12
11/18/91
M.Y.P. (1)

UCRL-ID-108260

Chernoff Bounds For Class-A Noise

Pamela A. Nielsen

August 12, 1991

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This report has been reproduced
directly from the best available copy.

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information
P.O. Box 62, Oak Ridge, TN 37831
Prices available from (615) 576-8401, FTS 626-8401

Available to the public from the
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd.,
Springfield, VA 22161

Chernoff Bounds for Class A Noise

Pamela A. Nielsen
Lawrence Livermore National Laboratory¹

August 12, 1991

¹The author is with the Dept. of ECE, University of Arizona, Tucson, Az.
85721

MASTER

Chernoff Bounds for Class A Noise

Pamela A. Nielsen
Lawrence Livermore National Laboratory*

August 12, 1991

Abstract

The goal is, using a very large passive array, to determine the performance limits of a detector. The signal of interest is narrowband with a Gaussian envelope, and the contaminating noise is multivariate Class-A. Two different multivariate models for the Class A family are presented. One of the models is appropriate for array processing applications. The data is spatially dependent and temporally independent. It is shown, in the spatially independent case, that the Chernoff approximation does closely approximate the performance of the optimal detector. It is shown that the approximation improves as the number of samples increases. Unfortunately, it is also shown that the Chernoff approximation requires numerical evaluation of a M -dimensional integral. For the application here, M may be as large as 150, ruling out this approach. Two alternative approaches are examined. First, approximating the Class A model by a Gaussian model is shown to result in a poor approximation. Second, the exact likelihood ratio is approximated by a piece-wise function. While the approximation can be done with very good accuracy, the bound must be evaluated numerically.

*The author is with the Dept. of ECE, University of Arizona, Tucson, Az. 85721

1 Introduction

The problem of interest is the passive detection of a subsurface moving source using a stationary, very large array. The source will radiate a wide band of low frequency components and is assumed to be in the far-field of the array. Since the source is in the far-field, the signal strengths will be small and the large array is used for coherent processing of the signal, e.g. some type of spatial processing or beamforming will be used. The goal is to evaluate different detection methods.

This problem is quite complex, and presumably it will take several years for the above questions to be solved to the accuracy desired. Unlike RADAR, the frequencies of interest are low, and the transmission medium, the ocean, will distort the transmitted signal in a more complicated manner. Hence, assumptions which are often acceptable in RADAR applications are not necessarily acceptable in this application.

The results presented in this paper are an outgrowth of the problem stated above. For an assumed noise model, the Chernoff bound was investigated and compared to simulated and theoretical performance measures. The goal of the investigation was to determine the usefulness of using the Chernoff bounds on the probability of false alarm and the probability of a miss, for a very large passive array.

2 Signal Model

The ultimate interest is in solutions for a wideband source. Since a wideband source can be modeled as a sum of narrowband sources, we constrain the analysis to the case of a narrowband source. The results can be generalized to the wideband case simply by parallel processing for the different frequency bands. The narrowband radiated signal can be represented as

$$s_t(t) = \alpha(t)\cos(\omega_c t + \phi(t)) \quad (1)$$

where the carrier frequency ω_c is known, and the envelope and phase functions, α and ϕ are random and slowly varying. Specifically, the functions α and ϕ are approximately constant as the wave sweeps across the array. The phase is assumed to be uniformly distributed on the interval $[0, 2\pi)$ and the envelope is a zero mean Gaussian random variable. The variance of α

is dependent on the carrier frequency, and is given as $\sigma^2(\omega_c) = \sigma^2$. For convenience, the complex analytic representation of the signal will be used, e.g.

$$\tilde{s}_i(t) = \alpha(t)e^{j\phi(t)}e^{i\omega_c t} \quad (2)$$

Since the signal is narrowband and the ocean, in deep water, can be accurately modeled as a waveguide, the normal model approximation for the solution to the full wave equation can be used. Denoting the source position as $\underline{\theta}$ (this is a three-dimensional vector, either in rectangular or spherical coordinates) and the position of the i th sensor as $\underline{\chi}_i$, then, from Clay and Medwin [1] the received acoustic pressure wave can be written as

$$\begin{aligned} \tilde{s}(t) &= \alpha(t)e^{j\phi(t)}e^{i\omega_c t} \frac{\rho(\underline{\chi}_i)}{\sqrt{r(\underline{\theta}, \underline{\chi}_i)}} \sum_m q_m Z_0(\underline{\theta}) Z_m(\underline{\chi}_i) e^{-j\kappa_m r(\underline{\theta}, \underline{\chi}_i)} \\ &= \alpha(t)e^{j\phi(t)}e^{i\omega_c t} e(\omega_c, \theta) \end{aligned} \quad (3)$$

