Coherence and Time Delay Estimation

G. CLIFFORD CARTER, SENIOR MEMBER, IEEE

This paper presents a tutorial review of work in coherence and time delay estimation. A review of coherence research and development is presented. A derivation of the ML estimator for time delay is presented together with an interpretation of that estimator as a special member of a class of generalized cross correlators. The performance of the estimator is given for both high and low signal-to-noise ratio cases. The proposed correlator is implemented and stimulated with synthetic data. The results are compared with performance predictions and found to be in good agreement.

INTRODUCTION

This paper is a summary of work done by the author and several coauthors over more than a decade. The paper is organized into five sections. Section I is a review of coherence. Section II is a review of the generalized framework for coherence estimation. Section III is a summary of statistics of the MSC estimator and has five subsections discussing the probability density function, experimental results, bias, receiver operating characteristics, and confidence bounds. Section IV discusses time-delay estimation. Finally, Section V discusses the focused time-delay beamformer form of passive ranging.

I. COHERENCE

This section discusses application of the coherence function. Much of the material in this section is a summary of the work by Carter and Knapp [13]. The coherence function between two wide-sense stationary random processes x and y is equal to the cross power spectrum $G_{xy}(f)$ divided by the square root of the product of the two auto power spectra. Specifically, the complex coherence is defined by

$$\gamma_{xy}(f) = \frac{G_{xy}(f)}{\sqrt{G_{xx}(f)\, G_{yy}(f)}} \tag{1a}$$

where f denotes the frequency of interest and where the complex cross power spectrum

$$G_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{j2\pi f\tau}\, d\tau \tag{1b}$$

is the Fourier transform of the cross correlation function

$$R_{xy}(\tau) = E[x(t)\, y(t+\tau)]. \tag{1c}$$

Manuscript received October 17, 1985; revised August 12, 1986. Portions of this paper were presented at the Advanced Study Institute on Signal Processing, Les Hoches, France, August 12-September 6, 1986.

The author is with the Naval Underwater Systems Center, New London, CT 06320, USA.

Here x and y are real and E denotes the mathematical expectation (for ergodic random processes the ensemble average can be replaced by a time average). The coherence is a normalized cross-spectral density function; in particular, the normalization constrains (1) so that magnitude-squared coherence (MSC) defined by

$$C_{xy}(f) \triangleq |\gamma_{xy}(f)|^2 \tag{2a}$$

lies in the range

$$0 \le C_{xy}(f) \le 1 \tag{2b}$$

for all frequencies. Throughout the text we use C and $|\gamma|^2$ interchangeably.

The coherence function has uses in numerous areas, including system identification, measurement of signal-to-noise ratio (SNR), and determination of time delay. The coherence—in particular, magnitude-squared coherence (MSC)—can only be put to use when its value can be accurately estimated. Indeed, it is highly desirable to understand the statistics of the estimator. Therefore, this section addresses interpretations of the coherence function. Following sections address procedures for properly estimating the MSC and statistics of the estimator.

One interesting interpretation of coherence—particularly MSC—is that it is a measure of the relative linearity of two processes. To illustrate this, consider Fig. 1 in which a

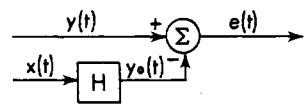


Fig. 1. Model of error resulting from linearly filtering $x(t)$ to match any desired signal $y(t)$.

sample function $y(t)$ of an arbitrary stationary random process consists of the response $y_0(t)$ of a linear filter plus an error component $e(t)$. When the linear filter is chosen to minimize the mean-square value of $e(t)$, i.e., the area under the error spectrum, then $y_0(t)$ becomes that part of $y(t)$ linearly related to $x(t)$. The spectral characteristics of $e(t)$ are given by

$$G_{ee}(f) = G_{yy}(f) + G_{xx}(f)\, |H(f)|^2 - H(f)\, G_{xy}^*(f) - H^*(f)\, G_{xy}(f) \tag{3}$$

where * indicates complex conjugation and $H(f)$ is the fil-

ter transfer function. The error spectrum

$$G_{ee}(f) = G_{xx}(f)\left|H(f) - \frac{G_{xy}(f)}{G_{xx}(f)}\right|^2 + G_{yy}(f)\,[1 - C_{xy}(f)]. \quad (4)$$

Hence, the optimum filter is given by

$$H_o(f) = \frac{G_{xy}(f)}{G_{xx}(f)}. \quad (5)$$

Note that the coherence is related to the optimum linear filter according to

$$\gamma_{xy}(f) = H_o(f)\sqrt{\frac{G_{xx}(f)}{G_{yy}(f)}} \quad (6a)$$

and

$$C_{xy}(f) = |H_o(f)|^2 \frac{G_{xx}(f)}{G_{yy}(f)}. \quad (6b)$$

These results apply regardless of the source of $y(t)$. When the linear filter is optimum in the mean-square sense, the error is uncorrelated with $x(t)$, i.e.,

$$G_{xy_0}(f) = H_o(f)\,G_{xx}(f) = G_{xy}(f). \quad (7)$$

Furthermore, the minimum value of $G_{ee}(f)$ is given by

$$G_{ee}(f) = G_{yy}(f)\,[1 - C_{xy}(f)] \quad (8)$$

and

$$G_{y_0y_0}(f) = |H_o(f)|^2\,G_{xx}(f) = G_{yy}(f)\,C_{xy}(f). \quad (9)$$

From the identity

$$G_{yy}(f) = C_{xy}(f)\,G_{yy}(f) + [1 - C_{xy}(f)]G_{yy}(f) \quad (10a)$$

we can show that

$$G_{yy}(f) = G_{y_0y_0}(f) + G_{ee}(f) \quad (10b)$$

indicating that the MSC is the fraction of $G_{yy}(f)$ contained in the linear component of $y(t)$, and $1 - C_{xy}(f)$ is the proportion of $G_{yy}(f)$ contained in the error, or nonlinear component of $y(t)$. These results can be applied to the configurations shown in Figs. 2 and 3.

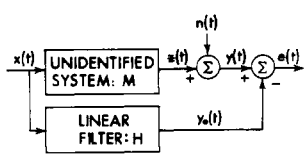


Fig. 2. Model of error resulting from linear approximation of unidentified system.

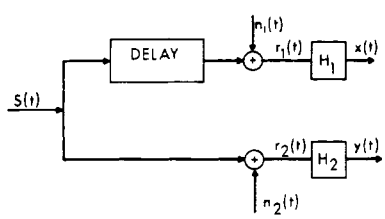


Fig. 3. Model of directional signal corrupted with additive noise and processed.

II. The Generalized Framework for Coherence Estimation

The purpose of this section is to review a generalized framework for power spectral estimation and to show how three estimation methods fit into this framework. (Beyond the scope of this paper are important methods of time-varying spectral estimation and parametric methods of spectral estimation. See for example IEEE PRESS books by Childers [21] and Kesler [44].)

In the generalized framework of Nuttall and Carter [59] we are concerned with both auto and cross (nonparametric) spectral estimation of stationary random processes; hence, we consider two discrete random processes. Auto spectral estimation is then a special case of cross spectral estimation. In particular, in auto spectral estimation we replace the second time series by a duplicate of the first time series. As is often the case in practice, we are limited to a single time-limited realization (TLR) of each random process. Within our generalized framework for power spectral estimation, we *first* partition each TLR into N segments, where the segments may be overlapped. *Second*, each segment is multiplied by a time-weighting function (the weighting function may be unity everywhere within the segment as, for example, rectangular weighting or it may be smooth as, for example, Hanning weighting). *Third* , the discrete Fourier coefficients (DFC) are computed for each weighted segment via an appropriate algorithm such as the FFT after each segment has been appropriately appended with zeros. *Fourth*, the DFCs for one TLR segment are multiplied by the complex conjugate of the DFCs for the other time synchronous segment (or same segment for auto spectra). As a matter of clarification both time series are presumed to be aligned, that is, in proper time register, so that the ith segment of one segment only interacts with the ith segment of the other time series. *Fifth*, the complex products are averaged over the N available segments (one segment if $N = 1$). *Sixth*, the resultant spectral estimates are Fourier transformed into the correlation or lag domain, where they are multiplied by a lag-weighting function (which may be unity). *Finally*, the results are transformed back into the frequency domain. (Alternatively, the last two steps can be replaced by a convolution in the frequency domain. Depending on the extent of the frequency-domain convolution, the former alternative may be computationally preferable over the latter.) Mathematical details are in the paper by Nuttall and Carter [60].

We now point out how three spectral analysis techniques fit into this generalized framework. All three achieve virtually the same statistical, in particular, mean and variance properties. First, the Blackman and Tukey [10] (BT) method allows for only one segment with rectangular time weighting over the entire record (from each TLR), and it applies a smooth lag-weighting function in the correlation domain, which goes to zero well before the end of the data record. (We note that, historically, the BT approach estimated the correlation function direction rather than first transforming into the frequency domain, but since it is faster than the original time-domain BT method, it is a viable and equivalent alternative approach.) By adjusting how quickly the lag weighting goes to zero, resolution and stability can be compromised. For example, if the weighting goes to zero quickly the spectral estimates will have coarse resolution and good stability compared with a lag weighting that does not go to zero quickly. The exact shape of the weighting will

influence the exact shape of the sidelobes and main lobe in the frequency domain.

The second method that falls within the generalized framework is the weighted overlapped-segment averaging or WOSA technique. The WOSA method is sometimes referred to as Welch's method based on the contribution of Welch [74]. We propose here that WOSA refer to Welch's overlapped-segment averaging method. In the WOSA method we apply a smooth multiplicative time weighting to each of a large number of segments, and average the DFC products from these overlapped segments to obtain a final spectral estimate, without employing additional lag weighting. See, for example, Fig. 4 for a graphical portrayal of a family of overlapped time-weighting functions. The time

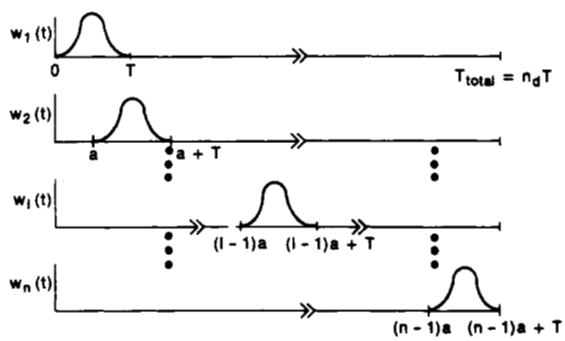


Fig. 4. Overlapped weighting functions.

weighting is typically a smooth weighting such as Hanning, in the WOSA method, because it yields good sidelobe behavior. Overlap is important in the WOSA method in order to realize maximum stability (that is, minimum variance) of the spectral estimate. The WOSA method is a statistically sound method widely in use today.

A third technique that falls within the generalized framework is the lag-reshaping method (see, e.g., Nuttall and Carter [60]). The lag-reshaping method recognizes that the number of available data points may be so large as to preclude the normal BT method in practical situations. We segment the data without overlapping and apply unit gain rectangular time weighting to each segment (this rectangular weighting requires no time weighting multiplications). Later we will undo the bad sidelobe effects that this rectangular time weighting initially causes, and gain additional stability. Note the segment-averaged power spectrum will be transformed into the correlation (or lag) domain, where a smooth multiplicative lag-weighting function will be applied before transforming back into the frequency domain. The smooth lag weighting will be the product of two lag weightings one for the desired window and one for lag reshaping. This lag-domain "reshaping" is an ingenious method for almost completely undoing the bad sidelobe effects of rectangular time weighting.

