PULSE-TRAIN BASED TIME-DELAY ESTIMATION IMPROVES RESILIENCY TO NOISE

[1]Nicola Neretti, [1]Nathan Intrator and [1,2]Leon N Cooper
[1]Institute for Brain and Neural Systems, Brown University, Providence RI 02912.
[2]Physics Department, Brown University, Providence RI 02912.
{Nicola_Neretti, Nathan_Intrator, Leon_Cooper}@brown.edu

Abstract. Time-delay estimation accuracy in echolocating systems decays for increasing levels of noise until a breakpoint is reached, after which accuracy deteriorates by several orders of magnitude. In this paper we present a robust fusion of time-delay estimates from multiple pings that significantly reduces the signal-to-noise ratio corresponding to the accuracy breakpoint. We further show that a simple average of the time-delay estimates does not shift the breakpoint to a lower signal-to-noise ratio. The proposed fusion of multiple pings has the potential of improving the resilience to noise of echolocating systems such as sonar, medical ultrasound and geo-seismic surveys, hence increasing their potential operating range and reducing health/environmental hazards.

INTRODUCTION

The theory of optimal receivers shows that the *matched filter* receiver maximizes the output peak-signal-to-mean-noise (power) ratio [1], and is the optimum method for the detection of signals in noise. Information about the distance of the target is extracted by computing the time at which the cross- correlation between the echo and a replica of the pulse is a maximum. This delay is converted into a distance by means of the sound velocity in the particular medium in consideration (e.g. water or air). This type of receiver is generally referred to as a *coherent* receiver. The classical theory of optimal receivers describes the accuracy of time-delay estimation by match filtering via the well-known Woodward equation, which can be derived by using a variety of methods [1]. For small SNR's, one of the parameters in the classical equation – i.e. the bandwidth – has to be modified, and the receiver is then called *semicoherent*. In [2] it was shown that the transition between the two types of behaviors occurs at different SNR's depending on characteristics of the pulses such as bandwidth and center frequency. With this observation, a novel system based on an adaptive choice of the pulse was proposed [2] that can improve accuracy in the case of relatively low SNR, when ambiguity in the choice of the correct peak of the cross-correlation function cannot be avoided.

0-7803-8608-6/04/$20.00 ©2004 IEEE

Due to the nonlinear nature of the time-delay estimation problem, when the SNR drops below certain critical values *threshold effects* take place. Threshold effects can be characterized by a sharp deterioration of the time-delay estimator variance. Several statistical bounds have been used in the past to describe the accuracy of a matched filter receiver performance in between the above-mentioned SNR critical values. These include the Cramer-Rao lower bound [3], the Barankin bound [4], and the Ziv-Zakai bound [5-8]. Such bounds have been applied both to the problem of time-delay estimation [9-16] and of frequency estimation [17-19]. In particular, the Barankin bound has been used to define the SNR breakpoints corresponding to the change in behavior of the optimal receiver as the SNR decreases for the case of a single ping and single echo [10, 12, 13], a single ping and multiple echoes [15, 16], and multiple pings and a single echo [11, 14].
In this paper we analyze the SNR breakpoint, studying the probability of choosing the correct peak from the noisy cross-correlation function with a method similar to the one in [18], where the threshold effect was related to the existence of highly probable outliers far from the true time-delay value. This will enable us to extend the result to the case of multiple pings without *a priori* knowledge on the time-delay itself. This approach is different from that used in previous work on the multiple pings and single echo case [11, 14], where the multiple echoes for a single object are obtained artificially via multiple receivers and a unique ping. In fact, the bounds found in [11, 14] are valid only if the noise at the different receivers is totally uncorrelated or if the distance between transducer and receiver is constant, both conditions difficult to realize in practice.