The components of the function e are defined as follows: $\rho(\underline{\chi}_i)$ is the ambient density at position $\underline{\chi}_i$; m is the mode number (the values of m are dependent on the waveguide dimensions and the source frequency); q_m is the source excitation coefficient for the m th mode; κ_m is the horizontal wave number for the m th mode (the wave number is range dependent only); $r(\underline{\theta}, \underline{\chi}_i)$ is the range between the source position and the i th sensor; Z_m is the m th mode eigenfunction. It should be understood, that if the source frequency is changed, then the coefficient values in e will change.

Rewriting the received analytic signal in vector notation, yields

$$\tilde{\underline{s}}(t) = \alpha(t)e^{j\phi(t)}e^{i\omega_c t} \underline{e}(\omega_c, \theta) \quad (4)$$

where the dimension of the replicant vector $\underline{e}$ is M , the number of sensors. Note that the vector $\underline{e}$ can be viewed as a steering vector for a beamforming operation. The mean of the received signal, $\tilde{\underline{s}}(t)$, is zero, and the signal covariance is

$$K_s = E(\tilde{\underline{s}}(t)\tilde{\underline{s}}(t)^H) = \sigma^2(\omega_c) \underline{e}(\omega_c, \theta) \underline{e}(\omega_c, \theta)^H \quad (5)$$

The received signal covariance is dependent on the signal energy through σ^2 and on the spatial covariance, through $\underline{e}\underline{e}^H$. Note the familiar decoupling of the temporal and spatial components, which occurs in narrowband beamforming.

Since the envelope is Gaussian, the complex analytic signal is Gaussian [5] with a density function

$$p_{\underline{\mathbf{z}}(t)}(\underline{\mathbf{z}}) = \frac{1}{\pi^M |\det K_s|} e^{-\underline{\mathbf{z}}^H K_s^{-1} \underline{\mathbf{z}}} \quad (6)$$

3 Noise Model

Ocean noise has been well documented as being non-Gaussian. A few recent papers demonstrating this are [7], [8], [9], [10]. It has been suggested that an appropriate model for oceanic noise is Middleton's Class-A model [8]. The Class-A model is a three-parameter family of density functions, mathematically described as [3], [4]

$$p(x) = e^{-A} \sum_{m=0}^{\infty} \frac{A^m}{m!} \frac{1}{\sqrt{2\pi\sigma_m^2}} e^{-x^2/2\sigma_m^2} \quad (7)$$

Clearly, the Class-A family is a Poisson weighted infinite sum of Gaussian density functions. The parameter $A > 0$ is the "Overlap Index". The overlap index "measures the mean deviation of a typical noise source from the average.." [4]. For the Class-A family, the parameter A is small, usually in the range 10^{-4} to 0.5.

The second parameter of the Class-A model is the "Gaussian Factor" Γ . This parameter is defined as the ratio of the Gaussian variance σ_g^2 to that of the variance or intensity of the non-Gaussian component. Γ relates the overlap index A to the variance of the m th term through

$$\sigma_m^2 = \frac{m/A + \Gamma}{1 + \Gamma} \quad (8)$$

For the Class-A family, $\Gamma > 0$ and is typically 10^{-5} to 0.1.

The final parameter of the Class-A model is the intensity of the non-Gaussian component. The Class A family was originally derived to model the electromagnetic interference caused by man-made sources. In the development, the noise was modeled as having two components, a Gaussian component with variance σ_g^2 , and a non-Gaussian component (Class A). The third parameter measures the relative strengths of the Gaussian and the

non-Gaussian component. This parameter is implicit and is important when fitting experimental data to the family.

