CHARACTERIZATION OF MARINE NOISE USING BEAULIEU SERIES

Mrinal K. Mandal[1], *Aditya A. Saha*[2]

Department of Electrical and Computer Engineering[1], Department of Physics[2],
University of Alberta, Edmonton, Canada

ABSTRACT

This paper proposes a computational technique for estimating parameters of probability density functions (PDFs) governing marine noise, using Beaulieu series. The PDFs are assumed to be mixtures-of-Gaussians from the widely used Middleton's Class A model. The Beaulieu series for such PDFs are derived. Such an approach is orders of magnitude more efficient than approaches based on convolution integrals, and can be done in real-time. The computational complexity for the technique is then derived. Numerical simulations of these estimates show the method to be robust with respect to perturbations prevailing in the ocean. These results make it obvious that such a procedure can be used in conjunction with existing detectors on sonar platforms.

***Index Terms*—** Marine noise, PDF estimation, Beaulieu series, Asymptotic accuracy.

1. INTRODUCTION

The characterization of the probability density function(PDF) governing marine noise, in real-time, is a problem of immense interest in sonar signal processing. This stems from the fact that the system parameters and performance of many optimal [3] and sub-optimal nonlinear detectors [7] are dependent on the marine noise PDF. The estimation of the marine noise PDF from ocean acoustic time series, a functional optimization problem, is a non-trivial and computationally intensive task [2]. The efficiency of this procedure can be greatly improved by the use of a method developed by Beaulieu [5].

The Beaulieu series, refers to a method for the computation of the cumulative density function(CDF) and equivalently the PDF, for a sum of independent random variables. This method has been used in numerous problems in communication ([6] and references therein). However, to the author's knowledge, its application to sonar signal processing has not been considered before.

In this paper we derive the Beaulieu series for a large class of marine noise PDFs, and propose an iterative technique for estimating the parameters defining the PDF from sonar time-series. The remainder of this paper goes as follows: Section 2 presents some background information; Section 3 presents the proposed technique; Sections 4 and 5 contain estimates of the computational complexity and a few results on the numerical stability of this technique, respectively, and finally Section concludes the paper.

2. REVIEW OF BEAULIEU SERIES

Let $X_i, i = 1, \dots M$ denote independent R.V's each with a PDF $f(.)$. Denote the sum of the M Random Variables (RVs) by Y i.e. $Y = \sum_{i=1}^{M} X_i$ and its PDF and CDF by $g(.)$ and $G(.)$, respectively. From elementary probability theory, the PDF of Y can be expressed as the convolution of M individual PDFs, as follows

$$g(y) = f * f * \cdots * f(y), \tag{1}$$

where the symbol $*$ represents a convolution operation. The cumulative distribution function (CDF), denoted by $G(.)$ can be found by integrating $g(.)$. The compuation of the integral for $g(.)$ in Eq.(1) and $G(.)$, while feasible, represent a computationally exorbitant task. A great simplification can be achieved by using an infinite series representation for the CDF $G(.)$ developed by Beaulieu [5]. It can be shown that the CDF of the sum $G(.)$ can be written as follows

$$G(y) = \frac{1}{2} - \sum_{n=1,odd}^{\infty} \frac{1}{n\pi j} \times$$
$$\prod_{i=1}^{M} E\left[e^{jn\omega(X_i - y/M)}\right] - \prod_{i=1}^{M} E\left[e^{-j\omega n(X_i - y/M)}\right] \tag{2}$$

The Expectation $E(.)$ is carried over the probability measure $f(x)dx$. In Eq.(2) $\omega = 2\pi/T$, where the term T represents a bound: $T = 2\max(|B - Y|)$. The Bound B is chosen such that the error resulting from truncation, denoted here by $\Delta = Prob(|Y| > B)$ is "negligible". For bounded variables, it can be chosen as the sum $B = Mb$, where b denotes the upper bound of the support of $f(.)$. In such a case Δ is trivially zero. For unbounded variables such as those considered in this paper, $\Delta = G(-B) + 1 - G(B)$. which can be upper bounded using the Chernoff Bound (or others). The equality in (2) is in the mean square sense and is guranteed by a Fourier theorem [5]. The utility of Beaulieu series representation for the CDF lies in the fact that $M-1$ successive integrals by quadrature, have been replaced by a single super-exponentially convergent summation.