SINGLE PING BREAKPOINT

Model for the autocorrelation function

In order to compute the probability of outliers, we introduce a model for the noiseless cross-correlation function envelope in the context of a semicoherent receiver. In this model the autocorrelation function is approximated by a piecewise constant function, with amplitude equal to A within the central interval I_Δ of length Δ , and zero elsewhere. When white Gaussian noise is added to the echo the cross correlation envelope vector has a multidimensional Gaussian distribution with centers at zero for all values that are outside of the central interval, and equal to A inside that interval. We also need to consider the width of the *a priori* window of the cross-correlation. The width of the window corresponds to the echolocating range. While the potential error in delay estimation is reduced when the width is reduced, so is the range. If the *a priori* window has a length of $2L$ and the sampling frequency is f_s, then there will be $N=N_A+N_0=2L{\cdot}f_s$ points, $N_A=\Delta{\cdot}f_s$ of which will be within the central bin ("correct bin"), and N_0 outside the central bin but within the *a priori* window. Hence, the probability of selecting a given time location in the cross-correlation function is given by α within the central interval I_Δ and $(1-\alpha)/N_0 =\beta/N_0$ elsewhere (Figure 1, top).

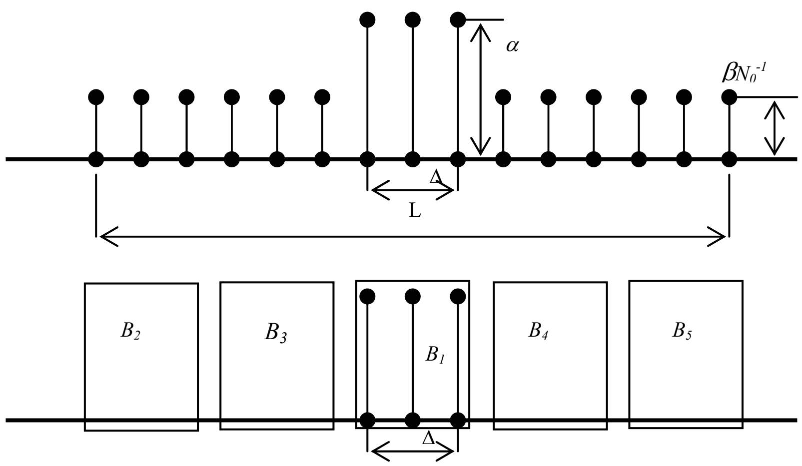

Figure 1 - The noiseless cross-correlation function is approximated by a piecewise constant function such that the probability of selecting a given time location in the cross-correlation function is given by α within the central interval I_Δ of length Δ and $(1-\alpha)N_0^{-1}$ elsewhere (top figure). We then divide the a priori window into intervals of length Δ equal to the size of the correct bin, to obtain m intervals $B_1, B_2, \ldots, B_m$, with $B_1 = I_\Delta$ representing the correct bin (bottom figure).

Without loss of generality, we consider the case where N_A and N_0 are integers. We define a random vector such that the first N_A random variables correspond to the amplitudes of the points within the correct bin, while the last N_0 correspond to the amplitudes of the points outside. For a white Gaussian noise, the joint probability density function for the vector of n random variables is then given by:

$$p\left(x_1, x_2, \ldots, x_N\right) = \frac{1}{\left(\sigma\sqrt{2\pi}\right)^N} e^{-\frac{\left\|\vec{x}-\vec{A}\right\|^2}{2\sigma^2}}, \quad \vec{v} = \Big[\underbrace{A, A, \ldots, A}_{N_A}, \underbrace{0, 0, \ldots, 0}_{N_0}\Big]^T \tag{1}$$

The desired probability of time-delay estimation within the correct bin of the center of the cross-correlation function is given by

$$\begin{aligned} \alpha &= P\left(\underset{1\le j\le N}{\arg\max}\left\{X_j\right\} \in I_\Delta\right) = \sum_{i=1}^{N_a} P\left(X_i > X_k, k \ne i\right) = \\ &= \sum_{i=1}^{N_a} \int_{-\infty}^{+\infty} dx_i \int_{-\infty}^{x_i} dx_{k_1} \int_{-\infty}^{x_i} dx_{k_2} \cdots \int_{-\infty}^{x_i} dx_{k_{N-1}} \frac{1}{\left(\sigma\sqrt{2\pi}\right)^N} e^{-\frac{\left\|\vec{x}-\vec{v}\right\|^2}{2\sigma^2}} \qquad (2) \\ &= \frac{N_A}{2^{N-1}\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\left(x - A/\sqrt{2}\sigma\right)^2} \left[1 + erf\left(x - A/\sqrt{2}\sigma\right)\right]^{N_a - 1} \left[1 + erf(x)\right]^{N_0} dx \end{aligned}$$