The Class A density function is symmetric about its mean of zero. As the parameter A tends to infinity, the model approaches a Gaussian density. The variance is a monotonic increasing function of the index m . It can be easily shown that

$$\frac{\sigma_{m+1}^2}{\sigma_m^2} = 1 + \frac{1}{A\Gamma} \quad (9)$$

Working with an infinite series is typically not appealing or convenient. We therefore wish to truncate the infinite series to two terms. Spaulding [6], has shown that the series, for small A and Γ , can be truncated to two terms with less than a 2% error as

$$p_T(x) = e^{-A} \left\{ \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-x^2/2\sigma_0^2} + \frac{e^A - 1}{\sqrt{2\pi\sigma_1^2}} e^{-x^2/2\sigma_1^2} \right\} \quad (10)$$

where

$$\sigma_0^2 = \frac{\Gamma}{1 + \Gamma} \quad ; \quad \sigma_1^2 = \left(1 + \frac{1}{A\Gamma}\right) \sigma_0^2 \quad (11)$$

The standard Gauss-Gauss mixture model can be written as

$$p_m(x) = \frac{\epsilon}{\sqrt{2\pi\sigma_{m0}^2}} e^{-x^2/2\sigma_{m0}^2} + \frac{1 - \epsilon}{\sqrt{2\pi\sigma_{m1}^2}} e^{-x^2/2\sigma_{m1}^2} \quad (12)$$

which is often used in the literature, and which some claim is a truncation of the Class-A model. Typically, the mixture parameter $\epsilon \in (0, 1)$, $\sigma_0^2 < 1$, and $\sigma_1^2 > 1$. The total variance is $\sigma^2 = \epsilon\sigma_{m0}^2 + (1 - \epsilon)\sigma_{m1}^2$. The difference between p_m and p_T is illustrated in Figure 1. The Gaussian curve has a variance of 0.04762. The Class A noise has parameters $A = 0.1$ and $\Gamma = 0.05$. Therefore, $e^{-A} = 0.905$; $\sigma_0^2 = 0.04762$; $\sigma_1^2 = 9.5714$, and the variance is 0.9523. For the mixture model, the mixture parameter $\epsilon = 0.905$. There are many choices of variances, σ_{m0} and σ_{m1} which will result in the total variance being 0.9523, the same as the Class-A truncated model. The plotted curve has $\sigma_{m0} = 0.2$ and $\sigma_{m1} = 8.11$. The mixture model and the Class-A models may be the same, but there are more constraints on the parameter choices for the Class-A model.

Figure 1: Comparison of Class A and Mixture Density Functions

3.1 Multivariate Model

The data received by the array will be spatially and possibly temporally correlated. For one snap-shot in time, the data will have a spatial multivariate density function. There are at least two ways that the scalar version of the Class-A model can be modified to handle dependent data.

Case I: Set $A > 0$ but set Γ to be a $M \times M$ matrix. Then the m th covariance matrix can be written as

$$K_m = (I + \Gamma)^{-1} \left(\frac{m}{A} I + \Gamma \right) \quad (13)$$

and the multivariate density written as

$$p(\underline{x}) = e^{-A} \sum_{m=0}^{\infty} \frac{A^m}{m!} \frac{1}{\sqrt{(2\pi)^{M/2} |\det K_m|}} e^{-\underline{x}^T K_m^{-1} \underline{x} / 2} \quad (14)$$

The multivariate density has a mean of zero and a covariance given by

$$E \underline{x} \underline{x}^T = e^{-A} \sum_{m=0}^{\infty} \frac{A^m}{m!} K_m. \quad (15)$$

The problem with this approach is the interpretation of Γ .

Case II: An alternative to the above approach is to let both A and Γ be scalars, as before. Hence, the covariance of the m th term is

$$K_m = \sigma_m^2 K_0 \quad (16)$$

with the scalar's σ_m^2 defined as in (8). K_0 could be interpreted as the spatial covariance matrix, e.g. the covariance of the inputs to the different sensors *at the same time*. The multivariate density can be written as

$$p(\underline{x}) = e^{-A} \sum_{m=0}^{\infty} \frac{A^m}{m!} \frac{1}{\sqrt{(2\pi)^{M/2} |\det K_0| \sigma_m}} e^{-\underline{x}^T K_m^{-1} \underline{x} / 2} \quad (17)$$

