the example found in [1, pg. 81, Fig. 2.16] to account for a clutter-dominated background when $P_d = 0.5$ and 0.9, respectively, with $P_f = 10^{-4}$. These results illustrate, as might be expected, that the increase in the design SNR with clutter tail heaviness persists as M increases. The dots in the figures are the precise result obtained through numerical inversion using the techniques described in Sects. 4.2.2 & 4.2.4. The lines exploit an empirical approximation to the decision threshold for $M \geq 2$ that is described in Sect. 4.2.3 (see App. C.4 for a MATLAB® implementation) and the benign-background approximations to P_d described in Sect. 4.2.1. The combined approximations are very simple to evaluate and yet have maximum absolute error less than $0.1\,\text{dB}$ in the cases shown in Figs. 12 and 13, where their efficacy is readily apparent.

4.2.1 Design SNR using a Gaussian-noise approximation

The approach presented in Sect. 4.1.3 for approximating P_d in a clutter-dominated background by that obtained for a benign background can also be applied to a detector integrating independent and identically distributed intensities. In the benign background, the decision statistics become proportional to central or non-central chi-squared random variables with $2M$ degrees of freedom. Similar to the single-intensity case presented earlier, an exact solution can be obtained for the Gaussian-fluctuating signal and Hmam's equation can be used for the deterministic signal.

The design SNR for a Gaussian-fluctuating signal in a benign background when summing M

---

[7]In addition to accuracy issues, approximations for the GPD obtained by moment matching up to the kth intensity moment restrict γ to be $< 1/k$.
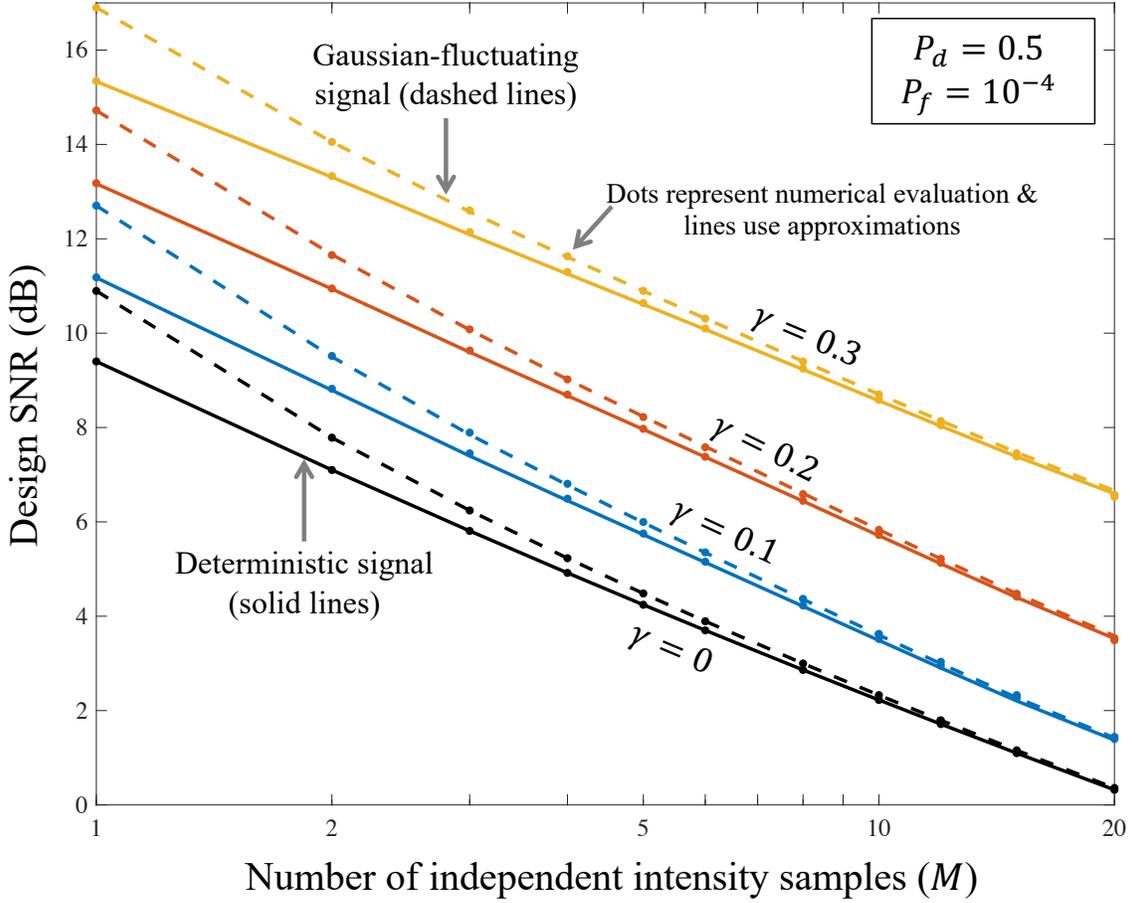
Figure 12: Design SNR as a function of the number of independent intensity samples being integrated (M) for for a deterministic signal in GPD clutter with different shape parameters when $P_d = 0.5$ and $P_f = 10^{-4}$.

independent and identically distributed instantaneous-intensity samples is

$$\mathsf{DT_g} = 10 \log_{10}\left[\frac{\gamma_f}{\gamma_d} - 1\right] \qquad \text{[units: dB]}, \tag{57}$$

from [1, pg. 77, eq. 2.81] where

$$\gamma_f = \tilde{\gamma}^{-1}(1 - P_f; M) \quad \text{and} \quad \gamma_d = \tilde{\gamma}^{-1}(1 - P_d; M) \tag{58}$$

are formed from the functional inverse of the normalized incomplete gamma function.[8] When $M = 1$, these simplify to $\gamma_f = -\log P_f$ and $\gamma_d = -\log P_d$.

Noting that the detector decision threshold in a benign background is $h_g = \gamma_f$, the approximation to the design SNR in a clutter-dominated background is simply

$$\mathsf{DT_c} = 10 \log_{10}\left[\frac{h_c}{\gamma_d} - 1\right] \qquad \text{[units: dB]}, \tag{59}$$

---

[8]See [1, pg. 77, eq. 2.80 and pg. 293] for more details on the normalized incomplete gamma function and its functional inverse. The functional inverse can be evaluated in MATLAB® using the functions `gammaincinv(1-Pd,M)` or `gaminv(1-Pd,M,1)`, for example, for γ_d. A useful approximation is also presented in (81).
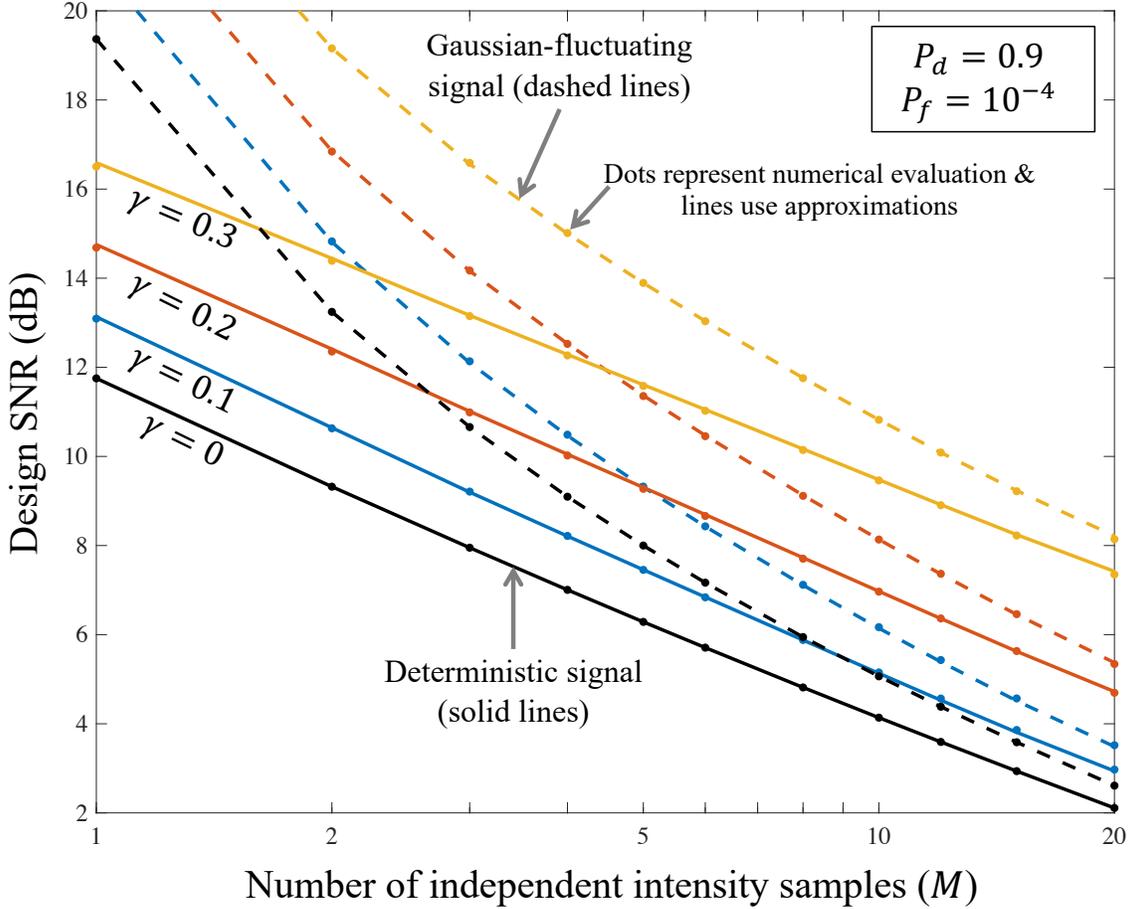
Figure 13: Design SNR as a function of the number of independent intensity samples being integrated (M) for for a deterministic signal in GPD clutter with different shape parameters when $P_d = 0.9$ and $P_f = 10^{-4}$.

where h_c is the linear-intensity decision threshold applied to U in (56).

Using Hmam's equation [21] as found in (51) with the appropriate decision threshold produces the approximation to the design SNR for a deterministic signal in GPD clutter when summing M independent intensities,

$$\mathsf{DT_c} \approx -10\log_{10} M + 10\log_{10}\left[\left(\sqrt{h_c - \frac{M}{2} + 0.25} - A\right)^2 - \frac{M}{2} + 0.25\right] \quad \text{[units: dB]}, \qquad (60)$$

where the dependence on P_d enters through A as shown in (52) & (53). Note that both (59) and (60) describe the SNR in a single intensity sample.

The simplicity of these two approximations belies the complexity encountered when applying them, primarily owing to the difficulty in accurately obtaining the detector decision threshold as a function of P_f. The integral equations of the forward model relating h_c to P_f found in Sect. 4.2.2 must be evaluated numerically to obtain h_c as a function of the P_f specification. This functional inversion was avoided in the following assessment of the accuracy of (59) and (60) by using the forward model to obtain a number of pairs (h_c, P_f) for which P_f was on the interval $[10^{-6}, 10^{-2}]$

with γ taking values every 0.05 from 0.05 to 0.45 and M taking even values from 2 to 20. The values of h_c were then used to approximate the design SNR using (59) or (60), which then initialized a search to find the precise value computed from the integral equations for P_d described in Sect. 4.2.4. The signal was modeled as either deterministic ($\rho = 0$ in the Rician model used in Sect. 4.2.4) or Gaussian-fluctuating ($\rho = 1$) and the design values of P_d were 0.5 or 0.9. Plots of the design SNR and its approximation as a function of P_f were similar to those seen Figs. 7–10.