3. PROPOSED TECHNIQUE

In section 2 we presented a brief review of Beaulieu series. In this section we develop a technique to apply it to estimating marine noise PDFs. Our proposed technique has three main components. These are:

- Derivation of the *Beaulieu series* for mixture-of-Gaussian PDFs which largely govern marine noise;
- Calculation of a cost functional which relates the theoretical estimate of CDF $G_{\alpha,\beta}(.)$ from Beaulieu series to that observed from sonar data; and lastly
- Calculation of the optimal values for the parameters α and β which the define marine noise PDF.

In the following sections these components are described in succession.

3.1. Beaulieu series for marine noise PDFs

Phenomenological studies of marine environments [8] show the PDF governing ocean acoustic noise can be modelled as a weighted sum of two Gaussians and is given by

$$f(y) = \frac{c}{\sqrt{2\pi}}\left[\alpha e^{-(cy)^2/2} + \frac{1-\alpha}{\beta}e^{-(cy)^2/2\beta^2}\right], \qquad (3)$$

where α and β denote the mixing parameter and the ratio of the deviations of the component PDFs, respectively, and

$$c = [\alpha + (1-\alpha)\beta^2]^{\frac{1}{2}}.$$

In Eq.(2) it can be observed that the series involves the expectation over complex exponentials. Using Euler's formula, each individual term within the product in the *Beaulieu series* can be written in terms of the quantities $E\left[\cos\left(n\omega(X_i - y/M)\right)\right]$ and $E\left[\sin\left(n\omega(X_i - y/M)\right)\right]$. They are evaluated to be

$$\begin{aligned} &E\left[\cos\left(jn\omega(X_i - y/M)\right)\right] \\ &= \left[\alpha e^{-n^2\omega^2/2c^2} + (1-\alpha)e^{-n^2\omega^2\beta^2/2c^2}\right]\cos\left(n\omega y/M\right) \\ &E\left[\sin\left(jn\omega(X_i - y/M)\right)\right] \\ &= \left[\alpha e^{-n^2\omega^2/2c^2} + (1-\alpha)e^{-n^2\omega^2\beta^2/2c^2}\right]\sin\left(n\omega y/M\right). \end{aligned} \qquad (4)$$

Combining Eqs.(2) and (4) the *Beaulieu series* for mixture-of-Gaussian PDFs can be expressed as

$$G_{\alpha,\beta}(y) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1,odd}^{\infty}\frac{\sin\left(n\omega y/M\right)}{n}\times \left[\alpha e^{-n^2\omega^2/2c^2} + (1-\alpha)e^{-n^2\omega^2\beta^2/2c^2}\right]. \qquad (5)$$

A few plots of $G_{\alpha,\beta}(y)$ for the typical values of parameters (α,β) are shown in Fig. 1.

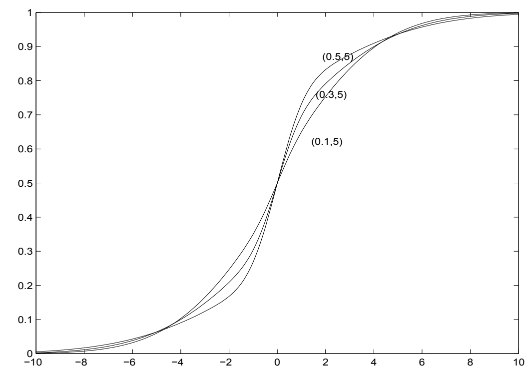

Fig. 1. Plots of $G_{\alpha,\beta}(y)$ using *Beaulieu series* for representative values of (α,β) and with $M=10$. We assume the values $T = 196.8$ and where $\omega = \pi/98.4$.