Accuracy breakpoint

Let T be the random variable (RV) whose probability distribution is given by equation (2). Let σ_Δ be the standard deviation (STD) of the distribution of the echo location in the central (correct) bin, and let σ_0 be the STD of the distribution which is outside the central bin. Now, suppose we sample from the original distribution whose cumulative function is given by equation (2). For n observations, a fraction of α of them falls in the correct bin on average, while βn fall outside. The standard deviation of the distribution will then be given by:

$$\mathrm{std}^2(T) = \alpha\ \mathrm{std}^2(T^\Delta) + \beta\ \mathrm{std}^2(T^0) = \alpha\sigma_\Delta^2 + \beta\sigma_0^2 \qquad (3)$$

We define the breakpoint (BP) as the level of noise for which the contribution of T^0 to the total error becomes dominant. Thus, the root-mean-square error (RMSE) will be significantly larger than the one given by the uniform distribution on I_Δ alone when $\alpha<\sigma_0^2/(\sigma_\Delta^2+\sigma_0^2)$. We then define the probability breakpoint to be:

$$\alpha_0 = \left(\sigma_0/\sigma_\Delta\right)^2 \Big/ \left[1+\left(\sigma_0/\sigma_\Delta\right)^2\right] \qquad (4)$$

It is possible to find the SNR breakpoint as the SNR value for which equation (2) equals the value in (4).

WHY DOES THE MEAN FAIL

Since the measurements from different pings are independent and identically distributed, the central limit theorem (CLT) implies that the standard deviation (error) of the averaged RV should be $n^{1/2}$ times smaller than the error made by each of the n measurements separately. This is indeed the case before the breakpoint. However, this process does not improve the situation after the breakpoint and, in particular, does not shift the breakpoint to lower SNR's. Thus, while averaging improves accuracy, it does not increase noise tolerance. Below we provide a mathematical analysis which explains why the breakpoint does not change.

The measurement process described above is equivalent to sampling form a uniform distribution F_Δ on the interval I_Δ with probability α and from a uniform distribution F_0 on the interval I_0, with a central gap corresponding to I_Δ, with probability β. Suppose we sample n times to obtain $T_1, T_2, \ldots, T_n$ and use the sample mean as our estimate for the delay. On average, αn values will be in the correct bin, $T_1^\Delta, T_2^\Delta, \ldots, T_{\alpha n}^\Delta$, while βn will be sampled from the uniform distribution, $T_1^0, T_2^0, \ldots, T_{\beta n}^0$. Then the sample mean can be decomposed into two parts:

$$\bar{T} = n^{-1}\sum_{i=1}^{n} T_i = n^{-1}\left(\sum_{i=1}^{\alpha n} T_i^\Delta + \sum_{i=1}^{\beta n} T_i^0\right) \qquad (5)$$

If σ_Δ is the standard deviation of the distribution F_Δ, and σ_0 is the standard deviation of the distribution F_0, then, applying the CLT to the two sums in equation (5) we obtain:

$$\left(\text{std} \sum_{i=1}^{\alpha n} T_i^\Delta\right)^2 = \alpha n \sigma_\Delta^2 \ , \qquad \left(\text{std} \sum_{i=1}^{\beta n} T_i^0\right)^2 = \beta n \sigma_0^2 \tag{6}$$

The root-mean-square error will be significantly larger than the one given by the distribution F_Δ alone when $\beta n \sigma_0^2 > \alpha n \sigma_\Delta^2$, i.e. when $\alpha < \sigma_0^2/(\sigma_\Delta^2 + \sigma_0^2)$. It is observed that this bound does not improve with the number of pings and is equal to the bound found for a single ping. This explains why averaging the time-delay estimates from multiple pings have not been found useful for very low levels of SNR. It should be noted that it is possible to shift the breakpoint by estimating the time delay form the averaged cross-correlation functions of all the observations, a process that would require an extremely good alignment of the cross-correlation functions. This is basically equivalent to reducing the noise level by averaging the echoes [17], which is not realistic in a situation where the sonar is not completely still with respect to the target, or where it is not practical to store the entire echo waveforms for off line processing. However, a successful implementation of this averaging can be achieved by using multiple receivers [11, 14].

