Non-Gaussian Target Detection in Sonar Imagery Using the Multivariate Laplace Distribution

Nick Klausner, *Student Member, IEEE*, and Mahmood R. Azimi-Sadjadi, *Senior Member, IEEE*

***Abstract*—This paper introduces a new non-Gaussian detection method for complex-valued synthetic aperture sonar (SAS) imagery. The detection method is based on a multivariate extension of the Laplace distribution derived using a scale mixture of Gaussian distributions. A goodness-of-fit test in the form of a likelihood ratio is then conducted on a sonar imagery data set consisting of high-frequency (HF) and broadband (BB) images coregistered over the same region on the seafloor showing the proposed model's applicability in sonar imagery. Detection based on testing the equality of parameters from two populations is then implemented on a database containing actual SAS images of the seafloor with synthetically generated targets inserted into the images and compared to a similar non-Gaussian technique. Detection performance in this paper is given in terms of receiver-operator characteristic (ROC) curve attributes, probability of detection, and average false alarm rate.**

***Index Terms*—Binary hypothesis testing, multivariate Laplace, non-Gaussian detection, synthetic aperture sonar (SAS), underwater target detection.**

I. Introduction

THE studies on bottom return statistics reveal [1]–[3] that the distribution of the envelope of the matched filtered output is dependent on the frequency, grazing angle, range, and roughness properties of the bottom. Rough surface measurements made using high-resolution sonar have indicated [2] that the envelope amplitude distributions can often deviate from a Rayleigh distribution and be better modeled using log-normal, Weibull, or other more complex distributions. This suggests that the underlying complex data are not Gaussian and that the first- and second-order moments will not be sufficient for detection and classification purposes. Thus, the standard tests that rely on assumptions of normality can fail to capture the true properties of the background clutter, hence increasing the incident of false alarms per image.

Statistically based detection techniques rely on the evaluation of likelihood ratios which are dependent on the validity of the model one assumes for the family of distributions used to characterize the observations as well as the parameters used to describe those distributions. For many binary hypothesis tests, especially those that rely on multivariate observations, the family of distributions is taken to be Gaussian because of its practicality and ability to model a wide range of problems. In this paper, we will consider an alternative multivariate distribution based on the Laplace distribution and investigate its practicality in synthetic aperture sonar (SAS) imagery.

Current research into non-Gaussian detection techniques typically look to accurately model the envelope or magnitude of the matched filtered output in sonar. In [4], Abraham discusses extensions to the detection threshold (DT) for the envelope of the output from a matched filter when the statistical distribution of the noise deviates from the Gaussian distribution. Here, the DT describes the signal-to-noise ratio needed to achieve a specific detection probability for a given probability of false alarm. When the noise or reverberation can be modeled using a Gaussian distribution, there exist easily evaluated approximations for determining the DT but these approximations become inaccurate when the background has a heavy-tailed probability density function (pdf). In this reference, the author shows that one can achieve very low approximation error in the DT when the envelope of the noise is modeled with a Weibull or K-distribution [3]. In [5], a detection method for sonar imagery was developed assuming K-distributed background clutter. The method uses a finite mixture model of K-distributions (FMMK) to represent the data under the $\mathcal{H}_1$ hypothesis. An optimum method for estimating the parameters of the FMMK and a generalized likelihood ratio test (GLRT) was also proposed. The detector was compared to the corresponding counterparts derived for Gaussian and Rayleigh distributions [6]. Test results of the proposed method on a data set of SAS images were also presented which illustrated the effectiveness of the FMMK-distributed detector.

The main issue with the aforementioned methods is that they model only the magnitude of the individual pixels in the image and hence they ignore the phase information of the data and also lack the ability to model spatial relationships among the pixels in higher dimensions as well as interchannel relationships in multichannel problems. Previously in [7], we developed a complex-valued multivariate statistical model based on the Laplace distribution which is derived using a scale mixture of Gaussian distributions. In this paper, we extend the theoretical foundations of this probabilistic model and apply it to the problem of underwater target detection in sonar imagery. Using derived expressions for the pdf of this statistical model, maximum likelihood estimates are found recursively using the expectation–maximization (EM) algorithm [8]. To compare the modeling capabilities of the Laplace distribution with that of the complex Gaussian distribution, a goodness-of-fit test in the

Manuscript received July 16, 2013; revised March 11, 2014; accepted May 24, 2014. Date of publication June 24, 2014; date of current version April 10, 2015. This work was supported by the U.S. Office of Naval Research (ONR) under Contract N00014-09-1-0087.

Associate Editor: J. Cobb.

The authors are with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523-1373 USA (e-mail: nklausne@engr.colostate.edu; azimi@engr.colostate.edu).

Digital Object Identifier 10.1109/JOE.2014.2328211

form of a likelihood ratio is conducted on a data set consisting of high-frequency (HF) and broadband (BB) sonar images of real targets. Using the parameters estimated from the EM algorithm, a detector that tests for equality in mean and covariance from two multivariate populations is then extended and tested on a data set consisting of HF and broadband BB SAS images of real background along with simulated targets of various geometrical shapes inserted into the image. Results show that, at an equal false alarm rate, the detection method based on the Laplace distribution outperforms its classical counterpart which relies on assumptions of normality.

This paper is organized as follows. Section II derives the complex-valued Laplace distribution and discusses how to estimate the parameters that describe this distribution using the EM algorithm. Section III then investigates the Laplace distribution's ability to model data of varying clutter difficulty through likelihood ratio measurements. Section IV reviews the hypothesis test used to discriminate target from background and provides results for both the Gaussian and Laplacian methods as well as a similar technique based on the K-distribution. Finally, Section V makes concluding remarks and discusses future lines of research.

II. The Complex-Valued Multivariate Laplace Distribution

A. Review and Density Characterization

In this section, we give a review of the multivariate Laplace distribution and derive its extension to complex-valued observations. For a detailed review of the methods discussed in this section, the reader is referred to [9]–[12]. In many practical statistical applications, e.g., sampling from particular distributions in Monte Carlo trials, it can be useful to generate realizations of random variables of the form $X = Z/Y$ where Y is a random variable with arbitrary distribution on the interval $[0, \infty)$ and Z is distributed according to a standard Gaussian distribution, i.e., $Z \sim N(0,1)$, and drawn independently of Y. Such a model is typically referred to as a scale mixture of Gaussian distributions. In [10], it was shown that, for such a representation to exist, it is both necessary and sufficient that the pdf of X, $f_X(x)$, be symmetric about the origin and have derivatives satisfying

$$(-1)^k \frac{d^k}{dx^k} f_X(\sqrt{x}) \geq 0 \qquad \forall x > 0. \tag{1}$$

Moreover, it was shown [10] that the distribution function of the random variable Y is very closely related to, and can be determined from, the Laplace transform of $f_X(x)$. One distribution that satisfies the requirement given in (1) is the Laplace distribution with pdf

$$f_X(x) = \frac{1}{2} \exp\{-|x|\}. \tag{2}$$

Through the use of the Laplace transform of $f_X(x)$, one finds that $1/2Y^2$ is exponentially distributed with unity rate parameter [10]. In fact, if we consider the random variable $X = \mu + \sqrt{Y} Z$ where Z is again standard Gaussian and Y is exponentially distributed with arbitrary rate parameter λ, denoted $Y \sim \text{Exp}(\lambda)$, then the same procedure can be used to show that the pdf of X is given to be [9]

$$f_X(x) = \frac{1}{2}\sqrt{2\lambda} \exp\left\{-\sqrt{2\lambda}|x-\mu|\right\} \tag{3}$$

i.e., a Laplace distribution centered at μ with scaling parameter $\sqrt{2\lambda}$.

When extending the Laplace distribution to multiple dimensions (see, for instance, [9] and [11]), we can simply replace the scalar random variable Z with the random vector $\mathbf{Z} \sim \mathcal{N}_d(\mathbf{0}, I)$ where $\mathcal{N}_d(\boldsymbol{\mu}, R)$ denotes a d-dimensional multivariate Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix R. In this case, we now generate multivariate realizations according to the model $\mathbf{X} = \boldsymbol{\mu} + \sqrt{Y}\Gamma^{1/2}\mathbf{Z}$ where $Y \sim \text{Exp}(\lambda)$, $\boldsymbol{\mu} \in \mathbb{R}^d$, and Γ is any symmetric, positive–definite (PD) matrix with $\Gamma^{1/2}\Gamma^{T/2} = \Gamma$. Again, random variable Y is assumed to be independent of $\mathbf{Z}$. As shown in [9], the multivariate Laplace pdf of $\mathbf{X}$, assuming that $\det(\Gamma) = 1$ (this will be discussed at the end of Section II-A), can be written

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{2\lambda}{(2\pi)^{\frac{d}{2}}} \frac{K_{\left(\frac{d}{2}\right)-1}\left(\sqrt{2\lambda q(\mathbf{x})}\right)}{\left(\sqrt{\frac{q(\mathbf{x})}{2\lambda}}\right)^{\left(\frac{d}{2}\right)-1}} \tag{4}$$

where $q(\mathbf{x})$ is the quadratic form

$$q(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu})^T \Gamma^{-1} (\mathbf{x}-\boldsymbol{\mu}) \tag{5}$$

and $K_\nu(x)$ denotes a modified Bessel function of the second kind [13]

$$K_\nu(x) = \frac{1}{2}\left(\frac{x}{2}\right)^\nu \int_0^\infty t^{-\nu-1} \exp\left\{-\left(t + \frac{x^2}{4t}\right)\right\} dt. \tag{6}$$

Fig. 1(a) displays the contour levels of a bivariate ($d = 2$) Gaussian pdf with zero mean and arbitrary covariance matrix. Likewise, Fig. 1(b) shows the same contour levels for the bivariate Laplace pdf given in (4) using the same mean and covariance matrix and with $\lambda = 1/4$. From these figures, one can see that both densities exhibit the same elliptical symmetry due to the functional dependence on the quadratic $q(\mathbf{x})$. However, one can also see that the density in (4) exhibits a more exponential decay from the mean and heavier tails, characteristics consistent with its univariate counterpart.