To summarize the errors, the design SNRs were interpolated onto a common set of P_f values in the range $[10^{-6}, 10^{-2}]$. The maximum absolute errors observed in the analysis are then shown in Fig. 14 (blue lines) as a function P_f. When using the exact threshold, the worst error occurred when $\gamma = 0.45$ and $P_f = 10^{-2}$. Although the errors are larger than those typically seen in Sect. 4.1.3 for a single intensity, restricting the analysis to cases for which $\gamma \leq 0.4$ and $P_f \leq 10^{-3}$ keeps the error under 0.1 dB.
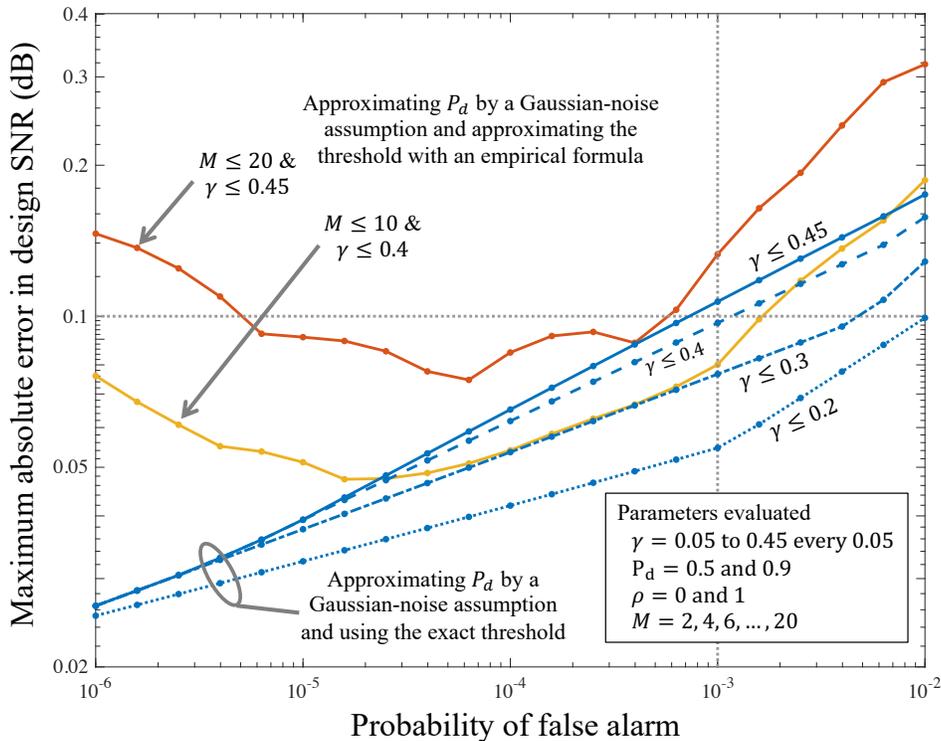
Figure 14: Maximum absolute error in the design SNR predicted using a benign-background assumption for P_d over the test cases evaluated with an exact detector decision threshold (blue lines) or an approximate threshold.

The other lines in Fig. 14 characterize the maximum error observed when incorporating an additional approximation for the detector decision threshold (h) using the empirical formula described in Sect. 4.2.3. Approximating the decision threshold clearly increases the maximum error to a few tenths of a decibel, although it remains below 0.1 dB for $P_f \in [10^{-6}, 10^{-3}]$, $\gamma \leq 0.4$, and $M \leq 10$. Given the simplicity of the combined approximation, this level of accuracy will generally be a worthwhile compromise when compared to the effort required to achieve the more accurate result.

4.2.2 Integral equations for P_f

In general, the PDF and CDF of the sum of scaled F-distributed random variables (for which the GPD is a special case) can be described using complicated functions [22] that might be difficult to evaluate numerically (e.g., multiple infinite summations, potentially with alternating signs). The approach taken here is to describe P_f through a single integral when $M = 2$, a double integral when $M = 3$, and an integral of a function containing an integral for $M > 3$. Evaluation of P_f can then be accomplished with no more than the equivalent of a two-dimensional numerical integral, irrespective of M.

The PDF of the sum of independent random variables is the convolution of the PDFs of the summands. The CDF of the sum can similarly be shown to be the convolution between the CDF of one of the random variables and the PDFs of the others. For example, if $U = X + Y$ and X and Y are independent, then the CDF of U is

$$F_U(u) = \int F_X(u - y) f_Y(y) \, dy = F_X(u) * f_Y(u), \tag{61}$$

where $*$ is the convolution operator. When summing two independent and identically distributed GPD random variables with shape γ and scale λ, this results in

$$P_f = 1 - F_U(h) = \frac{1}{\left(1 + \frac{\gamma h}{\lambda}\right)^{\gamma^{-1}}} \left[1 + \int_0^{\gamma h/\lambda} \frac{\gamma^{-1}}{\left(1 - \frac{y}{1 + \gamma h/\lambda}\right)^{\gamma^{-1}} (1 + y)^{\gamma^{-1}+1}} \, dy \right], \tag{62}$$

which is straightforward to evaluate numerically.

The CDF of the sum $U = X + Y + Z$ of three independent random variables can be obtained by iterating (61),

$$F_U(u) = F_X(u) * f_Y(u) * f_Z(u) = \iint F_X(u - y - z) f_Y(y) f_Z(z) \, dy \, dz. \tag{63}$$

Note that for non-negative random variables, the integral over y is restricted to the interval $(0, u)$ and the integral on z to $(0, u - y)$ in order for the arguments of the PDFs and CDF to be ≥ 0.

The convolution approach, which requires an $M - 1$ dimensional integral when summing M independent intensities, is clearly not reasonable for $M > 3$. In these scenarios P_f can be obtained using characteristic functions. The characteristic function of a GPD random variable is

$$\Phi_Y(\omega) = E\left[e^{j\omega Y}\right] = \int_0^\infty e^{j\omega y} f_Y(y) \, dy$$

$$= \int_0^\infty \frac{e^{j\omega y}}{\lambda \left(1 + \frac{\gamma}{\lambda} y\right)^{\gamma^{-1}+1}} \, dy \tag{64}$$

$$= j \int_0^\infty \frac{e^{-\omega r}}{\lambda (1 + jr\gamma/\lambda)^{\gamma^{-1}+1}} \, dr \tag{65}$$

using the PDF in (2). The change of variables from y to $r = y/j$ (with $j = \sqrt{-1}$) in (65) converts the complex envelope in the numerator to an exponential decay, which simplifies numerical evaluation.

Recalling that convolution in one domain of a Fourier transform equates to a product in the other domain, it can be seen that summing M independent and identically distributed GPD intensity samples as done in (56) equates to raising the characteristic function in (65) to the Mth power,

$$\Phi_U(w) = [\Phi_Y(w)]^M. \tag{66}$$

This form is particularly useful when the intensity sum is formed using correlated observations and leads to a non-integer value of M.

The exceedance distribution function (EDF) of the random variable U, which is one minus the CDF, is then obtained from its characteristic function using [1, pg. 653, eq. 9.101],

$$\Pr\{U > u\} = 1 - F_U(u) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Imag}\left\{ \frac{\Phi_U(\omega)}{\omega} e^{-j\omega u} \right\} \, \mathrm{d}\omega. \tag{67}$$

The combination of (67) and (65) has the computational requirements of a two-dimensional integral, as is the case for $M = 3$ in (63). This demanding computational requirement in the forward model of P_f presents the largest impediment to implementing the detector at a given P_f and subsequently finding the design SNR achieving the P_d specification.

It is also possible to exploit a discrete-Fourier-transform (DFT) approximation to the integral in (67), as done in [23] (also see [1, Sect. 9.2.7.2 & App. 9.A]). This approach is most likely less computationally intensive than using a generic numerical integration routine. However, the generic approach avoids the need to check that each of numerous variations evaluated here (across γ and M) did not suffer from the aliasing at large values that can affect the DFT-based approach. When evaluating a specific scenario or if smaller values of P_f are of interest, the DFT-based routine is recommended or [24] when P_f is extremely small.

MATLAB® code for evaluating P_f using the techniques described in this section can be found in App. C.3.

4.2.3   Empirical inversion of P_f to find the decision threshold

Given the computational effort required to invert the relationship between the decision threshold and the probability of false alarm for an integrated-intensity detector in heavy-tailed noise, a simple approximation to the inversion was obtained empirically through a least-squared-error (LSE) fit to a linear model based on the modified parameters: $x_1 = \Phi^{-1}(1 - P_f)$, $x_2 = \gamma$, and $x_3 = \log M$. The linear model was applied to the function

$$y = \log\left(\frac{h_M}{h_1} - 1 \right), \tag{68}$$

where h_M is the linear intensity threshold when summing M independent intensities. This mapping compresses the increase in the threshold from $M = 1$ when it is large and emphasizes it for small M. The modified parameters (x_1, x_2, x_3) and a sufficient number of their powers and cross terms were used in a Maclaurin-series expansion of the multi-dimensional function to achieve an estimate with an acceptable maximum absolute error over a useful range of input parameters. For some values of P_f and M, the function was not monotonic in γ, so it was necessary to apply the LSE model to four intervals spanning $\gamma \in [0, 0.45]$. This allowed keeping the maximum absolute error over the cases evaluated below $0.15\,\mathrm{dB}$ while accounting for rounding the coefficients to four significant digits. MATLAB® code implementing the approximation is found in App. C.4.

The range of validity for the approximation is:

$$P_f \in [10^{-6}, 10^{-2}], \quad \gamma \in [0, 0.45], \quad \text{and} \quad M \in [2, 20]. \tag{69}$$

The LSE analysis was performed within these intervals for 21 equally-spaced values in x_1, for $M = 2, 3, 4, 5$ and even numbers up to $M = 20$, and for 19 values of γ every 0.025. The accuracy of the approximation is illustrated in Fig. 15 for $\gamma = 0.2$ where the maximum absolute error is less than 0.1 dB, even for the values of $M = 7$, 9, 13, and 19, which were not part of the LSE fitting. For a scenario where the decibel change in the decision threshold is indicative of the increase in the design SNR (see Sect. 4.3), the error in the approximation is representative of the additional error in DT. The smoothness of the function as the modified parameters vary between the samples used in the LSE analysis (e.g., see Fig. 15 for variation with P_f) suggests the errors on the interior of the above regions will be similar in size to those at the points evaluated. Use of the approximation outside of the intervals noted in (69) is not recommended.
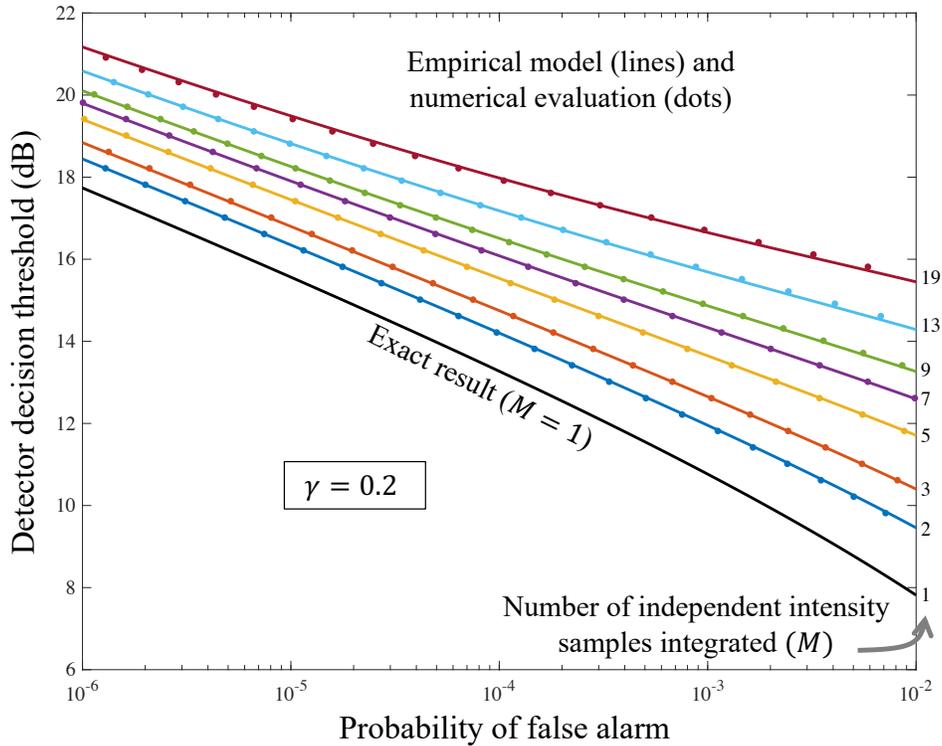
Figure 15: Detector decision threshold as a function of the probability of false alarm when integrating M independent intensity samples dominated by GPD clutter with shape parameter $\gamma = 0.2$.

4.2.4 Integral equation for P_d

Extending the definition of Y in (39) to represent a Rician signal in GPD clutter leads to its characterization as the scale of a non-central chi-squared-distributed random variable,

$$T = \frac{2}{\mu + \rho s} Y \sim \chi^2_{2,\delta}, \tag{70}$$

when conditioned on W, where s is the total SNR, ρ is the fraction of random signal power, $\mu = (1 - \gamma)/W$, and the non-centrality parameter is

$$\delta = \frac{2(1-\rho)s}{\mu + \rho s}. \tag{71}$$

Setting $\rho = 1$ yields the Gaussian-fluctuating signal, whereas $\rho = 0$ produces a deterministic signal. Letting $s = 0$ produces the noise-only distribution.