USING THE MODE

Suppose we divide the *a priori* window into intervals of length Δ equal to the size of the correct bin, to obtain $m = \lfloor 2L/\Delta \rfloor$ intervals $B_1, B_2, \ldots, B_m$, with $B_1 = I_\Delta$ representing the correct bin (Fig. 1, bottom). If $p_1, p_2, \ldots, p_m$ are the probabilities for an estimate to fall in each of the intervals and $Y_1, Y_2, \ldots, Y_m$ are random variables representing the number of estimates falling in each interval, then $Y_1 + Y_2 + \ldots + Y_m = n$, and $p_1 + p_2 + \ldots + p_m = 1$. The joint probability distribution for the number of estimates in each bin is given by the multinomial distribution

$$P\left(Y_1 = k_1, Y_2 = k_2, \ldots, Y_m = k_m\right) = \frac{n!}{k_1!k_2!\cdots k_m!} p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m} \tag{7}$$

The probability of choosing the correct bin using the mode is the probability that the number of estimates falling in the correct bin k_1, is greater than the number of estimates falling in any other bin k_i, $i \neq 1$:

$$P_{\text{correct}} = P\left(Y_1 > Y_j, \quad \forall j \neq 1\right) = \sum_{\substack{k_1, k_2, \ldots, k_m \\ k_1 > k_j, j \neq 1, \sum_{j=1}^{m} k_j = n}} \frac{n!}{k_1!k_2!\cdots k_m!} \prod_{l=1}^{m} p_l^{k_l} \tag{8}$$

The sum in equation (8) can be decomposed into two parts: 1) the probability $P_{>50\%}$ that more that half of the n points fall into the correct bin; 2) the probability $P_{<50\%}$ that even if less than half of the n points fall in the correct bin, the number of points in it is greater than that of any other bin, such that *P(correct bin)*$=P_{>50\%}+P_{<50\%}$. The probability $P_{>50\%}$ can be written as:

$$P_{>50\%} = \sum_{\substack{k_1,k_2,\ldots,k_m \\ k_1>n/2,\sum_{j=1}^{m} k_j=n}} \frac{n!}{k_1!k_2!\cdots k_m!} \prod_{l=1}^{m} p_l^{k_l} = \sum_{k_1>n/2} \binom{n}{k_1} p_1^{k_1} \left(\sum_{j=2}^{m} p_j \right)^{n-k_1} \tag{9}$$

The probability p_1 of an estimate to fall outside the correct bin is uniform over the a priori window with probability $\beta=(1-\alpha)$, so that the probability for it to fall in any interval of size Δ is $p_j=(1-\alpha)/(m-1),\ j\neq 1$. Substituting these values into equation (9) we obtain

$$P_{>50\%} = \sum_{k>n/2} \binom{n}{k} \alpha^k (1-\alpha)^{n-k}\ , \tag{10}$$

where α is a function of the SNR. The computation of $P_{<50\%}$ is more complicated. However, it is possible to derive an upper bound ($SNR_{>50\%}$) on the SNR breakpoint for the time-delay accuracy computed by using the mode of n estimates, as the SNR for which $P_{>50\%}=\alpha_0$, where α_0 is given in equation (4):

$$\sum_{k>n/2} \binom{n}{k} \alpha\left(SNR_{>50\%}\right)^k \left[1-\alpha\left(SNR_{>50\%}\right)\right]^{n-k} = \alpha_0 \tag{11}$$