To extend this result to complex-valued data, we will consider the same generative model as before but simply replace the multivariate Gaussian random vector $\mathbf{Z}$ with the one which is now distributed according to a circular symmetric complex Gaussian distribution [14] with mean $\mathbf{0}$ and covariance matrix I, denoted $\mathbf{Z} \sim \mathcal{CN}_d(\mathbf{0}, I)$. That is, we now assume that realizations of our observation arise from the model $\mathbf{X} = \boldsymbol{\mu} + \sqrt{Y}\Gamma^{1/2}\mathbf{Z}$ where Y is again exponentially distributed with parameter λ and independent of $\mathbf{Z}$, $\boldsymbol{\mu} \in \mathbb{C}^d$; Γ is any Hermitian, PD matrix with $\Gamma^{1/2}\Gamma^{H/2} = \Gamma$; and $\mathbf{Z}$ is distributed according to the pdf

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{\pi^d} \exp\left\{-\|\mathbf{z}\|^2\right\}. \tag{7}$$

Using the integral representation of the Bessel function given in (6) and taking advantage of the fact that $\mathbf{X}|Y \sim \mathcal{CN}_d(\boldsymbol{\mu}, y\Gamma)$,

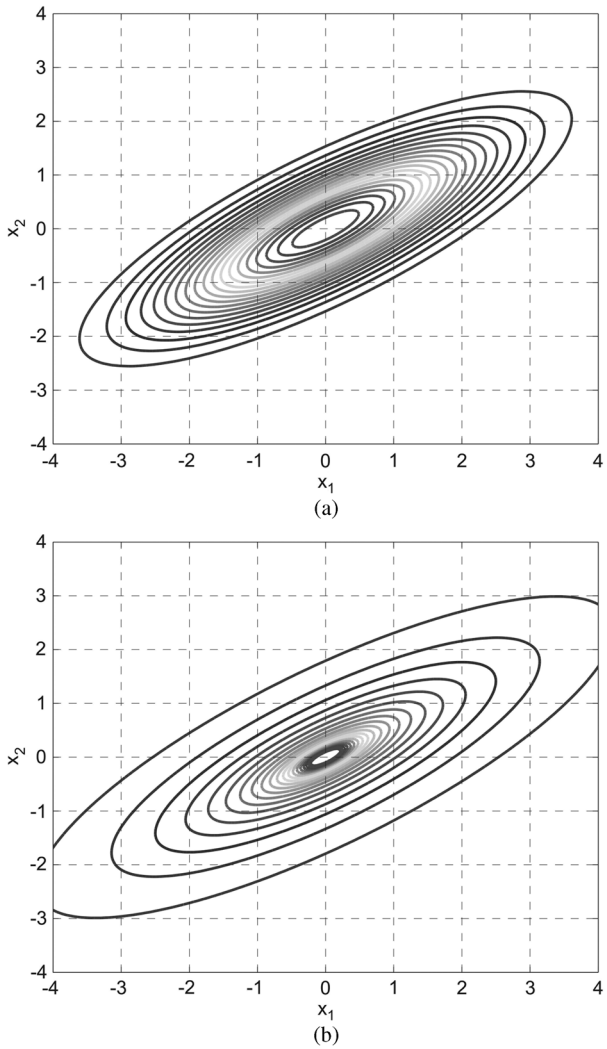


Fig. 1. Comparison of the pdf contours for a bivariate ($d = 2$) Gaussian and Laplacian distributions. Both distributions exhibit elliptical symmetry but the Laplace distribution is heavier tailed. (a) Bivariate Gaussian. (b) Bivariate Laplace.

the multivariate density of our observation can be derived by averaging the conditional density of $\mathbf{X}|Y$ over Y

$$f_{\mathbf{X}}(\mathbf{x}) = \int_0^\infty f_{\mathbf{X}|Y}(\mathbf{x}|y) f_Y(y) dy \tag{8}$$

$$= \frac{\lambda}{\pi^d \det(\Gamma)} \int_0^\infty y^{-d} \exp\left\{ -\left(\lambda y + \frac{q(\mathbf{x})}{y} \right) \right\} dy \tag{9}$$

$$= \frac{2\lambda^d}{\pi^d \det(\Gamma)} \frac{K_{d-1}\left(2\sqrt{\lambda q(\mathbf{x})} \right)}{\left(\sqrt{\lambda q(\mathbf{x})} \right)^{d-1}} \tag{10}$$

where $q(\mathbf{x})$ is again the quadratic form

$$q(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu})^H \Gamma^{-1} (\mathbf{x} - \boldsymbol{\mu}). \tag{11}$$

An interesting property of the random vector $\mathbf{X}$ is that its probability law remains invariant to positive, real-valued scaling in the parameters λ and Γ. That is, if we let $f_{\mathbf{X}}(\mathbf{x}; \lambda, \boldsymbol{\mu}, \Gamma)$ denote the pdf of $\mathbf{X}$ parametrized by λ, $\boldsymbol{\mu}$, and Γ, then for any $\alpha \in (0, \infty)$, we have $f_{\mathbf{X}}(\mathbf{x}; \lambda, \boldsymbol{\mu}, \Gamma) = f_{\mathbf{X}}(\mathbf{x}; \alpha\lambda, \boldsymbol{\mu}, \alpha\Gamma)$. This is easily confirmed by substituting $\alpha\lambda$ and $\alpha\Gamma$ for λ and Γ, respectively, into the expression given in (10). Thus, there is no loss in generality to assume that $\det(\Gamma) = 1$ because, if it were not the case, we could always choose $\alpha = (\det\Gamma)^{-1/d}$ to make it so. Note that this is by no means the only restriction we could make on these parameters. For example, we could choose λ to be equal to 1. However, in this paper, we will assume that $\det(\Gamma) = 1$ in which case the pdf of our observation becomes

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{2\lambda^d}{\pi^d} \frac{K_{d-1}\left(2\sqrt{\lambda q(\mathbf{x})} \right)}{\left(\sqrt{\lambda q(\mathbf{x})} \right)^{d-1}}. \tag{12}$$

B. Parameter Estimation

In this section, we discuss the parameter estimation method developed in [9] and its application to the complex-valued multivariate Laplace distribution given in (12). Given the generative model of our observation and the independence assumptions in the variables Y and $\mathbf{Z}$, it is simple to see that the observation $\mathbf{X}$ has first- and second-order moments

$$\boldsymbol{\mu}_X = E[\mathbf{X}] = \boldsymbol{\mu} + E[\sqrt{Y}]\Gamma^{\frac{1}{2}} E[\mathbf{Z}] = \boldsymbol{\mu} \tag{13}$$

$$R_X = E\left[(\mathbf{X} - \boldsymbol{\mu}_X)(\mathbf{X} - \boldsymbol{\mu}_X)^H \right] \tag{14}$$

$$= E[Y]\Gamma^{\frac{1}{2}} E[\mathbf{Z}\mathbf{Z}^H]\Gamma^{\frac{H}{2}} = \frac{1}{\lambda}\Gamma. \tag{15}$$

Therefore, given M independent realizations of the random vector $\mathbf{X}$, $\{\mathbf{x}_i\}_{i=1}^M$, and the following sample estimates of these moments

$$\hat{\boldsymbol{\mu}}_X = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i \tag{16}$$

$$\hat{R}_X = \frac{1}{M} \sum_{i=1}^M (\mathbf{x}_i - \hat{\boldsymbol{\mu}}_X)(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_X)^H \tag{17}$$

simple method-of-moment-type estimators of the parameters would be the following:

$$\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}_X \tag{18}$$

$$\hat{\lambda} = (\det \hat{R}_X)^{-1/d} \tag{19}$$

$$\hat{\Gamma} = \hat{\lambda} \hat{R}_X. \tag{20}$$

Note that the estimates given above enforce the constraint that $\det(\hat{\Gamma}) = 1$.

Extracting maximum likelihood estimates of the parameters in closed-form would no doubt be a challenge in this case due to the pdf's functional dependence on a modified Bessel function. Fortunately, an alternative for achieving this when our observation depends on one or more latent random variables is the EM algorithm [8], [9]. Here, the idea is to find the parameters which maximize the joint likelihood function for the random vector $\mathbf{X}$ and random variable Y, assuming that the latent random

variable Y is known, and subsequently replace it with its least squares estimate.

Using Baye's rule, the conditional distribution of $Y|\mathbf{X}$ becomes

$$f_{Y|\mathbf{X}}(y|\mathbf{x}) = \frac{f_{\mathbf{X}|Y}(\mathbf{x}|y)f_Y(y)}{f_{\mathbf{X}}(\mathbf{x})} \tag{21}$$

$$= \frac{\left(\sqrt{\frac{q(\mathbf{x})}{\lambda}}\right)^{d-1}}{2K_{d-1}\left(2\sqrt{\lambda q(\mathbf{x})}\right)} y^{-d}\exp\left\{-\left(y\lambda + \frac{q(\mathbf{x})}{y}\right)\right\} \tag{22}$$

from which we find the following least squares estimate of the random variable Y given the multivariate observation $\mathbf{x}_i$

$$\hat{y}_i = E[Y|\mathbf{X}=\mathbf{x}_i] = \sqrt{\frac{q(\mathbf{x}_i)}{\lambda}}\frac{K_{d-2}\left(2\sqrt{\lambda q(\mathbf{x}_i)}\right)}{K_{d-1}\left(2\sqrt{\lambda q(\mathbf{x}_i)}\right)}. \tag{23}$$

Assuming that we were given independent realizations of the latent random variable Y, $\{y_i\}_{i=1}^M$, the joint pdf of the observations and these latent random variables becomes

$$f_{\mathbf{X},Y}\left(\{\mathbf{x}_i\}_{i=1}^M, \{y_i\}_{i=1}^M\right)$$

$$= \prod_{i=1}^M f_{\mathbf{X}|Y}(\mathbf{x}_i|y_i)f_Y(y_i) \tag{24}$$

$$= \frac{\lambda^M}{\pi^{dM}}\prod_{i=1}^M y_i^{-d}\exp\left\{-\sum_{i=1}^M\left(\lambda y_i + \frac{q(\mathbf{x}_i)}{y_i}\right)\right\}. \tag{25}$$

Disregarding terms that are parameter independent, we find the following log-likelihood function:

$$L(\lambda,\boldsymbol{\mu},\Gamma) = \ln f_{\mathbf{X},Y}\left(\{\mathbf{x}_i\}_{i=1}^M, \{y_i\}_{i=1}^M\right) \tag{26}$$

$$= \ln\lambda^M - \lambda\sum_{i=1}^M y_i - \sum_{i=1}^M \frac{1}{y_i}(\mathbf{x}_i-\boldsymbol{\mu})^H\Gamma^{-1}(\mathbf{x}_i-\boldsymbol{\mu}). \tag{27}$$

Maximizing this expression with respect to λ results in the estimate

$$\hat{\lambda} = \left(\frac{1}{M}\sum_{i=1}^M y_i\right)^{-1} \tag{28}$$

i.e., the maximum likelihood estimate of the rate parameter for an exponentially distributed random variable. Likewise, maximizing the log-likelihood function with respect to $\boldsymbol{\mu}$ yields

$$\hat{\boldsymbol{\mu}} = \frac{\sum_{i=1}^M \frac{1}{y_i}\mathbf{x}_i}{\sum_{i=1}^M \frac{1}{y_i}} \tag{29}$$

which is simply a weighted sample mean.