All three of these techniques that fall within the generalized framework have good statistical properties. The Blackman and Tukey [10] method attains minimum variance spectral estimates and is the benchmark against which other techniques have been measured. The key assumptions for a statistical investigation of any spectral analysis technique are stationarity, Gaussian random process, and a large product of observation time and desired resolution bandwidth. Since we are interested in spectral estimates, second-order stationarity is required. Since one must investigate variances of second-order quantities for stability determination, fourth-order moments of the random process are required; hence, for mathematical tractability, the Gaussian assumption is needed. And in many practical cases, to have meaningful spectral estimates with any method, large observation-time resolution-bandwidth products are required. These assumptions are the practical essentials of any mathematical analysis of spectral estimation methods.

Under these assumptions, Nuttall [57] has shown that the WOSA method can achieve the same stability, in particular, the same number of equivalent degrees of freedom (EDF) as the BT spectral estimation method for both auto and cross spectral estimation if the proper overlap is used for each weighting, when both methods operate on the same amount of data and are constrained to the same frequency resolution. For many practical time weightings, most of the maximum EDF (minimum variance) can be attained by a computationally reasonable amount of overlap. For example, with Hanning weighting, 92 percent of the maximum EDF can be realized with 50-percent overlap. And for Parzen (cubic) weighting, 93 percent of the maximum EDF can be realized with 62.5-percent overlap. Furthermore, the number of FFTs required is virtually independent of the particular time weighting employed (with its optimum overlap), but depends only upon the observation-time resolution-bandwidth product.

For more complicated spectral measures, such as coherence, the analysis of WOSA becomes unwieldy, and one is driven to simulation. In particular, Carter, Knapp, and Nuttall [11] empirically investigated the effect of overlap on the variance of the coherence estimate via the WOSA method. There was a pronounced improvement (about a factor of two in variance reduction) with overlap as opposed to no overlap. In another experiment, Carter and Knapp [13] compared the use of Hanning and rectangular weightings for estimating coherence between a flat broad-band input to a second-order digital filter and its output. It was demonstrated that smooth weighting functions are required to obtain good coherence estimates with the WOSA method. (Recall that the WOSA method does not use additional lag shaping.)

The WOSA nonparametric spectral estimation method widely in use today with proper overlap can attain the EDF of the BT method, when both methods operate on the same amount of data and are constrained to have the same frequency resolution. Further, for good time weightings, reasonable amounts of overlap achieve most of the available EDF.

Based on analytic work by Nuttall and Carter [60], the lag-reshaping method can virtually (but not exactly) attain the EDF of the BT method and yield very good sidelobes through the use of unusual lag weighting. It appears highly certain that the lag-reshaping method requires fewer computations (perhaps by a factor of two) than the WOSA method in practice, and therefore deserves serious consideration as a replacement for (or variation on) the widely used WOSA method.

Table 1 Probability Density and Distribution Functions

Density Function	
$p(\lvert\hat{\gamma}\rvert^2 \mid n_d, \lvert\gamma\rvert^2) = (n_d - 1)(1 - \lvert\gamma\rvert^2)^{n_d}(1 - \lvert\hat{\gamma}\rvert^2)^{n_d-2}$	
$\cdot\ {}_2F_1(n_d, n_d; 1; \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2), \quad 0 \leqslant \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2 < 1$	(1a)
$= (n_d - 1)(1 - \lvert\gamma\rvert^2)^{n_d}(1 - \lvert\hat{\gamma}\rvert^2)^{n_d-2}$	
$\cdot\ (1 - \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2)^{1-2n_d}\,{}_2F_1(1 - n_d, 1 - n_d; 1; \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2)$	(1b)
$= (n_d - 1)\left[\dfrac{(1 - \lvert\gamma\rvert^2)(1 - \lvert\hat{\gamma}\rvert^2)}{(1 - \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2)^2}\right]^{n_d}$	
$\cdot\ \dfrac{(1 - \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2)}{(1 - \lvert\hat{\gamma}\rvert^2)^2}\,{}_2F_1(1 - n_d, 1 - n_d; 1; \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2).$	(1c)
Distribution Function	
$P(\lvert\hat{\gamma}\rvert^2 \mid n_d, \lvert\gamma\rvert^2) = \lvert\hat{\gamma}\rvert^2 \left(\dfrac{1 - \lvert\gamma\rvert^2}{1 - \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2}\right)^{n_d} \sum_{k=0}^{n_d-2} \left(\dfrac{1 - \lvert\hat{\gamma}\rvert^2}{1 - \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2}\right)^k$	
$\cdot\ {}_2F_1(-k, 1 - n_d; 1; \lvert\gamma\rvert^2 \lvert\hat{\gamma}\rvert^2).$	(1d)

III. Statistics of MSC Estimates Obtained Via the Welch's Overlapped Segment Averaging (WOSA) Method

A. Introduction

Much of the historical work on the statistics of the MSC estimates centers on the WOSA method; by proper interpretation of variables, these results also apply to the lag-reshaping method. Recall that the WOSA method consists of obtaining two finite-time series from the random processes being investigated. Each time series is partitioned into equal length segments and sampled at equally spaced data points. The segments are overlapped. However, the statistics are analytically developed for nonoverlapped segments. Empirical results are presented for overlapped segments. Samples from each segment are multiplied by a weighting function, and the FFT of the weighted sequence is performed. Then the Fourier coefficients for each weighted segment are used to estimate the auto- and cross-power spectral densities. The spectral density estimates thus obtained are used to form the MSC estimate.

Spectral resolution of the estimates varies inversely with the segment length T. Proper weighting or "windowing" of the T-second segment is also helpful in achieving good sidelobe reduction. On the other hand, for independent segments with ideal windowing, the bias and the variance of the MSC estimate vary inversely with the number of segments n. Therefore, to generate a good estimate with limited data, one may be faced with conflicting requirements on n and T. Segment overlapping can be used to increase both n and T. When the segments are disjoint, that is, non-overlapping, we call the number of segments n_d. As the percentage of overlap increases, however, the computational requirements increases rapidly, while the improvement stabilizes owing to the greater correlation between data segments (see Carter, Knapp, and Nuttall [11]).

B. Probability Density for the Estimate of the MSC

The first-order probability density and distribution functions for the estimate of MSC, given the true value of MSC, and the number of independent segments n_d are given in Table 1. Notationally, recall $|\gamma|^2 = C$. Equations (1b) and (1c) in that table are useful because the ${}_2F_1$ hypergeometric function is an $(n_d - 1)$st-order polynomial.

Figs. 5 and 6 illustrate the probability density and distribution functions for several cases, as computed using (1b) and (1d) from Table 1. It is evident from Fig. 6 that the variance of the MSC estimate decreases when n_d is increased.

The bias and variance of the MSC estimate can be evaluated using a general expression for the mth moment of the MSC estimate (see Carter, Knapp, and Nuttall [11]).

Bias and variance expressions obtained are summarized in Table 2. Approximations (2c) and (2d) are the result of truncating the series (2a) and (2b). Equations (2e) through (2g) then follow for large n_d; they indicate that the MSC estimate is asymptotically unbiased, and that for large n_d the following is true.

1) The bias is greatest, $1/n_d$, when the MSC equals zero and smallest, 0, when the MSC equals unity.

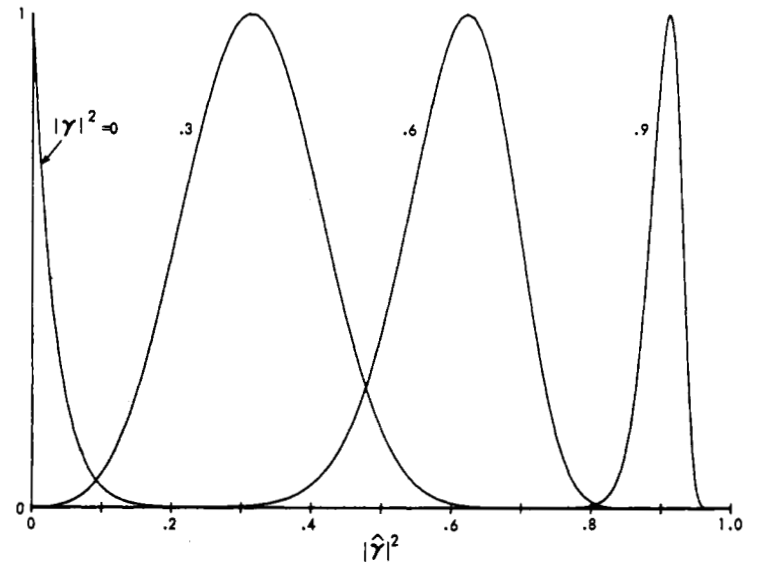
2) The variance is zero when the MSC equals unity and greatest, $(2/3)^3/n_d$, when the MSC equals one third.

3) The mean-square error from the true value is equal to the variance, provided the MSC is not zero.

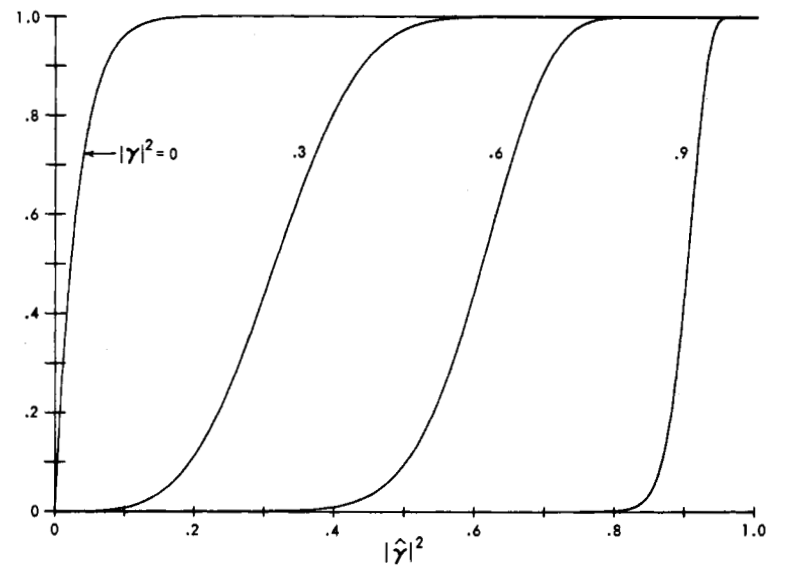
Figs. 7 and 8, respectively, show the bias and variance as functions of n_d and $|\gamma|^2$. For values of n_d in the range from 32 to 64, expressions (2e) through (2g) in Table 2 are good approximations; however, the curves in Figs. 7 and 8 were obtained using the exact formulas (2a) and (2b) of Table 2. Peaks in the variance curves when the MSC equals one third are evident. We note, however, there is an additional bias when our FFT size is too small. This second type of bias can be extremely important. It is the subject of texts including Koopmans [48] and Brillinger [6] and is discussed in Section III-D.

C. Experimental Investigation of Overlap Effects

An experimental study has been made of the effect of overlap of data on the MSC estimate. The analytical results presented earlier relate only to the case of independent seg-



(a)



(b)

Fig. 5. (a) Probability density functions. (Functions have been normalized by maximum values, which are 31.0, 4.13, 5.23, and 17.5.) (b) Distribution functions.

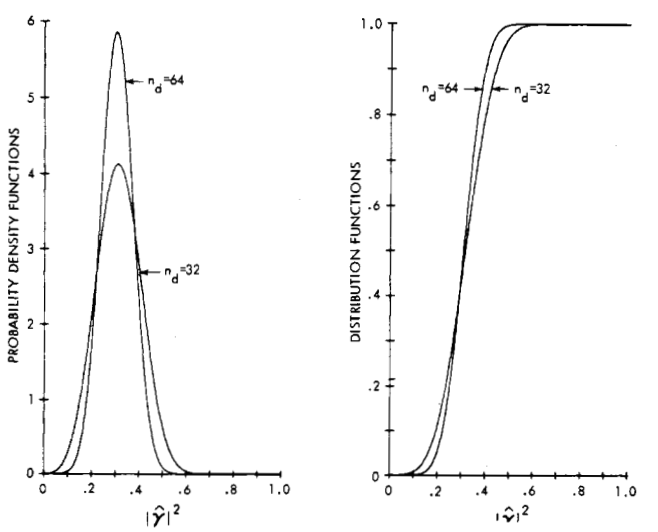


Fig. 6. Probability density and distribution functions of $|\hat{\gamma}|^2$ for $|\gamma|^2 = 0.3$.