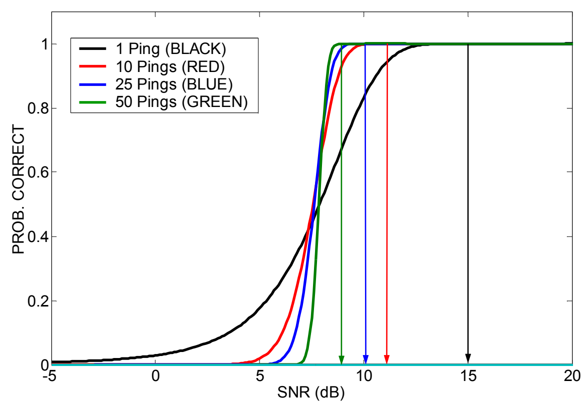

Figure 2 – Probability of making the correct choice as a function of SNR for different numbers of pings. The arrows indicate upper bounds on the SNR breakpoints.

The tighter bound corresponding to the total probability of choosing the correct bin will always be lower than the one derived from equation (11), i.e. $SNR_{BP}\leq SNR_{>50\%}$. The breakpoint which the mode can achieve will be for a lower SNR than the one calculated above, which is already significantly better (Figure 2) than the breakpoint achieved by either a single ping or by averaging of the echo delay estimates of multiple pings.

SIMULATION RESULTS

To test the mathematical results presented in the previous sections, we developed a set of Monte Carlo simulations using a cosine packet. We first analyzed the histograms of the errors in the delay estimate of the ideal receiver for different SNR's (figure 3). For high SNR (≥20dB) all the errors are small and follow the Woodward equation that corresponds to values within the central bin in figure 3a. As the level of the noise increases, large errors in the estimates appear. The errors are uniformly distributed over the entire a-priori window, and the relative ratio between the correct estimates (central bin) and the level of the uniform distribution decreases with SNR (figures 3b, 3c, and 3d). However, even for high levels of noise the central peak is significantly larger that the rest of the distribution.

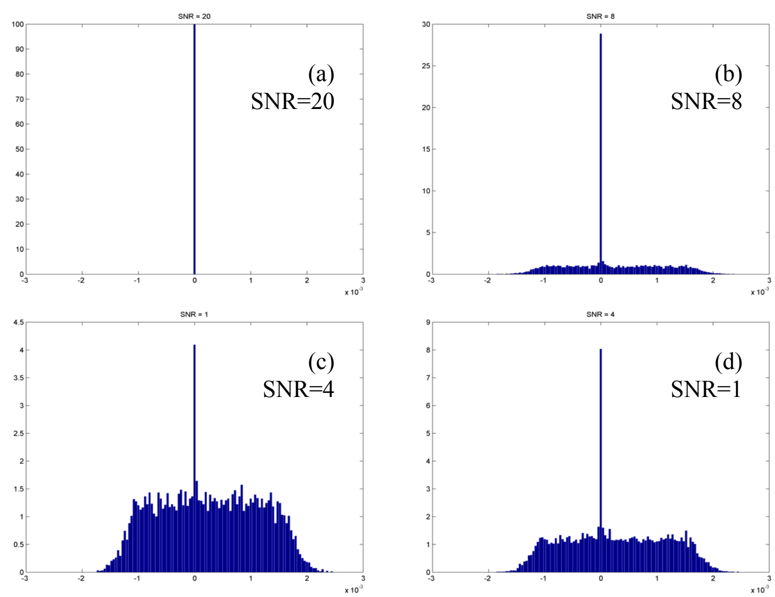

Figure 3 – Histograms of the errors in the delay estimate in a Monte Carlo simulation for different SNR's. For high SNR (≥20dB) all the errors are small and follow the Woodward equation that corresponds to values within the central bin in figure (a). As the level of the noise increases, large errors in the estimates appear. The errors are uniformly distributed over the entire a-priori window, and the relative ratio between the correct estimates (central bin) and the level of the uniform distribution decreases with SNR, see figures (b), (c), and (d). However, even for high levels of noise the central peak is significantly larger that the rest of the distribution.

Figure 4a shows the performance of ideal receiver for a single ping. For high SNR the accuracy follows the Woodward equation corresponding to a coherent ideal as expected from the theory of optimal receivers. The performance breaks for low

SNR around 17 dB. Figures 4b, 4c and 4d show the analysis of the accuracy breakpoint for different number of pings, 10, 50 and 100 respectively.