Derivation of the estimate of Γ is a little more involved as we have to maximize likelihood over a matrix-valued parameter under the restriction that its determinant is unity. For this, we consider a similar technique developed in [9] where we note that since Γ is PD there exists an upper triangular matrix U of the form

$$U = \begin{bmatrix} 1 & -u_{21} & -u_{31} & \cdots & -u_{d1} \\ 0 & 1 & -u_{32} & \cdots & -u_{d2} \\ 0 & 0 & 1 & \cdots & -u_{d3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -u_{d,d-1} \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{30}$$

along with a PD diagonal matrix $\Sigma = \text{diag}\{\sigma_1^2,\ldots,\sigma_d^2\}$ such that $U^H\Gamma U = \Sigma$. One will note that these two matrices simply correspond to a Cholesky decomposition [15] of matrix Γ and, from this decomposition, one can see that enforcing the constraint that $\det\Gamma = 1$ is equivalent to enforcing the constraint $\det\Sigma = \prod_{k=1}^d \sigma_k^2 = 1$.

Let $\mathbf{u}_k = [u_{k1}\cdots u_{k,k-1}] \in \mathbb{C}^{k-1}$ denote the unknown values in the kth column of matrix U. Defining $\mathbf{w}_k \in \mathbb{C}^M$ and $A_k \in \mathbb{C}^{M\times k-1}$

$$\mathbf{w}_k = \left[\frac{1}{\sqrt{y_1}}(x_{1k}-\mu_k)\cdots\frac{1}{\sqrt{y_M}}(x_{Mk}-\mu_k)\right]^T \tag{31}$$

$$A_k = [\mathbf{w}_1\cdots\mathbf{w}_{k-1}] \tag{32}$$

where x_{ik} and μ_k denote the kth element of vectors $\mathbf{x}_i$ and $\boldsymbol{\mu}$, respectively, then the likelihood function given in (27) can be written

$$L(\lambda,\boldsymbol{\mu},\Gamma) = \ln\lambda^M - \lambda\sum_{i=1}^M y_i - \sum_{i=1}^M \frac{1}{y_i}(\mathbf{x}_i-\boldsymbol{\mu})^H U\Sigma^{-1}U^H(\mathbf{x}_i-\boldsymbol{\mu}) \tag{33}$$

$$= \ln\lambda^M - \lambda\sum_{i=1}^M y_i - \sum_{k=1}^d \frac{1}{\sigma_k^2}\|\mathbf{w}_k - A_k\mathbf{u}_k\|^2. \tag{34}$$

From this expression it is clear that maximizing likelihood with respect to $\mathbf{u}_k$ and σ_k^2 corresponds to a linear least squares problem. To be specific, we have the following maximum likelihood estimate of vector $\mathbf{u}_k$:

$$\hat{\mathbf{u}}_k = \left(A_k^H A_k\right)^{-1} A_k^H\mathbf{w}_k. \tag{35}$$

Likewise, finding the maximum likelihood estimate of σ_k^2 then involves computing the quadratic

$$\hat{\eta}_k^2 = \frac{1}{M}\mathbf{w}_k^H P_{A_k}^{\perp}\mathbf{w}_k \tag{36}$$

and normalizing to enforce the unit determinant constraint

$$\hat{\sigma}_k^2 = \frac{\hat{\eta}_k^2}{\left(\prod_{j=1}^d \hat{\eta}_j^2\right)^{\frac{1}{d}}}. \tag{37}$$

Here, the matrix $P_{A_k} = A_k(A_k^H A_k)^{-1}A_k^H$ denotes the orthogonal projection onto the linear subspace $\langle A_k\rangle$ and $P_{A_k}^{\perp}$ denotes its orthogonal complement.

In conclusion, the EM-based estimation technique works as follows. We begin by choosing suitable initial conditions $\hat{\lambda}^{(0)}$,

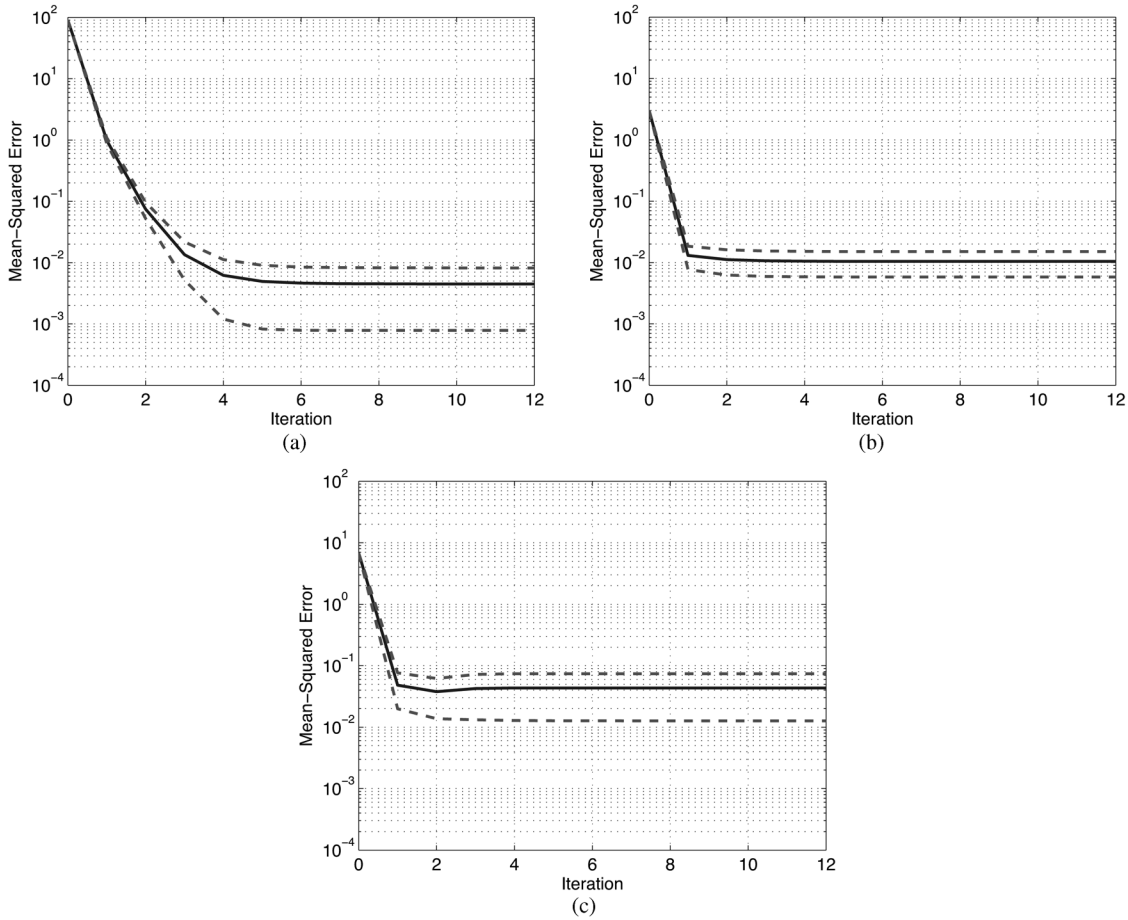


Fig. 2. Mean-squared error in parameter estimates versus iteration in the EM algorithm. Estimates for this example converge in good agreement with their true value within ten iterations. (a) Parameter estimate $\hat{\lambda}$. (b) Parameter estimate $\hat{\boldsymbol{\mu}}$. (c) parameter estimate $\hat{\Gamma}$.

$\hat{\boldsymbol{\mu}}^{(0)}$, and $\hat{\Gamma}^{(0)}$. Here, we suggest starting with the method-of-moment estimators given in (18)–(20). Given these initial parameter estimates, we then compute the least squares estimate $\hat{y}_i^{(1)}$ using the expression given in (23) and estimate the new parameters $\hat{\lambda}^{(1)}$ and $\hat{\boldsymbol{\mu}}^{(1)}$ using (28) and (29), respectively, replacing y_i with $\hat{y}_i^{(1)}$. The columns of U and diagonal elements of Σ are then computed using (35) and (37), respectively, replacing $\boldsymbol{\mu}$ with $\hat{\boldsymbol{\mu}}^{(1)}$ and y_i with $\hat{y}_i^{(1)}$. These matrices are then used to construct the covariance estimate $\hat{\Gamma}^{(1)} = U^{-H}\Sigma U^{-1}$. These updated parameters are then used to find the new least squares estimate $\hat{y}_i^{(2)}$, and the process is repeated until the change in the parameter estimates is sufficiently small. Note that the superscript notation used for the variables above denotes the values of those variables at a particular iteration in the algorithm.

C. Parameter Estimation Simulation

We conclude this section by considering a simple example to demonstrate the performance of the EM-based estimation technique through simulation. In this experiment, 500 samples of a 3-D observation are generated according to the model $\mathbf{X} = \boldsymbol{\mu}_{\text{true}} + \sqrt{Y}\Gamma_{\text{true}}^{1/2}\mathbf{Z}$ where $Y \sim \text{Exp}(\lambda_{\text{true}})$ and $\mathbf{Z} \sim \mathcal{CN}_3(\mathbf{0}, I)$. Here, we choose the rate parameter $\lambda_{\text{true}} = 1/2$, the mean vector $\boldsymbol{\mu}_{\text{true}} = \mathbf{0}$, and the following covariance matrix:

$$\Gamma_{\text{true}} = \begin{bmatrix} 0.97 & -0.49 + j0.42 & -0.36 - j0.10 \\ -0.49 - j0.42 & 2.08 & -0.33 + j0.23 \\ -0.36 + j0.10 & -0.33 - j0.23 & 0.9 \end{bmatrix}. \tag{38}$$

Confirmation of the fact that $\det(\Gamma_{\text{true}}) \approx 1$ is left to the reader. Estimates of these parameters are then computed using 12 iterations of the technique explained in Section II-B with the initial conditions $\hat{\lambda}^{(0)} = 10$, $\hat{\boldsymbol{\mu}}^{(0)} = \mathbf{1}$ (a 3-D all-one vector), and $\hat{\Gamma}^{(0)} = I$. Note that these initial parameter estimates were chosen far from their true values solely for the purposes of showing convergence. However, as explained earlier, we suggest using the method-of-moment estimators given in (18)–(20) as initial conditions.