Table 2 Bias and Variance Expressions

	Exact	
Bias	$B = E[\lvert\hat{\gamma}\rvert^2 \mid n_d, \lvert\hat{\gamma}\rvert^2] - \lvert\gamma\rvert^2 = \frac{(1 - \lvert\gamma\rvert^2)^{n_d}}{n_d} {}_3F_2(2, n_d, n_d; n_d + 1, 1; \lvert\gamma\rvert^2) - \lvert\gamma\rvert^2$	
	$= \frac{1}{n_d} + \frac{n_d - 1}{n_d + 1} \lvert\gamma\rvert^2 {}_2F_1(1, 1; n_d + 2; \lvert\gamma\rvert^2) - \lvert\gamma\rvert^2.$	(2a)
Variance	$V = E(\lvert\hat{\gamma}\rvert^4 \mid n_d, \lvert\gamma\rvert^2) - E^2(\lvert\hat{\gamma}\rvert^2 \mid n_d, \lvert\gamma\rvert^2)$	
	$= \frac{2(1 - \lvert\gamma\rvert^2)^{n_d}}{n_d(n_d + 1)} {}_3F_2(3, n_d, n_d; n_d + 2, 1; \lvert\gamma\rvert^2)$	
	$- \left[\frac{(1 - \lvert\gamma\rvert^2)^{n_d}}{n_d} {}_3F_2(2, n_d, n_d; n_d + 1, 1; \lvert\gamma\rvert^2)\right]^2.$	(2b)
	Approximate	
	$B_0 \cong \frac{1}{n_d} - \frac{2}{n_d + 1} \lvert\gamma\rvert^2 + \frac{1!(n_d - 1)}{(n_d + 1)(n_d + 2)} (\lvert\gamma\rvert^2)^2 + \frac{(n_d - 1)\,2!}{(n_d + 1)(n_d + 2)(n_d + 3)} (\lvert\gamma\rvert^2)^3; \quad B \cong \begin{cases} B_0, & B_0 \geqslant 0 \\ 0, & B_0 < 0 \end{cases}$	(2c)
	$V_0 \cong \frac{(n_d - 1)}{n_d(n_d + 1)} \left[\frac{1}{n_d} + 2\frac{n_d - 2}{n_d + 2} \lvert\gamma\rvert^2 - 2\frac{2n_d^3 - n_d^2 - 2n_d + 3}{(n_d + 1)(n_d + 2)(n_d + 3)} (\lvert\gamma\rvert^2)^2 \right.$	
	$+ 2\frac{n_d^4 - 6n_d^3 - n_d^2 + 10n_d - 8}{(n_d + 1)(n_d + 2)(n_d + 3)(n_d + 4)} (\lvert\gamma\rvert^2)^3$	
	$\left. + \frac{13n_d^5 - 15n_d^4 - 113n_d^3 + 27n_d^2 + 136n_d - 120}{(n_d + 1)(n_d + 2)^2(n_d + 3)(n_d + 4)(n_d + 5)} (\lvert\gamma\rvert^2)^4\right]; \quad V \cong \begin{cases} V_0, & V_0 \geqslant 0 \\ 0, & V_0 < 0 \end{cases}$	(2d)
	Approximation for Large n_d	
	$B \cong \frac{1}{n_d} [1 - \lvert\gamma\rvert^2]^2$	(2e)
	$\leqslant \frac{1}{n_d} [1 - \lvert\gamma\rvert^2]$	(2f)
	$V \cong \begin{cases} \frac{1}{n_d^2}, & \lvert\gamma\rvert^2 = 0 \\ \frac{2\lvert\gamma\rvert^2}{n_d} (1 - \lvert\gamma\rvert^2)^2, & 0 < \lvert\gamma\rvert^2 \leqslant 1. \end{cases}$	(2g)

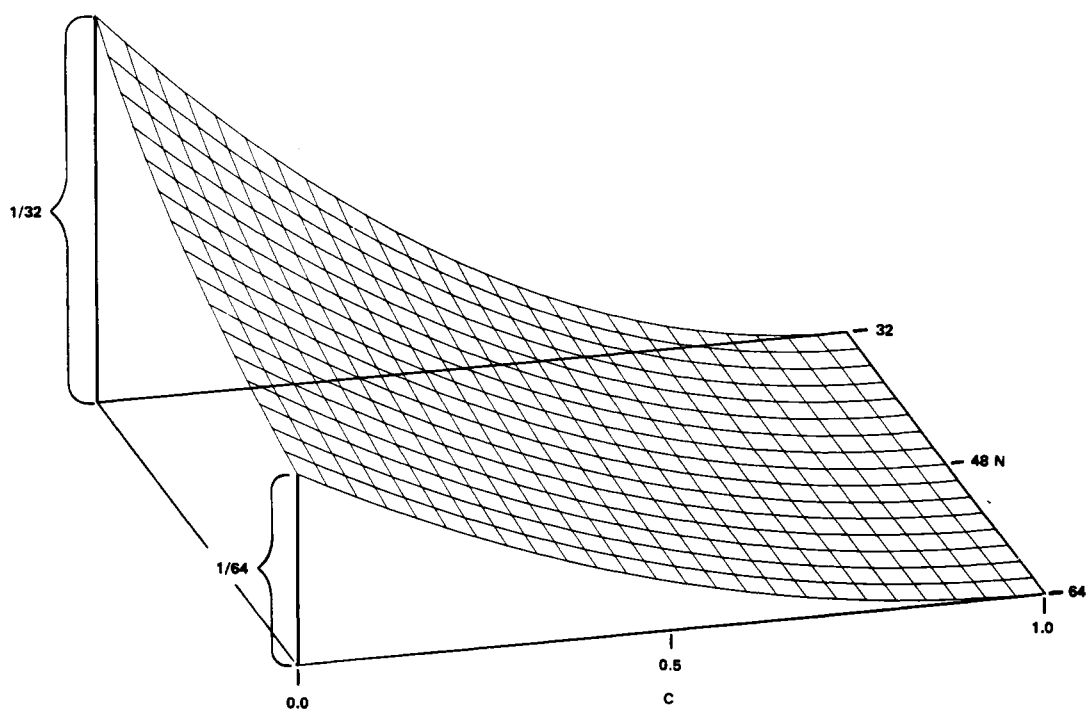


Fig. 7. Bias of $\lvert\hat{\gamma}\rvert^2$ versus $\lvert\gamma\rvert^2$ and n_d.

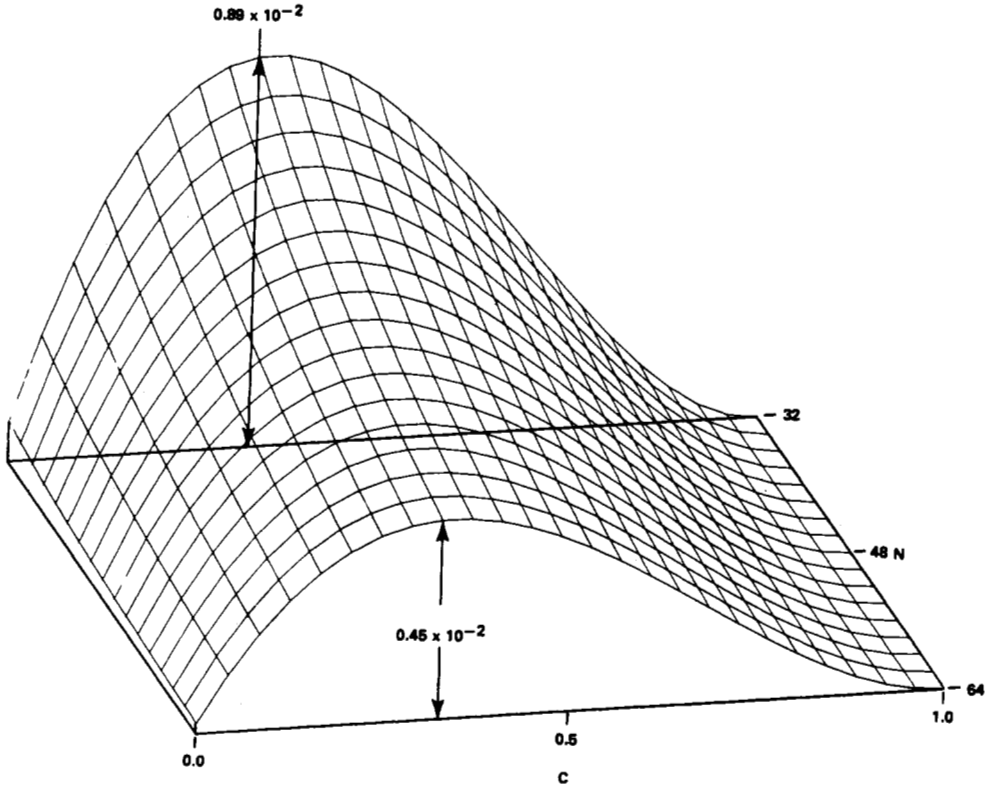


Fig. 8. Variance of $|\hat{\gamma}|^2$ versus $|\gamma|^2$ and n_d.

ments, that is, the case of zero overlap. Intuitively, the application of nonoverlapped smooth weighting functions does not make the most efficient use of the data when forming the MSC estimate. The experiment described herein examines this inefficiency, in terms of bias and variance of the MSC estimate as a function of different amounts of overlap.

The method of evaluating overlap is straightforward in concept. Data are generated with an accurately prespecified value of MSC that is independent of frequency. Then the sample mean and variance of the MSC estimate can be measured for the given overlap by averaging over frequency. Details are described in the paper by Carter, Knapp, and Nuttall [11].

Results of the experiment are summarized in Figs. 9 and 10. It is apparent from these results that the bias and variance can be reduced through overlapped processing. For example, when the MSC is 0.0, the variance of the estimator with a 50-percent overlap equals 31 percent of the variance of the estimator with no overlap. With a 50-percent overlap, the bias is 55 percent as large as with no overlap. Similarly, when the MSC = 0.3 and the overlap is 50 percent, the variance is 55 percent of the nonoverlapped estimator, and the bias is 50 percent as large. Observe that the bias and variance for zero overlap agree very well with (2e) through (2g) in Table 2. With a 62.5-percent overlap, or greater, the bias and variance achieve values corresponding to an effective n_d of about 64 with nonoverlapped processing. In the WOSA method, clearly, as the overlap increases, the computational cost must also increase. Increasing the overlap from 50 to 62.5 percent requires 32 percent more FFTs, but the variance of the MSC estimator decreases only from 80 to 95 percent of its value at a 50-percent overlap. It is doubtful, therefore, that the improvement in the WOSA method

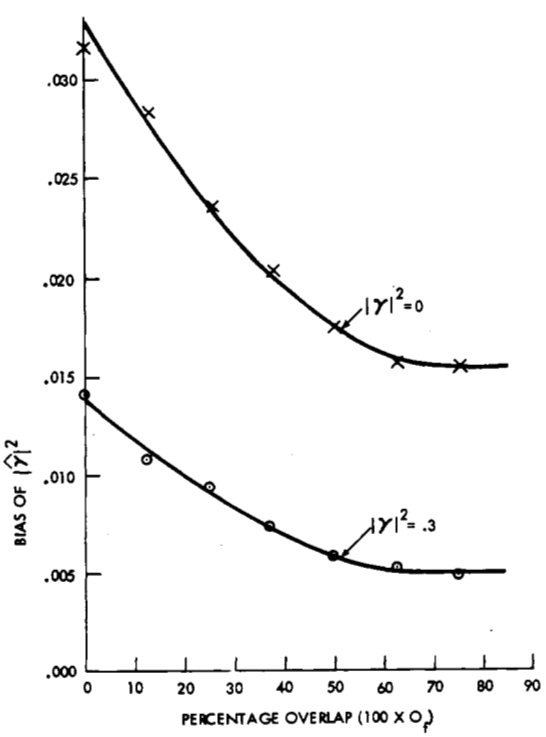


Fig. 9. Bias of $|\hat{\gamma}|^2$ when $n_d = 32$.

derived from using a 62.5-percent overlap, as opposed to a 50-percent overlap, will warrant the increased computational costs, except in unusual circumstances. Overlap percentages of 50 percent are quite reasonable and widely used. However, with the advent of the lag-reshaping

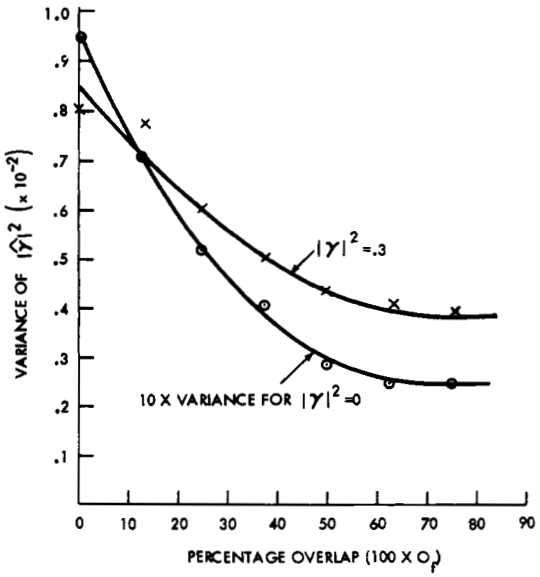


Fig. 10. Variance of $|\hat{\gamma}|^2$ when $n_d = 32$.

method other computational efficiencies need to be explored.