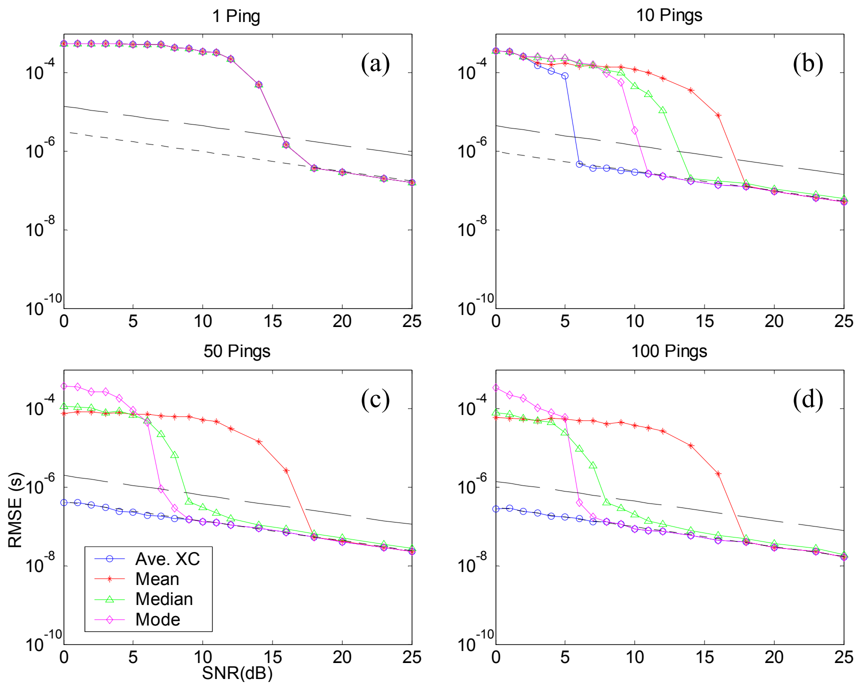

Figure 4 – RMSE as a function of SNR and number of pings – 1, 10, 50 and 100 respectively – for the Cosine Packet. Notice how the SNR breakpoint for the average of multiple pings (red line) does not decrease with the number of pings.

The blue line describes the optimal accuracy that can be achieved using cross-correlation from multiple pings. Its breaking point represents the optimal breaking point that could have been achieved using stationary sonar and target, and that could be predicted by using the Barankin bound as in [11, 14]. This breaking point however is not attainable, as it relies on careful registration of returns from different pings. Such careful registration can only be done if the distance between object to target is kept constant, or if it is known for each ping in advance. It can be seen that robust fusion of multiple pings based on the mode (light blue, and magenta lines) improves noise resiliency while retaining close to optimal achievable accuracy under multiple pings. In general there is no significant improvement in the resiliency to noise when a simple mean of the observations is used due the strong contamination of the distribution from outliers (red lines). This confirms the mathematical result presented in the preceding sections.

Figure 5, shows a summary of the results for the different methods. The breakpoint for the averaged cross-correlation function (blue squares) follows the ideal curve obtained by reducing the level of the noise (solid blue line) as described above. The breakpoint of the estimate obtained from the mean does not substantially change as the number of pings in increased. A more robust statistics such as the median improves the resiliency to noise as the number of pings increases (green triangles). The best results are obtained by using the mode of the estimates from the multiple pings (magenta diamonds).

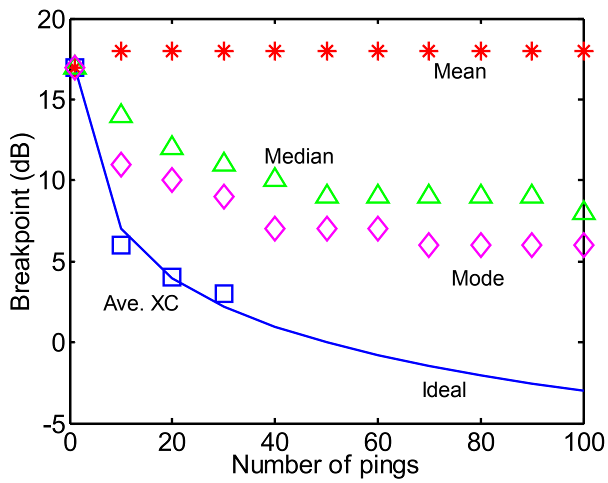

Figure 5 – Breakpoint in dB as a function of number of pings for different methods. The solid line corresponds to the ideal case of noise reduction by averaging the echoes.

