

Performance Prediction and Estimation for Underwater Target Detection Using Multichannel Sonar

Nick H. Klausner, *Member, IEEE*, and Mahmood R. Azimi-Sadjadi , *Life Member, IEEE*

Abstract—A critical need in the development of any automatic target recognition system is the ability to accurately predict and quantify the detection and classification performance under various operational and environmental conditions. In this paper, we propose new methods capable of predicting and estimating the performance of a multichannel detection system using multiple synthetic aperture sonar. Performance prediction and estimation is accomplished by analyzing the multichannel coherence statistics and characterizing the background conditions within the image. The ability of the method to provide an assessment of image complexity for various background conditions is studied. The saddlepoint approximation is employed to approximate the empirical null distribution of the test statistics for threshold selection and to achieve a prescribed false alarm rate. Test results on two real and one synthetic sonar imagery data sets with different target and background conditions are provided, which indicate the capability of the proposed methods in describing the distribution of the likelihood ratio and predicting the detector's performance in low to medium clutter environments.

Index Terms—Detection performance prediction, estimation, and optimization, generalized likelihood ratio test, multichannel coherence detector, saddlepoint approximation, synthetic aperture sonar, underwater target detection.

I. INTRODUCTION

THE task of automatic detection and classification of underwater objects in sonar imagery is a highly challenging problem. For such a problem, human operators typically outperform computer algorithms in moderate and difficult operating environments due to human proficiency in incorporating *in situ* information and context. Moreover, human operators are much more adept at characterizing the performance degradation risks when operating in a different environment. The performance of any sonar system invariably depends on the underwater environment and the interactions of the targets of interest with

that environment. Therefore, assessing the performance of the target detector and classifier in an automatic target recognition (ATR) system is of utmost importance for any fully autonomous sonar platform used for mine countermeasure operations to mitigate risk when operating in new environments. This involves developing methods capable of characterizing an algorithm's performance *a priori* before deployment (performance prediction), characterizing its performance as data are collected from the environment *in situ* (performance estimation), and possibly adapting the system to achieve a prescribed performance goal (performance optimization).

Sha and Nolte [1] developed closed-form expressions for the receiver operating characteristic (ROC) performance of the Bayesian detector in the presence of interference in uncertain environments. The results indicated the dependence of the detection performance on not only the signal-to-interference ratio, but also to the rank of the signal matrix that characterizes the degree of environmental uncertainty. Liu *et al.* [2] analyzed the performance of three commonly used adaptive detectors namely the subspace-based generalized likelihood ratio test (GLRT), subspace-based adaptive matched filter, and adaptive subspace detector in the presence of mismatch between the assumed signal subspace and the actual signal subspace. They obtained analytical expressions for the probabilities of detection and false alarm and showed how the detection performance of these detectors is impacted by a distance measure between the actual and the assumed signal subspaces.

As for sonar data, the work in [3] presented a performance prediction tool for target classification using sidescan sonar, which uses k -distribution to represent a wide variety of background statistics depending on the complexity of the environment. The *in situ* measurements of the parameters can be combined with other prior specifications to offer good performance estimation. In [4], a performance analysis and training tool module was used that allows the user to perform detection risk analysis by inserting a simulated target into sonar data for various background types. Controlled what if scenarios can then be studied by varying sonar range, target shapes, types, and orientations. Midtgaard *et al.* [5] proposed to use spatial coherence measure estimated from interferometric processing of two sonar receivers data with a displacement. They showed how the coherence can be converted into signal-to-noise ratio (SNR), which is indicative of performance based upon the local properties of the data.

Manuscript received April 13, 2018; revised September 14, 2018 and October 29, 2018; accepted November 8, 2018. Date of publication January 8, 2019; date of current version April 14, 2020. This work was supported by the U.S. Office of Naval Research under Contracts N00014-12-C-0017 and N00014-15-C-5054. (Corresponding author: Mahmood R. Azimi-Sadjadi.)

Associate Editor: T.-O. Sæbø.

N. H. Klausner was with Information System Technologies, Inc., Fort Collins, CO 80525 USA (e-mail: nklausne@gmail.com).

M. R. Azimi-Sadjadi is with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523 USA (e-mail: azimi@engr.colostate.edu).

Digital Object Identifier 10.1109/JOE.2018.2881527

Williams [6] proposed a fast method for adaptive detection of proud underwater targets that uses environmental characteristics collected *in situ*. A comprehensive study was conducted on a large database to determine detection performance as a function of target type, aspect, range, image quality, and seabed characteristics. The purpose was to identify the key factors that have the largest impact on detection performance.

The aforementioned references identified a variety of factors that can affect the performance of any ATR algorithm for synthetic aperture sonar (SAS) imagery among which are image *quality* and *complexity* [7]. Image quality [8], [9] refers to the perceived degradation of the image that can be influenced by things such as low SNR, an unknown sound-speed profile, vehicle motion, as well as a variety of other environmental factors. The work in [8] gave a model for SAS performance evaluation based upon local image quality. The developed model has modules for SNR, geometric resolution, radiometric resolution, and image sharpness. Image complexity [10], [11], on the other hand, refers to the difficulty of the natural scene owing to various types of bottom conditions and the presence of background clutter. To this end, the main goal of this paper is to investigate means of not only identifying various background conditions, but also quantifying their complexity by analyzing the effects on target detection performance.

The theme of this paper differs, in many ways, from those in all the above-mentioned references. First, a multichannel detection method [12] is proposed that can be applied to detect targets from complex-valued (e.g., output of the beamformer) data collected from multiple (two or more) sonar systems with disparate frequency, bandwidth, and/or location. In this paper, the channels correspond to vectorized blocks of data taken from two coregistered high-frequency (HF) and broadband (BB) SAS images. Our detection principle relies on the assumption that the presence of targets will lead to a higher degree of *multichannel coherence* than when those image blocks contain background alone. Second, under the null hypothesis (i.e., background noise alone) the multichannel coherence statistic has a distribution, which is theoretically equivalent to a product of independent but not identically distributed beta random variable [12]. The significance of this result is that this explicit null distribution can serve as a yardstick with which one can benchmark the statistical behavior of the multichannel coherence in different background conditions to predict the performance. More specifically, analyzing the deviation of the coherence statistics from those under the null hypothesis gives an assessment of image complexity and hence the difficulty of the background condition. Third, depending on the image complexity measure, the saddlepoint method [13] may be applied to produce accurate approximations of the distribution function of the likelihood ratio for the purposes of predicting the false alarm performance. This is because it is difficult to derive closed-form expressions for the likelihood ratio's distribution function especially for large number of channels and/or high dimensional channels. Finally, to optimize the detection performance in high cluttered environments an *implicit* saddlepoint approximation method was suggested, which uses the estimated cumulants to adapt the detection threshold to maintain a desired false alarm performance. The

applicability of the proposed methods in various background and target conditions is demonstrated on actual HF and BB sonar images collected from the Small Synthetic Aperture Minehunter (SSAM) I system.

This paper is organized as follows. Section II gives a review of the multichannel GLRT-based detector and its efficient implementation using the broadband coherence test, discusses the null distribution of the test statistic, and develops a method for approximating the null distribution using the saddlepoint approximation method. Section III gives a description of the data sets used in this study and an overview of their properties. Section IV gives the results of using coherence statistics for image complexity assessment and the saddlepoint approximation for performance prediction and estimation. The ability of the detector to detect targets from background under various conditions including aspect angle and range is also demonstrated. Finally, Section V provides concluding remarks on this paper.

II. GLRT-BASED DETECTION IN MULTISONAR IMAGERY

A. Review of the Multichannel GLRT Detector

Consider a set of L complex-valued data vectors $\{\mathbf{x}_i\}_{i=1}^L$ with $\mathbf{x}_i = [x_i[0] \cdots x_i[N-1]]^T \in \mathbb{C}^N$. In this particular application, the random vector \mathbf{x}_i represents a vectorized block extracted from the i th sonar (channel) image in which case $x_i[n]$ corresponds to the n th pixel within that block. The data from all L sonar channels can be concatenated to form the composite vector $\mathbf{z} = [\mathbf{x}_1^T \mathbf{x}_2^T \cdots \mathbf{x}_L^T] \in \mathbb{C}^{LN}$. Assuming zero-mean condition, the composite vector \mathbf{z} exhibits the block-structured covariance matrix $\mathbf{R} \in \mathbb{C}^{LN \times LN}$

$$\mathbf{R} = E[\mathbf{z}\mathbf{z}^H] = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1L} \\ \mathbf{R}_{12}^H & \mathbf{R}_{22} & \cdots & \mathbf{R}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{1L}^H & \mathbf{R}_{2L}^H & \cdots & \mathbf{R}_{LL} \end{bmatrix} \quad (1)$$

where $\mathbf{R}_{i\ell} = E[\mathbf{x}_i \mathbf{x}_\ell^H] = \mathbf{R}_{\ell i}^H \in \mathbb{C}^{N \times N}$ is the cross-covariance matrix (autocovariance when $i = \ell$) between channels i and ℓ . Without making any *a priori* assumption about the signal observed by each sensor, one intuitive way of determining if a common signal exists among all L channels is to test for deviations from statistical independence. If the set of random vectors $\{\mathbf{x}_i\}_{i=1}^L$ is jointly proper complex normal, testing for independence among all L channels becomes a test of whether $\mathbf{R}_{i\ell} = \mathbf{O}$ for all $i \neq \ell$. Consequently, the null hypothesis (i.e. no target present) \mathcal{H}_0 is the hypothesis that \mathbf{R} is block diagonal versus the alternative (i.e. target present) \mathcal{H}_1 that it is not. This implies that under the null hypothesis all L channels are spatially uncorrelated yet possibly temporally correlated versus the alternative that they are both spatially and temporally correlated.