Repeating this experiment 500 times, Fig. 2(a)–(c) displays the mean-squared error for the estimates of $\hat{\lambda}$, $\hat{\boldsymbol{\mu}}$, and $\hat{\Gamma}$, respectively, as a function of the iteration for the EM algorithm. Note that the mean-squared error for matrix Γ is measured using the Frobenius norm, i.e., $\|\Gamma_{\text{true}} - \hat{\Gamma}\|_F^2$. The dashed lines in each of these figures show one standard deviation above and below the mean-squared error. From the figure, we can observe that all parameters tend to converge in good agreement after only about five to ten iterations even with ill-chosen initial conditions.

III. Validating the Multivariate Laplace Distribution in SAS Imagery

As far as sonar data are concerned, the first and most obvious question is whether the multivariate Laplace distribution provides substantial improvement in modeling accuracy when compared to the Gaussian distribution. To try and answer this question, a test was applied to a multisonar data set collected using the small synthetic aperture minehunter (SSAM) [16]. SSAM is a dual-frequency band SAS system capable of producing high-resolution images of the seafloor. With its use of dual-frequency bands, SSAM provides detection and classification capabilities against both proud and shallow buried objects. These two frequency bands are both used to construct one HF high-resolution sonar image as well as one BB sonar image coregistered over the same region on the seafloor. Coregistration in these images is achieved as the two sonar systems are mounted on the same autonomous underwater vehicle (AUV) and use the same receive hydrophone array. The pinging for both HF and BB systems is done simultaneously as they are sufficiently far apart in frequency such that their returns are easily separable. Both HF and BB sonar images in this data set are complex valued and generated using the k-space or wave number beamformer [17].

The image database used in this study contains 61 pairs of HF and BB sonar images (both port and starboard-side images) containing 77 targets with some images containing none, one, or multiple targets. When processing the images, each pair is partitioned, with 50% overlap in both along-track and range, into regions of interest (ROIs) of size 84×144 and 28×288 for the HF and BB sonar images, respectively. The ROI size was chosen based on the average size of a target and the discrepancy among the two is due to differences in spatial resolution among HF and BB sonar. Each ROI is then partitioned with a rectangular blocking scheme with no overlap using block sizes of 6×6 and 2×12 for the HF and BB ROIs, respectively. Each pair of blocks is then vectorized and concatenated to form 336 realizations ($M = 336$) of a 60-dimensional ($d = 60$) complex-valued random vector $\mathbf{X}$.

In sonar imagery, one can often categorize realizations of background into three types: smooth, rippled, and difficult. Of course, there will always exist exceptions to this taxonomy but here we will only consider these three. Fig. 3(a)–(c) shows examples of HF ROIs corresponding to all three background types. For the test, approximately 1000 ROIs from each background type were extracted as well as all 77 ROIs containing a target.

Using all 336 realizations from each ROI, the maximum likelihood estimates $\hat{\lambda}$, $\hat{\boldsymbol{\mu}}$, and $\hat{\Gamma}$ were extracted according to the technique explained in Section II-B and used to evaluate the following log-likelihood function associated with the Laplace distribution:

$$\ell_L(\hat{\lambda}, \hat{\boldsymbol{\mu}}, \hat{\Gamma}) = \ln \prod_{i=1}^{M} f(\mathbf{x}_i; \hat{\lambda}, \hat{\boldsymbol{\mu}}, \hat{\Gamma}) \tag{39}$$

where $f(\mathbf{x}; \lambda, \boldsymbol{\mu}, \Gamma)$ denotes the multivariate Laplace pdf given in (12) and $\mathbf{x}_i$ denotes the ith pair of concatenated blocks extracted from each pair of ROIs. Likewise, the sample estimates

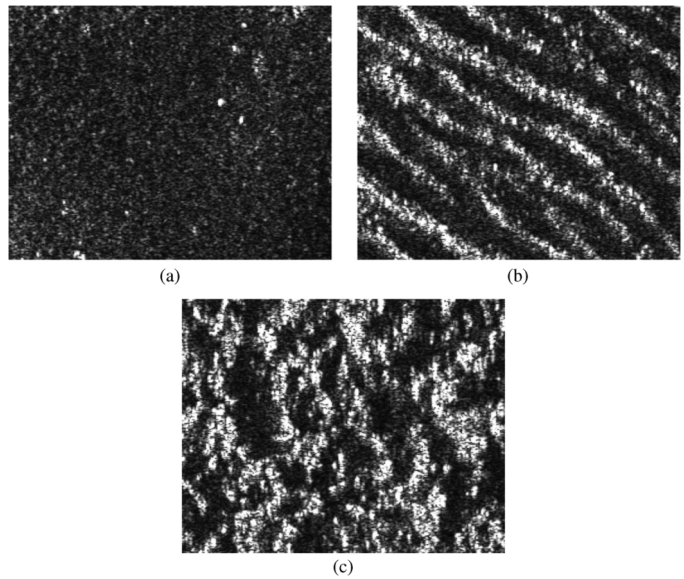


Fig. 3. Many different background textures can be encountered in sonar imagery. Several examples include smooth, rippled, and difficult clutter types. (a) Smooth texture. (b) Rippled texture. (c) Difficult texture.

$\hat{\boldsymbol{\mu}}_X$ and $\hat{R}_X$ given in (16) and (17), respectively, were also extracted and used to evaluate the following log-likelihood function associated with the Gaussian distribution:

$$\ell_N(\hat{\boldsymbol{\mu}}_X, \hat{R}_X) = \ln \prod_{i=1}^{M} f(\mathbf{x}_i; \hat{\boldsymbol{\mu}}_X, \hat{R}_X). \tag{40}$$

Here, $f(\mathbf{x}; \boldsymbol{\mu}_X, R_X)$ denotes the following multivariate complex Gaussian pdf:

$$f(\mathbf{x}; \boldsymbol{\mu}_X, R_X) = \frac{1}{\pi^d \det(R_X)} \exp\left\{-(\mathbf{x} - \boldsymbol{\mu}_X)^H R_X^{-1} (\mathbf{x} - \boldsymbol{\mu}_X)\right\}. \tag{41}$$

These values represent how well each respective model fits the data within that ROI and the two are subtracted to yield the log-likelihood ratio

$$\eta = \ell_L(\hat{\lambda}, \hat{\boldsymbol{\mu}}, \hat{\Gamma}) - \ell_N(\hat{\boldsymbol{\mu}}_X, \hat{R}_X). \tag{42}$$

Note that the objective here is not to discriminate between target and background but rather to decide which model has most likely generated the data we have observed. The larger the value of η, the more evidence exists in support of the conclusion that the multivariate Laplace more accurately describes the data.

Fig. 4 displays histograms of η for all three background types. From the figure, we can see that there is little, if any, advantage over Gaussian when modeling smooth background types with the Laplace distribution as the likelihood ratio tends to be small and even in some cases negative. However, as we move to more complicated background textures such as rippled and difficult, the Laplace distribution tends to model the data with higher accuracy as evidenced by the fact that the likelihood ratio tends to be larger.

Using the same background realizations, Fig. 5 displays the average value of η as a function of range for these three background textures. Results of this study seem to suggest that the

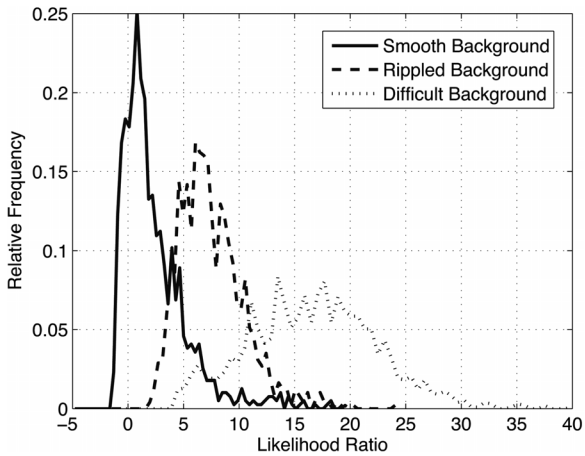


Fig. 4. Comparison of histograms of the improvement in likelihood versus background texture. The largest improvement tends to occur in rippled and difficult clutter environments.

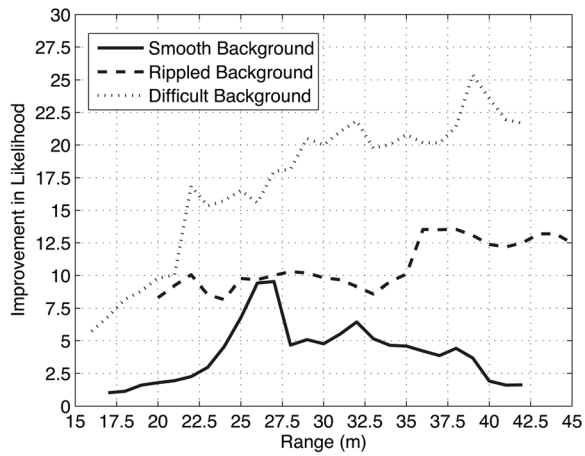


Fig. 5. Average improvement in likelihood versus range from sonar platform for all three background textures. Improvement for smooth and rippled background is fairly independent of range. However, one can see that the improvement in likelihood is directly proportional to range.

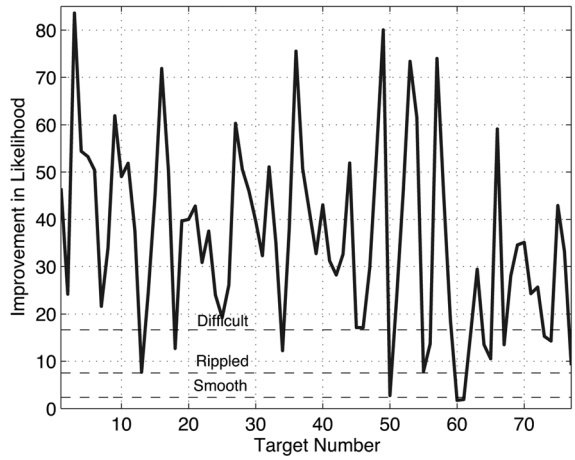


Fig. 6. Improvement in likelihood for all targets in the sonar imagery data set. Compared to the average improvement observed in background, one can see that the Laplace distribution tends to bring even larger improvements for targets.