CONCLUSIONS

In summary, we have demonstrated that multiple pings are useful for improving the accuracy and in particular the resilience of time-delay estimation to background noise. In particular, we have demonstrated, that a robust statistics such as the mode of the distribution of echo delays which is obtained from multiple pings, significantly decreases the signal-to-noise ratio breakpoint. We have further shown that the mean of this distribution has the same breakpoint as a single ping, thus not contributing at all to the resilience to noise.

Acknowledgements

This work was supported in part by ARO (DAAD 19-02-1-0403), and ONR (N00012-02-C-02960).

REFERENCES

[1] M. I. Skolnik, *Introduction to Radar Systems*, 1st ed: McGraw-Hill, 1962.

[2] N. Neretti, N. Intrator, and L. N. Cooper, "Adaptive pulse optimization for improved sonar range accuracy," *IEEE Signal Processing Letters*, vol. 11(4), April 2004.

[3] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, vol. 37, pp. 81-91, 1945.

[4] E. W. Barankin, "Locally best unbiased estimates," *Ann. Math. Stat.*, vol. 20, pp. 477-501, 1946.

[5] S. Bellini and G. Tartara, "Bounds on error in signal parameter estimation," *IEEE Trans. Commun.*, vol. COM-22, pp. 340-342, 1974.

[6] D. Chazan, M. Zakai, and J. Ziv, "Improved lower bound on signal parameter estimation," *IEEE Trans. Information Theory*, vol. IT-21, pp. 90-93, 1975.

[7] L. P. Seidman, "Performance limitations and error calculations for parameter estimation," *Proc. IEEE*, vol. 58, pp. 644-652, 1970.

[8] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Trans. Information Theory*, vol. IT-15, pp. 386-391, 1969.

[9] D. Slepian, "Estimation of Signal Parameters in the Presence of Noise," *IRE Trans.*, vol. PGIT-3, pp. 68-89, 1954.

[10] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Trans. Information Theory*, vol. 43, pp. 1084-1093, 1997.

[11] S.-K. Chow and P. M. Schultheiss, "Delay estimation using narrow-band processes," *IEEE Trans. ASSP*, vol. ASSP-29, pp. 478-484, 1981.

[12] R. J. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Trans. Information Theory*, vol. IT-17, pp. 669-676, 1971.

[13] R. J. McAulay and L. P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis," *IEEE Trans. Information Theory*, vol. IT-15, pp. 273-279, 1969.

[14] J. Tabrikian and J. L. Krolik, "Barankin bounds for source localization in an uncertain ocean environment," *IEEE Trans. Signal Processing*, vol. 47, pp. 2917 -2927, 1999.

[15] A. Zeira and P. M. Schultheiss, "Realizable lower bounds for time delay estimation: Part 2 - Threshold phenomena," *IEEE Trans. Signal Processing*, vol. 42, pp. 1001-1007, 1994.

[16] A. Zeira and P. M. Schultheiss, "Realizable lower bounds for time delay estimation," *IEEE Trans. Signal Processing*, vol. 41, pp. 3102-3113, 1993.

[17] L. Knockaert, "The Barankin bound and threshold behavior in frequency estimation," *IEEE Trans. Signal Processing*, vol. 47, pp. 2398-2401, 1997.

[18] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Information Theory*, vol. IT-20, pp. 591-598, 1974.

[19] B. James, B. D. O. Anderson, and R. C. Williamson, "Characterization of threshold for single tone maximum likelihood frequency estimation," *IEEE Trans. Signal Processing*, vol. 43, pp. 817-821, 1995.