D. MSC Bias

One type of bias, derived under simplifying assumptions including that each data segment is sufficiently long to ensure adequate spectral resolution, has been shown by Nuttall and Carter [58] from Table 2 to be

$$E[\hat{C}] - C \cong \frac{1}{n_d}(1 - C)^2\left(1 + \frac{2C}{n_d}\right) \tag{11}$$

where E denotes the expected value. Equation (11) corroborates the observation by Bendat and Piersol [4] that more than one segment must be used to estimate MSC. It can be shown that, for $n_d = 1$, the estimated MSC equals unity regardless of the true value of MSC; however since (11) is an approximation it cannot be used to prove this relationship.

However, there is a second type of bias, described by Koopmans [48] that can be extremely serious. This is the bias due to misalignment, or rapidly changing phase. In particular, Koopmans [11] notes that if the phase angle of the cross power spectrum is a rapidly varying function of frequency at the frequency that the coherence is to be estimated, the estimated coherence (in particular, MSC) can be biased downward to such an extent that a strong coherence is masked. An expression for the bias as a function of the first derivative of the phase spectrum was given by Jenkins and Watts [42]. Based on results which follow here, Koopmans' [11] statement is correct; however, the MSC bias results from the widely used benchmark text by Jenkins and Watts' [42] are quantitatively incorrect for the application here. A brief derivation of the effect of misalignment akin to rapidly changing phase will be given later. These results compare favorably with analytical results by Halvorsen and Bendat [33], with empirical results by Carter and Knapp [13], and with empirical results to be presented here.

Rapidly varying phase as a function of frequency is caused by a time delay. One way to see this is to consider that the units of the slope of the phase are radians divided by radians per second, or simply seconds. The data can be realigned to compensate for a time delay. As stated by Brillinger [5], the importance of some form of prefiltering cannot be overemphasized, the simplest form being to lag one time series relative to the other (we note here, this prefiltering importance is also true for time-delay estimation). This procedure for coherence estimation has been suggested by others, including Akaike and Yamanouchi [2], Jenkins and Watts [42], and Koopmans [48]. Important to the concept of prefiltering two time series before estimating the MSC is that (unlike the estimated value of MSC) the (true value of) MSC is invariant under the linear filtering of the two series, as shown, for example, by Carter, Knapp, and Nuttall [11], and Koopmans' [48].

The effect of misalignment can be seen in the correspondence by Carter [19]. The results of that work show the magnitude of the cross-power spectrum (and cross correlation) is decreased by a constant factor, depending on the ratio of the delay misalignment to the FFT time duration. Note, though, that the average phase estimate remains unaltered. Further, we note that the constant degradation factor will not appear in either of the auto-power spectral densities. Thus the complex coherence is degraded by the same factor as the cross spectrum and the MSC is degraded by the square of this factor. That is,

$$E[\hat{C}(f)] \cong \left(1 - \frac{|D|}{T}\right)^2 C(f) \tag{12}$$

which agrees with Halvorsen and Bendat [33]. Heuristically, this makes sense because for no delay there is no degradation and for a delay equal to, or greater than, the FFT size, the estimated MSC is zero. Note that the bias due to misalignment D, with FFT time duration T, is

$$E[\hat{C}] - C \cong \frac{-2|D|}{T} C + \left(\frac{|D|}{T}\right)^2 C \tag{13a}$$

$$\cong \frac{-2|D|}{T} C, \quad |D| \ll T. \tag{13b}$$

For example, if $|D|/T = 0.25$, the expected value of the estimated MSC is about one-half of its true value. Clearly, the effect is important. Indeed, empirical results bear this out.

One of the results of Carter and Knapp [13] was the demonstration of the need to make the FFT (or equivalent transform) size larger. Empirically, large T was observed to reduce the bias in MSC estimation. This is consistent with (13). Looking at these data again (Carter and Knapp [13, fig. 5]) we see that the phase appears to have undergone a 1.5-radian change in 10 Hz, or a 24-ms delay was encountered in the band. Since each FFT was 500 ms, the estimated value of MSC should be about 0.91 of its true value. Indeed, this is what was indicated in Carter and Knapp [13, fig. 7]).

One practical means of reducing the bias due to a single path misalignment is to realign the two time series under investigation before estimating the MSC. The effects of misalignment were evident in a recent empirical investigation in which a broad-band underwater acoustic signal was transmitted through a direct path from a submerged transmitter to a submerged receiver. The recorded signals were processed with a number of different bulk time delays inserted before estimating the MSC. The bulk time delays were quantized to 250 ms; the FFT size was 1.0 s. The effect

of degraded MSC estimation is evident in Figs. 11 and 12, computed from 16 independent FFTs. In Fig. 11, $D = 0.18$ and the estimated MSC appears to be about 0.45 at 250 Hz. In Fig. 12 introducing another 250-ms bulk delay moves the generalized cross correlation SCOT peak from −0.18 to 0.07. (The SCOT or smoothed-coherence transform is the Fourier transform of the complex coherence; it is therefore the Fourier transform of the cross-power spectrum after weighting with a particular frequency function. See, for example, Carter, Nuttall, and Cable [12], Knapp and Carter [46], and Kuhn [49] for a discussion of the SCOT.) Now the estimated MSC is about 0.7 at 250 Hz in Fig. 12. There is a notable and predictable increase in the MSC estimate due to realignment.

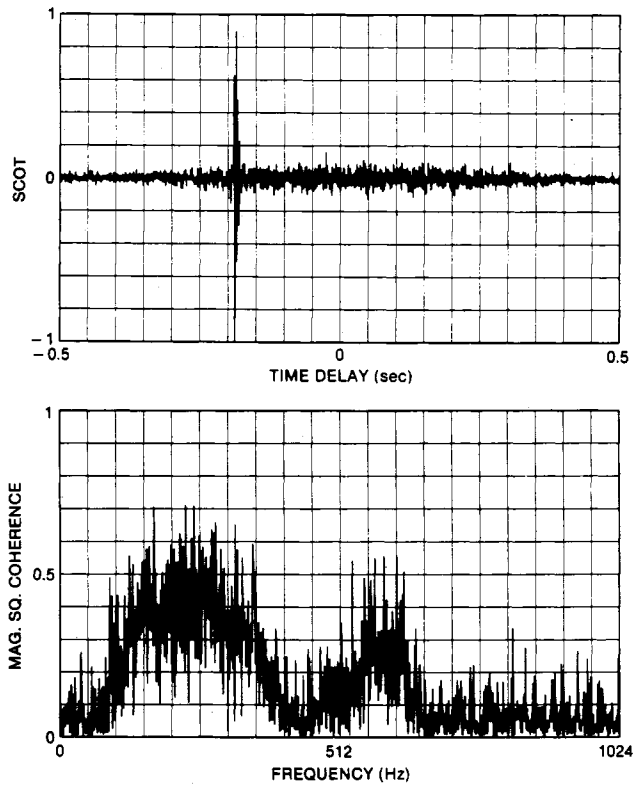


Fig. 11. SCOT estimate showing a −180-ms delay and corresponding MSC estimate of 0.45 at 250 Hz with a −180-ms delay.

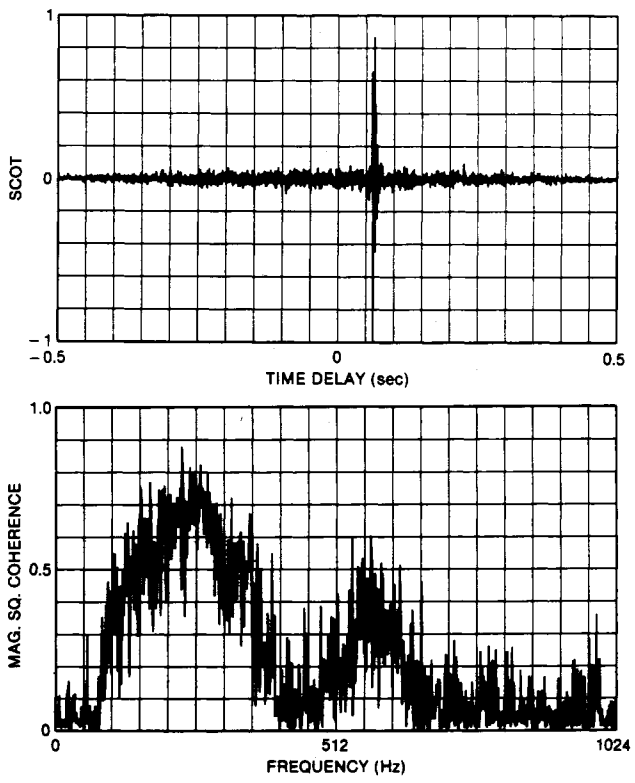


Fig. 12. SCOT estimate showing a 70-ms delay and corresponding MSC estimate of 0.7 at 250 Hz with a 70-ms delay.

Thus we see that even with a large number of FFT segments, estimates of the MSC can be significantly biased downward by unremoved bulk time delays, giving an erroneous indication of the value of the coherence. When the data are realigned and processed, estimates of the coherence are informative descriptors of the extent to which the ocean channel can be modeled by a linear time-invariant filter.

E. Receiver Operating Characteristics for a Coherence Detector

An algorithm for computing the receiver operating characteristics (ROC) or the probability of detection P_D versus the probability of false alarm P_F for a linearly thresholded coherence estimation detector is presented together with an example of an ROC table. More details can be found in Carter [15]. An article by Gevins *et al.* [27] presents results on using linearly thresholded coherence estimates to detect biomedical phenomena. We present an algorithm for computing P_D versus P_F for a specified amount of averaging and underlying coherence. Under simplifying assumptions, the probability density function of $\hat{C}$, when $C = 0$, is obtained from Table 1. In particular

$$p(\hat{C} \mid n_d, C = 0) = (n_d - 1)(1 - \hat{C})^{(n_d - 2)}. \tag{14}$$

Hence, the probability of false alarm is

$$P_F = 1 - \int_0^E (n_d - 1)(1 - \hat{C})^{(n_d - 2)}\, d\hat{C} \tag{15a}$$

and the threshold

$$E = 1 - (P_F)^{1/(n_d - 1)} \tag{15b}$$

that is, for a specified P_F we establish a threshold according to (15b). Now the computationally more complex question is: What probability of detection is achieved for this threshold value E? The answer is

$$P_D = \int_E^1 p(\hat{C} \mid n_d, C)\, d\hat{C} = 1 - P(C \le E \mid n_d, C) \tag{16}$$

where $P(\hat{C} \le E/n_d, C)$ is the cumulative distribution function (CDF). The CDF is given in Table 1. An example is illustrative.