We now assume we are given an experiment producing M iid realizations $\{\mathbf{x}_i[m]\}_{m=1}^M$ of the random vector from each channel i . The composite vectors $\mathbf{z}[m] = [\mathbf{x}_1^T[m] \cdots \mathbf{x}_L^T[m]]^T$

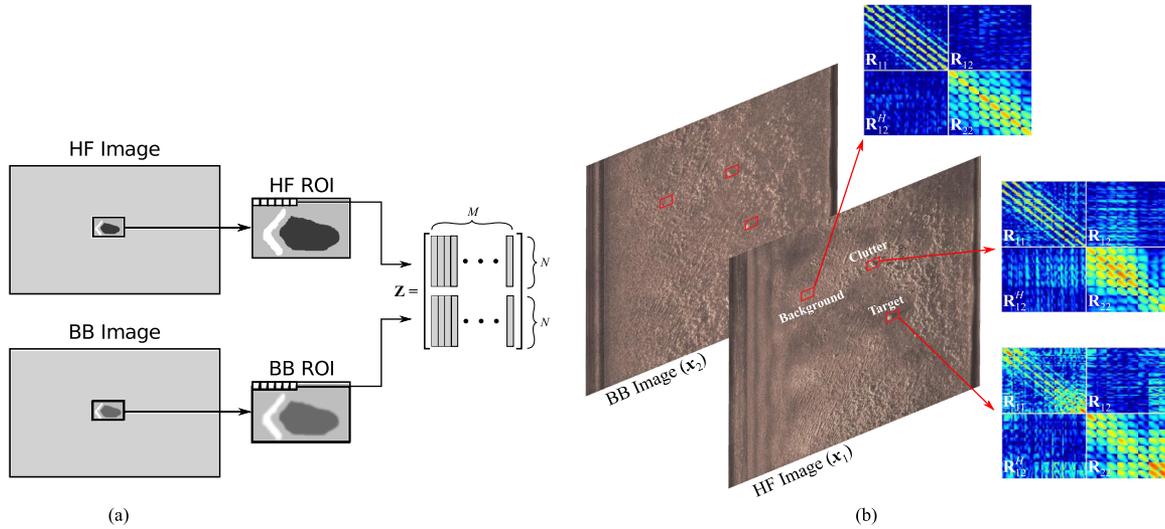


Fig. 1. (a) Data matrix \mathbf{Z} is formed by partitioning each pair of coregistered HF and BB ROIs into non-overlapping blocks and vectorizing to yield $M = 180$ realizations of an $N = 64$ dimensional observation per channel. (b) Several pairs of HF and BB ROIs and their corresponding composite covariance matrices.

are organized into the data matrix

$$\mathbf{Z} = [\mathbf{z}[1] \cdots \mathbf{z}[M]] = \begin{bmatrix} \mathbf{x}_1[1] & \cdots & \mathbf{x}_1[M] \\ \vdots & \ddots & \vdots \\ \mathbf{x}_L[1] & \cdots & \mathbf{x}_L[M] \end{bmatrix} \in \mathbb{C}^{LN \times M}. \quad (2)$$

The data matrix \mathbf{Z} has the probability density function (PDF) [12], [14]

$$\begin{aligned} f(\mathbf{Z}; \mathbf{R}) &= \prod_{m=1}^M f(\mathbf{z}[m]; \mathbf{R}) \\ &= \frac{1}{\pi^{LN} \det(\mathbf{R})^M} \exp \left\{ -M \text{tr} \left(\mathbf{R}^{-1} \hat{\mathbf{R}} \right) \right\}. \end{aligned}$$

In this expression, $\hat{\mathbf{R}}$ is the maximum-likelihood (ML) estimate of the composite covariance matrix

$$\hat{\mathbf{R}} = \frac{1}{M} \mathbf{Z} \mathbf{Z}^H = \begin{bmatrix} \hat{\mathbf{R}}_{11} & \hat{\mathbf{R}}_{12} & \cdots & \hat{\mathbf{R}}_{1L} \\ \hat{\mathbf{R}}_{12}^H & \hat{\mathbf{R}}_{22} & \cdots & \hat{\mathbf{R}}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{R}}_{1L}^H & \hat{\mathbf{R}}_{2L}^H & \cdots & \hat{\mathbf{R}}_{LL} \end{bmatrix} \quad (3)$$

and $\hat{\mathbf{R}}_{i\ell}$ is an M sample ML estimate of the $N \times N$ cross-covariance matrix $\mathbf{R}_{i\ell}$. The GLRT for this problem decides to accept or reject the null hypothesis \mathcal{H}_0 by computing the likelihood ratio [12], [14]

$$\Lambda = \left(\frac{\sup_{\mathbf{R} \in \mathcal{R}_0} f(\mathbf{Z}; \mathbf{R})}{\sup_{\mathbf{R} \in \mathcal{R}} f(\mathbf{Z}; \mathbf{R})} \right)^{1/M} = \frac{\det \hat{\mathbf{R}}}{\prod_{i=1}^L \det \hat{\mathbf{R}}_{ii}} = \det \hat{\mathbf{C}} \quad (4)$$

where \mathcal{R} denotes the set of all positive-definite Hermitian matrices and \mathcal{R}_0 denotes the set of all matrices in \mathcal{R} , which are block diagonal. Also, the matrix $\hat{\mathbf{C}}$ is referred to as the *multi-channel coherence matrix* [12], [14] as it measures the cross-correlation between the “whitened” random vectors $\hat{\mathbf{R}}_{ii}^{-1/2} \mathbf{x}_i$

for $i = 1, \dots, L$. That is, $\hat{\mathbf{C}} = \hat{\mathbf{D}}^{-1/2} \hat{\mathbf{R}} \hat{\mathbf{D}}^{-H/2}$ where $\hat{\mathbf{D}} = \text{blkdiag}[\hat{\mathbf{R}}_{11}, \dots, \hat{\mathbf{R}}_{LL}]$ is the ML estimate of \mathbf{R} under the null hypothesis. Note that the matrix $\hat{\mathbf{D}}^{1/2}$ is a square-root matrix of $\hat{\mathbf{D}}$ such that $\hat{\mathbf{D}} = (\hat{\mathbf{D}}^{1/2})(\hat{\mathbf{D}}^{1/2})^H$. Then, $\hat{\mathbf{D}}^{-1/2} = (\hat{\mathbf{D}}^{1/2})^{-1}$ and $\hat{\mathbf{D}}^{-H/2} = (\hat{\mathbf{D}}^{-1/2})^H$.

The likelihood ratio given in (4) is referred to as a *generalized Hadamard ratio* [14]. A measure of *bulk coherence* among all the channels may be written as follows:

$$\gamma^2 = 1 - \frac{\det \hat{\mathbf{R}}}{\prod_{i=1}^L \det \hat{\mathbf{R}}_{ii}} = 1 - \det \hat{\mathbf{C}} \quad (5)$$

which takes a value of $\gamma^2 = 0$ when $\hat{\mathbf{R}}$ is block diagonal (i.e., null hypothesis). An important and useful property of the test statistic in (4) or (5) is its invariance to channelwise nonsingular linear transformation, i.e., $\mathbf{T}_i \mathbf{x}_i$, of the data including scaling, filtering, unitary transformation (e.g., discrete Fourier transform), and also permutation or ordering of the channel index [12].

For a dual-channel sonar ($L = 2$) consisting of HF and BB sonar imagery data coregistered over the seafloor the data preparation process is depicted in Fig. 1(a). In this particular application, each sonar image is first partitioned, with 50% overlap in both the along-track and range dimensions, into regions of interest (ROIs) of sizes 80×144 . Each ROI is subsequently partitioned into nonoverlapping blocks of size 8×8 and vectorized to form the random vector $\mathbf{x}_i[m]$, $m \in [1, M]$. This processing yields $M = 180$ realizations of an $N = 64$ dimensional observation for each channel. All M realizations from each channel are finally concatenated to form the data matrix $\mathbf{Z} \in \mathbb{C}^{128 \times 180}$ given in (2) and used to form the likelihood ratio given in (4). Fig. 1(b) shows three pairs of ROIs, one for a target, one for a structured clutter, and one for the background, chosen from a pair of coregistered HF and BB sonar images and processed using this data formation scheme to generate the corresponding composite covariance matrix \mathbf{R} . These matrices shown in the

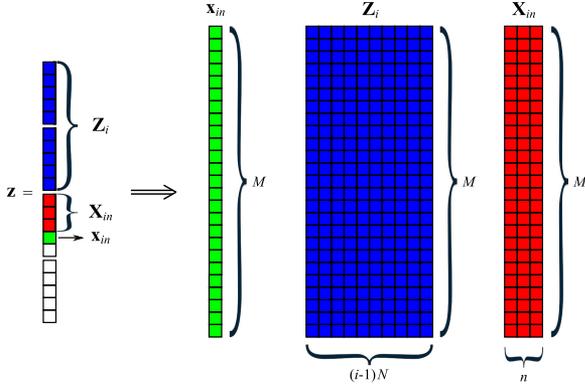


Fig. 2. Partitioning the composite vector \mathbf{z} into $M = 25$ realizations associated with the vector \mathbf{x}_{in} and data matrices \mathbf{Z}_i and \mathbf{X}_{in} for $L = 4$ channels each with $N = 5$ dimensions.

same figure illustrate the fact that for low to medium density clutter background this matrix is nearly block diagonal.