The multivariate density has a zero mean and a covariance of

$$E \underline{x} \underline{x}^T = K_0 e^{-A} \sum_{m=0}^{\infty} \frac{A^m}{m!} \sigma_m^2 \quad (18)$$

This model can be truncated to two terms, in the manner of Spaulding [6] yielding the following Class-A model

$$p_T(\underline{x}) = e^{-A} \left\{ \frac{1}{\sqrt{(2\pi)^{M/2} |\det K_0| \sigma_0}} e^{-\underline{x}^T K_0^{-1} \underline{x} / (2\sigma_0^2)} + \frac{e^A - 1}{\sqrt{(2\pi)^{M/2} |\det K_0| \sigma_1}} e^{-\underline{x}^T K_0^{-1} \underline{x} / (2\sigma_1^2)} \right\} \quad (19)$$

with σ_0^2, σ_1^2 defined as before, and K_0 the spatial covariance.

4 Hypothesis Problem

The detection problem can be modeled as the following simple binary hypothesis test [2].

$$\begin{aligned} H_0 &: \underline{X} = \underline{N} \\ H_1 &: \underline{X} = \underline{N} + \underline{S} \end{aligned} \quad (20)$$

$\underline{N}$ contains the M -samples of the noise, and $\underline{S}$ contains the M -samples of the signal due to the source, received at the array at some time t .

The noise is multivariate Class-A model. Using the complex analytic representation,

$$p_T(\underline{x}|H_0) = e^{-A} \left\{ \frac{1}{\pi^M |\det K_0| \sigma_0^2} e^{-\underline{x}^H K_0^{-1} \underline{x} / (\sigma_0^2)} + \frac{e^A - 1}{\pi^M |\det K_0| \sigma_1^2} e^{-\underline{x}^H K_0^{-1} \underline{x} / (\sigma_1^2)} \right\} \quad (21)$$

For the alternative hypothesis, the conditional density function is the convolution of the density of the signal with that of the noise. On recognizing that the convolution of two Gaussian densities yields a Gaussian density, the conditional density given that the signal is present can be quickly written as

$$p_T(\underline{x}|H_1) = e^{-A} \left\{ \frac{1}{\pi^M |\det(K_0 \sigma_0^2 + K_s)|} e^{-\underline{x}^H (K_0 \sigma_0^2 + K_s)^{-1} \underline{x}} + \frac{e^A - 1}{\pi^M |\det(K_0 \sigma_1^2 + K_s)|} e^{-\underline{x}^H (K_0 \sigma_1^2 + K_s)^{-1} \underline{x}} \right\} \quad (22)$$

This density is not Class-A because the variances do not follow the update rule defined by (8).

These expressions for the conditional density functions are, of course, for a single snap-shot in time. In practice data will be collected over a time interval, hence the exact density representation will be multivariate spatially and temporally. If the samples can be taken far enough apart, then the samples will be uncorrelated and possibly can be considered temporally independent. Let N_t be the number of time snap-shots. Assuming independence, the conditional density under each hypothesis is the product of the conditional densities above, e.g.

$$p_T(\underline{x}|H_0) = \prod_{l=1}^{N_t} \left(\frac{e^{-A}}{\pi^M |\det K_0| \sigma_0^2} e^{-\underline{x}_l^H K_0^{-1} \underline{x}_l / (\sigma_0^2)} + \frac{1 - e^{-A}}{\pi^M |\det K_0| \sigma_1^2} e^{-\underline{x}_l^H K_0^{-1} \underline{x}_l / (\sigma_1^2)} \right) \quad (23)$$

Here, $\underline{x}$ is a $M \times N_t$ matrix, and $\underline{x}_l$ is the l th row. For the alternative hypothesis, the conditional density function can be similarly written.