Using the characteristic function of the non-central chi-squared distribution from [1, pg. 297] for T, it is straightforward to show that the characteristic function of Y conditioned on W is

$$\Phi_{Y|W}(\omega|W) = \frac{1}{1 - j\omega\left(\frac{1-\gamma}{W} + \rho s\right)} \exp\left\{\frac{j\omega(1-\rho)s}{1 - j\omega\left(\frac{1-\gamma}{W} + \rho s\right)}\right\}. \tag{72}$$

Removing the conditioning on W then produces the characteristic function of Y,

$$\Phi_Y(\omega) = E\left[e^{j\omega Y}\right] = \int_0^\infty \Phi_{Y|W}(\omega|W = r) f_W(r) \, dr, \tag{73}$$

where the PDF of W is as shown in (43) and r is used as a variable of integration to avoid confusion between the lower-case w and ω, the latter of which is the argument of the characteristic function.

The probability of detection is then obtained by raising $\Phi_Y(\omega)$ to the Mth power, as done in (66), and then using it in (67). Although a single evaluation of P_d is not prohibitive, the effort required to solve for the design SNR generally is. Fortunately, the approximations presented in Sect. 4.2.1 are accurate enough in most cases to not require any direct evaluation of P_d in practice.

4.3 Adjusting a measured design SNR to account for clutter

In some applications the design SNR is obtained through the analysis of OOI measurements from a particular system operating in a benign background and it is desired to adjust it to account for a background dominated by clutter. Application of the techniques described in Sects. 4.1 & 4.2 requires the (P_d, P_f) specification corresponding to the system's detector decision threshold and the design-SNR analysis. If the analysis was performed using the minimum-detectable-level (MDL) criteria, then $P_d = 0.5$. However, the corresponding value of P_f is often not known (e.g., see the auditory detection approach to defining a design SNR in [25, Sect. 4.1.3]). Assuming statistical models for the signal and noise, as was done in [26], provides a means to obtain an equivalent P_f given P_d and the design SNR. It is then straightforward to use the previously described techniques to obtain the design SNR in a clutter-dominated background. Although this approach is straightforward, it still requires accuracy in the models of the noise background and estimation of the GPD shape parameter.

Given the ubiquity of clutter compared to the dearth of OOI measurements, an alternative approximate approach can be obtained by assessing how much the detector decision threshold must be increased in a clutter-dominated background in order to maintain the same probability of false alarm as that observed in the original system configuration in the background conditions for which it was designed. The impact of the clutter is dictated by the data in terms of how the new detector decision threshold affects P_d. As seen here, the decibel change in the detector decision threshold is

also a useful approximation to the clutter-induced increase in the design SNR for an MDL scenario under both the deterministic and Gaussian-fluctuating signal models.

The analysis presented in this section exploits an approximation to detection threshold that decomposes it into the sum of components related to the P_f and P_d specifications. For example, from (57), the detection threshold for an integrated-intensity detector with a Gaussian-fluctuating signal in Gaussian noise is

$$\mathsf{DT}_\mathrm{g} = 10\log_{10}\left[\frac{\gamma_f}{\gamma_d} - 1\right] \approx 10\log_{10}(\gamma_f) - 10\log_{10}(\gamma_d) \quad \text{[units: dB]}. \tag{74}$$

The approximation in (74) assumes that the detector decision threshold ($h_\mathrm{g} = \gamma_f$) is large compared to γ_d, which equates to $P_d \gg P_f$. For a single intensity sample, this simplifies to

$$\mathsf{DT}_\mathrm{g} \approx 10\log_{10}(-\log P_f) - 10\log_{10}(-\log P_d) \quad \text{[units: dB]}. \tag{75}$$

The first term in the approximation is precisely the logarithmic-quantity decision threshold for a detector with perfect normalization. When the approximation is accurate, a change in P_f is reflected in DT_g as an equivalent decibel change in the decision threshold.[9] Given that heavy-tailed background distributions require an increase in the detector decision threshold to maintain P_f, the approximation also applies to clutter-dominated scenarios.

4.3.1 Approximation for a Gaussian-fluctuating signal

Using (57) to solve for γ_d, letting $h_\mathrm{g} = \gamma_f$ be the decision threshold for a benign background, and using these in (59), it can be seen that the design SNR in a clutter-dominated background is

$$\mathsf{DT}_\mathrm{c} = 10\log_{10}\big(\bar{S}_\mathrm{g} + 1\big) + \Delta_h \tag{76}$$

$$\approx \mathsf{DT}_\mathrm{g} + \Delta_h \quad \text{[units: dB]} \tag{77}$$

where $\bar{S}_\mathrm{g} = 10^{\mathsf{DT}_\mathrm{g}/10}$ [unitless] is the linear-quantity design SNR for a benign background and

$$\Delta_h = 10\log_{10}\left(\frac{h_\mathrm{c}}{h_\mathrm{g}}\right) \quad \text{[units: dB]} \tag{78}$$

is the increase in the decision threshold (in decibels) required to maintain the same P_f when clutter dominates. The approximation shown in (77) will be accurate when the design SNR is large ($\bar{S}_\mathrm{g} \gg 1$). When the design SNR is small, the approximation implied by (59) is unlikely to be accurate. The efficacy of this approach can be seen in Figs. 7 & 8 (dots in color) for a wide range of scenarios.

4.3.2 Approximation for a deterministic signal

Consider Hmam's equation [21] from (51) in Sect. 4.1.3, which relates the decision threshold, P_d, and M to DT. Using the decision thresholds h_c and h_g, the design SNR for a deterministic signal in clutter can be approximated as that for the benign background plus the decibel difference

---

[9]The author thanks G. Wadsworth for pointing out this phenomena.

in the decision thresholds (Δ_h),

$$\mathsf{DT}_\mathrm{c} \approx \mathsf{DT}_\mathrm{g} + \Delta_h + 10 \log_{10} \left\{ \frac{\left[\sqrt{1 - \frac{2M-1}{4h_\mathrm{c}}} - \frac{A}{\sqrt{h_\mathrm{c}}} \right]^2 - \frac{2M-1}{4h_\mathrm{c}}}{\left[\sqrt{1 - \frac{2M-1}{4h_\mathrm{g}}} - \frac{A}{\sqrt{h_\mathrm{g}}} \right]^2 - \frac{2M-1}{4h_\mathrm{g}}} \right\} \tag{79}$$

$$\approx \mathsf{DT}_\mathrm{g} + \Delta_h \quad \text{[units: dB]}. \tag{80}$$

The approximation in (80) will be valid when the decision thresholds are large compared to $(2M - 1)/4$ and P_d is near 0.5 so A from (52) is near zero.

Noting that the decision threshold increases with the clutter severity, it suffices to consider when h_g will be large compared to $(2M - 1)/4$. A surprisingly accurate approximation to $h_\mathrm{g} = \gamma_f = \tilde{\gamma}^{-1}(1 - P_f; M)$ can be found in

$$h_\mathrm{g} \approx M - 1 + \left(\sqrt{M} - 1 \right) \Phi^{-1}(1 - P_f) - \log(P_f) \quad \text{[unitless: normalized power]}, \tag{81}$$

which was obtained by altering the threshold obtained from a Gaussian approximation to the gamma distribution (which is accurate at large M) to be precise when $M = 1$.[10] This illustrates that h_g increases with M and is always larger than its value at $M = 1$ (with $P_f < 0.5$), which is $-\log P_f$. However, $(2M - 1)/(4h_\mathrm{g})$ is an increasing function of M. As such the approximation in (80) may not be accurate when M is large. This is essentially the same restriction as that for the Gaussian-fluctuating signal, where the SNR (per intensity sample) must be large yet it decreases as M increases.

Some of these expectations can be seen in Figs. 9 & 10 (dots in color): the approximation is very accurate when $P_d = 0.5$ for $M = 1$ and less so when P_d increases to 0.9.

5 J-divergence detection currency for signals in GPD clutter

An alternative performance measure to the traditional (P_d, P_f) metrics more suitable to combination across multiple measurements can be found in the J-divergence detection currency (JDC). As described in [10], JDC is the logarithmic quantity of the J-divergence between the PDFs of the detector decision statistic under the noise-only and signal-plus-noise hypotheses, respectively, $f_0(y)$ and $f_1(y)$.

The linear-quantity J-divergence [27] is

$$J = \int_{-\infty}^{\infty} [f_1(y) - f_0(y)] \log \left[\frac{f_1(y)}{f_0(y)} \right] \, \mathrm{d}y, \tag{83}$$

---

[10]The maximum absolute error of the approximation in (81) is less than $0.2\,\mathrm{dB}$ for $P_f \in [10^{-16}, 10^{-2}]$ and $M \leq 100$, with the worst errors occurring for the smallest values of P_f. It can be improved to be below $0.1\,\mathrm{dB}$ by including an additional multiplicative term as follows,

$$h_\mathrm{g} \approx \left[M - 1 + \left(\sqrt{M} - 1 \right) \phi_f - \log(P_f) \right] 10^{-(1-1/M)\left(\phi_f - 3.2 - 0.8/\sqrt{M} \right)/35}, \tag{82}$$

where $\phi_f = \Phi^{-1}(1 - P_f)$.

which is also the difference in the means of the log-likelihood ratio under the two hypotheses. For a Gaussian-fluctuating signal in a benign background, the J-divergence is simply $J = s^2/(1+s)$ where s [unitless] is the linear-quantity SNR. From this example, and more generally when working with intensities, the linear-quantity J-divergence is seen to be similar to a ratio of squared intensities. As such, detection currency is obtained by converting J to decibels according to

$$\mathrm{JDC} = 5 \log_{10} J \quad [\text{units: dB}]. \tag{84}$$

The PDF under the noise-only hypothesis for a single GPD intensity sample is simply

$$f_0(y) = \frac{1}{\lambda\left(1 + \frac{\gamma}{\lambda}y\right)^{\gamma^{-1}+1}} \ \text{ for } \ y \geq 0 \tag{85}$$

from (2). However, presence of a signal leads to an integral definition similar to those obtained for P_d in Sect. 4.1.2 for deterministic and Gaussian-fluctuating signals. These are generalized to account for a Rician signal in Sect. 5.1 and a numerical integration is prescribed in Sect. 5.2 for evaluating detection currency. The Rician signal is a combination of the deterministic and Gaussian-fluctuating signals. In addition to the SNR, it is defined by a fraction $\rho \in [0,1]$ describing the proportion of the Gaussian-fluctuating signal component. Setting $\rho = 0$ produces a deterministic signal, whereas $\rho = 1$ results in a Gaussian-fluctuating signal. It is particularly useful when propagation is expected to induce some randomness in an echo, but not enough to produce a Gaussian-fluctuating signal.

An example of the J-divergence detection currency obtained for an 18-dB SNR Rician signal in GPD clutter is shown in Fig. 16 as a function of γ for values of ρ ranging from a deterministic signal to a Gaussian-fluctuating one. In each case, increasing tail heaviness of the clutter (i.e., increasing γ) leads to a loss in performance. As described in [10], 5 dB of detection currency represents a minimum detectable level (MDL), whereas 10 dB provides a high-quality operating point. The results in Fig. 16 demonstrate how a very-heavy-tailed clutter background can reduce detection performance from close to a high-quality operating point to a low-quality one.

The J-divergence detection currency quantifies the difference between the PDFs of the decision statistic at all values. When combining detection currency across multiple measurements (where the linear quantities simply add for independent measurements), it is often desirable to account for the thresholding operation commonly imposed on individual measurements. The numerical integration described in Sect. 5.2 can be modified to handle the thresholding. However, approximations to the distribution of the excess over the threshold enable a usefully accurate and straightforward analytical evaluation of JDC. The J-divergence after thresholding and its approximation for signals in GPD clutter are presented in Sect. 5.3.