B. Null Distribution of the Multichannel GLRT Detector

One of the most difficult tasks in implementing any likelihood ratio test is defining an appropriate detection threshold for making decision. This requires the knowledge of the distribution of the likelihood ratio under the null hypothesis \mathcal{H}_0 . Fortunately, in the case of the generalized Hadamard ratio in (4) it is known that, under the null hypothesis that $\mathbf{R} \in \mathcal{R}_0$, i.e., every channel is truly independent of one another, the likelihood ratio is distributed as a product of independent but *not* identically distributed beta random variables [12]. Having such knowledge is useful not only for defining detection threshold to yield a desired false alarm probability, but also for predicting the false alarm performance of the detector.

To come up with a description of the likelihood ratio under the null hypothesis, we begin by noting that, for any $i \geq 2$ and any $n = 0, \dots, N - 1$, those random variables that precede $x_i[n]$ in the composite vector \mathbf{z} can be partitioned based on those that correspond to sonar channel i and those that do not. Namely, let the data vector $\mathbf{x}_{in} \in \mathbb{C}^M$ contain all M realizations of the random variable $x_i[n]$, the data matrix $\mathbf{X}_{in} \in \mathbb{C}^{M \times n}$ contain all M realizations of random variables $x_i[0], \dots, x_i[n - 1]$ along its rows, and the data matrix $\mathbf{Z}_i \in \mathbb{C}^{M \times (i-1)N}$ contain all M realizations of the random variables in $\mathbf{x}_1, \dots, \mathbf{x}_{i-1}$ along its rows. Fig. 2 illustrates an example of this process for a composite vector consisting of $L = 4$ channels each with an $N = 5$ dimensional observation. For a given choice in i and n , one may partition those random variables that precede $x_i[n]$ (shown in green in Fig. 2) and use those that pertain to channel i to form \mathbf{X}_{in} (shown in red) and those that do not pertain to channel i to form \mathbf{Z}_i (shown in blue). Given the definition of these matrices, it was shown in [12] that the likelihood ratio given in (4) could be expressed as follows:

$$\Lambda = \prod_{i=2}^L \prod_{n=0}^{N-1} \frac{\mathbf{x}_{in}^H \mathbf{P}_{\mathbf{Z}_i}^\perp \mathbf{x}_{in}}{\mathbf{x}_{in}^H \mathbf{P}_{\mathbf{Z}_i}^\perp \mathbf{x}_{in} + \mathbf{x}_{in}^H \mathbf{P}_{\mathbf{X}_{in}}^\perp \mathbf{x}_{in}} = \prod_{i=2}^L \prod_{n=0}^{N-1} \eta_{in} \quad (6)$$

where η_{in} is the *coherence feature* of element n of channel i with respect to the previous channels. Here, $\mathbf{P}_{\mathbf{Z}_i}$ denotes the orthogonal projection matrix onto $(i - 1)N + n$ dimensional subspace spanned by the data matrix $[\mathbf{Z}_i \mathbf{X}_{in}]$, and $\mathbf{P}_{\mathbf{X}_{in}}^\perp$ is the orthogonal projection onto the $(i - 1)N$ dimensional subspace spanned by the matrix $\mathbf{P}_{\mathbf{Z}_i}^\perp \mathbf{Z}_i$. Note that the notation $\mathbf{P}_A^\perp = \mathbf{I} - \mathbf{P}_A$ denotes the projection onto the orthogonal complement of matrix \mathbf{A} . As discussed in [12], each term within the product of this expression represents the ratio of the estimated variance of a residual from two different linear least squares estimators. The numerator is found by regressing $x_i[n]$ onto all the random variables previous to it (i.e., $\mathbf{x}_1, \dots, \mathbf{x}_{i-1}$ and $x_i[0], \dots, x_i[n - 1]$) while the denominator is found by regressing $x_i[n]$ onto only those random variables associated with channel i (i.e., $x_i[0], \dots, x_i[n - 1]$). Thus, the quadratic form $\mathbf{x}_{in}^H \mathbf{P}_{\mathbf{X}_{in}}^\perp \mathbf{x}_{in} = \|\mathbf{P}_{\mathbf{X}_{in}}^\perp \mathbf{x}_{in}\|^2$ in (6) essentially represents the increase in mean-squared error incurred by excluding channels $\mathbf{x}_1, \dots, \mathbf{x}_{i-1}$ from the estimation problem. The smaller this term is relative to the quadratic form $\mathbf{x}_{in}^H \mathbf{P}_{\mathbf{Z}_i}^\perp \mathbf{x}_{in} = \|\mathbf{P}_{\mathbf{Z}_i}^\perp \mathbf{x}_{in}\|^2$, the more evidence in support of the null hypothesis of independence as it indicates that one can just as accurately estimate $x_i[n]$ by ignoring the previous channels $1, \dots, i - 1$.

The benefit of the expression in (6) compared to the likelihood ratio in (4) is that it shows that the likelihood ratio can be expressed as a product of scalars whose null distribution can easily be derived. Namely, if the observation \mathbf{x}_{in} is an M -dimensional, zero-mean complex normal random vector then the two quadratic forms $\mathbf{x}_{in}^H \mathbf{P}_{\mathbf{Z}_i}^\perp \mathbf{x}_{in}$ and $\mathbf{x}_{in}^H \mathbf{P}_{\mathbf{X}_{in}}^\perp \mathbf{x}_{in}$ in (6) are independently chi-squared distributed with degrees of freedom dependent on the rank of their respective projection matrix. Moreover, if two random variables X and Y are independently chi-squared then the ratio $X/(X + Y)$ is beta distributed. Thus, under the null hypothesis \mathcal{H}_0 that $\mathbf{R} = \text{blkdiag}[\mathbf{R}_{11}, \dots, \mathbf{R}_{LL}]$, the likelihood ratio in (4) exhibits the following stochastic representation [12]:

$$\Lambda | \mathcal{H}_0 \stackrel{d}{=} \prod_{i=2}^L \prod_{n=0}^{N-1} \eta_{in} \quad (7)$$

where each random variable in the product is independently distributed as $\eta_{in} \sim \text{Beta}(\alpha_{in}, \beta_i)$ with parameters

$$\alpha_{in} = M - (i - 1)N - n \quad (8)$$

$$\beta_i = (i - 1)N. \quad (9)$$

That is, under the null hypothesis the likelihood ratio is distributed as a product of independent but *not* identically distributed beta random variables. Moreover, this stochastic representation is only dependent on the number of channels (L), the channel dimension (N), and the number of independent samples (M) but completely independent of the intrachannel correlation matrices \mathbf{R}_{ii} for $i = 1, \dots, L$. This means that, as long as the covariance matrix \mathbf{R} is truly block diagonal, the stochastic representation given in (7) will hold regardless of the specific correlation structure exhibited by each channel.

Fig. 3(a) gives an example of one HF SAS image with several ROIs highlighted throughout the image: ROIs 1–3 correspond

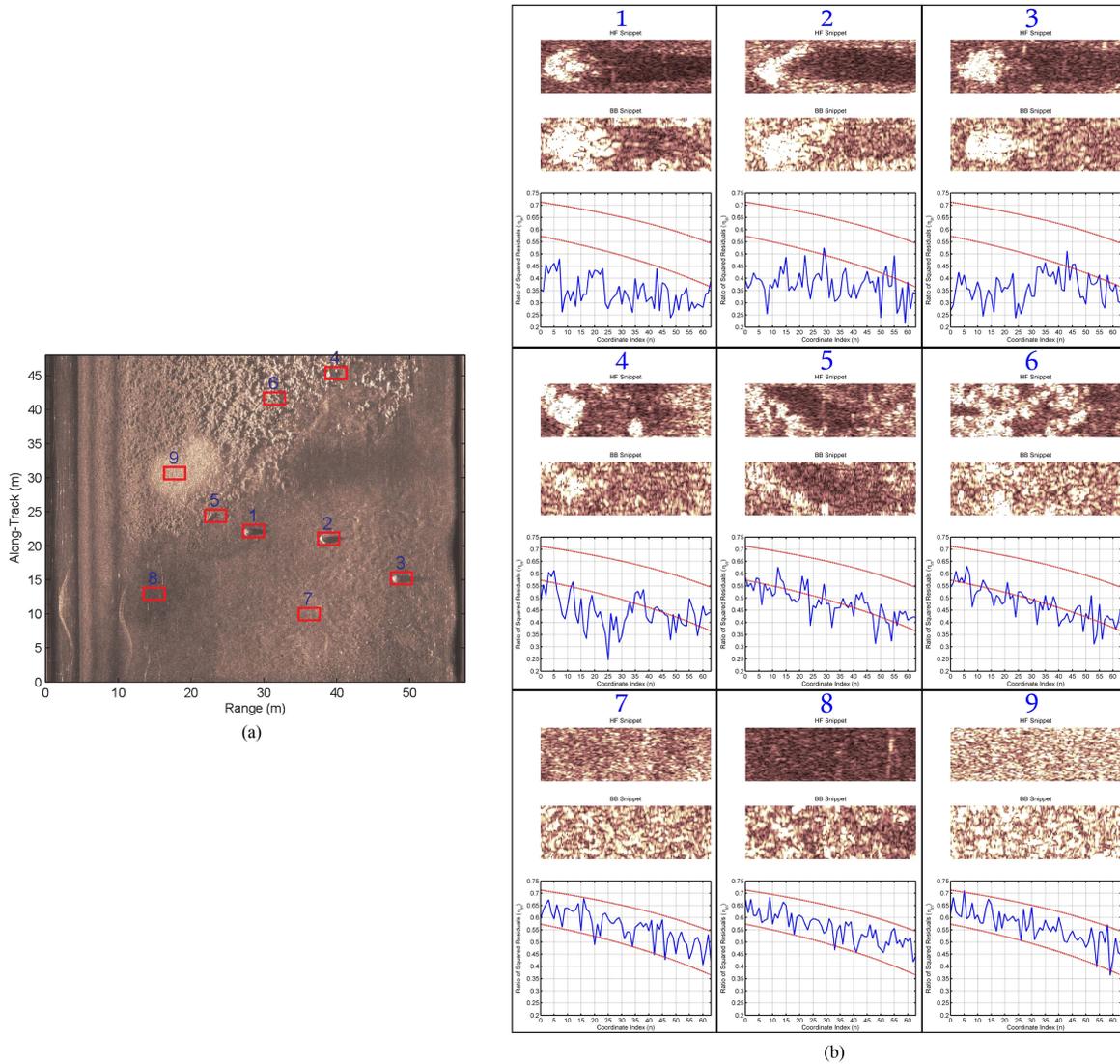


Fig. 3. Comparison of the coherence features η_{2n} in (6) to their probabilistic bounds under the null hypothesis for several ROIs extracted throughout a sonar image. (a) HF sonar image. (b) Coherence statistics η_{2n} and their probabilistic bound.