3.2. The cost functional

Observational estimates of the PDF $g_{obs}(y)$ and the CDF $G_{obs}(y)$ can be calculated as follows. Let $\{X_n\}_{n=1}^{M}$ denote M independent realizations of ambient oceanic noise. In practice, they refer to the time series recorded on different hydrophones or different windowed segments from the same hydrophone [2]. Given sufficiently many samples, a histogram of the RV $Y = \left\{\sum_{n=1}^{M} X_n\right\}$ can be constructed by bin allocation techniques [2]. This histogram, when normalized and integrated, yields $g_{obs}(y)$ and $G_{obs}(y)$ respectively.

The mean squared error (MSE) between the theoretical and empirical estimates of the CDF for particular values of (α,β) is then given by

$$MSE(\alpha,\beta) = \int_{y=-\infty}^{\infty}|G_{\alpha,\beta}(y) - G_{obs}(y)|^2 dy. \qquad (6)$$

A surface plot of $MSE(\alpha,\beta)$ is shown in Fig. 2. An synthetic estimate of $G_{obs}(.)$ is generated using the Beaulieu series with assumed values of the mixing parameters $(\alpha,\beta) = (0.3, 5)$. It is observed that the functional is largely unimodal in nature.

3.3. Optimization technique

It can be shown that the set of mixture-of-Gaussian PDFs in Eq.(3) and their Beaulieu series in Eq.(5) are not convex (the proof is provided in the Appendix). As a result of this claim, numerical techniques usually used for such functional optimization problems, such as those based on variational principles or projection-onto-convex-sets (POCS) are untenable. Therefore, we have used a less elegant and more intensive

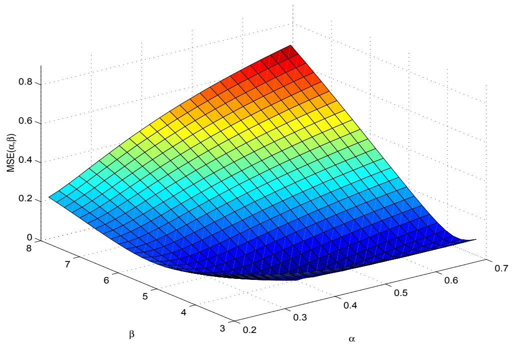

Fig. 2. A two-dimensional plot of $MSE(\alpha, \beta)$ with assumed values for $G_{obs}(y)$.

technique for our problem. The optimal estimates $(\alpha_{opt}, \beta_{opt})$ are given by the values which minimize the MSE as follows:

$$(\alpha_{opt}, \beta_{opt}) = Argmin_{\alpha,\beta} MSE(\alpha, \beta). \tag{7}$$

These estimates can be found by a classical, deterministic algorithm developed by Brent [1]. Brent's algorithm, unlike gradient based techniques, only requires the evaluation of the function being mimimized (here $MSE(\alpha, \beta)$) and not its derivatives. In addition, this benchmark algorithm is guaranteed to converge to the global optimum in a finite number of steps. The proposed technique can then be summarized as follows:

- Minimize $MSE(\alpha, \beta)$ defined in Eq.(6) using Brent's algorithm;
- At each iteration evaluate $G(\alpha, \beta)$ using Beaulieu series given in Eq.(5).

The stopping criterion is achieved when within 0.1% of the global minimum of $MSE(\alpha, \beta)$.

4. COMPUTATIONAL COMPLEXITY

The computational complexity of the proposed algorithm is evidently the product of the complexities of (i) the Beaulieu series summation given by Eq.(5); (ii) the computation of the integral in Eq.(6) by quadrature and lastly (iii) Brent's algorithm.

The complexity of the Beaulieu series summation in step (i), denoted by C_1 is proportional to N_1 which denotes the number of terms after which the series in Eq.(5) is truncated. Therefore,

$$C_1 = kN_1 \tag{8}$$

where k is a proportionality constant denoting the number of arithmetic operations per term in the series with the numerical value $k = 4$. This necessitates an appropriate value for N_1 which will now be derived. It can be shown [5] that the error incurred by truncating the summation at the N^{th} term, denoted here by R_N is upper bounded by: $|R_{N_1}| \leq \left\{[\sup f(x)]\left[\frac{T}{4} - \frac{2T}{\pi^2}\sum_{n=1,odd}^{N_1}\frac{1}{n^2}\right]\right\}^{1/2}$. From the bound given above, for the chosen value of $T \sim 200$ and a suitable error margin we get $N_1 = 500$.