5.1 Intensity PDF for a Rician signal in GPD clutter

In contrast to the design-SNR analysis of Sect. 4, which was limited to deterministic and Gaussian-fluctuating signals, the approximation employed in Sect. 5.3.1 to obtain JDC is based on the first two moments of the intensity and is therefore straightforward to apply to the more general case of a Rician signal in GPD clutter. Similar to the development in Sects. 4.1.2 & 4.2.4, the intensity for this scenario can be characterized by describing Y from (39) conditioned on W as being proportional to a non-central-chi-squared-distributed random variable. Using (70) and (71),
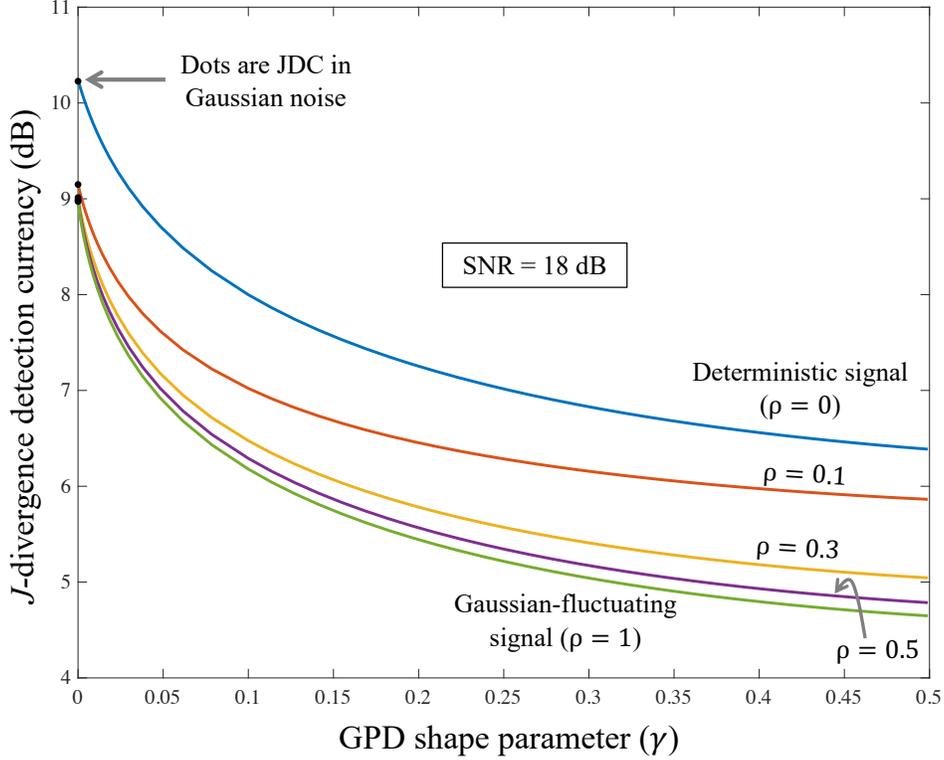
Figure 16: J-divergence detection currency (JDC) for various Rician signals in GPD clutter as a function of the GPD shape parameter (γ) for a fixed SNR of $18\,\mathrm{dB}$.

the PDF of Y for a Rician signal in GPD clutter is obtained by removing the conditioning on W,

$$f_1(y) = \int_0^\infty \frac{2w}{1 - \gamma + \rho sw} \cdot f_{\chi^2_{2,\delta}}\left(\frac{2wy}{1 - \gamma + \rho sw}\right) f_W(w)\,\mathrm{d}w, \tag{86}$$

where $f_{\chi^2_{2,\delta}}(\cdot)$ is the non-central chi-squared PDF with two degrees of freedom and non-centrality parameter δ.

This illustrates that evaluation of J-divergence for a Rician signal in a GPD background requires computational effort equivalent to a two-dimensional integral. The following section describes how this can be accomplished efficiently.

5.2 Numerical evaluation of J-divergence

If direct evaluation of the PDF of a decision statistic is computationally intensive, as can be the case when the signal or noise is heavy tailed, it is prudent to implement an economical numerical approximation to the J-divergence integral in (83). In most cases, the integrand in (83) is smooth enough to use a trapezoidal-rule [28, Sect. 7.1] approximation. Owing to the potential problems when evaluating the PDFs at the upper and lower limits, however, the integral should be restricted to the interval $[y_a, y_b]$ where $-\infty < y_a < y_b < \infty$ are chosen to limit the approximation error (applying this to an envelope or intensity decision statistic implies $y_a > 0$). Once the interval is defined, the final requirement is to determine how frequently to sample the integrand.

Hard-limiting the decision statistic to be on the interval $[y_a, y_b]$ results in a mixed continuous and discrete random variable. Because the hard-limiting discards information, the J-divergence of the hard-limited decision statistic acts as a lower bound on the unlimited decision statistic. It is straightforward to show that the hard limiting results in

$$J \geq (p_a - q_a) \log\left(\frac{p_a}{q_a}\right) + \int_{y_a}^{y_b} [f_1(y) - f_0(y)] \log\left[\frac{f_1(y)}{f_0(y)}\right] \, \mathrm{d}t + (q_b - p_b) \log\left(\frac{q_b}{p_b}\right), \tag{87}$$

where

$$p_a = F_0(y_a), \ \ q_a = F_1(y_a), \ \ p_b = 1 - F_0(y_b), \ \ \text{and} \ \ q_b = 1 - F_1(y_b) \tag{88}$$

define the probabilities of observing $Y \leq y_a$ or $Y > y_b$ under the two hypotheses from the CDFs $F_0(y)$ and $F_1(y)$. If the presence of signal has the effect of moving the distribution to the right (i.e., the CDFs satisfy $F_1(y) \leq F_0(y)$ for all y), then $q_a < p_a$ and $q_b > p_b$. By choosing y_a and y_b so that p_a and q_b are sufficiently small, the contributions in (87) arising from the discrete events (i.e., the first and last terms) can be ignored to produce the approximation

$$J \approx \int_{y_a}^{y_b} [f_1(y) - f_0(y)] \log\left[\frac{f_1(y)}{f_0(y)}\right] \, \mathrm{d}t, \tag{89}$$

which is easily evaluated using a trapezoidal-rule numerical integral [28, Sect. 7.1] and the sampling paradigm described below. In most cases, y_a can be set so $p_a = F_0(y_a) \approx \epsilon$ and y_b so that $q_b = 1 - F_1(y_b) \approx \epsilon$ with $\epsilon = 10^{-4}$. In rare circumstances (e.g., a very heavy-tailed signal in a benign background), the other probabilities can force the use of a smaller value of ϵ or inclusion of the binary terms via (87). Solving the above equations for y_a and y_b requires a functional inversion of the CDF under H_0 and H_1 (e.g., using a Newton-Raphson iteration). Although this is generally not an issue under H_0 (even for heavy-tailed noise), it can be burdensome under H_1 when each CDF evaluation requires its own numerical integral.[11] An alternative can be found in using an appropriate approximation to the distribution of the decision statistic under H_1 via moment matching; q_b does not need to be precisely equal to ϵ, it only needs to be sufficiently small.

An advantage of this implementation is that it is straightforward to adapt when evaluating the J-divergence after thresholding (see Sect. 5.3) by letting $y_a = h$ and retaining the binary component associated with the event $T \leq y_a$ (i.e., the first term in (87)). Noting that $p_a = 1 - P_f$ and $q_a = 1 - P_d$, the result is seen to match that described in (90).

The sampling of the integrand in (87) or (89) should be done logarithmically to account for large SNRs. Spacing the samples every half decibel was generally found to be sufficient for testing single-intensity decision statistics. However, when accounting for thresholding, it may be necessary to enforce a minimum number of samples (e.g., 20).

5.3 *J*-divergence detection currency after thresholding

A useful attribute of JDC is that optimal combination of multiple independent measurements results in addition of the linear-quantity J-divergences. In many systems where multiple measurements are combined, they are first subjected to their own detection processing (e.g., as in a

---

[11]Note that evaluating $1 - F_1(y)$ may require averaging the conditional exceedance distribution directly, rather than evaluating $F_1(y)$ and subtracting it from one.

distributed system where individual sensors must detect a signal before conveying the information to a fusion center). To account for this in the evaluation of J-divergence, the decision statistic is hard-limited to remain at or above the detector decision threshold (h), which produces a mixed continuous and discrete random variable. If $P_d = \Pr\{Y \geq h | H_1\} = 1 - F_1(h)$ is the probability of detection and $P_f = \Pr\{Y \geq h | H_0\} = 1 - F_0(h)$ is the probability of false alarm, then the linear-quantity J-divergence after thresholding is

$$ J = [P_d - P_f] \log\left[\frac{1 - P_f}{1 - P_d}\right] + \int_h^\infty [f_1(y) - f_0(y)] \log\left[\frac{f_1(y)}{f_0(y)}\right] \, \mathrm{d}y. \tag{90} $$

By rewriting the integral in (90) in terms of the PDFs of the excess over the threshold, it can be phrased as

$$ J = [P_d - P_f] \log\left[\frac{P_d[1 - P_f]}{P_f[1 - P_d]}\right] + P_d \tilde{I}_{1:0} + P_f \tilde{I}_{0:1} \tag{91} $$

where $\tilde{I}_{1:0}$ and $\tilde{I}_{0:1}$ are the Kullback-Liebler (KL) divergences,

$$ \tilde{I}_{i:j} = \int_{-\infty}^\infty \tilde{f}_i(z) \log \frac{\tilde{f}_i(z)}{\tilde{f}_j(z)} \, \mathrm{d}z, \tag{92} $$

for the excess over the threshold $Z = Y - h$ given $Y \geq h$ under H_0 and H_1. The PDF of Z under H_i is denoted by $\tilde{f}_i(z)$ and is related to the PDF of Y according to

$$ \tilde{f}_i(z) = \frac{f_i(z + h)}{\Pr\{Y \geq h | H_i\}}. \tag{93} $$

The tilde notation will be used to represent functions and parameters associated with the excess over the threshold. When evaluating J-divergence through numerical integration, (90) is (marginally) easier to implement than (91). However, the form in (91) can be easier to work with analytically when the PDFs of Z or their approximations are easily described.

As noted in Sect. 2.3, the excess over the threshold under the noise-only case for GPD clutter is another GPD model with the same shape parameter ($\tilde{\gamma} = \gamma$) and a scale parameter $\tilde{\lambda} = \lambda + \gamma h = 1 + (h - 1)\gamma$ (recall that $\lambda = 1 - \gamma$ for perfect normalization). Unfortunately, the signal-present hypothesis does not have a simple closed-form solution. However, as seen in the following section, it can be approximated accurately enough by a gamma distribution to be useful for evaluating JDC. This requires the KL divergences between the gamma and GPD models, which are presented in Sect. 5.3.3.

An example evaluation of JDC after thresholding is shown in Fig. 17 for a Rician signal in GPD clutter with $\rho = 0.3$, an SNR of $18\,\mathrm{dB}$, and various levels of thresholding. As might be expected, the thresholding operation results in a loss in detection currency that increases with the threshold. The black dots use the aforementioned gamma approximation to the excess over the threshold when signal is present and the lines use the numerical integration described in Sect. 5.2. Although the approximation is good in these examples (the maximum absolute error in JDC was less than $0.08\,\mathrm{dB}$), the accuracy is degraded in general when JDC is small (e.g., $< 0\,\mathrm{dB}$) or for very low thresholds and extremely heavy-tailed clutter.
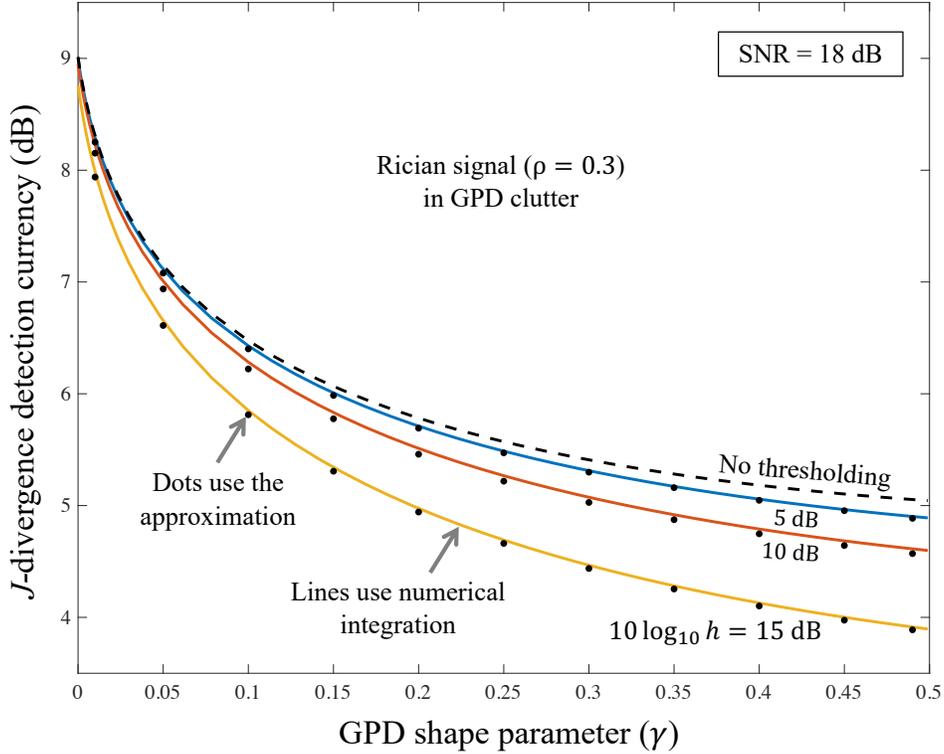
Figure 17: *J*-divergence detection currency (JDC) for a Rician signal ($\rho = 0.3$) in GPD clutter after various levels of thresholding as a function of the GPD shape parameter (γ) for a fixed SNR of 18 dB.