to targets, ROIs 4–6 contain cluttered background, and ROIs 7–9 contain uncluttered background. For each selected ROI, the corresponding tiles on the right-hand side of Fig. 3(b) display both the HF (top) and corresponding BB (middle) snippet images. For each pair of HF and BB snippets shown, the plot immediately below them shows the plot of the coherence features, η_{2n} , in (6) versus the index n for $n = 0, \dots, 63$ (block size 8×8 with no overlap). Note that in this case since $L = 2$, the expression in (6) can be rewritten as $\Lambda = \prod_{n=0}^{N-1} \eta_{2n}$, i.e., the product of blue values given in this plot yields the likelihood ratio measurement for that pair of snippets. Knowing that each of these values should be beta distributed under the null hypothesis according to (7), the two red lines in this plot give a 95% confidence interval for each coherence feature. More specifically, these two red lines correspond to the interval $[a, b]$ such that $P[a \leq \eta_{2n} \leq b | \mathcal{H}_0] = 0.95$, i.e., under the null hypothesis each value in blue should fall between the two red lines with 95% probability.

Recalling that snippets 1–3 contain targets, one can see that the coherence features for these three examples clearly fall outside the interval, which is indeed desirable given that the presence of a target represents a deviation from the null hypothesis. However, what is more interesting from the perspective of this test is that the coherent features for snippets 7–9, which contain background alone signifying the null hypothesis, do in fact fall within the interval with high probability. This lends credence to the notion that the stochastic representation should be capable of giving a probabilistic description of the likelihood ratio when applied to background. Finally, snippets 4–6 containing structured natural clutter, albeit of high density, are also situations where the coherence features deviate from the confidence interval though some values do indeed fall within that interval. To physically interpret these results in terms of image complexity one should note that the coherence features defined in (6) are indeed dependent on the relation of the intrachannel coherence to the combined inter- and intrachannel coherence. In absence

of any target or structured clutter there is little interchannel coherence, hence resulting in larger values of coherence features especially for smaller n . Thus, the deviations of these coherence features from those under the null hypothesis is a measure of clutter density (image complexity) and hence the detector's performance.

C. Multichannel Broadband Coherence Detector

The analysis used to derive the likelihood ratio in (4) is general in that no specific structure is imposed on the composite covariance matrix of (1) other than it is block diagonal under the null hypothesis. However, doing so requires a data rich environment in which $M \gg LN$. In cases of data poverty, one can impose additional structure on the covariance matrix, solve the ML estimates under that model, and construct the likelihood ratio. Such is the case for the broadband coherence detector where we assume that the random vectors from each channel are realizations from two-dimensional (2-D) wide-sense stationary (WSS) processes. That is, we assume that the data in the images are WSS in both the along-track and range dimensions leading to a Toeplitz structure in each block $\mathbf{R}_{i\ell}$ of (3). Results on asymptotically large Toeplitz matrices show that, as the length of each channel N grows large, the likelihood ratio in (4) converges to the *broadband coherence test* [12], [14]

$$\Lambda \xrightarrow{N \rightarrow \infty} \exp \left\{ \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln \frac{\det \hat{\mathbf{S}}(e^{j\theta}, e^{j\phi})}{\prod_{i=1}^L \hat{S}_{ii}(e^{j\theta}, e^{j\phi})} \frac{d\theta d\phi}{4\pi^2} \right\}. \quad (10)$$

In this expression, $\hat{\mathbf{S}}(e^{j\theta}, e^{j\phi}) \in \mathbb{C}^{L \times L}$ is an estimated cross-spectral matrix with $\hat{S}_{i\ell}(e^{j\theta}, e^{j\phi})$ denoting the cross-spectral density (power spectral density when $i = \ell$) between channels i and ℓ at frequency θ and wavenumber ϕ . Thus, the likelihood ratio becomes the log of a frequency/wavenumber dependent Hadamard ratio integrated over the Nyquist band. This likelihood ratio can then be compared to a preselected threshold to decide whether targets are present. Unlike the generalized Hadamard ratio in (4) which requires manipulation on $LN \times LN$ covariance matrix \mathbf{R} , the broadband coherence test statistic in (10) would involve $L \times L$ matrix, $\hat{\mathbf{S}}(e^{j\theta}, e^{j\phi})$, and hence is not plagued by data poverty issues. Moreover, the test statistic in (10) is computationally efficient as the estimated cross-spectral matrix can be computed using Welch's method [15] by partitioning each pair of ROIs into 32×32 blocks similar to that in Fig. 1 but with a 50% overlap in both dimensions. Each block is subsequently windowed using a separable 2-D Hamming window. A 2-D fast Fourier transform (FFT) is then applied and the corresponding periodograms for each block are averaged over the number of overlapped blocks to produce the estimate $\hat{\mathbf{S}}(e^{j\theta}, e^{j\phi})$. That is, if we let $X_i^{(m)}(e^{j\theta}, e^{j\phi}) \in \mathbb{C}$ be the 2-D FFT of the m th block at θ and ϕ , then the cross-power spectrum between channels i and ℓ may be computed as follows:

$$\hat{S}_{i\ell}(e^{j\theta}, e^{j\phi}) = \frac{1}{M} \sum_{m=1}^m X_i^{(m)}(e^{j\theta}, e^{j\phi}) \left(X_\ell^{(m)}(e^{j\theta}, e^{j\phi}) \right)^*. \quad (11)$$

D. Approximating the Null Distribution Using the Saddlepoint Method

Although the stochastic representation given in (7) does indeed give some insight about how the likelihood ratio behaves probabilistically under the null hypothesis, it is very difficult to derive the distribution of the likelihood ratio Λ itself as it would involve having to find the distribution of multiple beta random variables. This becomes difficult especially in situations where L or N is large. However, there do exist a number of techniques that can be used to approximate the distribution of any random variable with known characteristic function. Here, we focus on the use of saddlepoint approximations [13], [16] for obtaining accurate approximations of the likelihood ratio's density and distribution functions. Given any scalar random variable X with cumulant generating function (CGF) $\psi(t) = \ln E[e^{tX}]$, the saddlepoint approximation to the PDF, $f(x)$, is defined as [16], [17]

$$\hat{f}(x) = \frac{1}{\sqrt{2\pi\psi''(\hat{t})}} \exp \{ \psi(\hat{t}) - \hat{t}x \} \quad (12)$$

where the saddlepoint, \hat{t} , denotes the value of t such that $\psi'(t) = x$. Note that the notation ψ' and ψ'' denote the first- and second-order derivatives of the CGF, respectively. Likewise, an approximation formula for the cumulative distribution function (CDF) is given by [18]

$$\hat{F}(x) = \begin{cases} \Phi(\hat{w}) + \phi(\hat{w}) \left(\frac{1}{\hat{w}} - \frac{1}{\hat{u}} \right), & \text{for } x \neq \mu \\ \frac{1}{2} + \frac{\psi'''(0)}{6\sqrt{2\pi}(\psi''(0))^{3/2}}, & \text{for } x = \mu \end{cases} \quad (13)$$

where ϕ and Φ represent the PDF and CDF of a standard normal random variable, respectively, and ψ''' is the third-order derivative of the CGF. Also, the values \hat{w} and \hat{u} in (13) are given by

$$\begin{aligned} \hat{w} &= \text{sgn}(\hat{t}) \sqrt{2(\hat{t}x - \psi(\hat{t}))} \\ \hat{u} &= \hat{t} \sqrt{\psi''(\hat{t})} \end{aligned}$$

and μ denotes the mean of random variable X .