Theoretical estimates of the complexities of parts (ii) and (iii) denoted by C_2 and C_3 are found [1] and [4] to be:

$$\begin{aligned} C_2 &= 2B^2\frac{\sqrt{U_\alpha U_\beta}}{\epsilon}\log_2\left(2B\sqrt{\frac{U_\alpha}{2\epsilon}}\right)\log_2\left(2B\sqrt{\frac{U_\beta}{2\epsilon}}\right), \\ C_3 &= \left[\frac{2}{3}B^3U_y\right]^{1/2}\epsilon^{-1/2}, \end{aligned} \tag{9}$$

where U_α, U_β and U_y denote upper bounds on various second order partial derivatives of the functional being optimized given by $Sup|\frac{\partial^2 MSE(\alpha,\beta)}{\partial\alpha^2}| \leq U_\alpha$, $Sup|\frac{\partial^2 MSE(\alpha,\beta)}{\partial\beta^2}| \leq U_\beta$ and $Sup|\frac{\partial^2 MSE(\alpha,\beta)}{\partial y^2}| \leq U_y$, respectively.

The total computational complexity of the procedure C_{total} is then given by

$$\begin{aligned} C_{total} &= C_1 \times C_2 \times C_3, \\ &= k\left[(8/3)^{1/2}B^{7/2}(U_yU_\alpha U_\beta)^{1/2}\right]N_1\epsilon^{-3/2} \\ &\quad \log_2\left(2B\sqrt{\frac{U_\alpha}{2\epsilon}}\right)\log_2\left(2B\sqrt{\frac{U_\beta}{2\epsilon}}\right) \end{aligned} \tag{10}$$

It can be inferred for sufficiently high accuracy i.e. as $\epsilon \to 0$, the complexity is proportional to

$$C_{total} \propto N_1\epsilon^{-3/2}(\log_2\epsilon)^2 \tag{11}$$

For a typical marine ocean acoustic environment, this translates to $\sim 10^7$ elementary arithmetic operations for the entire optimization routine. On a cluster of 4 Pentium-5 160Mhz computers this would require $0.01 - 0.02$ secs, and $0.02 - 0.04$ secs on a Texas Instruments(TI) DSP platform (Model: TMS320C6713).

5. ACCURACY AND STABILITY

In this section, we will evaluate the stability of these estimates to perturbations in the oceanic environment. In the absence of ground-truthed sonar data, this will be done using synthetic data. In order to quantify the stability of the estimates in Eq.(7) with respect to errors in measurement and modelling, we introduce an error term η in the summed RV as follows

$$\widetilde{Y} = \sum_{n=1}^{M} X_n + \eta. \tag{12}$$

This error, by assumption, can be modelled by a zero-mean RV of unknown variance σ^2 governed by the following Gaussian PDF

$$f_\eta(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-y^2/2\sigma^2}. \quad (13)$$

The CDF of the composite variable $\widetilde{Y}$ can be expressed as

$$\widetilde{G}_{\alpha,\beta}(y) = G_{\alpha,\beta} * f_\eta(y). \quad (14)$$

By extension, the recomputed values of the MSE and the optimal estimates in Eqs.(6) and (7) can then be denoted by $\widetilde{MSE}(\alpha, \beta)$ and $(\widetilde{\alpha}_{opt}, \widetilde{\beta}_{opt})$ respectively.

In the following example, we assume the ideal, unknown mixing parameters to be $(\alpha, \beta) = (0.3, 5)$. The recomputed estimates $(\widetilde{\alpha}_{opt}, \widetilde{\beta}_{opt})$ calculated for increasing values of σ^2, signifying an increasing deviation from the assumed model of a mixture-of-Gaussian PDF are given in Table 1.