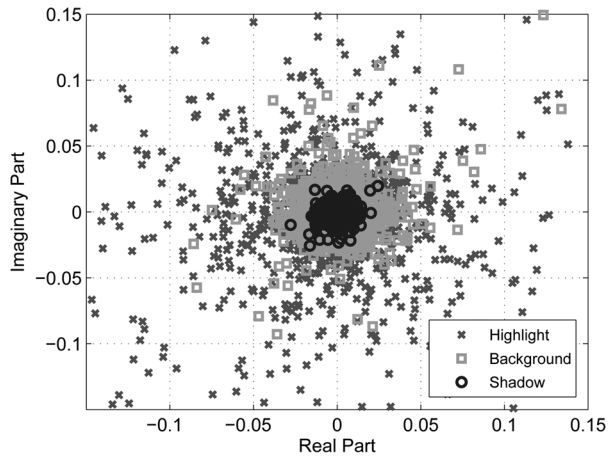


Fig. 7. Scatter plot of the highlight, shadow, and background pixels in one HF ROI containing a target. Distribution exhibits a large peak at the origin due to shadow and heavy tails due to highlight, two attributes better modeled using the Laplace distribution.

Laplace distribution is best at modeling smooth background at medium ranges and that the improvement in likelihood for rippled background is fairly independent of range. For difficult background, however, one sees a large increase in the average value of η with increasing range, which may be due to the shallower grazing angles observed at high range.

Fig. 6 displays the log-likelihood ratio given in (42) for all 77 targets as well as the average likelihood ratio for each background type which is shown with dotted lines for comparison. From this figure, we can see that the likelihood ratio for targets tends to be even larger than the average background likelihood ratio, leading us to conclude that, under most circumstances, the Laplace distribution provides a more accurate model for ROIs that contain targets as well. Note that targets like #50, #60, and #61 faintly show up in the images and thus serve as poor examples of target realizations.

Typically, when observing an ROI that contains an object of interest lying on the seafloor, one notices two defining features: a highlight region corresponding to strong sonar return from the object itself followed by a shadow region as the object blocks the sonar return immediately behind it. To investigate this phenomenon's effect on the distribution of the data and try and answer why a Laplace distribution would model target ROIs better than Gaussian, the pixels of the HF ROI corresponding to the target with the highest likelihood in Fig. 6 (target #3) were partitioned into those pixels corresponding to highlight, shadow, and background. Fig. 7 displays a scatter plot of the real and imaginary parts of the pixel values for this ROI. From here, it is clear that the distribution exhibits a significant peak at the origin due in large part to the shadow but at the same time exhibits heavy tails due solely to the highlight, which are two characteristics that are better modeled using a Laplace distribution. Much of the same argument could probably be made for complicated background textures like difficult and possibly even rippled.

IV. Detection in Sonar Imagery

In this section, we review the hypothesis test used to discriminate target from background and discuss its extension to the multivariate Laplace distribution. A simulation study is subsequently conducted to observe if and what improvement this detection strategy can bring over the classical test based on the Gaussian distribution. Finally, the performance of the detector

is evaluated on an SSAM I data set, different from that used in Section III, and compared to a similar detection technique based on the K-distribution.

A. Testing the Equality of Two Gaussian Populations and Extension to the Laplace Distribution

Consider a pooled data set of sample support $M_1 + M_2$, $\{\mathbf{x}_i\}_{i=1}^{M_1+M_2}$, that can be partitioned into two subsets, one of size M_1 and the other of size M_2. Here, it is assumed that $\mathbf{x}_i \overset{\text{iid}}{\sim} \mathcal{CN}_d(\boldsymbol{\mu}_1, R_1)$ for $i = 1, \ldots, M_1$ and

$$\mathcal{H}_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{\mu}, \quad R_1 = R_2 = R \tag{43}$$

$$\mathcal{H}_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2, \quad R_1 \neq R_2. \tag{44}$$

Implementing the generalized likelihood ratio test (GLRT) [18] involves forming sample estimates of the covariance matrices over both subsets as well as the pooled data set

$$\hat{\boldsymbol{\mu}}_1 = \frac{\sum_{i=1}^{M_1} \mathbf{x}_i}{M_1}, \quad \hat{R}_1 = \frac{\sum_{i=1}^{M_1} (\mathbf{x}_i - \hat{\boldsymbol{\mu}}_1)(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_1)^H}{M_1} \tag{45}$$

$$\hat{\boldsymbol{\mu}}_2 = \frac{\sum_{i=M_1+1}^{M_1+M_2} \mathbf{x}_i}{M_2}, \quad \hat{R}_2 = \frac{\sum_{i=M_1+1}^{M_1+M_2} (\mathbf{x}_i - \hat{\boldsymbol{\mu}}_2)(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_2)^H}{M_2} \tag{46}$$

$$\hat{\boldsymbol{\mu}} = \frac{\sum_{i=1}^{M_1+M_2} \mathbf{x}_i}{M_1 + M_2}, \quad \hat{R} = \frac{\sum_{i=1}^{M_1+M_2} (\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^H}{M_1 + M_2} \tag{47}$$

and computing the following likelihood ratio [19] (with larger values bringing more evidence in support of $\mathcal{H}_0$)

$$\Lambda = \frac{(\det \hat{R}_1)^{\beta} (\det \hat{R}_2)^{1-\beta}}{\det \hat{R}} \tag{48}$$

where $\beta = M_1/(M_1 + M_2)$ is the proportion of the samples in the pooled data set associated with the first subset. However, performing a minor amount of algebra, it is relatively simple to show that the covariance estimate for the pooled data set can actually be written as a convex combination of the covariance estimates from each individual subset plus a rank-one correction

$$\hat{R} = \beta \hat{R}_1 + (1 - \beta)\hat{R}_2 + \beta(1 - \beta)\boldsymbol{\Delta}_{\mu}\boldsymbol{\Delta}_{\mu}^H \tag{49}$$

where $\boldsymbol{\Delta}_{\mu} = \hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2$. Substituting this into (48) gives us the following likelihood ratio representing a simple test for the equality in both mean and covariance of two Gaussian populations

$$\Lambda = \frac{(\det \hat{R}_1)^{\beta} (\det \hat{R}_2)^{1-\beta}}{\det \left[\beta \hat{R}_1 + (1 - \beta)\hat{R}_2 + \beta(1 - \beta)\boldsymbol{\Delta}_{\mu}\boldsymbol{\Delta}_{\mu}^H \right]}. \tag{50}$$

Looking closely at this likelihood ratio, we can see that under the $\mathcal{H}_0$ hypothesis (i.e., background only) we are maximizing likelihood under the restriction that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ and $R_1 = R_2$ while under the $\mathcal{H}_1$ hypothesis (i.e., target plus background) no restrictions are made on these parameters.

When trying to extend the same test to the Laplace distribution, one could undoubtedly start from scratch, make the same assumptions, and simply replace the Gaussian pdf with the Laplace pdf given in (12). However, doing so would clearly lead to unappealing likelihood forms which would be difficult to compute. Additionally, it would also require extracting maximum likelihood estimates of the parameters over all $M_1 + M_2$ samples as it is doubtful that one can decompose the pooled estimates in terms of the individual ones, as seen earlier for the Gaussian distribution. That being said, one alternative would simply be to stick with the same likelihood function given in (50) but instead extract first- and second-order moments using the EM-based method discussed in Section II-B. More specifically, having estimated the parameters $\hat{\lambda}_i$, $\hat{\boldsymbol{\mu}}_i$, and $\hat{\Gamma}_i$ for $i = 1, 2$, using the EM method and recalling the fact that our observation has covariance matrix $(1/\lambda)\Gamma$ with $\det\Gamma = 1$, we consider the following likelihood ratio as an alternative:

$$\Lambda = \frac{\left(\hat{\lambda}_1^{\beta} \hat{\lambda}_2^{1-\beta}\right)^{-d}}{\det \left[\frac{\beta}{\hat{\lambda}_1}\hat{\Gamma}_1 + \frac{1-\beta}{\hat{\lambda}_2}\hat{\Gamma}_2 + \beta(1 - \beta)\boldsymbol{\Delta}_{\mu}\boldsymbol{\Delta}_{\mu}^H \right]}. \tag{51}$$

Note that it is important that we scale the matrix Γ by $1/\lambda$ as doing so leaves the second-order moment invariant to positive, real-valued scaling in the parameters λ and Γ. Again, this is exactly the same likelihood form derived from a Gaussian distribution with the only thing different being the method we use to extract estimates of the parameters. This may be unnerving as it is not in keeping with the standard theory on the optimality of likelihood ratios. However, as we will show in Section IV-B, if the data are well modeled with a Laplace distribution, we can still expect to see moderate increases in detection performance when using (51) as opposed to (50).

B. Detection Simulation

In this simulation, we study the improvement in detection performance that (51) can bring over (50) when the data are truly Laplace distributed as well as what is lost by not using the correct distribution in the construction of the likelihood ratio. For this experiment, 100 samples of a 3-D observation are generated according to the model $\mathbf{X} = \boldsymbol{\mu} + \sqrt{Y}\Gamma^{1/2}\mathbf{Z}$ with $Y \sim \text{Exp}(\lambda)$ and $\mathbf{Z} \sim \mathcal{CN}_3(\mathbf{0}, I)$. Under $\mathcal{H}_0$, all 100 samples are identically distributed with the rate parameter $\lambda = 1/2$, the mean vector $\boldsymbol{\mu} = \mathbf{0}$, and the same Γ used in the previous simulation experiments given in (38). However, under the alternative hypothesis, only the first 75 samples ($\beta = 3/4$) are generated using these parameters with the last 25 samples generated using the same rate parameter of $\lambda = 1/2$ and the same mean vector $\boldsymbol{\mu} = \mathbf{0}$ but with a perturbed covariance matrix

$$\tilde{\Gamma} = \Gamma + \epsilon \begin{bmatrix} 1 & 1+j & 1+j \\ 1-j & 1 & 1+j \\ 1-j & 1-j & 1 \end{bmatrix}. \tag{52}$$

The test described in Section IV-A is then used to decide whether the parameters of the first 75 samples match those of the last 25 samples.