The first step to build a saddlepoint approximation of (4) under the null hypothesis is to derive its CGF. Using properties of beta distributions, it is straightforward to show that each random variable within the product given in (7) has k th raw moment

$$E[\eta_{in}^k] = \frac{\Gamma(\alpha_{in} + k)\Gamma(\alpha_{in} + \beta_i)}{\Gamma(\alpha_{in})\Gamma(\alpha_{in} + \beta_i + k)} \quad (14)$$

where the parameters α_{in} and β_i are given in (8) and (9), respectively, and $\Gamma(\cdot)$ denotes the gamma function [19]. Using (14) along with the fact that these random variables are independent, one can see that the likelihood ratio in (4) has the following k th

order moment under the null hypothesis

$$E[\Lambda^k | \mathcal{H}_0] = \prod_{i=2}^L \prod_{n=0}^{N-1} E[\eta_{in}^k] \\ = \prod_{i=2}^L \prod_{n=0}^{N-1} \frac{\Gamma(\alpha_{in} + \beta_i) \Gamma(\alpha_{in} + k)}{\Gamma(\alpha_{in}) \Gamma(\alpha_{in} + \beta_i + k)}. \quad (15)$$

One can then use this expression for the moments to show that, for any real-valued scalar c , the random variable $W = c \ln \Lambda$ has the CGF

$$\psi_W(t) = \ln(E[e^{tW}]) = \ln(E[\Lambda^{ct} | \mathcal{H}_0]) \\ = \sum_{i=2}^L \sum_{n=0}^{N-1} \left[\ln \frac{\Gamma(\alpha_{in} + \beta_i)}{\Gamma(\alpha_{in})} + \ln \frac{\Gamma(\alpha_{in} + ct)}{\Gamma(\alpha_{in} + \beta_i + ct)} \right]. \quad (16)$$

As discussed above, the next step in building a saddlepoint approximation involves solving for the saddlepoint. Letting $\gamma_k(z)$ denote the polygamma function of order k

$$\gamma_k(z) = \frac{d^{(k+1)} \ln \Gamma(z)}{dx^{(k+1)}}$$

the CGF given in (16) has k th-order derivative

$$\frac{d^k \psi_W(t)}{dt^k} = c^k \sum_{i=2}^L \sum_{n=0}^{N-1} [\gamma_{k-1}(\alpha_{in} + ct) \\ - \gamma_{k-1}(\alpha_{in} + \beta_i + ct)]. \quad (17)$$

For any argument x in (12), solving for the saddlepoint in this particular case involves finding the value of t such that

$$\psi'_W(t) = c \sum_{i=2}^L \sum_{n=0}^{N-1} [\gamma_0(\alpha_{in} + ct) - \gamma_0(\alpha_{in} + \beta_i + ct)] = x. \quad (18)$$

Unfortunately, there is no simple closed-form solution to the problem given in (18). In lieu of this fact, one must resort to numerical root-finding algorithms such as the Newton–Raphson algorithm [19] to find the saddlepoint.

Thus, the process of the saddlepoint approximation involves the following.

- 1) For every value of the dependent variable x in (12) and (13), find the saddlepoint \hat{t} by solving for the value of t that satisfies (18).
- 2) Substitute the saddlepoint \hat{t} into the CGF and its higher order derivatives given in (16) and (17), respectively.
- 3) Compute the saddlepoint approximation of the PDF \hat{f} or the CDF \hat{F} at that given value of x using (12) and (13), respectively.
- 4) Repeat this process for every value of x needed to compute the approximation given the implicit dependence on this variable in (18).

TABLE I
PROPERTIES OF THE SSAM I DATA SETS

		Clutter Difficulty	# of Images	# of Targets
Dataset 1	—	Medium/Hard	122	77
Dataset 2	Day 1	Easy	180	4
	Day 2	Hard	136	17
	Day 3	Easy	142	47
	Total	—	458	68

III. SONAR DATA SETS

In this paper, several real and synthetic sonar data sets consisting of pairs of HF and BB sonar images are used. These are briefly described next.

A. SSAM I Data

The performance evaluation of the proposed methods in different background and environmental conditions was conducted based on two SSAM I data sets that contain real contacts and seafloor clutter. The first data set was collected in one environment in the Gulf of Mexico while the second data set was collected in several different environments in the Ligurian Sea. The properties of these data sets in terms of the total number of sonar images, the number of targets, and the clutter difficulty are given in Table I. The first data set has medium to high clutter density because many images contain large coral reef or rocky clutter fields. Days 1–3 of the second data set were collected at different locations. Two of the three environments in the second data set are easy, as they have a sandy seafloor, whereas one environment (Day 2) is very difficult, as it contains large ripple fields and seagrass.

B. SAS900A Data

The SAS900A sonar image database consists of a set of 143 pairs of HF and BB sonar images characterized by their background and object configuration. Each image pair uses one of 11 different backgrounds that are actual images of the seafloor with varying clutter density. Each pair of images also contains the synthetically generated signatures of six different objects that are superimposed on the background image according to one of 13 different configurations that define the types, orientations, and relative positions of the objects. The number of each object type at a single resolution is listed in Table II. The sonar returns from each object are synthetically generated using version 10.0 of the Shallow Water Acoustic Toolset [20] and represent the following object types: block, cone, sphere, cylinder, torus, and pipe. The aspect angles of the objects range from 0° to 120° and objects are placed far enough away from each other such that their signatures are always completely disjoint.

IV. PERFORMANCE PREDICTION AND ESTIMATION

The results in Sections IV-A and IV-B that involve predicting and estimating the false alarm rates of the multichannel detector under various background conditions were generated using the

TABLE II
NUMBER OF EACH OBJECT TYPE AT A SINGLE RESOLUTION (HF OR BB)

Object Type	Number of Objects
Block	88
Cone	66
Sphere	66
Cylinder	308
Torus	22
Pipe	22

two real SSAM I sonar data sets described in Section III-A. While those in Section IV-D that revolve around empirically characterizing how the likelihood ratio of the detector behaves under various target conditions including aspect angle and range variation are produced using the model-generated data described in Section III-B. However, only sonar returns corresponding to the cylindrical object were used owing to the fact that this particular object type has a greater variation in aspect and range in this data set. In fact, one can see from Table II that over 50% of the targets correspond to this object type.

A. Image Complexity Assessment Using Multichannel Coherence Statistics

As illustrated in Section II-B, the multichannel coherence test could be decomposed into a product of independent coherence statistics, η_{in} , which are beta distributed under the null hypothesis. The significance of this result is that the explicit theoretical distribution of each coherence statistic can serve as a yardstick with which one can benchmark their statistical behavior to predict how the multichannel detector can perform in different environments. The examples in Section II-B illustrated that by analyzing and measuring the deviation of these coherence features from those of the null hypothesis one can gain an assessment of image complexity and hence the type of the background condition. More specifically, as shown in the examples in Fig. 3 higher complexity in the sonar image leads to a larger deviation from the theoretical null distribution.

Fig. 4 shows the scatter plot of the first two principal components (out of 64) of the coherence statistics/features, η_{2n} , $n = 0, \dots, 63$, for four different background conditions, namely a muddy bottom type (blue dots), seagrass vegetation (red dots), sand ripples (green dots), and a rocky or corral background (magenta dots) taken from both SSAM I data sets. For several examples of each background type indicated by arrows in this figure, the left-hand subplot displays the HF image snippet and its corresponding coherence features that are plotted in blue in the subplot on the right. As stated before, the two red lines in each subplot on the right denote a 98% confidence interval for their null distribution. First, from the scatter plot we can observe that these coherence statistics/features in fact do a fairly good job of separating each background type. The first two principal components for muddy background ROIs tend to cluster tightly together and widely separated from those of the others. Moreover, as one moves from right to left in this figure,

TABLE III
FALSE ALARM PROBABILITIES VERSUS CLUTTER DIFFICULTY FOR BOTH DATA SETS

		Clutter Difficulty	# of Images	Estimated P_{FA}	Avg # of False Alarms/Image
Dataset 1		Easy/Medium	40	0.042	12.6
		Medium	27	0.157	
		Hard	55	0.413	
Dataset 2	Day 1	Easy	180	0.007	8.6
	Day 2	Hard	136	0.426	
	Day 3	Easy	142	0.011	

one can observe a general increase in background complexity as there is good agreement between the coherence statistics and the null distribution for the featureless muddy bottom (i.e., the coherence statistics tend to fall within the confidence interval), a small deviation for the fairly complicated seagrass background, and much larger deviations for the more complicated sand ripple and rocky background cases. Nevertheless, their clusters are for the most part nonoverlapping with the exception of some seagrass ROIs.

B. Performance Prediction and Estimation

To investigate the theory's ability to predict the false alarm rate produced by the detector, the saddlepoint approximation given in (13) was used to determine the threshold, ξ , needed to approximately achieve a false alarm probability of $P_{FA} = 0.01$, i.e., a threshold such that only up to 1% of ROIs containing background would be falsely detected. This was accomplished by finding the value of ξ such that $\hat{F}(\xi) = 1 - 0.01$ using the CDF approximation in (13). The Newton–Raphson algorithm was used to solve the saddlepoint equation in (18). The detector was then applied to each image pair from both SSAM I data sets described in Section III-A and the percentage of background ROIs whose likelihood value exceeded the threshold ξ was used as an estimate of the false alarm probability.