5.3.1 Gamma approximation to the excess over the threshold

An advantage of working with JDC after thresholding is that the approximations to the distribution of the excess over the threshold can yield accurate representations without the need for numerical evaluation of the *J*-divergence integral. As argued in Sect. 4.1.3, the signal-present distribution of Y at high SNR can be approximated by that for the signal in a benign background. For the Rician signal, this results in a non-central chi-squared distribution for the instantaneous intensity, which can be difficult to work with analytically. An alternative can be found in the gamma distribution, which is often used as an approximation to the non-central chi-squared distribution [29, Sect. 29.8].

The gamma approximation to the intensity PDF when signal is present is accomplished by matching the moments of the two distributions. The mean and variance of an instantaneous intensity comprising a Rician signal in GPD clutter are, respectively,

$$\mu_Y = E[Y|H_1] = 1 + s \quad \text{[unitless: normalized power]} \tag{94}$$

for the mean and

$$\sigma_Y^2 = E\big[(Y - \mu_Y)^2|H_1\big] = 2s + \rho(2-\rho)s^2 + \frac{1}{1-2\gamma} \quad \text{[unitless: normalized power}^2] \tag{95}$$

for the variance. These moments can be obtained using the formulas found in [1, Sect. 7.5.6.2] (in particular, eqs. 7.241 and 7.242 coupled with the Rician-signal entry in Table 7.8 and the GPD

entry in Table 7.9). Equating these to the moments of a gamma distribution results in a shape parameter

$$\alpha = \frac{(1+s)^2(1-2\gamma)}{1 + s[2 + \rho(2-\rho)s]\,(1-2\gamma)} \tag{96}$$

and a scale parameter

$$\beta = \frac{1 + s[2 + \rho(2-\rho)s]\,(1-2\gamma)}{(1+s)(1-2\gamma)} = \frac{1+s}{\alpha}. \tag{97}$$

Given that the intensity Y is gamma distributed with shape α and scale β, the distribution of $Z = Y - h$ given $Y \geq h$ is simply

$$\tilde{f}(z) = \frac{f_{\mathrm{G}}(z+h; \alpha, \beta)}{1 - F_{\mathrm{G}}(h; \alpha, \beta)} \ \ \text{for} \ \ z \geq 0, \tag{98}$$

where $f_{\mathrm{G}}(y; \alpha, \beta)$ and $F_{\mathrm{G}}(y; \alpha, \beta)$ are, respectively, the PDF and CDF of the gamma distribution. When $\alpha = 1$, the gamma distribution simplifies to the exponential for which the distribution of Z is also exponential. This justifies use of a gamma distribution to describe Z when α is near one, which occurs when ρ is near one. When h is large enough, the exponential form of the limit distribution of the excess over the threshold for a non-central chi-squared distribution (see Sect. 2.4) also suggests the gamma model might be appropriate even for small values of ρ.

The gamma approximation to the excess over the threshold is obtained by matching the moments of Z. The first two moments of Z can be shown to be

$$\mu_1 = E[Z|H_1] = \alpha\beta \frac{[1 - F_{\mathrm{G}}(h; \alpha+1, \beta)]}{[1 - F_{\mathrm{G}}(h; \alpha, \beta)]} - h \tag{99}$$

and

$$\mu_2 = E[Z^2|H_1] = \alpha(\alpha+1)\beta^2 \frac{[1 - F_{\mathrm{G}}(h; \alpha+2, \beta)]}{[1 - F_{\mathrm{G}}(h; \alpha, \beta)]} - 2h\alpha\beta \frac{[1 - F_{\mathrm{G}}(h; \alpha+1, \beta)]}{[1 - F_{\mathrm{G}}(h; \alpha, \beta)]} + h^2. \tag{100}$$

Equating (99) and (100) with the moments of a gamma distribution produces the shape parameter

$$\tilde{\alpha} = \frac{\mu_1^2}{\mu_2 - \mu_1^2} \tag{101}$$

and scale parameter

$$\tilde{\beta} = \frac{\mu_2 - \mu_1^2}{\mu_1}. \tag{102}$$

These are used with the characterization of the excess over the threshold under the noise-only hypothesis as GPD with shape $\tilde{\gamma} = \gamma$ and scale $\tilde{\lambda} = 1 + (h-1)\gamma$ to obtain the KL divergences used in (91).

To summarize, the technique proposed for obtaining JDC for a Rician signal in GPD clutter is to make a sequence of approximations, with each having weaknesses. These include:

1. approximating the intensity distribution of the Rician signal in GPD clutter by that for a benign background [weakness: high γ or low SNR, which equates to low P_d or JDC],

2. approximating the resulting non-central chi-squared distribution by a gamma distribution [weakness: low ρ], and

3. approximating the excess over the threshold of a gamma distribution by another gamma model [weakness: low threshold and low ρ].
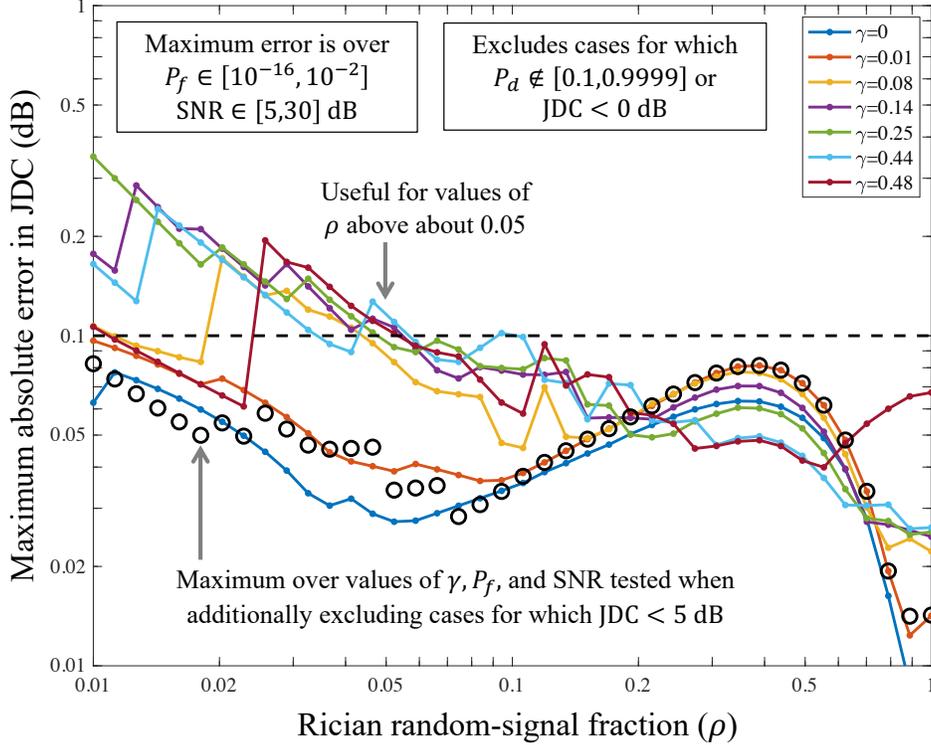
Figure 18: Maximum absolute error in the J-divergence detection currency (JDC) approximation over the cases evaluated as a function of the Rician random-signal fraction (ρ) for various GPD clutter shape parameters.

5.3.2 Accuracy of the detection-currency approximation

To assess the accuracy of the gamma-based approximation to JDC over a relevant set of scenarios, the parameters were quantized to values of

- P_f at each decade from 10^{-16} to 10^{-2},

- ρ for 40 values logarithmically spaced from 0.01 to 1,

- $\gamma = 0$, 0.01, 0.08, 0.14, 0.25, 0.44, and 0.48, and

- SNR every decibel from 5 dB to 30 dB.

The values of $\gamma \in [0.08, 0.44]$ were taken from Table 1 and then supplemented with more and less extreme values closer to the boundaries of the full $[0, 0.5)$ domain. Cases for which $P_d \notin [0.1, 0.9999]$ were discarded, as were any where the numerically evaluated JDC < 0 dB. The maximum absolute

error in JDC over the P_f and SNR parameters is then shown in Fig. 18 as a function of ρ with one line for each value of γ. The results indicate that the error is close to or below $0.1\,\mathrm{dB}$ when $\rho \geq 0.05$, $P_f \in [10^{-16}, 10^{-2}]$, and the SNR is large enough to achieve the constraints on P_d and JDC. This corroborates the expectation that the gamma approximation degrades as ρ decreases.

The effect of SNR was assessed by restricting the analysis to cases for which the JDC was at least $5\,\mathrm{dB}$. The results seen in Fig. 18 (black circles) indicate that the gamma approximation is useful for values of ρ as low as 0.01 at these higher quality operating points. In the majority of cases of practical interest, the gamma-based approximation to JDC (see Sect. D.2 for MATLAB®) code) will be accurate to within a tenth of a decibel.

5.3.3   Kullback-Liebler divergences between the gamma and GPD models

Recall that the approximation in Sect. 5.3.1 employs a gamma distribution with shape $\tilde{\alpha}$ and scale $\tilde{\beta}$ for the excess over the threshold when a Rician signal is present. As seen in (91), evaluation of J-divergence then requires the Kullback-Liebler (KL) divergences between the gamma and GPD models, with the latter representing the noise-only case and having shape $\tilde{\gamma}$ and scale $\tilde{\lambda}$. If Z is gamma distributed with shape $\tilde{\alpha}$ and scale $\tilde{\beta}$, then its PDF is

$$\tilde{f}_1(z) = \frac{z^{\tilde{\alpha}-1} e^{-z/\tilde{\beta}}}{\Gamma(\tilde{\alpha})\tilde{\beta}^{\tilde{\alpha}}} \quad \text{for } z > 0. \tag{103}$$

The log-likelihood-ratio (LLR) between the Gamma$(\tilde{\alpha}, \tilde{\beta})$ and GPD$(\tilde{\gamma}, \tilde{\lambda})$ PDFs is

$$l(z) = \log\frac{\tilde{f}_1(z)}{\tilde{f}_0(z)} = (\tilde{\alpha}-1)\log(z) - \frac{z}{\tilde{\beta}} - \log\Gamma(\tilde{\alpha}) - \tilde{\alpha}\log\tilde{\beta} + \log\tilde{\lambda} + \left(1 + \frac{1}{\tilde{\gamma}}\right)\log\left(1 + \frac{\tilde{\gamma}z}{\tilde{\lambda}}\right). \tag{104}$$

The KL divergences are obtained by taking the expectation of $l(Z)$ in (104) under H_1 to obtain $\tilde{I}_{1:0}$ and H_0 to yield $-\tilde{I}_{0:1}$. The former results in

$$\tilde{I}_{1:0} = E[l(Z)|H_1] = \log\left(\frac{\tilde{\lambda}}{\tilde{\beta}\Gamma(\tilde{\alpha})}\right) - \tilde{\alpha} + (\tilde{\alpha}-1)\psi(\tilde{\alpha}) + \left(1 + \frac{1}{\tilde{\gamma}}\right)g_{\tilde{\alpha}-1}\left(\frac{\tilde{\gamma}\tilde{\beta}}{\tilde{\lambda}}\right) \tag{105}$$

where $\psi(\cdot)$ is the digamma function, the solution to $E[\log Z|H_1] = \log\tilde{\beta} + \psi(\tilde{\alpha})$ is obtained via [30, eq. 4.352-1], and the function $g_\kappa(\delta)$ is defined as

$$g_\kappa(\delta) = \int_0^\infty \frac{z^\kappa e^{-z}}{\Gamma(\kappa+1)}\log(1 + \delta z)\,\mathrm{d}z. \tag{106}$$

When κ is an integer (k), [30, eq. 4.337-5][12] can be used to evaluate it via

$$g_k(\delta) = \sum_{j=0}^{k}\frac{1}{(k-j)!}\left[-(-\delta)^{j-k}e^{1/\delta}E_i\left(-\delta^{-1}\right) + \sum_{l=1}^{k-j}(l-1)!(-\delta)^{l+j-k}\right], \tag{107}$$

where $E_i(x)$ is the exponential integral function (EIF). The EIF is defined as $E_i(x) = \int_{-\infty}^{x}t^{-1}e^t\,\mathrm{d}t$ and can be obtained in Matlab for $x < 0$ via `-expint(-x)`. For the Gaussian-fluctuating signal, $\tilde{\alpha} =$

---

[12]Note that there are two missing minus signs in [30, eq. 4.337-5].