For models of the form

$$x(t) = z_1(t) + n_1(t) \tag{17a}$$

$$y(t) = z_2(t) + n_2(t) \tag{17b}$$

where $z_i(t)$ is the output of a linear filter $H_i(f)$ excited by $s(t)$, $i = 1, 2$, and the noises are mutually uncorrelated and uncorrelated with the signal, then it can be shown that

$$C_{xy}(f) = C_{sx}(f)\, C_{sy}(f) \tag{18}$$

that is, the coherence between two receivers is the product of the coherence between the source and each of the indi-

vidual receivers for the model (17). Substituting results in

$$\frac{G_{z_1z_1}(f)}{G_{n_1n_1}(f)} \cdot \frac{G_{z_2z_2}(f)}{G_{n_2n_2}(f)} = \frac{C_{xy}(f)}{[1 - C_{sx}(f)][1 - C_{sy}(f)]}. \quad (19)$$

Now if $C_{sx}(f) = C_{sy}(f) = [C_{xy}(f)]^{1/2}$, then it follows that

$$\left[\frac{G_{z_1z_1}(f)\, G_{z_2z_2}(f)}{G_{n_1n_1}(f)\, G_{n_2n_2}(f)}\right]^{1/2} = \frac{\sqrt{C_{xy}(f)}}{1 - \sqrt{C_{xy}(f)}}. \quad (20)$$

Hence, to study the 0-dB (or equal signal-to-noise) case, we must select

$$10 \log_{10}\left[\frac{\sqrt{C}}{1 - \sqrt{C}}\right] = 0 \quad (21)$$

which implies $C = 0.25$. The ROC curves for $|\gamma|^2 = C = 0.25$ and $N = n_d = 4$, 8, and 16 independent data segments are given in Fig. 13. As seen in Fig. 13, performance can be

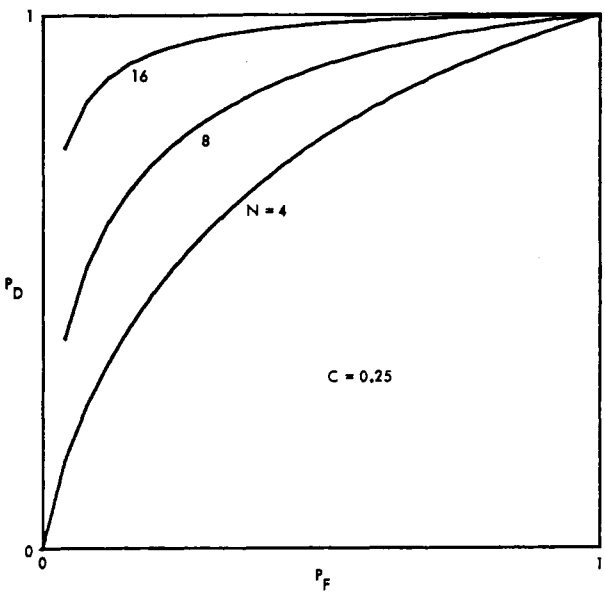


Fig. 13. ROC curves for $|\gamma|^2 = C = 0.25$, $n_d = N = 4, 8, 16$.

improved by increasing the number of disjoint data segments n_d if a sufficient amount of stationary data exists; if not, n_d can only be increased at the expense of degrading the frequency resolution with its inherent difficulties. If n_d is fixed, performance is determined by the underlying coherence or, equivalently, the signal-to-noise ratio (SNR) (see, e.g., Fig. 14). For many particular problems, the performance will be desired for different values of n_d and C. Because of the large number of possible choices for these parameters, we will not present an exhaustive series of results; a basic computer program listing is available in Carter [14].

F. Confidence Bounds for Magnitude-Squared Coherence Estimates

In many applications, two received signals are digitally processed to estimate coherence. Results of computing coherence estimate confidence bounds for stationary Gaussian signals are presented. Computationally difficult examples are given for 80- and 95-percent confidence with

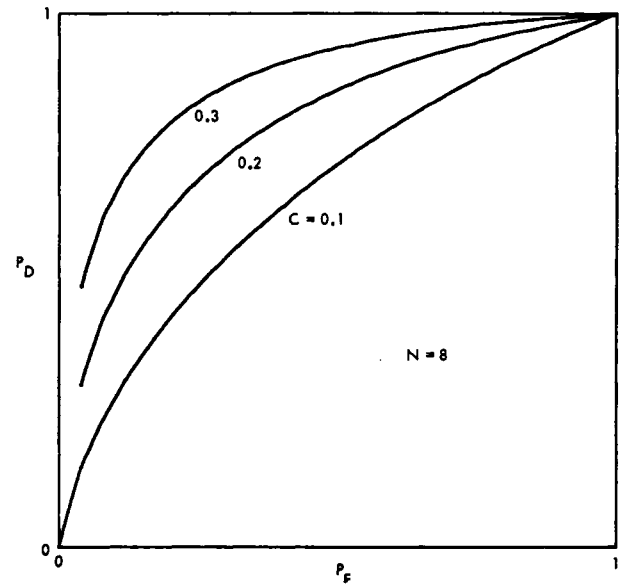


Fig. 14. ROC curves for $n_d = N = 8$, $|\gamma|^2 = C = 0.1, 0.2, 0.3$.

independent averages of 8, 16, 32, 64, and 128. A more complete discussion can be found in Scannell and Carter [65].

The MSC is useful in detection and is also of value in estimating the amount of coherent power common between two received signals. Therefore, it would be desirable, having estimated a particular value of MSC, to state with certain confidence that the true coherence falls in a specified interval. (A general discussion of confidence intervals is available in Cramér [24].) Early attempts to present 95-percent confidence were accomplished by Haubrich [39] who apparently used precomputed CDF curves and used a different method of presentation than the one used here. Related confidence work for the magnitude coherence (MC) is presented by Koopmans [48]. Empirical results for 95-percent confidence are given by Benignus [7]. The confidence limits given here appear to agree with approximate results in Bendat and Piersol [4], and Enochson and Goodman [26], and with results of Brillinger [5] from tabulated densities. Gosselin [28] compared MSC detectors with other detectors using the notion of ROC curves.

A computer program has been written (see Scannell and Carter [65]) to evaluate the CDF and confidence limits. Recall the CDF is a finite sum of ${}_2F_1$ hypergeometric functions, each of which is a polynomial, as given in Table 1. When C equals zero or unity, CDF values can be computed in closed form.

Let C be the true value of an unknown parameter and let $\hat{C}$ be its estimate. $\hat{C}$ is a random variable (RV) with a known probability density function (pdf) $p(\hat{C} \mid C)$. (The conditioning on C indicates that the shape of the pdf of $\hat{C}$ depends on the exact (unknown) value of C.)

Suppose we choose $A_L(C)$ and $A_u(C)$ such that

$$\Pr\,(A_L(C) < \hat{C} \mid C) = \int_{A_L(C)}^{\infty} dC\, p(\hat{C} \mid C) = 0.95 \text{ (say)} \quad (22a)$$

$$\Pr\,(\hat{C} < A_u(C)\, C) = \int_{\infty}^{A_u(C)} d\hat{C}\, p(\hat{C} \mid C) = 0.95 \text{ (say)}. \quad (22b)$$

Then the probability that RV $\hat{C}$ lies in the range $(A_L(C), A_u(C))$ is

$$\Pr\,(A_L(C) < \hat{C} < A_u(C) \mid C) = 0.90. \quad (23)$$

Now assume that $A_L(C)$ and $A_u(C)$ are monotonically increasing with C, and are continuous. Then there follows:

$$\Pr\left(A_u^{-1}(\hat{C}) < C < A_L^{-1}(\hat{C}) \mid C\right) = 0.90. \tag{24}$$

Therefore, the confidence interval for C is

$$\left(A_u^{-1}(\hat{C}),\ A_L^{-1}(\hat{C})\right), \text{ with confidence coefficient 0.90.} \tag{25}$$

Given a measurement $\hat{C}$, this interval can be computed once the functions $A_u^{-1}(\cdot)$ and $A_L^{-1}(\cdot)$ are known.

Fig. 15 presents computer generated 80- and 95-percent confidence limits (see Scannell and Carter [65]). The five

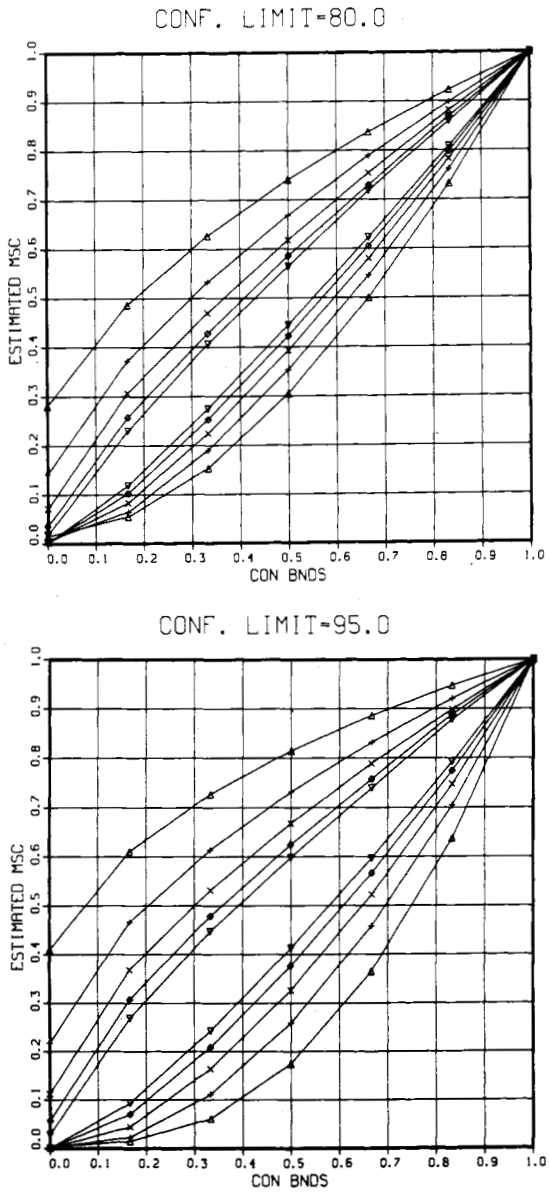


Fig. 15. 80-percent and 95-percent MSC estimate confidence bounds for N = 8, 16, 32, 64, and 128.

pairs of curves in each figure are for n_d = 8, 16, 32, 64, and 128 from outer to inner, respectively. If we make many estimates of MSC and keep applying confidence rules (whether or not C is random or constant), we will correctly include the true value of C in the determined interval the specified percent of the time. Sometimes the method of applying confidence rules is in doubt; for example in Fig. 15 if the estimate comes out to be 0.3 and n_d = 8, then a horizontal line does not intersect the upper confidence limit unless we extrapolate it backwards. Thus we could say that with 95-percent confidence the true MSC is in the region (−0.1, 0.62). Since we know *a priori* that the true value of C is nonnegative, we could just as easily say (but with no more confidence) that with 95-percent confidence (for n_d = 8 and $\hat{C}$ = 0.3) the true MSC falls in the region (0.0, 0.62). It is interesting to note that larger values of N do not always result in the upper confidence bound being lower. This also occurs not only in MSC but also in MC estimate confidence limits (see Koopmans [48]). It is also interesting to note that while increasing n_d is desirable, the confidence bounds for n_d = 128 are still very large.