With a value of $\epsilon = 0.25$, Fig. 8 displays the receiver-operator characteristic (ROC) curves for the likelihood ratio given in (50) (denoted "Gaussian"), the likelihood ratio given in (51) (denoted "quasi-Laplacian"), as well as the true GLRT for the Laplace distribution which uses the density given in

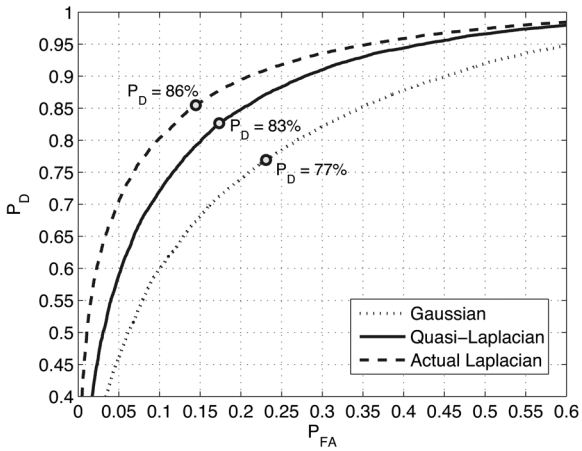


Fig. 8. Simulated ROC curves demonstrating the improvement that (51) brings over (50) when the data are truly Laplace distributed. Although there is improvement when using (51), one can see that there is a slight sacrifice in performance by not using Laplacian density function when constructing the likelihood ratio.

(12) to construct the likelihood ratio (denoted "actual Laplacian"). The knee point of each ROC curve, i.e., where $P_D + P_{FA} = 1$, is also shown with circles in this figure. As can be seen, extracting parameters using the EM algorithm and considering the test statistic given in (51) provides substantial improvement over the Gaussian case with an increase in knee point P_D of 6%. However, comparing (51) to the actual Laplacian GLRT, results show that one does indeed sacrifice some performance by not building the likelihood ratio using the density given in (12).

C. Detection Results on Dual-Channel SAS Imagery

When detecting proud objects in sonar imagery, they will often exhibit sonar returns that significantly differ from their immediate surroundings in a statistical sense. On the other hand, realizations of ROIs that solely contain background will often exhibit statistical behavior that more closely mimics that of the background that immediately surrounds it. As a consequence of these observations, our detection hypothesis for this problem is that the presence of a target in any particular ROI creates large discrepancies among the parameters that describe the data in the ROI and those that describe the data outside the ROI. However, when that same ROI only contains background we expect to observe parameters that are more closely matched with those of the surrounding data. Thus, the choice of likelihood ratio in Section IV-A is now clear as we wish to determine when the statistical characterization of a particular ROI significantly deviates from its surroundings. In doing so, the detector simply becomes one that searches for local statistical anomalies in the sonar image.

To test and compare the detection methods described in Section IV-A, the detection strategy was applied to an SSAM I image database different from that used in Section III. The image database considered here contains actual SAS images of the seafloor with synthetically generated targets inserted into the images [20]. The images are generated by inserting the simulated sonar returns from the target into actual returns from the seafloor before the beamforming process. The data set

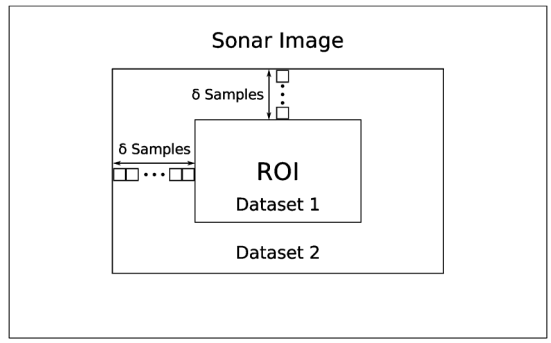


Fig. 9. Formation of the pooled data sets used for target detection. Data set 1 is formed using the data within the ROI while the second is formed using the data immediately outside that ROI.

consists of a high-resolution HF sonar image as well as one BB sonar image coregistered over the same region of the seafloor. The processing of the images to prepare them for detection is the same as that explained in the beginning of Section III, i.e., both the HF and BB images are partitioned into ROIs, each pair of ROIs is partitioned into blocks with no overlap, and the blocks are vectorized and concatenated to form multiple realizations of a complex-valued multivariate observation. For this data set, the HF and BB images were of the same resolution with both using ROI sizes of 90×144 with 6×6 block sizes resulting in 360 realizations of a 72-dimensional complex-valued observation per ROI. However, for every particular ROI, additional δ samples (blocks) are extracted in both the along-track and cross-track dimensions. Fig. 9 displays this process with the realizations from the ROI itself forming the first data set $(M_1 = 360)$ and the realizations surrounding the ROI forming the second data set. Here we choose $\delta = 4$ in which case we extract 376 samples $(M_2 = 376)$ resulting in $\beta = 0.489$.

The image database used for this test consists of 145 pairs of HF and BB sonar images with each pair corresponding to a particular data collection scenario characterized by its background and object configuration. Each image pair uses one of 29 different backgrounds that are real images of the seafloor and contains the synthetically generated signatures of four different geometrical shapes that are inserted into the image according to one of five different configurations that define the orientations and relative positions of the objects with respect to the sonar. Each pair of HF and BB images contains one of each of the following four target types: block, cone, sphere, and cylinder. The objects in these images are placed at ranges or cross-track locations varying from 12 to 40 m and at aspect angles varying from 25° to 75°. Fig. 10(a) and (b) shows examples of the HF and BB images, respectively, for one background and object configuration. All of the four objects in this image are outlined with a white box. Fig. 11 displays η, i.e., the improvement in likelihood given in (42), for every pair of ROIs within these images. Recalling the discussion in Section III, we can again see that the Laplace distribution tends to bring larger improvement for targets and more complicated clutter.

As the methods in this paper revolve around the use of complex-valued data, one important question is if there exists any useful information in the phase of the images and whether it is important for detection. To try and answer this question, a

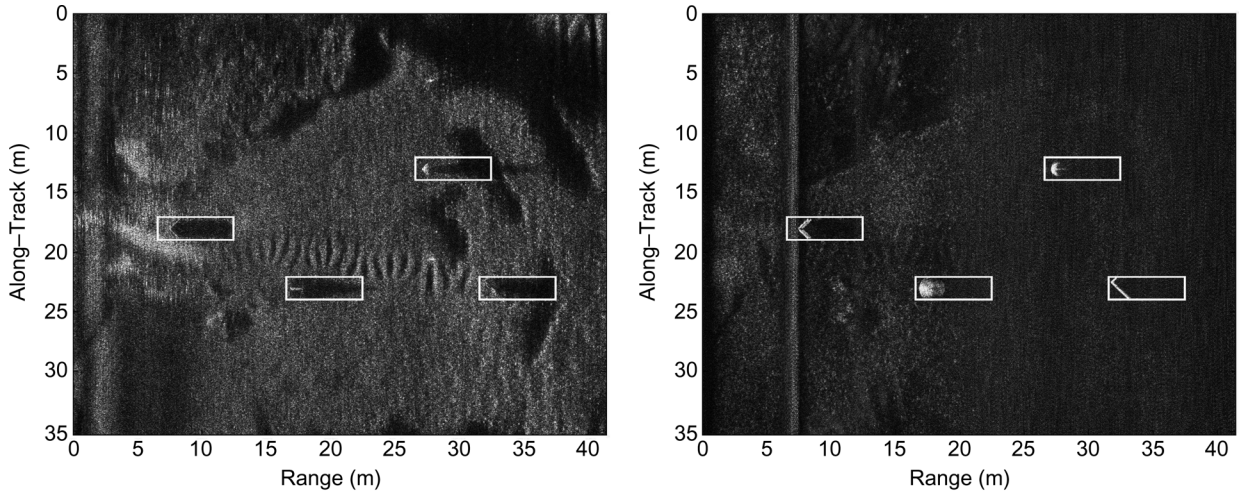


Fig. 10. Examples of a corresponding pair of HF and BB sonar images used for the detection studies conducted here. In these images, targets are outlined with a white box. (a) HF image. (b) BB image.

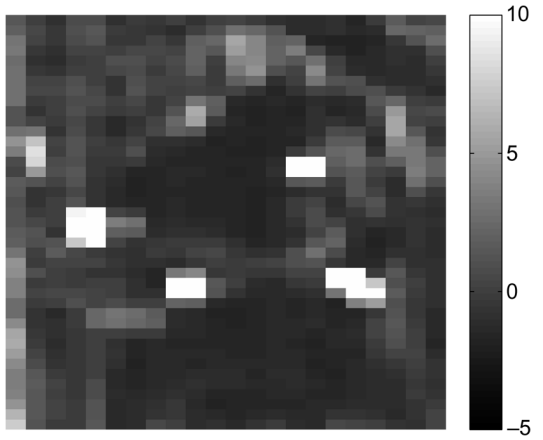


Fig. 11. Improvement in likelihood for the images shown in Fig. 10. Similar to the conclusions made in Section III, the greatest improvement in likelihood comes from targets and regions with more complicated clutter.

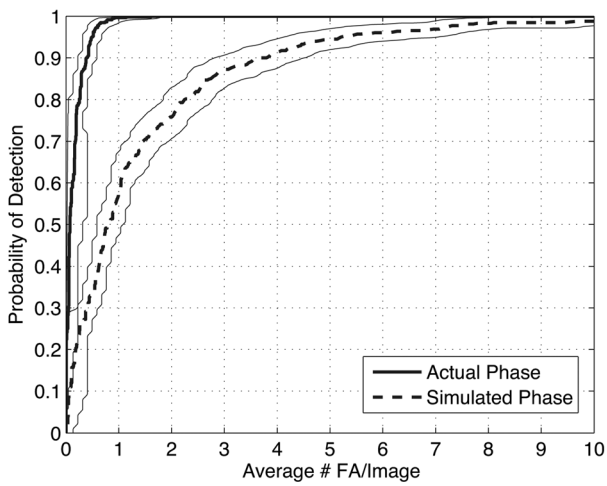
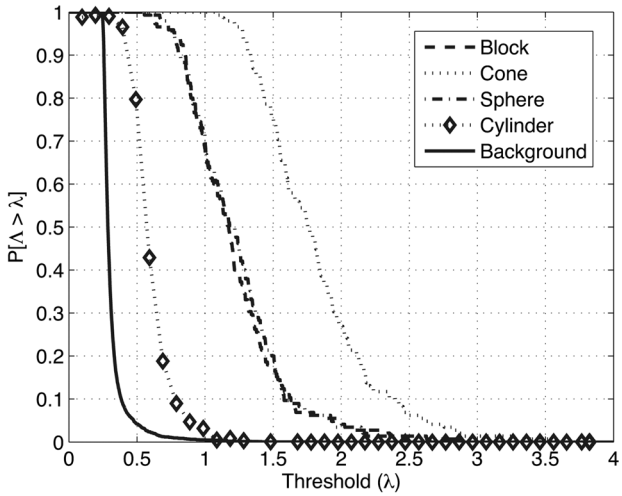


Fig. 12. ROC curves comparing the Laplacian detector using both actual and simulated phases. The results of this study suggest that the information contained in the phase of the images is useful for detection.