The blue lines in Fig. 5(a) and (b) plot the empirical probability of false alarm for every image of the first and second SSAM I data sets, respectively. The horizontal dashed line in both of these plots show the desired false alarm probability of $P_{FA} = 0.01$. To observe how different clutter densities affect the false alarm rate, the results of Fig. 5(a) have been partitioned based on the density of the clutter in the image. Moreover, the results of Fig. 5(b) have been partitioned based on the day the data were collected. Recall that the environments corresponding to Days 1 and 3 are considered relatively easy as they contained low to medium clutter density while Day 2 was collected in an environment with high clutter density. Table III gives the average false alarm probability observed in each subset of the two SSAM I data sets together with the average number of false alarms per image for the entire first and second data sets. From the results of Fig. 5 and Table III, one can see that there is a fairly good correspondence between the desired false alarm probability and that which is actually observed in environments with low to medium clutter density. However, similar to the conclusions drawn from Fig. 3, one can see that the estimated false alarm rate deviates from what is desired in images containing high clutter density. Thus, one can conclude that the

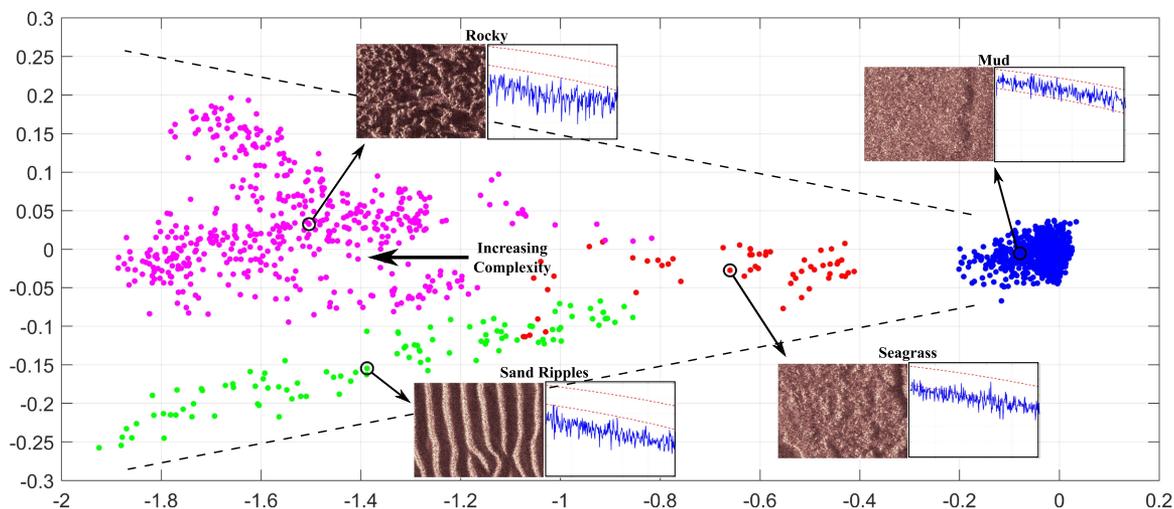


Fig. 4. Scatter plot of the first two principal components of the coherence statistics for several background conditions. Examples of the HF snippet and actual coherence statistics for several examples demonstrate the increased deviation from the null hypothesis (denoted by a pair of red lines in each subplot) with increasing background complexity.

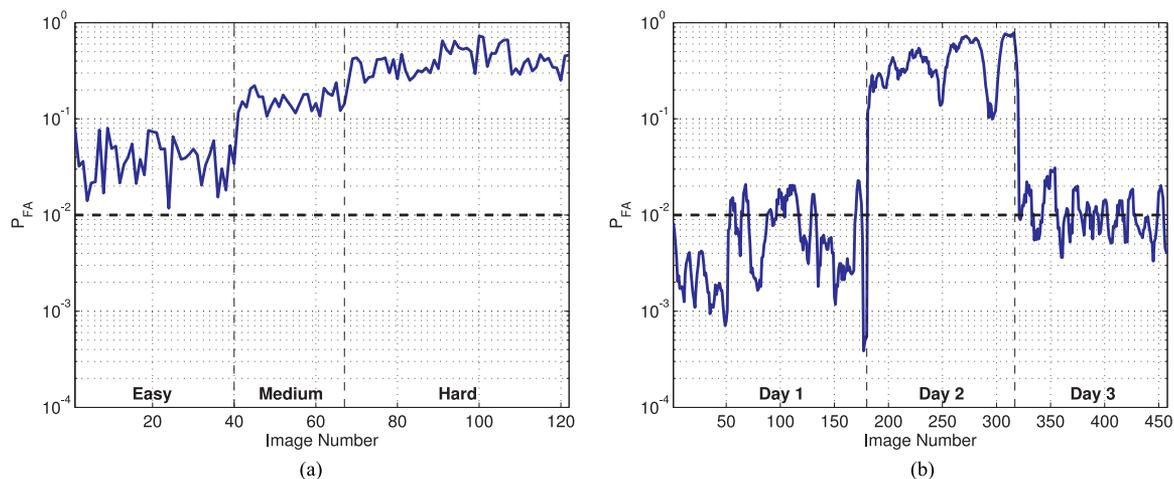


Fig. 5. Per image false alarm probability versus environment/clutter difficulty for both SSAM I data sets. (a) Data set 1. (b) Data set 2.

proposed theoretical model can predict false alarm performance fairly accurately in environments with low to medium clutter density. Additionally, from the deviations of the distributions of the coherence statistics from those of the null hypothesis one can determine which environments would require performance optimization using the empirical data collected *in situ*.

As a final demonstration of this theory and its applicability to this problem, the top images of Fig. 6(a)–(d) gives the examples of several HF images with starkly different clutter densities. Fig. 6(a) and (b) gives the examples of images with relatively low clutter while Fig. 6(c) and (d) contains difficult clutter in certain areas of each image. Note that the targets in these images have been outlined using a green box. The images in the middle of Fig. 6 show the likelihood ratio map corresponding to their respective image shown above it. That is, each pixel in this image gives the likelihood ratio Λ in (10) (more specifically the scaled logarithm of the likelihood ratio) for each pair of HF and

BB ROIs extracted throughout the image. Regions in each of these likelihood images that produce values, which exceed the threshold ξ obtained based upon the saddlepoint approximation, are outlined by a red box. Finally, the bottom plots of Fig. 6 give the saddlepoint approximation, $\hat{f}(x)$, in (12) (shown in red) and compare that to the histogram of the likelihood ratio for each image (shown in blue). Looking at the likelihood images in Fig. 6, one can see that although the targets are successfully detected in each image, the more difficult environments in sonar images of Fig. 6(c) and (d) clearly produce a larger number of false alarms. The effect of this high clutter density on the distribution of the likelihood ratio and its deviation from the theoretical distribution generated using (12) is clear from the histograms at the bottom of Fig. 6(c) and (d) where one can see a significantly heavier-tailed distribution compared to the theoretical saddlepoint approximation. However, for the cases with low to medium clutter like those in Fig. 6(a) and (b), one can

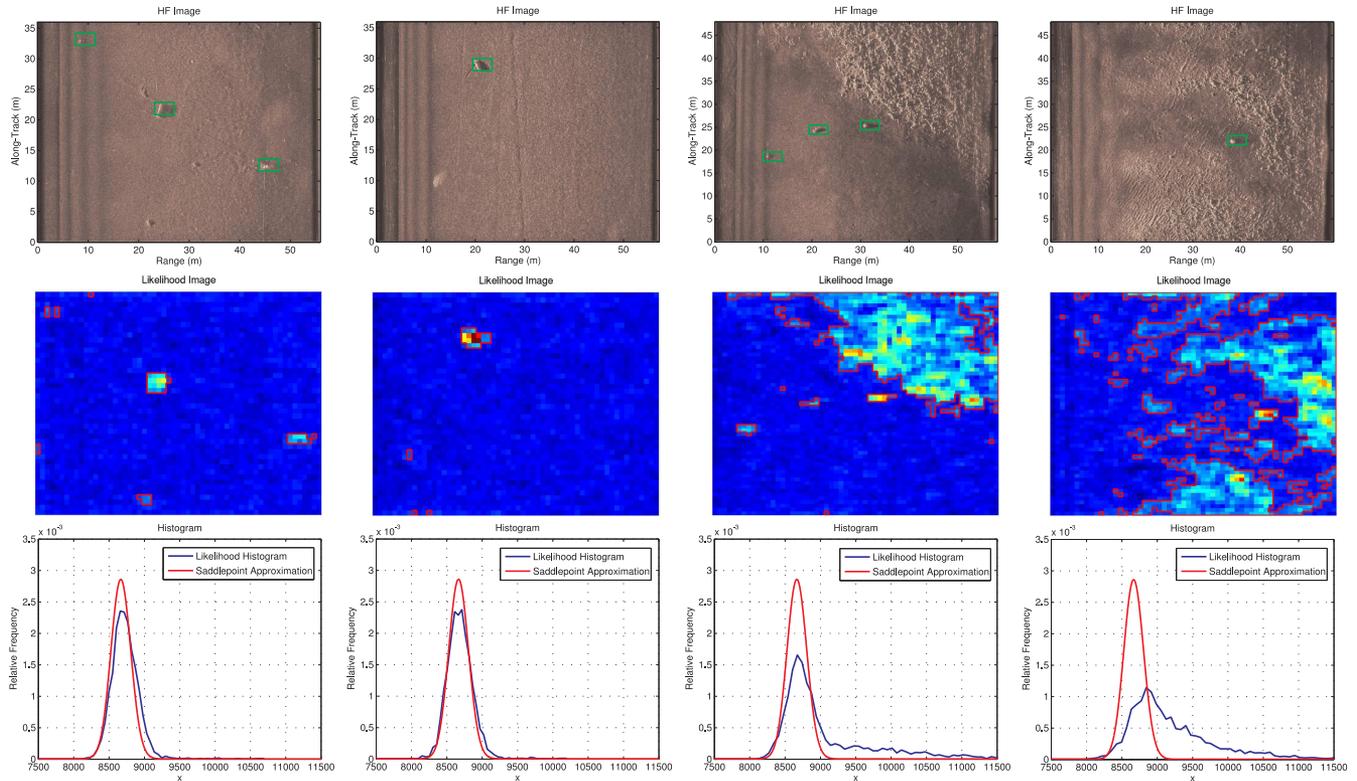


Fig. 6. Examples of several images consisting of uncluttered and cluttered background, their likelihood images, and a comparison of the likelihood ratio's histogram to the saddlepoint approximation in each case. (a) Uncluttered background. (b) Uncluttered background. (c) Cluttered background. (d) Cluttered background.

see a much better correspondence between the theoretical distribution given by the saddlepoint approximation and the actual histogram of the likelihood ratio. This shows once again that the theory is sufficiently capable of describing the distribution of the likelihood ratio over background. However, if that background contains significant amounts of clutter then one will more than likely see a large deviation from the theory.