IV. Time Delay Estimation

A. Introduction

A coherent signal emanating from an underwater acoustic source and monitored in the presence of noise at two spatially separated sensors can be mathematically modeled in the direct path as

$$x_1(t) = s(t) + n_1(t) \tag{26a}$$

$$x_2(t) = s(t - D) + n_2(t) \tag{26b}$$

where $s(t)$, $n_1(t)$, and $n_2(t)$ ar real, jointly stationary random processes. Signal $s(t)$ is assumed to be uncorrelated with noise $n_1(t)$ and $n_2(t)$ and one desires to estimate the unknown time delay D. We note that in the literature the received signals in noise are sometimes denoted by $r_1(t) = x_1(t)$ and $r_2(t) = x_2(t)$.

There are many applications in which it is of interest to estimate the time delay D. This section reviews the derivation of a maximum likelihood (ML) estimator given by Knapp and Carter [46]. While the model of the physical phenomena presumes stationarity, the techniques to be developed herein are usually employed in slowly varying environments where the characteristics of the signal and noise remain stationary only for finite observation time T. Studies of more complex effects are given by Bjørnø [9]; Tacconi [70]; Griffiths, Stocklin, and VanSchooneveld [29]; and Chan [20]. Further, the time delay D may also change slowly requiring time-varying or adaptive techniques such as those of Griffiths [30]; Owsley [62]; Picinbono [63]; Meyr [55]; and Lindsey and Meyr [50]. Other investigations of the time-varying case are studied by Knapp and Carter [47]; Adams, Kuhn, and Whylund [1]; Carter and Abraham [18]; and Schultheiss and Weinstein [69].

Another important consideration in estimator design is the available amount of prior knowledge of the signal and noise statistics. In many problems, this information is negligible. For example, in passive detection, unlike the usual communications problems, the source spectrum is unknown or only known approximately.

B. Derivation of the ML Estimator

The ML estimator is derived as follows. Assume that signals and noises are Gaussian. Denote the Fourier coefficients of $x_i(t)$ by

$$X_i(k) = \frac{1}{T}\int_{-T/2}^{T/2} x_i(t)e^{-jkt\omega}\,\Delta dt \tag{27a}$$

where

$$\omega_\Delta = \frac{2\pi}{T}. \tag{27b}$$

Note that the linear transformation $X_i(k)$ is Gaussian since $x_i(t)$ is Gaussian. Further, as T goes to infinity and k goes to infinity such that $k\omega_\Delta = \omega$ is constant

$$\tilde{X}_i(\omega) = \lim_{T\to\infty} TX_i(k) \tag{28a}$$

$$= \int_{-\infty}^{\infty} x_i(t)e^{-j\omega t}\,dt \tag{28b}$$

where $\tilde{X}_i$ is the Fourier transform of $x_i(t)$. A more complete discussion on Fourier transforms and their convergence is given in standard texts. When the observation time is large compared to the correlation time of the signal plus the magnitude of the delay then

$$E[X_1(k)\,X_2^*(\ell)] \cong \begin{cases} \frac{1}{T}G_{x_1x_2}(k\omega_\Delta), & k = \ell \\ 0, & k \neq \ell \end{cases}. \tag{29}$$

Now let the vector

$$\boldsymbol{X}(k) = [X_1(k),\, X_2(k)]' \tag{30}$$

where ' denotes transpose. Define the power spectral density matrix Q such that

$$E[\boldsymbol{X}(k)\,\boldsymbol{X}^{*\prime}(k)] = E\begin{bmatrix} X_1(k)\,X_1^*(k) & X_1(k)\,X_2^*(k) \\ X_2(k)\,X_1^*(k) & X_2(k)\,X_2^*(k) \end{bmatrix} \tag{31a}$$

$$= \frac{1}{T}\begin{bmatrix} G_{x_1x_1(k\omega_\Delta)} & G_{x_1x_2(k\omega_\Delta)} \\ G^*_{x_1x_2(k\omega_\Delta)} & G_{x_2x_2(k\omega_\Delta)} \end{bmatrix} \tag{31b}$$

$$\triangleq \frac{1}{T}Q(k\omega_\Delta). \tag{31c}$$

Recall from Section I that the magnitude-squared coherence (MSC)

$$C_{x_1x_2}(f) = \frac{|G_{x_1x_2}(f)|^2}{G_{x_1x_1}(f)\,G_{x_2x_2}(f)}. \tag{32}$$

Using the positive semidefinite properties of the spectral density matrix Q one can readily show that C is bounded by zero and unity. The vectors $\boldsymbol{X}(k)$, $k = -N, -N+1, \cdots, N$ are uncorrelated Gaussian (hence, independent) random variables. More explicitly, the probability density function for $\boldsymbol{X} = [\boldsymbol{X}(-N), \boldsymbol{X}(-N+1), \cdots, \boldsymbol{X}(N)]$, given the power spectral density matrix Q (or the delay, and spectral characteristics of the signal and noises necessary to determine Q) is

$$p(X \mid Q) = p(X \mid G_{ss}, G_{n_1n_1}, G_{n_2n_2}, G_{n_1n_2}, D) \tag{33a}$$

$$= c_p \exp(-\tfrac{1}{2}J_1) \tag{33b}$$

where

$$J_1 = \sum_{k=-N}^{N} \boldsymbol{X}^{*\prime}(k)Q^{-1}(k\omega_\Delta)\,\boldsymbol{X}(k)\,T \tag{33c}$$

and c_p is a function of the determinant of the spectral density matrix $|Q(k\omega_\Delta)|$. With proper substitution we obtain

$$J_1 = \sum_{k=-N}^{N} \tilde{X}^{*\prime}(k\omega_\Delta)Q^{-1}(k\omega_\Delta)\,\tilde{X}(k\omega_\Delta)\,\frac{1}{T}. \tag{33d}$$

For ML estimation it is desired to choose D to maximize $p(X \mid Q, D)$.

In general, the parameter D affects both c_p and J_1 in $p(\cdot)$. However, under certain simplifying assumptions, c_p is constant or is only weakly related to the delay. Specifically, suppressing the frequency argument

$$Q = (G_{ss} + G_{n_1n_1})(G_{ss} + G_{n_2n_2}) - (G_{n_1n_2} + G_{ss}e^{-j2\pi fD})(G^*_{n_1n_2} + G_{ss}e^{+j2\pi fD}) \tag{34}$$

which is independent of D if $G_{n_1n_2} = 0$ (i.e., the noises are uncorrelated).

For large observation times we have

$$J_1 \cong \int_{-\infty}^{\infty} \tilde{X}^{*\prime}(f)Q^{-1}(f)\,\tilde{X}(f)\,df. \tag{35}$$

The inverse of the spectral density matrix is given by

$$Q^{-1}(f) = \frac{\begin{bmatrix} G_{x_2x_2}(f) & -G_{x_1x_2}(f) \\ -G^*_{x_1x_2}(f) & G_{x_1x_1}(f) \end{bmatrix}}{G_{x_1x_1}(f)\,G_{x_2x_2}(f) - |G_{x_1x_2}(f)|^2}. \tag{36}$$

When the noises are uncorrelated

$$G_{x_1x_1}(f) = G_{ss}(f) + G_{n_1n_1}(f) \tag{37a}$$

$$G_{x_2x_2}(f) = G_{ss}(f) + G_{n_2n_2}(f) \tag{37b}$$

$$G_{x_1x_2}(f) = G_{ss}(f)\,e^{-j2\pi fD} \tag{37c}$$

and it follows that

$$J_1 = \int_{-\infty}^{\infty} \tilde{X}^{*\prime}(f)\,Q^{-1}(f)\,\tilde{X}(f)\,df = J_2 + J_3 \tag{38a}$$

where

$$J_2 = \int_{-\infty}^{\infty}\left[\frac{|\tilde{X}_1(f)|^2}{G_{x_1x_1}(f)} + \frac{|\tilde{X}_2(f)|^2}{G_{x_2x_2}(f)}\right]\cdot\frac{1}{[1 - C_{12}(f)]}\,df \tag{38b}$$

$$-J_3 = \int_{-\infty}^{\infty} A(f) + A^*(f)\,df \tag{38c}$$

$$A(f) = \tilde{X}_1(f)\,\tilde{X}_2^*(f)\cdot\frac{G_{ss}(f)\,e^{j2\pi fD}}{G_{x_1x_1}(f)\,G_{x_2x_2}(f)[1 - C_{12}(f)]}. \tag{38d}$$

Note that for real signals and noise $A^*(f) = A(-f)$ and it follows that

$$-J_3 = \int_{-\infty}^{\infty} A(f)\,df + \int_{-\infty}^{\infty} A(-f)\,df = 2\int_{-\infty}^{\infty} A(f)\,df. \tag{39}$$

Letting $T\hat{G}_{x_1x_2}(f)$ be defined as $\tilde{X}_1(f)\,\tilde{X}_2^*(f)$, we have

$$-J_3 = 2T\int_{-\infty}^{\infty} \hat{G}_{x_1x_2}(f)\,\frac{1}{|G_{x_1x_2}(f)|}\,\frac{C_{12}(f)}{[1 - C_{12}(f)]}\cdot e^{j2\pi fD}\,df. \tag{40}$$

Notice that the ML estimator for D will minimize $J_1 = J_2 + J_3$, but the selection of D has no effect on J_2. Thus D should maximize $-J_3$. Equivalently, the ML estimator selects as the estimate of delay the value of τ at which

$$\hat{R}_{y_1y_2}(\tau) = \int_{-\infty}^{\infty} \hat{G}_{x_1x_2}(f) \frac{1}{|G_{x_1x_2}(f)|} \frac{C_{12}(f)}{[1 - C_{12}(f)]} \cdot e^{j2\pi f\tau}\, df \tag{41}$$

achieves a peak.

These results compare favorably with closely related work of MacDonald and Schultheiss [53]; Hannan and Thomson [34]–[36]; Hahn and Tretter [31]; and Cleveland and Parzen [23]. A more complete discussion of the derivation and related references can be found in Knapp and Carter [46].

C. Interpretation of the ML Estimator

One common method of determining the time delay D is to compute the standard cross-correlation function

$$R_{x_1x_2}(\tau) = E[x_1(t)\, x_2(t - \tau)] \tag{42}$$

where E denotes expectation. The argument τ that maximizes (42) provides an estimate of delay. Because of the finite observation time, however, $R_{x_1x_2}(\tau)$ can only be estimated. For example, for ergodic processes, an estimate of the cross correlation is given by

$$\hat{R}_{x_1x_2}(\tau) = \frac{1}{T - \tau} \int_{\tau}^{T} x_1(t)\, x_2(t - \tau)\, dt \tag{43}$$

where T represents the observation time. In order to improve the accuracy of the delay estimate $\hat{D}$, it is desirable to pre-filter $x_1(t)$ and $x_2(t)$ prior to cross correlation. We call this simple but very important process generalized cross correlation for lack of a better description. As shown in Fig. 16, x_i or r_i may be filtered through H_i to yield y_i for $i = 1, 2$.

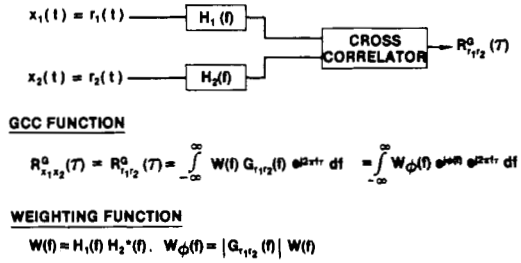


Fig. 16. Generalized cross-correlation (GCC) function.

The resultant y_i are cross correlated, that is, multiplied and integrated, for a range of hypothesized time delays or time shifts τ until the peak is obtained. We caution the reader that in the literature and even here we use x_i and r_i interchangeably to denote received signal. The time shift causing the peak is an estimate of the true delay $\hat{D}$. When the filters $H_1(f) = H_2(f) = 1$, for all f, the estimate D is simply the abscissa value at which the standard cross-correlation function peaks. Knapp and Carter [46] provides for a generalized cross correlation through the introduction of the filters $H_1(f)$ and $H_2(f)$ which, when properly selected, can significantly enhance the estimation of time delay.