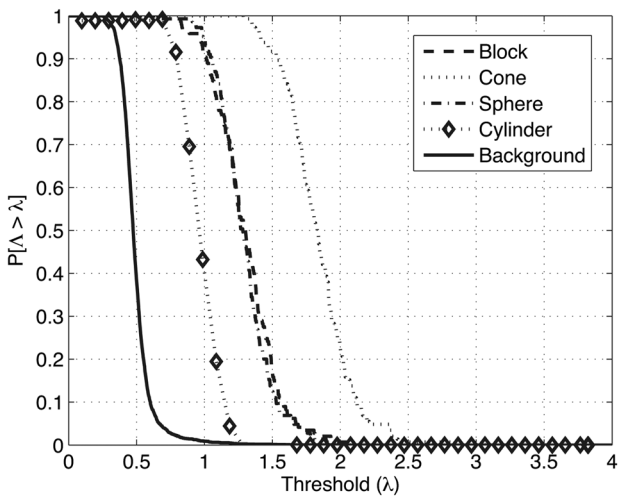
study was conducted by extracting all ROIs containing a target as well as a set of randomly extracted ROIs containing only background. Each of these ROIs is guaranteed to be distinct by ensuring that there is at least a 4-m separation among them. The detector given in (51) was then applied to this set of target and background realizations and compared to a situation where the magnitude of both HF and BB images is retained but the phase of each image is replaced with iid realizations of a random variable uniformly distributed over the interval $[-\pi, \pi]$. The purpose of this study is to see if there is any loss in performance when replacing the phase with unstructured randomness. Fig. 12 displays the ROC curves for both situations as well as a 95% confidence interval [21] around each curve. From this figure, it is clear that replacing the phase of the image with random values does have a negative impact on the detector. This suggests that, for this particular detector at least, there is indeed information in the phase which is useful for detection.

Both the Gaussian and Laplacian detectors given in (50) and (51), respectively, were then applied to the same target and background realizations used to generate Fig. 12. Note that the actual phase of the images is used throughout the rest of the paper. Fig. 13(a) and (b) displays right tail probabilities $P[\Lambda > \lambda]$, as a function of the threshold λ for the Gaussian and Laplacian detectors, respectively, and for all four object types along with background. In other words, for a given value of λ along the x-axis of these plots, each curve gives the percentage of ROIs containing either background or one of the four object types that fall above that threshold. Table I lists the average signal-to-clutter ratio (SCR) for each object type found by computing the variance of the pixels within each target ROI over the variance of the pixels surrounding that ROI. As the SCR can vary depending on the clutter within the image and where the target is inserted into the image, the values for SCR listed in this table represent the average over all 29 backgrounds and five configurations of the data set. From these figures and the SCR given in this table, it is clear that a cone is the easiest object type to detect in this data set with a block and a sphere coming next and more difficulty in detecting a cylinder.

Fig. 14 likewise displays ROC curves for both of these detection techniques [denoted "Gaussian" and "Laplacian (HF & BB)"] along with a 95% confidence interval. Note that these curves give a combined probability of detection over all four



(a)



(b)

Fig. 13. Detection rates for all four object types and background in the data set. These results show that it is easy to discriminate the cone object type from background with increasing difficulty in the sphere, block, and cylinder object types, respectively. (a) Gaussian. (b) Laplacian.

TABLE I
AVERAGE SCR VERSUS OBJECT TYPE

Block	Cone	Sphere	Cylinder
0.42 dB	1.44 dB	0.56 dB	-0.38 dB

object types. As discussed in Section I, many non-Gaussian models have been considered and shown to be adequate in modeling amplitude sonar data. One model in particular, which has shown great promise for high-resolution SAS systems, is the K-distribution [3], [5]. For this reason, these two methods were also compared to the likelihood ratio given in (60) of the Appendix (denoted "K-distribution" in Fig. 14) which relies on the same detection principle described in Section IV-A but fits a K-distribution to the magnitude of the pixels in the HF image only. Another theme in this paper is the use of both HF and BB images for target detection. To see if and what improvement the inclusion of the BB image brings to the detector, Fig. 14 also shows the results of using (51) with the HF image alone [denoted "Laplacian (HF only)"]. The general overview of the characteristics of these four detectors in terms of which

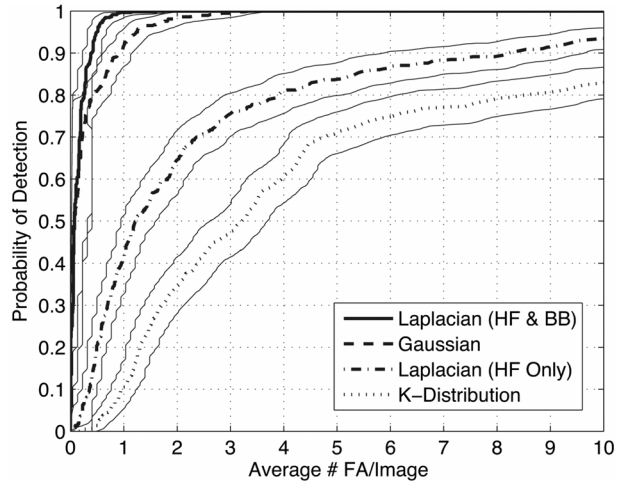


Fig. 14. Comparison of the ROC curves for each detection method with their associated confidence intervals. For this study, the methods employing dual-sonar, multivariate observations clearly outperform their alternatives.

TABLE II
CHARACTERISTICS OF EACH DETECTION METHOD

Detection Method	HF Image	BB Image	Complex Valued	Multivariate
Laplacian (HF & BB)	Yes	Yes	Yes	Yes
Laplacian (HF Only)	Yes	No	Yes	Yes
Gaussian	Yes	Yes	Yes	Yes
K-Distribution	Yes	No	No	No

images they use and whether they employ complex-valued data and multivariate observations is given in Table II. From the results of this figure, we can see that the likelihood ratio given in (51) for the Laplacian detector, and more importantly the estimation procedure that goes into the construction of this likelihood ratio, provides a significant improvement over the performance of the likelihood ratio given in (50). One can also see that these two detectors bring significant improvement over both the Laplacian detector employing the HF image only and the detector based on the K-distribution.

As a final evaluation of the detection methods considered here, a threshold was selected for all four detection methods such that each achieves an average of one false alarm per image. Each pair of HF and BB images is then partitioned into ROIs with 50% overlap and the likelihood ratio from each method is compared to its corresponding threshold. All overlapping ROIs that pass the test are accumulated into one detection or contact. If the location of a contact is within 2.5 m of the known location of a target, the contact is labeled a correct detection. Otherwise, the contact is labeled as a false alarm. With this strategy and the chosen thresholds, Table III gives the number of detected targets for each object type, the overall number of detected targets for each detection method, and the average SCR for each object type. Note that the numbers in parentheses at the top of this table show the total number of each object type as well as the total number of targets in the data set. From this table, we can again see that the Laplacian and Gaussian detectors employing both images outperform the other two detection methods considered here. For both Laplacian and Gaussian detectors, one can also see that all three methods tend to perform relatively well for block, cone, and sphere target types. The performance

TABLE III
OBJECT SPECIFIC AND OVERALL DETECTION RESULTS

Detection Method	Block (145)	Cone (145)	Sphere (145)	Cylinder (145)	Total (580)
Laplacian (HF & BB)	144	145	144	135	568
Laplacian (HF Only)	100	78	90	5	273
Gaussian	145	145	145	98	533
K-Distribution	40	67	6	68	181

for cylinder, which exhibits a relatively low SCR in Table I, is worse for these three methods. For the K-distribution, however, the detector performs roughly the same for block, cone, and cylinder types, but struggles with the sphere type. Looking at the images shown in Fig. 10, this may be because the sphere target has a relatively small highlight signature and narrow shadow which may present a greater difficulty for pixel-based methods. Overall, we can see that the detector based on the Laplace distribution performs extremely well on this data set in comparison to the classical test based on assumptions of normality as the Laplacian detector achieves $P_D = 98\%$ while the Gaussian detector only achieves $P_D = 92\%$ when comparing these two methods at a comparable false alarm rate.

V. CONCLUSION AND FUTURE WORK

In this paper, we have developed a complex-valued multivariate Laplace distribution for the purposes of more accurately modeling the complex-valued multivariate data in HF and BB sonar imagery and thus building detectors that are able to more accurately resolve two competing hypotheses. The detection method is based on a multivariate extension of the Laplace distribution derived using a scale mixture of Gaussian distributions. Deriving expressions for the pdf of this statistical model, maximum likelihood estimates are found recursively using the EM algorithm. A goodness-of-fit test in the form of a likelihood ratio is then conducted on a sonar imagery data set consisting of HF and BB images coregistered over the same region on the seafloor showing the proposed model's applicability in sonar imagery. The results from this test show that the Laplace distribution provides the greatest improvement in modeling accuracy for targets and more complicated clutter density. This is most likely due to highlight and shadow characteristics typically observed in sidescan sonar. Detection based on testing the equality of parameters from two populations is then extended to the Laplace distribution. A comparison study involving several different detection paradigms, including one that fits a K-distribution to the magnitude of the image, was implemented on a sonar imagery data set containing actual SAS images of the seafloor with synthetically generated targets inserted into the images. Setting thresholds such that each method achieves an average of one false alarm per image, the detection method based on the Laplace distribution was shown to improve the detection probability by approximately 6% compared to the Gaussian counterpart.