C. Performance Optimization

When operating in a cluttered environment with coherence statistics that deviate from those of the theory calls for a performance optimization procedure to maintain a prescribed false alarm rate. This threshold adaptation can be done in many ways using some limited background data collected in the new environment. One possible approach is to fit a parametric distribution to the test statistic produced from the data and then recursively update the parameters as new data are collected to give the detector the ability to adapt to the likelihood ratio's changing statistical behavior. Alternatively, this performance optimization can be done using the saddlepoint approximation in Section II-D. As shown before, the saddlepoint approximation is reliant on a theoretical CGF and hence, any deviation from this theoretical behavior can result in a false alarm that is drastically higher than what was originally desired. To adapt the saddlepoint approximation of the detection statistic in (12) and (13) to better match the statistics of the likelihood ratio observed in the actual limited data taken when operating in a new envi-

ronment, one can use an empirical CGF and sample cumulants known as k -statistics [21]. This empirical CGF is then made to match the estimated cumulants of the specific background type identified using the coherence-based image complexity results in Section IV.A by solving a constrained optimization problem. This empirical CGF is subsequently used in the saddlepoint approximation to obtain a better estimate of the likelihood ratio's distribution in that particular region and hence a new threshold that yields a desired false alarm performance. This is a topic for future research.

D. Detection Performance Versus Target Characteristics

In addition to being able to predict the detector's false alarm performance, another important aspect is being able to characterize the detector's ability to detect targets from background under various conditions such as target range and aspect angle. In theory, one could accomplish this by generalizing the results of Section II-B to the alternative hypothesis and studying how these various conditions affect the composite covariance matrix \mathbf{R} in (1). However, given the difficulty of this task, the results of this section rely on an empirical answer to this question.

For this experiment, synthetic data corresponding to a cylindrical object at various aspect angles and ranges were extracted from the SAS900A data set described in Section III and applied to the detector in (10). Multiple measurements of the likelihood ratio (formed from 11 different backgrounds) for a given configuration were then used to form an empirical confidence bound.

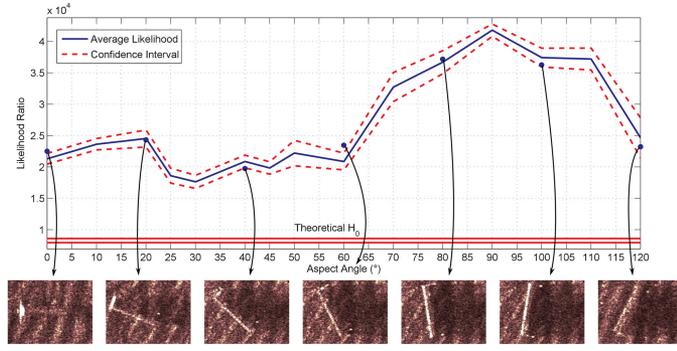


Fig. 7. Average likelihood produced by a synthetic cylindrical object versus aspect angle (blue solid line) and its 98% confidence interval (red-dashed line) versus a 98% confidence interval under the null distribution (red solid line). Examples of the object over various aspect angles are given immediately below the plot.

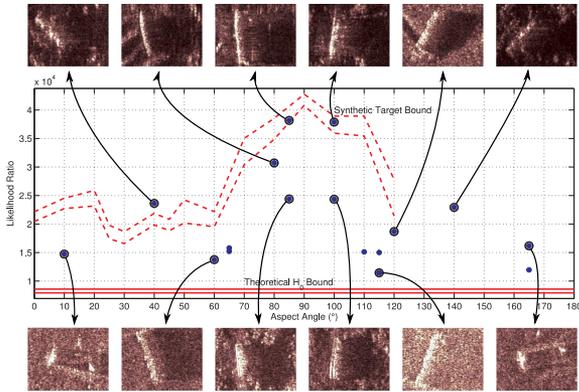


Fig. 8. Likelihood ratio values for several actual cylindrical targets at various aspect angles versus the 98% confidence interval constructed in Fig. 7. Several examples of these targets are displayed above and below the plot.

More specifically, it is assumed that the set of log-likelihood measurements $\{\lambda_i\}_{i=1}^M$ for a given target configuration are *iid* normally distributed with unknown mean μ and variance σ^2 . Letting $\hat{\mu}$ and $\hat{\sigma}^2$ denote the following sample estimates of these two quantities

$$\hat{\mu} = \frac{1}{M} \sum_{i=1}^M \lambda_i \quad (19)$$

$$\hat{\sigma}^2 = \frac{1}{M-1} \sum_{i=1}^M (\lambda_i - \hat{\mu})^2 \quad (20)$$

then it is well-known [19] that the ratio

$$T = \frac{\hat{\mu} - \mu}{\sqrt{\hat{\sigma}^2/M}}$$

follows a Student's t -distribution with $M - 1$ degrees of freedom. Given a confidence level α and critical value c such that $P[-c \leq T \leq c] = \alpha$, the interval $\mathcal{I} = \hat{\mu} \pm c\sqrt{\hat{\sigma}^2/M}$ gives one a confidence interval of size α for the unknown mean μ since $P[\mu \in \mathcal{I}] = \alpha$.

The log-likelihood ratio was then computed for all realizations of the synthetic cylindrical target over various aspect an-

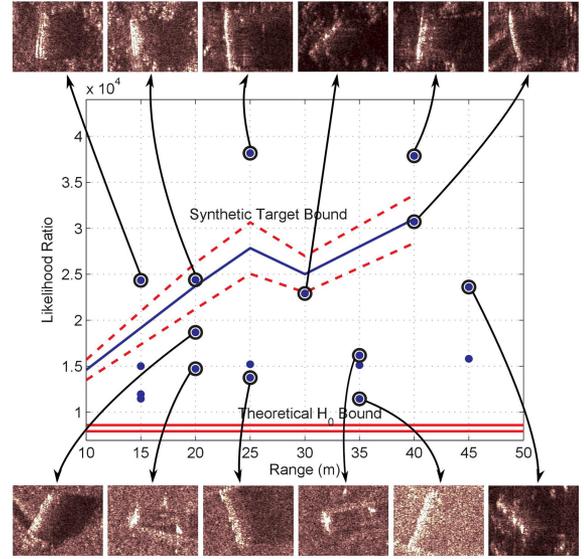


Fig. 9. Average likelihood ratio for synthetic data versus range and its 98% confidence interval. The likelihood ratio values produced by actual targets are denoted by blue dots in the plot.

gles ranging from 0° to 120° . Using these measurements to compute the estimates in (19) and (20), Fig. 7 plots the average likelihood ratio $\hat{\mu}$ in blue as a function of aspect angle while the red-dashed lines denote the upper and lower bounds of the confidence interval \mathcal{I} with probability $\alpha = 0.98$. Likewise, the red solid lines at the bottom of the plot denote a 98% confidence interval under the null hypothesis found using the saddlepoint approximation of CDF in (13). This theoretical confidence interval corresponds to the interval $[a, b]$ such that $P[a \leq \lambda_i \leq b | \mathcal{H}_0] = \hat{F}(b) - \hat{F}(a) = 0.98$, i.e., the likelihood ratio under the null hypothesis falls between the two red solid lines with 98% probability. Finally, the likelihood produced by several examples of a cylindrical target embedded in the same background are denoted in the plot using a blue dot and their corresponding HF snippets are also shown. Several important observations can be made from these plots. First, as can be seen, the average likelihood tends to be relatively low from 0° to 60° but increases significantly for aspects near 90° . This is to be expected given that a 90° aspect corresponds to a broadside view of the target. Moreover, not only the log-likelihood ratios of most of the synthetic targets fall within the confidence interval but also are substantially larger than those of the background alone.

To see how well this empirically based confidence interval is able to represent the behavior of the likelihood ratio for similar targets in SSAM I data, the likelihood ratio was applied to a set of cylindrically shaped targets from both data sets described in Section III. Fig. 8 plots the same pair of confidence intervals shown in Fig. 7 while each blue dot in the figure denotes the log-likelihood value produced by one of the actual targets in this set as a function of its approximate aspect angle. For several examples whose dots are circled in this plot, the HF snippet of the target is shown on the top and bottom as denoted by the arrow pointing to each snippet. Note that the synthetic confidence

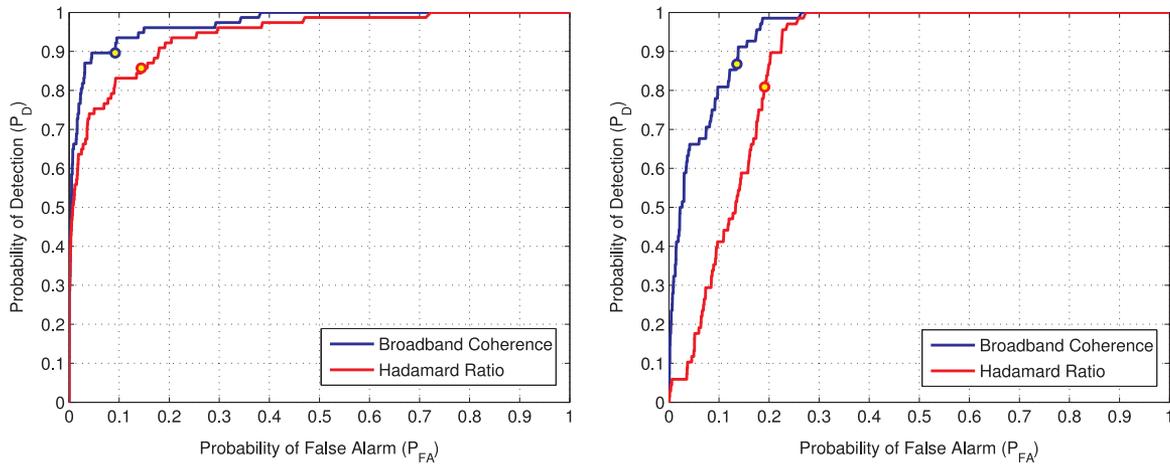


Fig. 10. Comparison of the ROC curves for the generalized Hadamard ratio and BB coherence statistics when applied to the SSAM I data sets. (a) First data set. (b) Second data set.

interval denoted with red-dashed lines stops at 120° as there was no synthetic examples of a cylinder in the SAS900A data set beyond that point. From these results, one can see that very few targets produced a log-likelihood ratio that actually fell within the confidence interval inferred from model-generated data. Given the dissimilarities observed between the snippets in Fig. 8 and those in Fig. 7, one expects to see this mismatch given how environmental effects and various target conditions can affect the data. Be that as it may, one can see that the synthetic bound does give one a fairly accurate prediction of the *relative* behavior of the likelihood ratio versus aspect as one tends to observe relatively larger values between 80° and 100° compared to other aspect angles. More importantly, as in the case of synthetic targets the likelihood ratio in (10) for actual targets produced detection statistics that consistently exceeded the null distribution bound over the entire range of aspect angles.