5 Chernoff Bounds

The usefulness of the Chernoff bound to determine upper bounds on the false alarm and miss probabilities was investigated for the problem described. The Chernoff bound [2] hinges on evaluating $\mu(s)$, where

$$\mu(s) = \log \left\{ \int_{-\infty}^{\infty} \left(\frac{p(\underline{x}|H_1)}{p(\underline{x}|H_0)} \right)^s p(\underline{x}|H_0) d\underline{x} \right\} \quad (24)$$

If μ exists, then the false alarm can be bounded above by

$$P_{fa} \leq e^{\mu(s)-s\gamma} \quad , \quad 0 < s < 1 \quad (25)$$

where γ is the threshold of the test. The miss probability can be bounded above by

$$P_m \leq e^{\mu(s)+(1-s)\gamma} \quad , \quad 0 < s < 1 \quad (26)$$

The best bound is found by optimizing over s , which yields

$$\begin{aligned} P_{fa} &\leq e^{\mu(s)-s\dot{\mu}(s)} \\ P_m &\leq e^{\mu(s)+(1-s)\dot{\mu}(s)} \quad , \quad 0 < s < 1 \end{aligned} \quad (27)$$

The value of s is chosen such that $\gamma = \dot{\mu}(s)$. It can be shown [2] that for a detector l ,

$$E(l(\underline{x}|H_0)) = \dot{\mu}(0) \quad E(l(\underline{x}|H_1)) = \dot{\mu}(1) \quad (28)$$

hence the choice of threshold must be between the means of the detector under the two different hypotheses.

If the detector can be written as

$$l(\underline{x}) = \sum_i \log(g(x_i)) \quad (29)$$

then the central limit theorem can be applied to yield the following approximate expressions for the error probabilities

$$\begin{aligned} P_{fa} &\approx \text{erfc}(s\sqrt{\ddot{\mu}(s)})e^{\mu(s)-s\dot{\mu}(s)+s^2\ddot{\mu}(s)/2} \\ P_m &\approx \text{erfc}((1-s)\sqrt{\ddot{\mu}(s)})e^{\mu(s)+(1-s)\dot{\mu}(s)+(1-s)^2\ddot{\mu}(s)/2} \end{aligned} \quad (30)$$

where $0 < s < 1$ and $erfc$ is the complimentary error function defined by

$$erfc(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-y^2/2} dy \quad (31)$$

Generally, the Chernoff bounds are loose for N_t small; however, as N_t grows then the bounds become tighter. The approximations are generally quite good when N_t is of moderate size, say 8 or so. Therefore, evaluation of the Chernoff bound requires evaluation of μ and its first two derivatives.

5.1 Multivariate Bound Evaluation

For the problem stated above, the first step to evaluating the Chernoff bound is to evaluate μ .

$$\begin{aligned} \mu(s) = \log \int_{-\infty}^{\infty} & \left(\frac{e^{-A}}{\pi^M |\det(K_0 \sigma_0^2 + K_s)|} e^{-\underline{x}^H (K_0 \sigma_0^2 + K_s)^{-1} \underline{x}} + \right. \\ & \left. \frac{1 - e^{-A}}{\pi^M |\det(K_0 \sigma_1^2 + K_s)|} e^{-\underline{x}^H (K_0 \sigma_1^2 + K_s)^{-1} \underline{x}} \right)^s \\ & \left(\frac{e^{-A}}{\pi^M |\det K_0| \sigma_0^2} e^{-\underline{x}^H K_0^{-1} \underline{x} / (\sigma_0^2)} + \frac{1 - e^{-A}}{\pi^M |\det K_0| \sigma_1^2} e^{-\underline{x}^H K_0^{-1} \underline{x} / (\sigma_1^2)} \right)^{(1-s)} d\underline{x} \quad (32) \end{aligned}$$

A cursory examination of the above integral results in the understanding that it cannot be evaluated analytically and must be evaluated numerically. Accurate numerical evaluation of integrals for $M > 2$ dimensions is very difficult. For our case, M could be quite large, say 100 to 150, and numerical integration is not feasible.

5.2 Univariate Bound Evaluation

As an alternative, we evaluate the univariate case. The intention is to determine if the bounds are accurate and if an alternative method for evaluating the μ function can be derived. Perhaps, some conclusions can be drawn about the bound for the multivariate case from the univariate case results. For $M = 1$ the scalar density functions will be used. Hence,

$$p(x|H_0) = e^{-A} \left\{ \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-x^2/2\sigma_0^2} + \frac{e^A - 1}{\sqrt{2\pi\sigma_1^2}} e^{-x^2/2\sigma_1^2} \right\} \quad (33)$$

$$p(x|H_1) = e^{-A} \left\{ \frac{1}{\sqrt{2\pi(\sigma_0^2 + \sigma_s^2)}} e^{-x^2/2(\sigma_0^2 + \sigma_s^2)} + \frac{e^A - 1}{\sqrt{2\pi(\sigma_1^2 + \sigma_s^2)}} e^{-x^2/2(\sigma_1^2 + \sigma_s^2)} \right\} \quad (34)$$