1 and $g_0(\delta) = -e^{1/\delta}E_i(-1/\delta)$ can be obtained exactly. Approximations to Rician and deterministic signals are likely to result in non-integer values of $\tilde{\alpha}$ and therefore κ. For these cases, the function in (106) was approximated by linear interpolation from the nearest integers to $\kappa = \tilde{\alpha} - 1$ using (107). When k was too large and/or δ too small (approximately when $\delta/k < 0.001$), numerical errors were encountered in the evaluation of (107). In these scenarios (high SNR with distributions approaching a deterministic signal in a benign background), a numerical evaluation of (106) was required.

The KL divergence $\tilde{I}_{0:1}$ for the gamma/GPD model pair can be simplified to

$$\tilde{I}_{0:1} = -E[l(Z)|H_0] = (\tilde{\alpha} - 1)\left[\psi\left(\frac{1}{\tilde{\gamma}}\right) + \mathbf{C}\right] + \frac{\tilde{\gamma}^2 - 1 + \tilde{\lambda}/\tilde{\beta}}{(1 - \tilde{\gamma})} + \log\left[\frac{\tilde{\beta}^{\tilde{\alpha}}\tilde{\gamma}^{\tilde{\alpha}-1}\Gamma(\tilde{\alpha})}{\tilde{\lambda}^{\tilde{\alpha}}}\right] \tag{108}$$

using [30, eqs. 4.293-14 & 4.293-9] and [31, eqs. 44:5:1 & 44:5:3] with $\mathbf{C} \approx 0.5772$ equal to Euler's constant.

6 Conclusions

The focus of this report has been on representing active-sonar clutter with the generalized Pareto distribution (GPD) and using it to model the degradation in detection performance of a sonar system as the clutter severity increases. In reviewing background information on the GPD model, the domain of its shape parameter (which controls severity in terms of the heaviness of the distribution tail) was partitioned into regions ranging from scarcely to extremely heavy tailed. These regions were related to potential sources of clutter through a mapping to the K-distribution shape parameter and inform which parameter values are appropriate for modeling different scenarios.

To determine how GPD model parameters should be obtained in practice, several estimation approaches were compared. A Bayesian technique previously developed for estimating the K-distribution shape parameter was applied to the GPD to force the envelope-based method-of-moments estimate to lie on a region representing physically realizable processes. Of the estimators evaluated, the most appropriate for use was seen to be a mixed estimator taking the envelope-based method-of-moments estimate when it is viable and otherwise resorting to the Bayesian adaptation.

Approximations to the design SNR required to meet a (P_d, P_f) operating point (i.e., the detection threshold (DT) term in the sonar equation) were developed for deterministic and Gaussian-fluctuating signals in GPD clutter for single- and integrated-intensity detectors. The approximations were seen to be accurate (absolute error less than 0.1 dB) over a wide range of operating points and clutter severity. The approach entailed using the detector decision threshold (h) required to meet the P_f specification in the clutter-dominated background, while representing the relationship mapping h and SNR to P_d with the Gaussian distribution chain for a benign background. Although the integrated-intensity detector required an additional approximation to obtain the detector decision threshold as a function of P_f, the clutter severity, and number of intensities being integrated, it was still usefully accurate over a wide range of scenarios. This type of analysis is representative of systems that adapt their decision threshold to the clutter background in order to maintain a constant probability of false alarm. A simple approximation to the increase in DT when the clutter background becomes more severe was shown to be the decibel increase in the intensity-based decision threshold required to maintain P_f. Such an approximation can be applied to determine the increase in DT empirically using data from a system in a cluttered environment.

Techniques for evaluating the J-divergence detection currency (JDC) for the basic sonar signal

models in GPD clutter were presented. JDC is useful when a system combines multiple measurements, such as those from different transmit waveforms or separate receivers. Although a numerical solution was required to obtain the JDC after matched filtering, an accurate approximation was developed when the analysis includes the thresholding process found in most detectors. Analysis of JDC illustrated how a high-quality operating point in a benign background can be a low-quality one in a severe-clutter scenario.

Although the deterministic signal exhibits similar performance to the Gaussian-fluctuating signal at low SNR in some scenarios, its consistent amplitude provides significant gains as SNR increases (e.g., see Fig. 13). However, the performance metrics (P_d or JDC) are known to be very sensitive at high SNR. This sensitivity was also observed when the background was dominated by clutter. When JDC is the performance metric, a reasonable alternative is to use a Rician signal with a small random-power fraction (e.g., $\rho = 0.1$). Unfortunately, there is no simple solution for DT other than to use the Gaussian-fluctuating signal, which is not always appropriate.

In order to facilitate application of the parameter estimation algorithms and performance modeling techniques, MATLAB$^{\circledR}$ functions implementing the key results can be found in the appendices.

References

[1] D. A. Abraham, *Underwater Acoustic Signal Processing: Modeling, Detection, and Estimation.* Springer, 2019.

[2] *Acknowledgement:* SCARAB 1997 Experiment, NATO SACLANT Undersea Research Centre. Scientist in charge: Dr. C. Holland.

[3] D. A. Abraham and C. W. Holland, "Statistical analysis of low-frequency active sonar reverberation in shallow water," in *Proceedings of 4th European Conference on Underwater Acoustics*, Rome, Italy, September 1998, pp. 795–800.

[4] T. J. Barnard and F. Khan, "Statistical normalization of spherically invariant non-Gaussian clutter," *IEEE Journal of Oceanic Engineering*, vol. 29, no. 2, pp. 303–309, April 2004.

[5] E. Conte and M. Longo, "Characterisation of radar clutter as a spherically invariant random process," *IEE Proceedings-F, Radar and Signal Processing*, vol. 134, no. 2, pp. 191–197, April 1987.

[6] B. R. La Cour, "Statistical characterization of active sonar reverberation using extreme value theory," *IEEE Journal of Oceanic Engineering*, vol. 29, no. 2, pp. 310–316, 2004.

[7] J. M. Gelb, R. E. Heath, and G. L. Tipple, "Statistics of distinct clutter classes in mid-frequency active sonar," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 2, pp. 220–229, April 2010.

[8] D. A. Abraham and A. P. Lyons, "Reliable methods for estimating the K-distribution shape parameter," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 2, pp. 288–302, April 2010.

[9] D. A. Abraham, "Detection-threshold approximation for non-Gaussian backgrounds," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 2, pp. 355–365, April 2010.

[10] ——, "Modeling sonar performance using J-divergence," in *Proceedings of the Underwater Acoustics Conference & Exhibition*, Kalamata, Greece, 2023.

[11] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. John Wiley & Sons, Inc., 1994, vol. 1.

[12] T. J. Barnard and D. D. Weiner, "Non-Gaussian clutter modeling with generalized spherically invariant random vectors," *IEEE Transactions on Signal Processing*, vol. 44, no. 10, pp. 2384–2390, October 1996.

[13] D. A. Abraham and A. P. Lyons, "Novel physical interpretations of K-distributed reverberation," *IEEE Journal of Oceanic Engineering*, vol. 27, no. 4, pp. 800–813, October 2002.

[14] D. A. Abraham, J. M. Gelb, and A. W. Oldag, "Background and clutter mixture distributions for active sonar statistics," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 2, pp. 231–247, April 2011.

[15] A. A. Balkema and L. de Haan, "Residual life time at great age," *The Annals of Probability*, vol. 2, no. 5, pp. 792–804, 1974.

[16] J. Pickands III, "Statistical inference using extreme order statistics," *The Annals of Statistics*, vol. 3, no. 1, pp. 119–131, 1975.

[17] R. N. Bracewell, *The Fourier Transform and Its Applications*, 2nd ed. New York: McGraw-Hill, Inc., 1986.

[18] I. R. Joughin, D. B. Percival, and D. P. Winebrenner, "Maximum likelihood estimation of K distribution parameters for SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, no. 5, pp. 989–999, September 1993.

[19] D. A. Abraham, "Statistical normalization of non-Rayleigh reverberation," in *Proceedings of Oceans 97 Conference*, Halifax, Nova Scotia, October 1997, pp. 500–505.

[20] W. J. Albersheim, "A closed-form approximation to Robertson's detection characteristics," *Proceedings of the IEEE*, vol. 69, no. 7, p. 839, 1981.

[21] H. Hmam, "SNR calculation procedure for target types 0, 1, 2, 3," *IEEE Transactions on Aerospace and Electronic System*, vol. 41, no. 3, pp. 1091–1096, 2005.

[22] H. Du, J. Zhang, J. Cheng, and B. Ai, "Sum of Fisher-Snedecor F random variables and its applications," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 342–356, 2020.

[23] A. H. Nuttall, "Accurate efficient evaluation of cumulative or exceedance probability distributions directly from characteristic functions," Naval Underwater Systems Center, New London, CT, Tech. Rpt. 7023, October 1983.

[24] ——, "Evaluation of small tail probabilities directly from the characteristic function," Naval Undersea Warfare Center, Newport, RI, Tech. Rpt. 10840, September 1997.

[25] "Recognition of underwater sounds," Office of Scientific Research and Development, Nat. Def. Res. Comm. Div. 6, Summary Technical Report 9, 1946.

[26] R. J. Urick, "Prediction methods for sonar systems," *Radio and Electronic Engineer*, vol. 25, no. 6, pp. 501–508, June 1963.

[27] H. Jeffreys, *Theory of Probability*, 3rd ed. Oxford University Press, 1948.

[28] J. H. Mathews and K. D. Fink, *Numerical Methods Using MATLAB*. Prentice Hall, 1999.

[29] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York: John Wiley & Sons, Inc., 1995, vol. 2.

[30] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 8th ed. Waltham, MA: Elsevier Academic Press, 2015, edited by D. Zwillinger.

[31] K. Oldham, J. Myland, and J. Spanier, *An Atlas of Functions*, 2nd ed. New York: Springer Science, 2009.

A   MATLAB® functions for intensity distributions in GPD clutter

A.1   Intensity PDF for GPD clutter

```
function f=gpdipdf(y,gam,lam)
% f=gpdipdf(y,gam,lam)
%   PDF of a generalized-Pareto-distributed intensity with mean = lam/(1-gam)
% Parameters: [mix of scalar and common-dimension matrix/vector]
%   f = GPD intensity PDF
%   y = PDF argument (linear-intensity)
%   (gam,lam) = GPD shape & scale parameters
%
[y,gam,lam]=input_par_std(y,gam,lam);
f=(1+gam.*y./lam).^(-1-1../gam)./lam;
f(gam==0)=exp(-y(gam==0)./lam(gam==0))./lam(gam==0);
f(y<0)=0;
```

A.2   Intensity CDF for GPD clutter

```
function F=gpdicdf(y,gam,lam)
% f=gpdicdf(y,gam,lam)
%   CDF of a generalized-Pareto-distributed intensity with mean = lam/(1-gam)
% Parameters: [mix of scalar and common-dimension matrix/vector]
%   F = GPD intensity CDF
%   y = CDF argument (linear-intensity)
%   (gam,lam) = GPD shape & scale parameters
%
[y,gam,lam]=input_par_std(y.*(y>0),gam,lam);
F=1-(1+gam.*y./lam).^(-1../gam);
F(gam==0)=1-exp(-y(gam==0)./lam(gam==0));
```

A.3   Inverse of GPD-intensity CDF

```
function h=gpdiinv(p,gam,lam)
% h=gpdiinv(p,gam,lam)
%   Functional inverse of the CDF of a generalized-Pareto-distributed
%   intensity with mean = lam/(1-gam)
% Parameters: [mix of scalar and common-dimension matrix/vector]
%   h = intensity yielding desired probability (p)
%   p = probability [CDF at output h is F(h)=p]
%   (gam,lam) = GPD shape & scale parameters
%
[p,gam,lam]=input_par_std(p,gam,lam);
h=((1-p).^-gam-1).*lam./gam;
h(gam==0)=-lam(gam==0).*log(1-p(gam==0));
```

A.4   Intensity CDF of a Rician signal in GPD clutter

```
function F=ricegpd_cdf(y,s,rho,gam)
% F = ricegpd_cdf(y,s,rho,gam)
%   Intensity CDF of a Rician signal in unit-power GPD clutter
% Parameters:
%   y = CDF argument (linear-intensity) [matrix/vector]
%   s = Rician signal SNR (linear) [scalar]
%   rho = Rician signal fraction of random signal power [scalar]
%   gam = GPD shape parameter [scalar]
%
if gam>0
  F=zeros(size(y));
  for i=1:numel(y)
    F(i)=integral(@(w) gampdf(w,1/gam,gam)...
    .*ncx2cdf(2*y(i)*w./(1-gam+rho*s*w),2,2*(1-rho)*s*w./(1-gam+rho*s*w)),0,inf);
  end
else, F=ncx2cdf(2*y./(1+rho*s),2,2*(1-rho)*s./(1+rho*s));
end
```