The detection methods considered in this paper revolved around the idea of using the same likelihood ratio function derived from assumptions of normality and simply considering alternative methods of estimating the first- and second-order moments of the observation. Although we have shown that this idea will produce detection methods that generally outperform the Gaussian detector when the data are well modeled with a Laplace distribution, future work should consider the development of efficient likelihood forms explicitly derived for the multivariate Laplace distribution. Another line of future research might also be the development of environmentally adaptive detection systems that are capable of deciding *in situ* which statistical model is more appropriate based on the background characteristics in the sonar image. One of the disadvantages of the Laplace distribution is the increase in computational burden as one must resort to iterative methods when extracting maximum likelihood estimates of the parameters. However, if environmental context is available, one can use the Laplacian-based detector in this paper only when it is needed.

APPENDIX

In this Appendix, we discuss a detection strategy similar to that presented in Section IV-A but for the K-distribution. Again, we assume that there exists a real-valued, scalar data set $\{x_i\}_{i=1}^{M_1+M_2}$ that can be partitioned into two subsets, each of size M_1 and M_2. Similar to the discussion given at the beginning of Section IV-C, the first subset represents the magnitude of the pixels within an ROI of the image while the second represents the magnitude of the pixels immediately surrounding that ROI, as shown in Fig. 9. Here, it is assumed that $x_i \overset{\text{iid}}{\sim} f(x_i;\nu_1,\alpha_1)$ for $i = 1,\ldots,M_1$ and $x_i \overset{\text{iid}}{\sim} f(x_i;\nu_2,\alpha_2)$ for $i = M_1+1,\ldots,M_1+M_2$, where $f(x;\nu,\alpha)$ denotes the pdf of a K-distribution [3], [5] with shape and scale parameters ν and α

$$f(x;\nu,\alpha) = \frac{2\alpha}{\Gamma(\nu)}\left(\frac{\alpha x}{2}\right)^{\nu} K_{\nu-1}(\alpha x), \qquad x_i \geq 0 \tag{53}$$

and $K_\nu(x)$ is the modified Bessel function given in (6). In keeping with the detection principle discussed in Section IV-C, our detection hypothesis is that the presence of a target in any particular ROI creates large discrepancies among the parameters that describe the data in the ROI and those that describe the data outside the ROI. However, when that same ROI only contains background, we expect to observe parameters that are more closely matched with those of the surrounding data. Thus, our objective is to determine whether the distribution of the data within the ROI deviates from the distribution describing the data outside the ROI by considering the hypothesis test

$$\mathcal{H}_0 : \nu_1 = \nu_2 = \nu, \quad \alpha_1 = \alpha_2 = \alpha \tag{54}$$

$$\mathcal{H}_1 : \nu_1 \neq \nu_2, \quad \alpha_1 \neq \alpha_2. \tag{55}$$

The kth subset for $k = 1, 2$ has the likelihood function

$$\ell_k(\nu_k,\alpha_k) = \prod_{i\in\mathcal{S}_k} f(x_i;\nu_k,\alpha_k) \tag{56}$$

$$= \left(\frac{\alpha_k^{\nu_k+1}}{2^{\nu_k-1}\Gamma(\nu_k)}\right)^{M_k} \prod_{i\in\mathcal{S}_k} x_i^{\nu_k} K_{\nu_k-1}(\alpha_k x_i) \tag{57}$$

where $\mathcal{S}_1 = \{1, \ldots, M_1\}$ and $S_2 = \{M_1 + 1, \ldots, M_1 + M_2\}$. Likewise, under the constraint that $\nu_1 = \nu_2 = \nu$ and $\alpha_1 = \alpha_2 = \alpha$, the entire data set has the combined likelihood function

$$\begin{aligned}\ell_{\text{total}}(\nu, \alpha) &= \prod_{i \in \mathcal{S}_1 \cup \mathcal{S}_2} f(x_i; \nu, \alpha) \\ &= \left(\frac{\alpha^{\nu+1}}{2^{\nu-1}\Gamma(\nu)} \right)^{M_1+M_2} \prod_{i \in \mathcal{S}_1 \cup \mathcal{S}_2} x_i^{\nu} K_{\nu-1}(\alpha x_i).\end{aligned} \tag{58}$$

Implementing the GLRT in this case involves evaluating (57) and (58) using ML estimates of the parameters ν and α and computing the likelihood ratio

$$\Lambda = \frac{\max\limits_{\nu,\alpha} \ell_{\text{total}}(\nu, \alpha)}{\left(\max\limits_{\nu_1,\alpha_1} \ell_1(\nu_1, \alpha_1) \right) \left(\max\limits_{\nu_2,\alpha_2} \ell_2(\nu_2, \alpha_2) \right)}. \tag{59}$$

Unfortunately, finding ML estimates of the parameters for the K-distribution is an intractable problem with no closed-form solution. As an alternative, the estimates $\hat{\nu}$ and $\hat{\alpha}$ are found using the Bayesian method-of-moments technique developed in [22] and used to compute the likelihood ratio

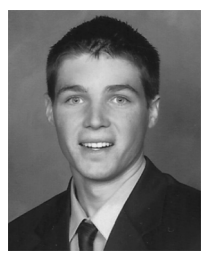
$$\tilde{\Lambda} = \frac{\ell_{\text{total}}(\hat{\nu}, \hat{\alpha})}{\ell_1(\hat{\nu}_1, \hat{\alpha}_1)\ell_2(\hat{\nu}_2, \hat{\alpha}_2)}. \tag{60}$$

ACKNOWLEDGMENT

The authors would like to thank the Naval Surface Warfare Center Panama City Division (NSWC PCD, Panama City, FL, USA) for support and providing the data used in this study.

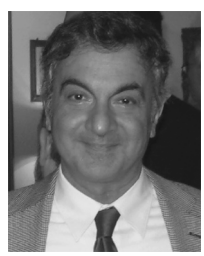
REFERENCES

[1] S. Stanic and E. Kennedy, "Reverberation fluctuation from a smooth seafloor," *IEEE J. Ocean. Eng.*, vol. 18, no. 2, pp. 95–99, Apr. 1993.
[2] S. Stanic, R. Goodman, K. Briggs, N. Choliros, and E. Kennedy, "Shallow-water bottom reverberation measurements," *IEEE J. Ocean. Eng.*, vol. 23, no. 3, pp. 203–210, Jul. 1998.
[3] J. Cobb, K. Slatton, and G. Dobeck, "A parametric model for characterizing seabed textures in synthetic aperture sonar images," *IEEE J. Ocean. Eng.*, vol. 35, no. 2, pp. 250–266, Apr. 2010.
[4] D. Abraham, "Detection-threshold approximation for non-Gaussian background," *IEEE J. Ocean. Eng.*, vol. 35, no. 2, pp. 355–365, Apr. 2010.
[5] J. Tucker, J. Cobb, and M. Azimi-Sadjadi, "Generalized likelihood ratio test for finite mixture model of K-distributed random variables," in *Proc. IEEE Digital Signal Process. Workshop*, 2011, pp. 443–448.
[6] J. D. Tucker and M. R. Azimi-Sadjadi, "Neyman Pearson detection of K-distributed random variables," *Proc. SPIE—Int. Soc. Opt. Eng.*, vol. 7664, pp. 1–12, Apr. 2010.
[7] N. Klausner and M. Azimi-Sadjadi, "Non-Gaussian target detection in sonar imagery using the multivariate Laplace distribution," in *Proc. OCEANS Conf.*, 2011, pp. 1–10.
[8] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc.*, vol. 39, no. 1, pp. 1–38, 1977.
[9] T. Eltoft, T. Kim, and T. Lee, "On the multivariate Laplace distribution," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 300–303, May 2006.
[10] D. Andrews and C. Mallows, "Scale mixtures of normal distributions," *J. Roy. Stat. Soc.*, vol. 36, pp. 99–102, 1974.
[11] D. Anderson, "A multivariate Linnik distribution," *Stat. Probab. Lett.*, vol. 14, pp. 333–336, 1992.
[12] S. Kotz, T. Kozubowski, and K. Podgorski, *The Laplace Distribution and Generalizations: A Revisit With Applications to Communications, Economics, Engineering, Finance*. Boston, MA, USA: Birkhauser, 2001, ch. 5.
[13] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York, NY, USA: Dover Publications, 1965, ch. 9.
[14] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Stat.*, vol. 34, no. 1, pp. 152–177, 1963.
[15] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: John Hopkins Univ. Press, 1996, ch. 4.
[16] D. Brown, D. Cook, and J. Fernandez, "Results from a small synthetic aperture sonar," in *Proc. OCEANS Conf.*, Sep. 2006, DOI: 10.1109/OCEANS.2006.306855.
[17] R. Bamler, "A comparison of range-Doppler and wavenumber domain SAR focusing algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 4, pp. 706–713, Jul. 1992.
[18] K. Mardia, J. Kent, and J. Bibby, *Multivariate Analysis*. New York, NY, USA: Academic, 1979, ch. 5.
[19] R. Muirhead, *Aspects of Multivariate Statistical Theory*. Hoboken, NJ, USA: Wiley, 2005, ch. 10.8.
[20] G. Sammelmann, "High-frequency images of proud and buried 3D-targets," in *Proc. OCEANS Conf.*, Sep. 2003, vol. 1, pp. 266–272.
[21] J. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 251–255, Apr. 2008.
[22] D. Abraham and A. Lyons, "Reliable methods for estimating the K-distribution shape parameter," *IEEE J. Ocean. Eng.*, vol. 35, no. 2, pp. 288–302, Apr. 2010.



Nick Klausner (S'08) received B.S. degrees in electrical engineering and economics from the Colorado School of Mines, Golden, CO, USA, in 2008 and the M.S. degree in electrical and computer engineering from Colorado State University, Fort Collins, CO, USA, in 2010, where he is currently working toward the Ph.D. degree in electrical and computer engineering under the direction of Dr. M. R. Azimi-Sadjadi.

His current research interests include statistical signal processing, digital signal processing, and digital image processing.



Mahmood R. Azimi-Sadjadi (M'81–SM'89) received the M.S. and Ph.D. degrees in electrical engineering with specialization in digital signal/image processing from the Imperial College of Science and Technology, University of London, London, U.K., in 1978 and 1982, respectively.

He is currently a Full Professor with the Electrical and Computer Engineering Department, Colorado State University (CSU), Fort Collins, CO, USA. He is also serving as the Director of the Digital Signal/Image Laboratory at CSU. He is the coauthor of the book *Digital Filtering in One and Two Dimensions* (New York, NY, USA: Plenum, 1989). His main areas of interest include statistical signal and image processing, target detection, classification and tracking, sensor array processing, learning systems, and distributed sensor networks.

Dr. Azimi-Sadjadi served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON NEURAL NETWORKS.