For the case when $N_t = 1$, the first and second derivatives are

$$\begin{aligned} \dot{\mu}(s) &= \frac{d}{ds} \mu(s) \\ &= e^{-\mu(s)} \int_{-\infty}^{\infty} \left(\frac{p(\underline{x}|H_1)}{p(\underline{x}|H_0)} \right)^s p(\underline{x}|H_0) \log \left(\frac{p(\underline{x}|H_1)}{p(\underline{x}|H_0)} \right) d\underline{x} \end{aligned} \quad (35)$$

and

$$\ddot{\mu}(s) = -[\dot{\mu}(s)]^2 + e^{-\mu(s)} \int_{-\infty}^{\infty} \left(\frac{p(\underline{x}|H_1)}{p(\underline{x}|H_0)} \right)^s p(\underline{x}|H_0) \log^2 \left(\frac{p(\underline{x}|H_1)}{p(\underline{x}|H_0)} \right) d\underline{x} \quad (36)$$

Numerical techniques are needed to evaluate μ and the first two derivatives. The extension to the situation when $N_t > 1$ and the data are temporally independent is straight forward on recognizing that the multivariate density function will be a product of univariate densities, and that the logarithm will convert the product to a sum. See [2] for examples.

6 Results

The Chernoff bound was evaluated numerically for the case when the noise was Class-A with parameters $A = 0.1$ and $\Gamma = 0.05$. The number of snapshots was taken to be $N_t = 1, 2, 5$, and the sample by sample SNR was either one or two. Using the expressions above bounds on the probability of false alarm and detection were obtained.

These bounds were compared to the theoretical performance and to performance determined via simulation. For theoretical performance, the optimal log-likelihood ratio detector was determined and the decision regions for the null and alternative hypothesis were determined. Then, the conditional density functions were numerically integrated over these regions to determine the error probabilities. Note that this technique can be used only when $N_t = 1$.

Figure 2: Comparison of theoretical and simulated performance.

The performance was simulated by generating data belonging to the Class-A family and the conditional density functions under the alternative hypothesis. The 4000 samples were generated using the rejection method. Comparison of empirical density functions of the data with theoretical density functions shows very good agreement.

Figure 2 plots the simulated performance as compared to the theoretical performance. The detection probability is simulated very accurately, while there is a small error in the false alarm probability. The ROC comparing the simulated and the theoretical probabilities is shown in Figure 3. The effect of the small error in the false alarm probability, is that for a fixed probability of detection, the simulation slightly over estimates the probability of false alarm.

Figures 4 plots the simulated performance as compared to the Chernoff bound. The results in the figures are typical. The Chernoff bound on the false alarm probability is much larger than the actual false alarm probability, and the lower bound on the probability of detection is much lower than the actual probability. The bounds are so loose that they are not useful and they are not a good indicator of the detectors performance. As N_t is increased, the bounds become tighter, but they are still too loose to be useful.

Figure 3: ROC of the simulated and theoretical performance.

Figure 4: Comparison of Chernoff bounds and simulated performance.

Figure 5: Comparison of Chernoff approximations and simulated performance.

Since the bound was so loose, and since we will use multiple snap shots ultimately, the approximate expressions for the error probabilities were evaluated. The results are shown in Figure 5. There is a surprisingly good agreement between the theoretical, simulated, and approximate expressions. As expected, the approximate expression becomes more accurate as the number of snap shots increases.

To summarize, we can evaluate the Chernoff bounds. In fact, the approximate expressions are reasonably good even for N_t small, but the extension to the case when the data is spatially dependent is not clear. Hence, the original problem has not been solved.