A.5   Intensity PDF of a Rician signal in GPD clutter

```
function f=ricegpd_pdf(y,s,rho,gam)
% f = ricegpd_pdf(y,s,rho,gam)
%   Intensity PDF of a Rician signal in unit-power GPD clutter
% Parameters:
%   y = PDF argument (linear-intensity) [matrix/vector]
%   s = Rician signal SNR (linear) [scalar]
%   rho = Rician signal fraction of random signal power [scalar]
%   gam = GPD shape parameter [scalar]
%
if gam>0
  f=zeros(size(y)); lam=1-gam;
  for i=1:numel(y)
    f(i)=integral(@(w) gampdf(w,1/gam,gam)...
    .*ncx2pdf(2*y(i).*w./(lam+rho.*s.*w),2,2*(1-rho).*s.*w./(lam+rho.*s.*w))...
    .*2.*w./(lam+rho.*s.*w),0,inf);
  end
else, f=ncx2pdf(2*y/(1+rho*s),2,2*(1-rho)*s/(1+rho*s))*2/(1+rho*s);
end
```

A.6   GPD random number generator

```
function y=gpdirnd(gam,lam,varargin)
% y = gpdirnd(gam,lam,M,N,...)
%   Simulate GPD random numbers
% Parameters:
%   (gam,lam) = GPD shape & scale parameters (scalars)
%   M,N,... = dimension of matrix to produce
%
MN=[varargin{:}]; y=exprnd(lam,MN);
if gam>0, y=y./gamrnd(1/gam,gam,MN); end
```

A.7 Support function

```
function varargout=input_par_std(varargin)
% [x1,x2,x3,...] = input_par_std(x1,x2,x3,...)
%   Finds the maximum dimension of the input variables (up to 2-D) and
%   fills scalar and vector inputs to be that size 2-D array in the output.
%
nrc=[cellfun('size',varargin,1)' cellfun('size',varargin,2)'];
varargout=cell(1,nargout);
for i=1:nargout, varargout{i}=repmat(varargin{i},max(nrc)-nrc(i,:)+1); end
```

B MATLAB® functions for estimating the GPD shape parameter

B.1 Example estimation of P_f for GPD clutter

```
% File test_gpd_est.m
% Example fitting of PFA through the GPD model
gam=0.2; lam=2.5; N=1e4; % True parameters & # observations
y=gpdirnd(gam,lam,N,1); % Simulate linear-quantity intensity data
gam_est=gpdmombaymix(y)   % Shape parameter estimate
lam_est=mean(y)*(1-gam_est)    % Scale parameter estimate
[f,h]=hist(y,200); F=1-cumsum(f)/N; Fge=1-gpdicdf(h,gam_est,lam_est);
figure(1); Fexp=exp(-h/mean(y));    % Exponential-intensity PFA
hf=semilogy(h,Fge,h,Fexp,h,F,'k-'); ylim([10/N 1]);   % Compare PFA via histogram
legend(hf,'GPD model','Exponential model','Data');
xlabel('Intensity (linear)'); ylabel('PFA');
```

B.2 Envelope method-of-moments estimator

```
function gam=gpdmomenv(x)
% gam = gpdmomenv(x)
%   GPD shape-parameter estimate via envelope method of moments
% Parameters:
%   x = envelope data (single vector or matrix with separate trials in each column)
% Note: when moment equations are not invertible, gam is set to zero
%
D=(mean(x)./std(x,1)).^2;  % Envelope moment ratio
gam=zeros(1,length(D)); % Initialize to default of gam=0 (Rayleigh envelope)
i=find(D<pi/(4-pi)); % For those within bound, map D to estimate of gamma
gam(i)=arrayfun(@(d)gpdmominv(d,0),D(i));
```

B.3 Intensity method-of-moments estimator

```
function gam=gpdmomint(y)
% gam = gpdmomint(y)
%   GPD shape-parameter estimate via intensity method of moments
% Parameters:
%   y = intensity data (single vector or matrix with separate trials in each column)
%
D=(mean(y)./std(y,1)).^2; % Intensity moment ratio
gam=(1-D)/2;
```

B.4 Functional inverse of the GPD moment-ratio equation

```
function gam=gpdmominv(D,qNR)
% gam = gpdmominv(D,[qNR])
%   Solution of the GPD envelope moment-ratio equation: g(gamma)=D
% Parameters:
%   D = envelope moment ratio (mean square envelope/envelope variance)
%   qNR = 0 for approximation or 1 [default] to refine via Newton-Raphson iteration
%
dmax=pi/(4-pi); r=D/dmax; if nargin==1, qNR=1; end
gam=(1-r).*(1-r*8/19+r.^2*26/53-r.^3*3/14); % Empirical approximation
if qNR % Refine via Newton-Raphson iteration
  qend=0;  TOL=1e-5;  i=0;  maxiter=20;
  while ~qend
    cg=exp(log(gam)+2*(gammaln(1../gam)-gammaln(1../gam-0.5)));
    f=pi*(1-gam)./(4*cg-pi*(1-gam));
    cgp=cg.*(gam-2*psi(1../gam)+2*psi(1../gam -0.5))./(gam.^2);
    fp=-4*pi*(cg+(1-gam).*cgp)./((4*cg-pi*(1-gam)).^2);
    g=gam-(f-D)./fp;
    if (max(abs((g-gam)./g))<TOL)||(i>maxiter), qend=1; end
    i=i+1; gam=g;
  end
  gam(D==dmax)=0; % Fix nan in iteration when D=dmax & gam=0
end
```

B.5 Bayesian method-of-moments estimator

```
function gam=gpdmombayes(y)
% gam = gpdmombayes(y)
%   GPD shape-parameter estimate via Bayesian method of moments
% Parameters:
%   y = intensity data (single vector or matrix with separate trials in each column)
%
x=sqrt(y); n=length(x); % Convert to envelope data
u=mean(x); v=mean((x-u).^2); k4=mean((x-u).^4)./(v.^2); % Sample moments
m1e=(1-1/n); m2e=(1-3/n+5/(n^2)+k4*(1-2/n)/n); % Variance moments
t=(u.^2)./v; m1=(t*m1e+1/n); m2=((t.^2).*m2e+6*t.*m1e/n+3/(n^2)); % Numerator moments
at=(m1.^2)./(m2-m1.^2); bt=(m2-m1.^2)./m1; % Fit moments to a gamma
dmax=pi/(4-pi); dmin=pi^2/(16-pi^2);
% Posterior mean on d
tmp=exp(at.*log(dmax)-dmax./bt-gammaln(at+1)-at.*log(bt)).*...
          (1-exp(at.*log(dmin/dmax)-(dmin-dmax)./bt));
d=at.*bt.*(1-tmp./(gamcdf(dmax,at,bt)-gamcdf(dmin,at,bt)));
gam=gpdmominv(d,0); % Transform back to g
```

B.6   Mixed Bayesian method-of-moments estimator

```
function gam=gpdmombaymix(y)
% gam = gpdmombaymix(y)
%   GPD shape-parameter estimate via envelope MoM if it is on (0,0.5)
%   and the Bayesian MoM if it is not
% Parameters:
%   y = intensity data (single vector or matrix with separate trials in each column)
%
if size(y,1)==1, y=y(:); end % Convert row vector to column
gam=gpdmomenv(sqrt(y));
ib=find((gam==0)|(gam>=0.5)); % Check validity
if any(ib), gam(ib)=gpdmombayes(y(:,ib)); end
```

B.7   Maximum likelihood estimator

```
function [gam,lam]=gpdmle(y,ginit)
% [gam,lam] = gpdmle(y,[ginit])
%   GPD parameter estimates via maximum likelihood
% Parameters:
%   [gam,lam] = GPD shape and scale parameters
%   y = intensity data (single vector)
%   ginit = optional initialization to MLE iteration [default uses intensity MoM]
%
arguments
  y double;
  ginit (1,1) double = gpdmomint(y);
end
gam=ginit; lam=mean(y)*(1-gam);
qend=0; TOL=1e-6; i=0; maxiter=200; gam0=inf; ymax=max(y);
while ~qend % Iterate on MLE with intensity MOM for initialization
  if lam<-gam*ymax   % Test for argument of logarithm being negative
    [gam,lam]=gpdmle(y,abs(ginit));  % if so re-run initializing at gam>0
  end
  lam1=(gam+1)*mean(y./(1+gam*y/lam));  gam=mean(log1p(gam*y/lam));  lam=lam1;
  if (abs(gam-gam0)<TOL)||(i>maxiter), qend=1; end
  i=i+1; gam0=gam;
end
```

C MATLAB® functions for design SNR in GPD clutter

C.1 Detection threshold (DT) in GPD clutter

```
function DT=dt_gpd(Pd,Pf,gam,M,Qsig)
% DT = dt_gpd(Pd,Pf,gam,M,Qsig)
%   Design SNR (detection threshold) in decibels (dB) for an integrated-intensity
%   detector with basic signals in GPD clutter
% Parameters: [mix of scalar and common-dimension matrix/vector]
%   Pd = probability of detection
%   Pf = probability of false alarm
%   gam = GPD shape parameter \in [0,0.5) (gam=0 is Gaussian noise)
%   M = number of independent intensities summed
%   Qsig = 'det' for deterministic signal
%        = 'gfs' for Gaussian-fluctuating signal
% Parameters Pd, Pf, gam, & M can be a mix of scalars & 2-D matrices
%
[Pd,Pf,gam,M,DT]=input_par_std(Pd,Pf,gam,M,nan); % Standardize input size
h=10..^(gpdiint_thr(Pf,gam,M)/10); % Get decision threshold
if strcmpi(Qsig,'det') %----Deterministic signal
  % Approximate SNR-PD relationship via Hmam's equation
  B=0.19./(sqrt(0.819025+1.5206*Pd.*(0.9998-Pd))-0.905);
  A=sign(0.5-Pd).*sqrt(0.85616*log(B));
  DT=-10*log10(M)+10*log10((sqrt(h-M/2+0.25)-A).^2-M/2+0.25);
elseif strcmpi(Qsig,'gfs') %----Gaussian-fluctuating signal
  gamd=gaminv(1-Pd,M,1); DT=10*log10(h./gamd-1);
end
```

C.2 Numerical evaluation of P_d for intensity integration in GPD clutter

```
function Pd=gpdiint_pdrice(hdb,gam,M,s,rho)
% Pd = gpdiint_pdrice(hdb,gam,M,s,rho)
%   The probability of detection for a Rician signal in GPD clutter
%   with a detector summing M unit-power intensity samples
% Parameters:
%   hdb = detector decision threshold [units: dB] (matrix/vector)
%   gam = GPD shape parameter (scalar)
%   M = number of independent GPD intensities integrated (scalar)
%   s = linear-quantity SNR (per intensity sample) (scalar)
%   rho = Rician fraction of random signal power (scalar)
%
ncxcfun=@(w,gam,lam,s,rho) ...
   exp(1j.*w.*(1-rho).*s./(1-1j.*w.*(rho.*s+lam)))./(1-1j.*w.*(rho.*s+lam));
gpd1cfun=@(w,gam,s,rho) integral(@(W) ...
   ncxcfun(w,gam,(1-gam)./W,s,rho).*gampdf(W,1/gam,gam),0,inf,'Waypoints',1);
cfun=@(w,gam,M,s,rho) gpd1cfun(w,gam,s,rho).^M;
h=10..^(hdb/10); Pd=zeros(size(h)); rTOL=1e-6;
for i=1:numel(h)
  if M==1, Pd(i)=1-ricegpd_cdf(h(i),s,rho,gam);
  else
    Pd(i)=0.5-integral(@(w) imag(exp(1j*w).*conj(cfun(w/h(i),gam,M,s,rho)./w))/pi,...
          0,inf,'ArrayValued',true,'RelTol',rTOL);
  end;
end
```