Next, we repeat the above study but for different target range from the sonar platform using the same synthetic data from the SAS900A data set as well as actual targets from the first and second SSAM I data sets. Fig. 9 plots the synthetic target and null distribution confidence intervals using dashed and solid red lines, respectively, as well as the log-likelihood ratio values produced by each actual target (blue dots). Once again for several circled choices in the plot an arrow points to the corresponding HF snippet image. In this case, one can see that the confidence interval generated using the synthetic data generally increases and widens with target range. This may be due in part to the increase in SNR associated with a shallower grazing angle at greater ranges. Looking at the log-likelihood values corresponding to actual targets (denoted by the blue points throughout the plot), it is difficult to see a similar pattern as the points are fairly scattered. Nevertheless, as before one can see that these points are consistently above the confidence interval corresponding to the null hypothesis demonstrating the likelihood ratio's ability to separate these targets from background. To further illustrate this detection performance Fig. 10(a) and (b) show the receiver operator characteristic (ROC) curves for the broadband coherence detector in (10) for the first and second SSAM I data sets,

respectively. These ROC curves were generated by extracting ROIs of all targets in each data set as well as equal number of randomly located background ROIs. The knee-point of each ROC curve (the point where $P_D + P_{FA} = 1$) shown in these plots using a small circle attest to the validity of the above statement. These figures also present the ROC curves for the generalized Hadamard ratio detector in (4) as a benchmark. As evident, the broadband coherence detector outperforms the generalized Hadamard ratio detector owing to the data poverty issues mentioned in Section II-C for the latter.

V. CONCLUSIONS AND FUTURE WORK

This paper addresses the problem of performance prediction and estimation for a multichannel coherence-based detector. This detection method discriminates targets of interest from background by looking for high levels of coherence among pairs of coregistered HF and BB sonar images using a generalized Hadamard ratio. This detection statistic arises as a GLRT that tests whether the composite covariance matrix of the data is block diagonal. Under the null hypothesis that the covariance matrix is truly block diagonal, it was shown that the generalized Hadamard ratio is statistically equivalent to a product of independent beta random variables. This stochastic representation of the null hypothesis is then used to derive a saddlepoint approximation of log-likelihood ratio's density and distribution functions. The ability of the saddlepoint approximation to accurately predict the distribution of the likelihood ratio when applied to background and achieve a predefined false alarm probability is tested using several SSAM I data sets. The results of the study suggest that the methods are indeed capable of describing the distribution of the likelihood ratio in situations where the background contains low to medium clutter but will exhibit mismatch in images with high clutter density. For the latter scenarios, a performance optimization procedure using an empirical saddlepoint approximation was recommended to adaptively select detection threshold depending on the environmental context. Moreover, a study was carried out to quantify the

detector's ability to detect target from background under various conditions, namely the aspect angle and range of the target. The results of this study were more empirically driven by relying on the use of synthetically generated data to compute a confidence bound for the average likelihood ratio values produced by the detector. Results generated from actual cylindrically shaped targets did seem to generally reflect a dependence on aspect similar to that of the synthetic data but results concerning dependence on range were less predictable. This is likely because the ability of the detector to separate a target from background depends on image quality and many other factors other than just the aspect angle and range of that target.

ACKNOWLEDGMENT

The authors would like to thank Mike Rowe at NSWC-Panama City for providing the data used in this study.

REFERENCES

- [1] L. Sha and L. W. Nolte, "Bayesian sonar detection performance prediction in the presence of interference in uncertain environments," *J. Acoust. Soc. Amer.*, vol. 117, pp. 1954–1964, Apr. 2005.
- [2] W. Liu, J. Liu, C. Zhang, H. Li, and X. Wang, "Performance prediction of subspace-based adaptive detectors with signal mismatch," *Signal Process.*, vol. 123, pp. 122–126, 2016.
- [3] J. Gazagnaire, P. Beaujean, and J. Stack, "Combining model-based and in situ performance prediction to evaluate detection and classification performance," in *Proc. Inst. Acoust.*, 2007, vol. 29, no. 6, pp. 11–18.
- [4] P. Y. Mignotte, J. Vazquez, J. Wood, and S. Reed, "PATT: A performance analysis and training tool for the assessment and adaptive planning of mine counter measures (MCM) operations," in *Proc. MTS/IEEE OCEANS Conf.*, Oct. 2009, doi: [10.23919/OCEANS.2009.5422291](https://doi.org/10.23919/OCEANS.2009.5422291).
- [5] Ø. Midtgaard, T. Sæbø, and R. Hansen, "Estimation of detection/classification performance using interferometric sonar coherence," in *Proc. 3rd Int. Conf. Underwater Acoust. Meas.*, 2009. [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.630.7309&rep=rep1&type=pdf>
- [6] D. P. Williams, "Fast target detection in synthetic aperture sonar imagery: A new algorithm and large-scale performance analysis," *IEEE J. Ocean. Eng.*, vol. 40, no. 1, pp. 71–92, Jan. 2015.
- [7] Ø. Midtgaard, I. Alm, T. Sæbø, M. Geilhufe, and R. Hansen, "Performance assessment tool for AUV based mine hunting," in *Proc. Inst. Acoust. Int. Conf. Synthetic Aperture Sonar Synthetic Aperture Rada*, 2014, vol. 36, no. 1.
- [8] R. E. Hansen and T. O. Sæbø, "Towards automated performance assessment in synthetic aperture sonar," in *Proc. MTS/IEEE OCEANS Conf.*, Jun. 2013, pp. 1–6.
- [9] J. Gazagnaire, "Survey of image quality metrics from the perspective of detection and classification performance," *Proc. SPIE*, vol. 10182, 2017, doi: [10.1117/12.2262665](https://doi.org/10.1117/12.2262665).
- [10] O. Daniel, Y. Petillot, and S. Reed, "Unsupervised sea-floor classification for automatic target recognition," in *Proc. Int. Conf. Underwater Remote Sens.*, 2012.
- [11] L. Picard, A. Baussard, I. Quidu, and G. L. Chenadec, "Seafloor description in sonar images using the monogenic signal and the intrinsic dimensionality," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 1–16, Apr. 2018.
- [12] N. Klausner, M. R. Azimi-Sadjadi, and L. L. Scharf, "Detection of spatially correlated time series from a network of sensor arrays," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1396–1407, Mar. 2014.
- [13] N. Klausner, M. R. Azimi-Sadjadi, and L. L. Scharf, "Saddlepoint approximations for correlation testing among multiple Gaussian random vectors," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 703–707, May 2016.
- [14] D. Ramirez, J. Via, I. Santamaria, and L. Scharf, "Detection of spatially correlated Gaussian time series," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5006–5015, Jun. 2010.
- [15] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, no. 2, pp. 70–73, Jun. 1967.
- [16] H. Daniels, "Saddlepoint approximations in statistics," *Ann. Math. Statist.*, vol. 25, no. 4, pp. 631–650, Dec. 1954.
- [17] R. Butler, *Saddlepoint Approximations With Applications*. New York, NY, USA: Cambridge Univ. Press, 2007.
- [18] R. Lugannani and S. Rice, "saddlepoint approximation for the distribution of the sum of independent random variables," *Adv. Appl. Probab.*, vol. 12, no. 2, pp. 475–490, Jun. 1980.
- [19] E. Kreyszig, *Advanced Engineering Mathematics*, 10th ed. Hoboken, NJ, USA: Wiley, 2011.
- [20] G. Sammelmann, "Simulation, beamforming, and visualization of bistatic synthetic aperture sonar," *Naval Surf. Warfare Center Panama City Div.*, Panama City, FL, USA, Tech. Rep. 9407, 2010, vol. 18.
- [21] C. Rose and M. D. Smith, *Mathematical Statistics With Mathematica*. New York, NY, USA: Springer-Verlag, 2002.



Nick H. Klausner (S'08–M'14) received the M.S. and Ph.D. degrees in electrical and computer engineering from Colorado State University, Fort Collins, CO, USA, in 2010 and 2014, respectively.

During 2014–2017, he was a Research Scientist with the Information System Technologies, Inc., Fort Collins, CO, USA. He is currently with the Numerica Corporation, Fort Collins, CO, USA. His main research interests include statistical signal and image processing.



Mahmood R. Azimi-Sadjadi (M'81–LM'18) received the M.S. and Ph.D. degrees from the Imperial College of Science and Technology, University of London, London, U.K., in 1978 and 1982, respectively, both in electrical engineering with specialization in digital signal/image processing.

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

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