6.1 Gaussian Approximation

Suppose the Class A noise is approximated as Gaussian. Will this cause a significant change in the Chernoff bounds and/or approximations? Unfortunately, it will. If the noise has a variance of σ_n^2 and the signal has a variance

Figure 6: Comparison of μ function for Gaussian and Class A noise.

of σ_s^2 , then for real data the μ function can be evaluated to be

$$\mu(s) = \frac{1}{2} \left\{ (1-s) \log(1 + \sigma_s^2/\sigma_n^2) - \log(1 + (1-s)\sigma_s^2/\sigma_n^2) \right\} \quad (37)$$

Similarly, for the complex case, using the notation defined earlier,

$$\mu(s) = (1-s) \log(1 + SNR) - \log(1 + (1-s)SNR) \quad (38)$$

where $SNR = \sigma_s^2 \underline{e}^H K_n^{-1} \underline{e}$. The factor of one-half disappears because the samples are complex. The SNR is essentially the output of a beamformer. The expressions above can be easily evaluated without any integration. The question is really whether the above μ function is a good approximation of the μ function for Class-A noise. A comparison is shown in Figure 6. Note that the shapes of the curves are significantly different, with the Gaussian μ function strongly asymmetric.

Figure 7 compares the approximate expressions for the probability of false alarm and detection for the Class A and the Gaussian cases. Note the difference in the SNR. The Gaussian noise was modeled as having a variance $\sigma_n^2 = \sigma_0^2$. The signal variance was chosen such that the sample-by-sample SNR for the Class A noise was unity. It is important to notice that the

Figure 7: Comparison of performance for Gaussian and Class A noise.

shapes of the two curves are different. This suggests that assuming the noise is Gaussian will not provide an accurate estimate of the error probabilities. The graph shown is typical, though comparisons were made for different SNR and different temporal sample lengths N_t .

An explanation for the difference can be found in the shape of the likelihood ratio, shown in Figure 8. The likelihood ratio for Class A noise censors some of the large values (it does increase towards infinity as the amplitude increases towards infinity). To conclude, using a simple Gaussian approximation is not appropriate and will provide results which are misleading.

6.2 Piece-wise Approximation

A better approach than using the Gaussian approximation, might be to approximate the likelihood ratio, $L(x) = p(x|H_1)/p(x|H_0)$ in the calculation of the μ function, with a simpler function. Several different functions were considered, but it was decided to break the log-likelihood ratio, $l(x) = \log(L(x))$, into three regions, as illustrated in Figure 9.

The log-likelihood ratio was approximated as follows. In region one, set

Figure 8: Comparison of log likelihood ratio for Gaussian and Class A noise.

Figure 9: Regions for the piece-wise approximation.

Figure 10: Piece-wise approximation of the likelihood ratio.

$\hat{l}(x)$ equal to the log-likelihood ratio of two Gaussian random variables, each with a mean of zero, and variances of $\sigma_s^2 + \sigma_0^2$ and σ_0^2 respectively. Therefore in region one, $\hat{l}(x) = a_1 x^2 + a_2$, with the constants a_1 and a_2 determined by the variances. The break point for region one is the value of x such that $\hat{l}(x) = \max(l(x))$. In region two the log-likelihood ratio is approximated by the line $\hat{l}(x) = mx + b$. In region three, $\hat{l}(x) = 0$. The break point between regions two and three determines the values of the slope and intercept. These two parameters were chosen such that the line starting at the end point of region one approximately passed through the inflection point of $l(x)$. Finally, we have $\hat{L} = \exp(\hat{l})$ completing the approximation. For the case when the $SNR = 1$, $A = 0.1$, and $\Gamma = 0.05$, an example of the piece-wise approximation is shown in Figure 10. The break points between the regions are ± 0.6755 and ± 3.755 .

The μ function using this approximation can be evaluated. Note that while the approximation is not differentiable in x , it does not imply that μ is not differentiable with respect to s . A comparison of the μ functions using the exact and the piece-wise approximation is shown in Figure 11. The curves are indistinguishable.