C.3 Numerical evaluation of P_f for intensity integration in GPD clutter

```
function Pf=gpdiint_pf(hdb,gam,M)
% Pf = gpdiint_pf(hdb,gam,M)
%   The probability of false alarm of the detector summing M unit-power
%   intensity samples of GPD clutter
% Parameters:
%   hdb = detector decision threshold [units: dB] (matrix/vector)
%   gam = GPD shape parameter (scalar)
%   M = number of independent GPD intensities integrated (scalar)
%
h=10..^(hdb/10);
if gam==0, Pf=gammainc(h,M,'upper');
else
  Pf=zeros(size(hdb)); lam=1-gam; gami=1/gam;
  rTOL=1e-8;  % Relative tolerance useful for Pf>=1e-6
  % Functions for characteristic function inversion (M~=1, 2 or 3)
  fcfun=@(w,g) 1j*integral(@(r)exp(-w*r).*gpdipdf(1j*r,g,1-g),0,inf,'RelTol',rTOL);
  cfun=@(w,gam,M) fcfun(w,gam).^M;
  % Loop over threshold
  for i=1:numel(h)
    if M==1
      Pf(i)=(1+gam.*h(i)./lam).^(-gami);
    elseif M==2
      c=gam*h(i)/lam;
      Pf(i)=(1+integral(@(z) (1../(1-z./(1+c)).^gami).*...
              (gami./(1+z).^(gami+1)),0,c))./(1+c).^gami;
    elseif M==3
      Pf(i)=1-integral2(@(y,z) gpdicdf(h(i)-z-y,gam,lam).*gpdipdf(y,gam,lam).*...
              gpdipdf(z,gam,lam),0,h(i),0,@(y)h(i)-y,'RelTol',rTOL);
    else
      Pf(i)=0.5-integral(@(w) imag(exp(1j*w).*conj(cfun(w/h(i),gam,M)./w))/pi,...
              0,inf,'ArrayValued',true,'RelTol',rTOL);
    end
  end
end
```

C.4 Empirical formula for the detector decision threshold (h) with intensity integration in GPD clutter

```
function hdb=gpdiint_thr(Pf,gam,M)
% hdb = gpdiint_thr(Pf,gam,M)
%   Empirical approximation to the decision threshold (in decibels)
%   achieving a desired probability of false alarm for a detector
%   integrating a number of independent GPD intensities
% Parameters: [mix of scalar and common-dimension matrix/vector]
%   hdb = detector decision threshold [units: dB]
%   Pf = desired probability of false alarm
%   gam = GPD shape parameter
%   M = number of independent GPD intensities integrated
% The error in the approximation is <0.15 dB for
%   Pf in [1e-6,1e-2]
%   M  in [2,20]
%   gamma in [0,0.45]
% Exact results are produced when M=1 or gamma=0
%
[Pf,M,gam,hdb]=input_par_std(Pf,M,gam,nan);
for i=1:numel(Pf)
  if gam(i)==0
    hdb(i)=10*log10(gaminv(1-Pf(i),M(i),1));
  elseif M(i)==1
    hdb(i)=10*log10((Pf(i)^-gam(i)-1)*(1-gam(i))/gam(i));
  else
    x1=sqrt(2)*erfinv(2*(0.5-Pf(i))); x2=gam(i); x3=log(M(i));
    A=[1 x1 x1^2 x1^3 x2 x2^2 x2^3 x3 x3^2 x3^3 x1*x2 x2*x3 x1*x3 x1^2*x2^2];
    [c,gambrk]=clsefun;
    ic=find(gam(i)>gambrk,1,'last');
    hdb(i)=10*log10(exp(A*c(:,ic))+1)+10*log10((Pf(i)^-gam(i)-1)*(1-gam(i))/gam(i));
  end
end
% Coefficients from LSE analysis
function [c,gambrk]=clsefun
gambrk=[0.000 0.100 0.200 0.325 0.500];
c=[    -1.38    -3.058   -0.9907      2.783
     -0.4208    0.6683   -0.3213     -2.438
      0.03421  -0.2158  -0.07045     0.3319
     -0.002611  0.01936  0.01319  -0.01695
         7.37     10.85   -0.9404     -9.392
        -9.575    -6.365    9.632      11.13
        46.49    -17.44    -14.65     -6.268
         2.36     2.591     2.593      2.367
       -0.5438   -0.5592   -0.5776   -0.5819
        0.07318   0.07612   0.07652   0.07498
        -2.641    -3.982   -0.8378     1.977
       -0.6256    -0.928   -0.3259    0.1923
       -0.0252  -0.07776  -0.09447  -0.06795
         0.781     1.975    0.5775   -0.1522];
```

D MATLAB® functions for evaluating JDC in GPD clutter

D.1 Numerical evaluation of JDC for a Rician signal in GPD clutter

```
function JDC=jdc_ricegpd_num(Sdb,rho,gam,hdb)
% JDC = jdc_ricegpd_num(Sdb,rho,gam,[hdb])
%   J-divergence detection currency (JDC) in decibels for a Rician signal
%   in unit-power GPD clutter after thresholding as obtained through
%   a trapezoidal-rule numerical integral
% Parameters: all scalar
%   (Sdb,rho) = Rician signal SNR (dB) and fraction of random signal power
%   gam = GPD shape parameter
%   hdb = normalized intensity threshold (dB) [omit or set to -inf for no thresholding]
%
if nargin==3, hdb=-inf; end
s=10^(Sdb/10); h=10^(hdb/10); lam=1-gam; e=1e-4;
z1init=ncx2inv(1-e,2,2*(1-rho)*s/(1+rho*s))*(1+rho*s)/2; % Initialize w/ Gaussian
z0=max(log10(h),log10(gpdiinv(e,gam,lam)));
z1=fzero(@(h) e-(1-ricegpd_cdf(h,s,rho,gam)),z1init); % iterate to improve
z1=max(log10(2*h),log10(z1)); % Make sure it's above threshold
Nz=max(20,ceil(20*(z1-z0))); z=logspace(z0,z1,Nz); % Every 0.5 dB
f0=gpdipdf(z,gam,lam); f1=ricegpd_pdf(z,s,rho,gam);
Jh=trapz(z,(f1-f0).*log(f1./f0)); % Approximate J-div above threshold
Jb=0;
if h>0 % Get discrete part if thresholding
  Pd=1-ricegpd_cdf(h,s,rho,gam); Pf=1-gpdicdf(h,gam,lam);
  Jb=(Pd-Pf)*log((1-Pf)/(1-Pd)); Jb(~isfinite(Jb))=0; % In case Pd=1
end
JDC=5*log10(Jh+Jb);
```

D.2 Approximating JDC after thresholding for a Rician signal in GPD clutter

D.2.1 JDC using the gamma approximation

```
function JDC=jdc_ricegpd_apx(Sdb,rho,gam,hdb)
% JDC = jdc_ricegpd_apx(Sdb,rho,gam,hdb)
%   J-divergence detection currency (JDC) in decibels for a Rician signal
%   in unit-power GPD clutter after thresholding as obtained through
%   gamma approximations
% Parameters: [mix of scalar and common-dimension matrix/vector]
%   (Sdb,rho) = Rician signal SNR (dB) and fraction of random signal power
%   gam = GPD shape parameter
%   hdb = normalized intensity threshold (dB)
%
dbinv=@(X) 10..^(X/10);
[s,rho,gam,h,I01,I10]=input_par_std(dbinv(Sdb),rho,gam,dbinv(hdb),0,0);
lam0=1-gam;  lam=lam0+gam.*h;
Pf=1-gpdicdf(h,gam,lam0); Pd=ones(size(s));
for i=1:numel(s)
  if h(i)>0, Pd(i)=1-ricegpd_cdf(h(i),s(i),rho(i),gam(i)); end
end
logPf=(-1../gam).*log(1+gam.*h./lam0);
logPf(gam==0)=-h(gam==0)./lam0(gam==0);
alp1=(1+s).^2.*(1-2*gam)./(1+s.*(2+rho.*(2-rho).*s).*(1-2*gam));
bet1=(1+s)./alp1;
[a,b]=gamgam_thresh(alp1,bet1,h);
for i=1:numel(gam)
  if gam(i)>0, [I10(i),I01(i)]=kld_gamgpd(a(i),b(i),gam(i),lam(i));
  else, [I10(i),I01(i)]=kld_gamgam(1,1,a(i),b(i)); end
end
Jex=(Pd-Pf).*(log(Pd)-logPf)+Pd.*I10+Pf.*I01;
Jb=(Pd-Pf).*log((1-Pf)./(1-Pd));
Jb(~isfinite(Jb))=0; % for h=0 where Pd=1
JDC=5*log10(Jb+Jex);
```

D.2.2 Gamma approximation to the excess over the threshold

```
function [alp1,bet1]=gamgam_thresh(alp0,bet0,h0)
% [alp1,bet1] = gamgam_thresh(alp0,bet0,h0)
%   Gamma approximation to excess over the threshold (EOT) of a gamma distribution
% Parameters: [mix of scalar and common-dimension matrix/vector]
%   (alp1,bet1) = output gamma distribution shape and scale parameters
%   (alp0,bet0) = input gamma distribution shape and scale parameters
%   h0 = threshold on input gamma distribution
%
Fh0=gamcdf(h0,alp0,bet0,'upper'); Fh1=gamcdf(h0,alp0+1,bet0,'upper');
Fh2=gamcdf(h0,alp0+2,bet0,'upper');
u=(alp0.*bet0.*Fh1-h0.*Fh0)./Fh0;
P=(alp0.*(alp0+1).*(bet0.^2).*Fh2-2*h0.*alp0.*bet0.*Fh1+h0.^2.*Fh0)./Fh0;
u2=u.^2; alp1=u2./(P-u2); bet1=(P-u2)./u;
```

D.3   KL divergence functions

D.3.1   KL divergence between gamma and GPD models

```
function [I10,I01]=kld_gamgpd(a,b,gam,lam)
% [I10,I01] = kld_gamgpd(a,b,gam,lam)
%   Kullback-Liebler divergence between gamma and GPD models
% Parameters: [mix of scalar and common-dimension matrix/vector]
%   (a,b) = gamma-distribution shape and scale parameters
%   (gam,lam) = GPD shape and scale parameters
%
[a,b,gam,lam]=input_par_std(a,b,gam,lam);
N=numel(a); g0=zeros(size(a)); g1=g0;
del=b.*gam./lam; a0=max(1,floor(a)); a1=ceil(a);
for i=1:N
  g0(i)=gnr43375(a0(i)-1,del(i));
  if a1(i)>a0(i), g1(i)=gnr43375(a1(i)-1,del(i));
  else, g1(i)=g0(i); end
end
Itmp1=g0+(a-a0).*(g1-g0);
I10=log(lam./b)-a-gammaln(a)+(a-1).*psi(a)+(1+1../gam).*Itmp1;
C=0.577215665; % Euler constant
I01=(a-1).*(psi(1../gam)+C)+(gam.^2-1+lam./b)./(1-gam)+a.*log(b./lam)...
        +(a-1).*log(gam)+gammaln(a);
%----------------------------------------------------------------
function g=gnr43375(ninp,ainp)
% g = gnr43375(n,a)
%   Evaluates g=integral(@(x)(x.^n).*exp(-x).*log(1+a*x),0,inf)/gamma(n+1);
%   -->When n is not an integer, it is rounded to the nearest one
%   Reference: Gradshteyn and Ryzhik, "Table of Integrals, Series, and Products,"
%              2015, Equation 4.337-5, page 576
%
[nv,av]=input_par_std(ninp,ainp); g=zeros(size(nv)); nmax=round(max(nv));
kf=gamma(2:nmax+1)';  eia=exp(1../av).*(-expint(1../av));
for i=1:length(g(:))
  n=round(nv(i));
  if av(i)/n<0.001 % Use an exact integral here
    g(i)=integral(@(x)exp(n*log(x)-x-gammaln(n+1)).*log(1+av(i)*x),0,inf);
  else
    if n==0, g(i)=-eia(i);
    else
      a=av(i); S=ones(n,1);
      for l=(n-2):-1:0
        S(l+1)=kf(n-l-1)-S(l+2)/a;
      end
      j=(0:n-1)'; t1=((-a).^(j-n))*eia(i);
      g(i)=-eia(i)+sum((-t1+S)./kf(n-j));
    end
  end
end
```

D.3.2 KL divergence between two gamma distributions

```
function [I10,I01]=kld_gamgam(a0,b0,a1,b1)
% [I10,I01] = kld_gamgam(a0,b0,a1,b1)
%   Kullback-Liebler divergence between two gamma distributions
% Parameters: [mix of scalar and common-dimension matrix/vector]
%   (a0,b0) & (a1,b1) = gamma-distribution (shape,scale) parameters
%
c=gammaln(a0)-gammaln(a1);
I10=(1../b0-1../b1).*a1.*b1+a0.*log(b0)-a1.*log(b1)+(a1-a0).*(log(b1)+psi(a1))+c;
I01=(1../b1-1../b0).*a0.*b0+a1.*log(b1)-a0.*log(b0)+(a0-a1).*(log(b0)+psi(a0))-c;
```