The cross correlation between $x_1(t)$ and $x_2(t)$ is related to the cross-power spectral density function by the well-known Fourier transform relationship

$$R_{x_1x_2}(\tau) = \int_{-\infty}^{\infty} G_{x_1x_2}(f)\, e^{j2\pi f\tau}\, df. \tag{44}$$

When $x_1(t)$ and $x_2(t)$ have been filtered with filters having transfer functions H_1 and H_2, respectively, as depicted in Fig. 16, the cross-power spectrum between the filter outputs is given by

$$G_{y_1y_2}(f) = H_1(f)\, H_2^*(f)\, G_{x_1x_2}(f) \tag{45}$$

(again we note $x_1 = r_1$ and $x_2 = r_2$).

Therefore, the generalized cross-correlation or GCC function between $x_1(t)$ and $x_2(t)$ is

$$R_{y_1y_2}(\tau) = \int_{-\infty}^{\infty} W(f)\, G_{x_1x_2}(f)\, e^{j2\pi f\tau}\, df \tag{46a}$$

where the generalized frequency weighting

$$W(f) = H_1(f)\, H_2^*(f). \tag{46b}$$

In practice, only an estimate of the cross-power spectral density can be obtained from finite observations of the received signals. Consequently, the integral

$$\hat{R}_{y_1y_2}(\tau) = \int_{-\infty}^{\infty} W(f)\, \hat{G}_{x_1x_2}(f)\, e^{j2\pi f\tau}\, df \tag{47}$$

is evaluated and used for estimating delay. Indeed, depending on the particular form of $W(f)$ and prior information available, it may also be necessary to estimate the generalized weighting. For example, when the role of the prefilters is to accentuate the signal passed to the correlator at those frequencies at which the coherence or signal-to-noise ratio (SNR) is highest, then $W(f)$ can be expected to be a function of the coherence or signal-and-noise spectra which must either be known or estimated. Besides the maximum likelihood (ML) weighting there are an entire family of generalized weightings. See Fig. 17 for some common GCC weightings. A discussion of when to use these *ad hoc*

METHOD	$W(f) = H_1(f)\, H_2^*(f)$	$W_\phi(f) = W(f)\,\lvert G_{r_1r_2}(f)\rvert$
SCC	1	$\lvert G_{r_1r_2}(f)\rvert$
ROTH	$1/G_{r_1r_1}(f)$	$\lvert G_{r_1r_2}(f)\rvert / G_{r_1r_1}(f)$
WIENER PROCESSOR	$C_{r_1r_2}(f)$	$C_{r_1r_2}(f)\,\lvert G_{r_1r_2}(f)\rvert$
SCOT	$1/\sqrt{G_{r_1r_1}(f)\, G_{r_2r_2}(f)}$	$\sqrt{C_{r_1r_2}(f)}$
PHAT	$1/\lvert G_{r_1r_2}(f)\rvert$	1
ML	$\frac{C_{r_1r_2}(f)}{[1 - C_{r_1r_2}(f)]\lvert G_{r_1r_2}(f)\rvert}$	$\frac{C_{r_1r_2}(f)}{1 - C_{r_1r_2}(f)}$

$$C_{r_1r_2}(f) = \frac{|G_{r_1r_2}(f)|^2}{G_{r_1r_1}(f)\, G_{r_2r_2}(f)}$$

Fig. 17. Various GCC functions.

weightings is beyond the scope of this paper; however, the reader should be aware that other weightings have been shown under certain conditions to offer attractive processing capabilities (e.g., tonal rejection via SCOT processing).

The ML weighting, that causes minimum variance TDE under Gaussian assumptions, is given by

$$W_{ML}(f) = \frac{1}{|G_{x_1x_2}(f)|} \cdot \frac{C_{12}(f)}{[1 - C_{12}(f)]}. \tag{48}$$

When $G_{x_1x_2}(f)$ and $C_{12}(f)$ are known, this is exactly the proper weighting. When these terms are unknown, they can be estimated via lag-reshaping spectral estimation techniques discussed earlier or using techniques of others, e.g., Marple [54]; or the WOSA method of Carter, Knapp, and Nuttall [11]; or classical methods, e.g., Jenkins and Watts [42] and Bendat and Piersol [6]. Substituting estimated weighting for true weighting is entirely a heuristic procedure whereby the ML estimator can approximately be achieved in practice. When the noises are uncorrelated but have the same power spectrum we can show

$$W_{ML}(f) = \frac{G_{ss}(f)/G_{nn}^2(f)}{[1 + 2\,G_{ss}(f)/G_{nn}(f)]}. \tag{49}$$

For small SNR we have

$$W_{ML}(f) \cong \frac{G_{ss}(f)}{G_{nn}^2(f)} \tag{50}$$

which is the well-known Eckart filter (see Eckart [25]) used in optimum signal detection at low SNR. Thus the prefilters used for optimum signal detection at low SNR are the same prefilters as used for minimum variance time-delay estimation at low SNR.

D. Fundamental Performance Limits

The Cramér–Rao lower bound (CRLB), is given by

$$\text{var} \geq \left. \frac{-1}{E\left\{\dfrac{\partial^2 \ln p(X|Q,\tau)}{\partial \tau^2}\right\}} \right|_{\tau = D}. \tag{51}$$

The only part of the log density which depends on τ, the hypothesized delay, is J_3. More explicitly

$$E\left\{\frac{\partial^2}{\partial \tau^2} \ln p(X \mid Q, \tau)\right\} = \frac{\partial^2}{\partial \tau^2} E\left(\frac{-1}{2} J_3\right). \tag{52}$$

If

$$G_{x_1x_2}(f) = |G_{x_1x_2}(f)|\, e^{-j2\pi fD}$$

then (since the complex cross-spectral estimator is unbiased) it follows that

$$E\left(\frac{-1}{2} J_3\right) = T \int_{-\infty}^{\infty} e^{j2\pi f(\tau - D)} \frac{C_{12}(f)}{[1 - C_{12}(f)]}\, df. \tag{53}$$

Hence the minimum variance of any time-delay estimator is

$$\min \text{var} = \left[T \int_{-\infty}^{\infty} (2\pi f)^2 \frac{C_{12}(f)}{[1 - C_{12}(f)]}\, df \right]^{-1}. \tag{54}$$

This is the minimum variance and that which the ML processor achieves asymptotically for sufficiently large T. For constant signal-and-noise power spectra

$$\frac{C_{12}(f)}{1 - C_{12}(f)} = \frac{(\text{SNR})^2}{1 + 2\text{SNR}} \tag{55}$$

so that

$$\min \text{var} = \frac{1}{2T \displaystyle\int_0^B (2\pi)^2 f^2 \left(\frac{\text{SNR}^2}{1 + 2\text{SNR}}\right) df} \tag{56a}$$

or simply

$$\sigma_{\text{CRLB}}^2 = \frac{3}{8\pi^2} \frac{(1 + 2\text{SNR})}{\text{SNR}^2} \cdot \frac{1}{B^3 T}. \tag{56b}$$

When the signal and noise have the same flat power spectra, as noted by Scarbrough, Tremblay, and Carter [67], for the time-delay estimation (TDE) problem, this form of the Cramér–Rao lower bound (CRLB) is commonly used as the performance standard. The CRLB yields a lower bound on the variance of any unbiased time-delay estimate as a function of several parameters (e.g., the signal-and-noise power spectra and the integration (observation) time). Part of the appeal of the CRLB is that for cases of practical interest, there is a theorem which states that the maximum likelihood (ML) estimate can be made arbitrarily close to the CRLB for sufficiently long integration times, see Van Trees [71]. However, the theorem does not specify how long the integration time must be. Thus while the CRLB sets a lower bound on the variance of the time-delay estimate, actual performance can be much worse for a given signal-to-noise (SNR) and observation time. This is corroborated by the simulation results of Scarbrough, Ahmed, and Carter [66], Scarbrough, Tremblay, and Carter [67], and Hassab and Boucher [37]. Several studies have been conducted to find a bound tighter than the CRLB which would predict performance more accurately. Work in this area has been done by Ianniello, Weinstein, and Weiss [41], Weiss and Weinstein [72] and [73], as well as Chow and Schultheiss [22]. Ianniello [40] has developed a correlator performance estimate (CPE). It has been shown via simulation that, for the cross-correlation technique of TDE, the CPE yields a more accurate estimate of performance than the CRLB, especially at low SNR (see Ianniello [40]). The following discussion presents some additional comparisons of the CPE and CRLB and discusses some implications of these comparisons. In particular, the behavior of the CPE and the CRLB is considered as a function of the observation time and SNR, and the implications of this behavior are considered as related to coherent and incoherent signal-processing techniques for time-delay estimation. In addition, simulation results are presented to support the inferences of the theoretical analysis.

Chow and Schultheiss [22], Scarbrough [68], Betz [8], and Johnson, Ohlms, and Hampton [43] have studied the low SNR problem. In the work here, consideration will be limited to signal-and-noise power spectra $G_{ss}(f)$ and $G_{nn}(f)$, respectively, which are flat (constant) over the frequency range $-B$ to $+B$ hertz and zero outside this range. B is then a measure of the source signal bandwidth. Additionally, it will be assumed that the bandwidth–observation time product BT is large (say $BT \geq 100$, e.g., for $B = 100$ Hz, $T \geq 1$ s).

In general, proper prefiltering prior to cross correlation is required to achieve the ML estimate of time delay. Under the special conditions here, the ML estimate of the time delay can be obtained by computing the cross-correlation function $R(\tau)$ between $x_1(t)$ and $x_2(t)$. The ML time-delay estimate is the value of τ which maximizes $R(\tau)$. For other power spectra, a generalized cross correlator with the proper prefilters is required.

For the signals of interest in this section, the minimum variance of the TDE for the CRLB can be expressed as (see Quazi [64])

$$\sigma_{\text{CRLB}}^2 = \frac{3}{8\pi^2} \cdot \frac{(1 + 2\text{SNR})}{\text{SNR}^2} \cdot \frac{1}{B^3 T} \tag{57}$$

where T is the observation time and SNR = $G_{ss}(f)/G_{nn}(f)$ = constant for $|f| < B$. Note that in this definition, SNR is the signal-to-noise ratio at the input of a single receiver.

The CPE was developed by Ianniello [40] to provide more accurate performance prediction for the cross-correlation technique for TDE in the presence of large estimation errors or anomalous estimates. The CPE assumes that the anomalous estimates will be uniformly distributed across the correlation window, say, from $-T_0$ to $+T_0$ seconds. The CPE yields the following estimate of the variance of the TDE error:

$$\sigma^2_{CPE} = PT_0^2/3 + (1 - P)\sigma^2_{CRLB} \tag{58}$$

where P is the probability of an anomalous estimate. For the signal-and-noise spectra considered here, P can be approximated as

$$P \simeq 1 - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(x - \alpha)^2\right] \cdot \left\{\int_{-\infty}^{\beta x} \frac{dy}{2\pi} \exp -\left[\frac{1}{2}y^2\right]\right\}^{M-1} dx \tag{59a}$$

where

$$\alpha = \frac{\sqrt{2BT}\,(\text{SNR})}{[\text{SNR}^2 + (1 + \text{SNR})^2]^{1/2}} \tag{59b}$$

$$\beta = \left[1 + \frac{\text{SNR}^2}{(1 + \text{SNR})^2}\right]^{1/2} \tag{59c}$$

and $M = 4BT_0$. The probability of anomaly must be evaluated numerically to obtain the probability of anomaly P for a given set of parameters B, T, T_0, and SNR.