To conclude, we can accurately approximate the likelihood ratio and eval-

Figure 11: Comparison of μ for exact and approximate likelihood ratio.

uate μ . There is no reason to believe that the first and second derivatives cannot be easily evaluated also. However, the original problem has not really been solved. That is, the solution method does not remove the need for evaluating a M -dimensional integral numerically. Why is this so can be illustrated as follows:

$$\mu(s) = \log \left\{ \int_{-\infty}^{\infty} L(x)^s p(x|H_0) dx \right\} \approx \log \left\{ \int_{-\infty}^{\infty} \hat{L}(x)^s p(x|H_0) dx \right\} \quad (39)$$

where

$$\hat{L}(x) = \begin{cases} 1 & |x| \geq 3.755 \\ e^{m|x|+b} & 0.6755 \geq |x| \geq 3.755 \\ \frac{\sigma_0}{\sqrt{\sigma_0^2 + \sigma_i^2}} \exp \frac{x^2 \sigma_i^2}{2\sigma_0^2(\sigma_0^2 + \sigma_i^2)} & 0.6755 \geq |x| \geq 3.755 \end{cases} \quad (40)$$

and $p(x|H_0)$ is the Class A noise density described earlier. In the Gaussian case the integral limits are $\pm\infty$, which allows the integral to be evaluated by a proper choice of constants. Here, the integral is broken into a sum where at least one limit is finite. Hence, the integral must still be evaluated

numerically and the original problem of evaluating the Chernoff bound for data which is temporally independent, but is spatially dependent has not been solved.

7 Conclusions

The original problem is the determination of the performance of a very large array. The signal was assumed to be narrowband with a Gaussian envelope, and the noise was modeled as being a member of the multivariate Class-A family. Two different multivariate models for the Class A family were presented. One of the models is appropriate for array processing applications. The data was assumed to be spatially dependent and temporally independent. It was found that the Chernoff approximation does closely approximate the performance for this model if the array dimension is one, that is, if the data is spatially independent. It was found that the approximation improves as the number of samples increases. Unfortunately, it was also found that direct calculation of the Chernoff approximation must be numerical, involving a M -dimensional numerical integral. For the application here, M may be as large as 150, ruling out this possibility. The bounds were compared to theoretical and simulated performance where possible.

Two alternative approaches were examined. First, approximating the Class A model by a Gaussian model was shown to result in a poor approximation. Second, the exact likelihood ratio was approximated by a piecewise function. The approximation can be done with very good accuracy, though the bound must be evaluated numerically. Therefore, the original problem has not been solved. The author would like to thank D. Harris and J. Candy. Additionally, the author is interested in hearing about alternative approaches.

References

- [1] C.S. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications*, John Wiley and Sons, New York, 1977.
- [2] H.L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I*, John Wiley and Sons, New York, 1968.

- [3] D. Middleton, *Introduction to Statistical Communication Theory*, Peninsula Press, Los Altos, CA, 1987.
- [4] D. Middleton, "Procedures for Determining the Parameters of the First-Order Canonical Models of Class A and Class B Electromagnetic Interference," *IEEE Trans. on EMC*, EMC-21(3), 190-208, 1979.
- [5] K. Miller, "Complex Gaussian Processes", *SIAM Review*, 11(4), 544-567, 1969.
- [6] A.D. Spaulding D. Middleton, "Optimum Reception in Impulsive Interference Environment," *OT Report 75-67*, June 1975. (Office of Telecom., U.S. Dept. of Commerce, 325 Broadway, Boulder, CO. 80303), see 193-195.
- [7] P.L. Brockett, M. Hinich, and G.R. Wilson, "Bispectral Characterization of Ocean Acoustic Time Series: Nonlinearity and Non-Gaussianity", in *Topics in Non-Gaussian Signal Processing*, Eds. E.J. Wegman, S.C. Schwartz, and J.B. Thomas, Springer-Verlag, New York, 2-16, 1989.
- [8] D.R. Powell and G.R. Wilson, "Class A Modeling of Ocean Noise Processes", in *Topics in Non-Gaussian Signal Processing*, Eds. E.J. Wegman, S.C. Schwartz, and J.B. Thomas, Springer-Verlag, New York, 17-28, 1989.
- [9] F.W. Machell, C.S. Penrod, and G.E. Ellis, "Statistical Characteristics of Ocean Acoustic Noise Processes," in *Topics in Non-Gaussian Signal Processing*, Eds. E.J. Wegman, S.C. Schwartz, and J.B. Thomas, Springer-Verlag, New York, 29-57, 1989.
- [10] P.A. Nielsen and J.B. Thomas, "A Comparison of Parametric and Non-Parametric Performance Levels for Underwater Noise", *J. Acous. Soc. Am.*, 225-236, 1990.

END

**DATE
FILMED**

12/19/91