Table 1: Effect of increasing modelling error on estimates (α, β).

σ	$\widetilde{\alpha}_{opt}$	$\widetilde{\beta}_{opt}$
0.2	0.32	5.5
0.5	0.34	6.1
1.0	0.4	7.0
2.0	0.5	8.0
5.0	0.54	9.0
10	0.54	10.0

These values indicate that the estimates remain stable and bounded even for very high error in modelling and measurements.

6. CONCLUSIONS

We develop a procedure for the application of Beaulieu series to the estimation of PDFs governing marine noise. The Beaulieu series for a large class of marine noise PDFs are first derived. An iterative procedure to compute the optimal estimates of the parameters defining the PDF, from sonar time-series is then formulated. Unlike existing methods, this procedure can be performed in real-time. Numerical simulations indicate that the resulting estimates are stable to perturbations prevailing in the marine environment. These results indicate that such a procedure can be used in tandem with nonlinear detectors with system parameters which depend on marine noise PDFs.

7. APPENDIX

Claim (Nonconvexity of Middleton's Class A model): *The set of all mixture-of-Gaussian PDFs defined in Eq(3) are not convex.*

Proof(*by counter-example*): Consider 2 PDFs $f_1(.)$ and $f_2(.)$ parametrized by the values $\alpha_1 = 1/2, \beta_1 = 3$ and $\alpha_2 = 5/6, \beta_2 = 5$ respectively. From (4) it is observed they both have the value $c = \sqrt{5}$. Therefore they can be written as

$$\begin{aligned} f_1(y) &= \sqrt{\frac{5}{2\pi}} \left[\frac{1}{2} e^{-5y^2/2} + \frac{1}{6} e^{-5y^2/18} \right], \\ f_2(y) &= \sqrt{\frac{5}{2\pi}} \left[\frac{5}{6} e^{-5y^2/2} + \frac{1}{30} e^{-y^2/10} \right]. \end{aligned} \quad (15)$$

Consider a weighted sum of the above 2 PDFs,

$$\begin{aligned} f_{mix}(y) &= \frac{1}{2} f_1(y) + \frac{1}{2} f_2(y), \\ &= \sqrt{\frac{5}{2\pi}} \left[\frac{2}{3} e^{-5y^2/2} + \frac{1}{12} e^{-5y^2/18} \right. \\ &\quad \left. + \frac{1}{60} e^{-y^2/10} \right]. \end{aligned} \quad (16)$$

It is evident that it cannot be represented as a sum of 2 Gaussian PDFs by redefining its parameters. Therefore, the space of all mixture-of-Gaussian PDFs cannot be convex (*QED*).

8. REFERENCES

[1] R.P. Brent, "Algorithms for minimization without derivatives," Dover (2002).

[2] W.S. Burdic, "Underwater acoustic system analysis," Second edition, Prentice-Hall, New Jersey (1991).

[3] S.M. Kay, "Fundamentals of statistical signal processing, Vol. II: Detection Theory," Prentice-Hall, New Jersey (1998).

[4] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, "Numerical recipes in C: The art of scientific computing," (Cambridge University Press, 1995).

[5] N.C. Beaulieu, "An infinite series for the computation of the complementary probability distribution function of the sum of independent random variables and its application to the sum of Rayleigh random variables," IEEE Trans. Comm. **38**, pp. 1463-1474, Sept. 1990.

[6] C. Tellambura, "Cochannel interference computation for arbitrary Nakagami fading," IEEE Trans. Veh. Technol., **48**, pp. 487-489, May 1999.

[7] A. A. Saha and G.V. Anand, "Design and analysis of detectors using stochastic resonance," Signal processing, **83**, 1193-1212 (2003).

[8] F.W. Machell, C.S. Penrod, and G.E. Ellis, in "Topics in non-Gaussian signal processing," edited by E.J. Wegman, S.C. Schwartz and J.B. Thomas, Springer-Verlag, New York (1989), Chapter 3, 29-57.