The CPE and CRLB are compared in Fig. 18 for the case of B = 100 Hz, T = 8 s, and $T_0 = \frac{1}{8}$ s relative to an assumed sampling frequency of 2048 Hz. Both curves in Fig. 18 are

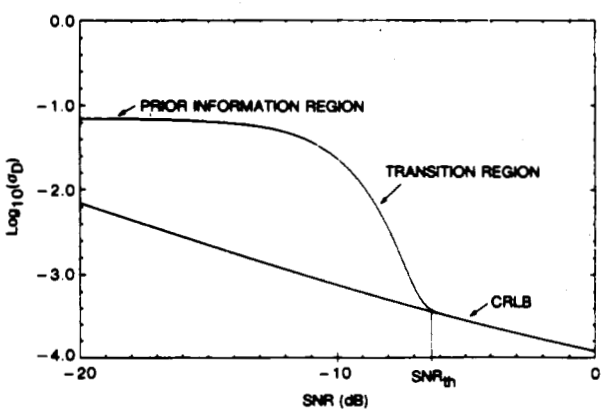


Fig. 18. Plot of CPE and CRLB versus SNR (dB) (T = 8 s, B = 100 Hz, $T_0 = \frac{1}{8}$ s).

plotted as the base-10 logarithm of σ_D, the standard deviation of the time-delay error, versus the input SNR in decibels. The upper curve is the CPE; the lower curve is the CRLB. The CPE is characterized by three regions: 1) at low SNR, there is a region where prior information limits the variance (e.g., the maximum observable delay T_0 is known); 2) at moderate SNR, there is a transition region from the prior information limit to the CRLB; and 3) at high SNR, the CPE coincides with the CRLB. The SNR at which the CPE begins to deviate from the CRLB is referred to as the threshold SNR (SNR_{th} in Fig. 18). We note also that while the CPE is not a bound, recent work has been done to derive a bound tighter than the CRLB. This new bound is close to the CPE, falling just below the CPE on Fig. 18. This new bound, the Ziv Zakai lower bound (ZZLB) is discussed by Ianniello, Weinstein, and Weiss [41].

E. Simulation Description and Results

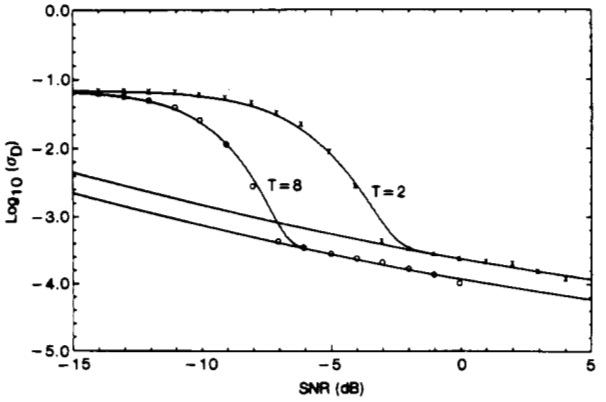
A computer simulation was conducted to corroborate the theoretical TDE performance predictions. Earlier simulations have been conducted by Hassab and Boucher [37]. Good agreement among many earlier simulation was shown in a comparison of a large number of simulation results by Kirlin and Bradley [45]. The cross correlation of simulated received sequences was computed using the FFT approach described by Oppenheim and Schafer [61]. The particular method is very similar to the lag-reshaping method of spectral estimation presented earlier. It is a little more complex in that zero filling on FFT segments is not necessarily equal. The details are not important to the discussion here but may be if attempting to reproduce the results. It is important that the mean and variance of all the noise-only correlation points be the same for all correlation lags. After cross correlation the estimate of the time delay was obtained by finding the delay value for which the cross correlation was a maximum. This approach yields the time-delay estimate quantized in units of the sampling interval.

The simulation was conducted for two integration times over a range of SNR values. Unlike many earlier simulations, effects of low SNR were studied. Different integration times are obtained by varying the number of cross power spectra which are averaged before taking the IFT (inverse Fourier transform) to obtain the estimate of the cross-correlation function. 8 and 32 data segments were processed coherently in the simulation to obtain integration times of 2 and 8 s, respectively, for the case of 512-point data segments and an assumed sampling frequency of 2048 Hz. A total of 2000 trials was conducted at each SNR to obtain the experimental time-delay variances. These results are plotted in Fig. 19(a) along with the corresponding theoretical curves for the CRLB and the CPE. The symbol size of the experimental points is indicative of the 90-percent confidence limits. The theory and results are in very close agreement and support the previous analysis. Of profound importance is that the CRLB is a poor predictor of performance at low-input SNR. This has significant implications for using techniques to increase coherent processing time as discussed by Scarbrough, Tremblay, and Carter [67] and Scarbrough [68]. These results are graphically portrayed in Fig. 19(b). Note now we plot the Ziv Zakai lower bound (ZZLB) as opposed to the CPE. Also we compare a coherent or tracking approach for T = 8 with an incoherent approach. (In the incoherent case, we first estimate the delay over short observation intervals and then average the time-delay estimates over the available intervals.) The threshold is clearly reduced by coherent processing.

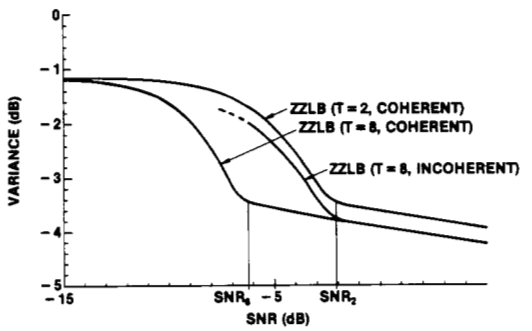
V. Focussed Time-Delay Estimates for Ranging

A. Introduction

This section discusses using focussed time-delay estimates to passively estimate range to a coherent source. Most of the work in this section is a summary of the work in Carter



(a)



(b)

Fig. 19. (a) Comparison of CPE, CRLB, and simulation results (B = 100 Hz, $T_0 = \frac{1}{8}$ s, T = 2, 8 s). (b) Plot of ZZLB and effects of coherent versus incoherent processing.

[15]. Other related references include Owsley [62], Lynch [51], and Lynch, Schwartz, and Sivaprosad [52].

The maximum-likelihood (ML) processor is presented for passively estimating range and bearing to an acoustic source. The source signal is observed for a finite-time duration at several sensors in the presence of uncorrelated noise. When the speed of sound in an isovelocity medium and the sensor positions are known, the ML estimator for position constrains the source-to-sensor delays to be focused into a point corresponding to a hypothesized source location. The variances of the range error and bearing error are presented for the optimum processor. It is shown that for bearing and range estimation, different sensor configurations are desirable. However, if the area of uncertainty is to be minimized, then the sensors should be divided into equal groups with one-third of the sensors in each group.

An underwater acoustic point source radiating energy to several collinear receiving sensors is shown in Fig. 20. The

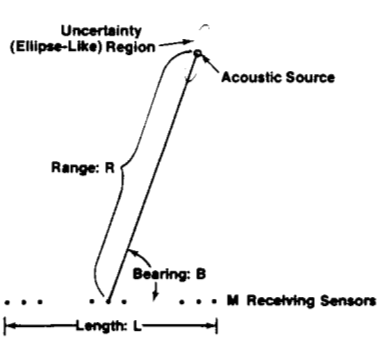


Fig. 20. Array geometry used to estimated source position.

position of the source in two space can be characterized by range R and bearing B from a given frame of reference. The particular geometry of interest is two-dimensional with an acoustic point source whose range and bearing are to be estimated by a fixed number of receivers. For the purposes of this work, we presume that the receiving hydrophones are collinear. However, regardless of the hydrophone positions, a fixed number of sensors have an inherent uncertainty in estimating source location. For our geometries here, this uncertainty region is nominally elliptical, so that by properly defining how range and bearing are measured, the estimation errors can be decoupled (see Fig. 20). For a collinear array of sensors, we measure the bearing as the angle between the line array and the major axis of the uncertainty region.

For a radiating source distant from an array, the uncertainty in measuring R and B is characterized by an extremely elongated elliptical uncertainty region. The problem addressed here is how to estimate range R and bearing B to a source when M sensors separated by a maximum of L (meters) have observed T seconds of received data. We will examine the maximum-likelihood (ML) technique for position estimation.

B. Mathematical Model

For our purposes we assume that each receiving sensor at the tth instant in time corresponds to a signal plus noise. Namely, the ith (of M) sensor outputs is characterized by

$$r_i(t) = s(t + D_i) + n_i(t), \quad i = 1, M, \quad 0 \le t \le T. \tag{60}$$

The signal and noises are uncorrelated and the noises are mutually uncorrelated. Without loss of generality $D_1 = 0$.

For a spatially stationary, that is, nonmoving, source the signal can be viewed as an attenuated and delayed source signal. However, it is felt the problem of estimating the position of a stationary source is considerably easier than that of a moving source. Thus the results here serve as a bound on performance; still we will see that it is extremely difficult in the best case to passively estimate source range.

C. Maximum Likelihood Performance

The maximum-likelihood (ML) estimate for the time-delay vector has been derived for stationary Gaussian processes by Hahn [32] and Carter [14]. It is not difficult to show that an ML estimate for range and bearing when the sensor (element) positions are known is achieved by a variation of Carter [14]. In particular, by focusing all the time-delay elements at many (hypothesized) range and bearing pairs and watching for the peak output of the ML time-delay vector system, the ML position estimate is observed. An ML system realization is shown in Fig. 21. This figure is sometimes referred to as a focused beamformer. More details are given in the paper by Carter [15]. Another way to look at this problem is that we want to maximize a quantity by adjusting a number of delay parameters subject to the constraint that all the delays must intersect in a single hypothesized position. Such a system and its variance has been examined by Bangs and Schultheiss [3].

For equal noise spectra, we should utilize a product of pre- and postfilters with magnitude-squared transfer function, given by Bangs and Schultheiss [3] (see also Knapp and Carter [46], Hannan and Tompson [35], MacDonald and

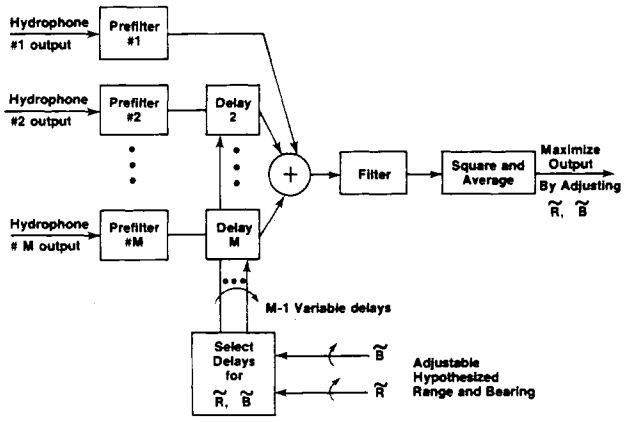


Fig. 21. Maximum-likelihood estimator for range and bearing.

Schultheiss [53], and Carter [14])

$$|H(\omega)|^2 = \frac{S(\omega)/N^2(\omega)}{1 + M[S(\omega)/N(\omega)]} \tag{61}$$

in order to minimize the variance of delay estimates (where S and N denote signal-and-noise-power spectral densities). For passive range and bearing estimation the variance of the parameter estimate Θ, where $\Theta = R$ or $\Theta = B$, is more complicated (see Carter [15]) and is summarized in the following subsection.

Having selected proper prefilters for a specified array geometry, the maximum likelihood estimates of range and bearing (that is, minimum variance estimates) is obtained by coherently processing the outputs of the sensing hydrophones. In particular, each hydrophone output is prefiltered to accentuate a high signal-to-noise ratio (SNR), then delayed and summed. The summed signal is fed to a filter, then squared and averaged for the observation time. The output of this network is maximized through the indirect adjustment of the delay parameters. The delay parameters are derived on the basis of two adjustable parameters: hypothesized bearing and hypothesized range. Thus an operator need only adjust his best estimate of bearing and range. From these two inputs, proper delays are inserted in each hydrophone receiving line. The process of delaying and summing is a focused beamformer where the delays used cause the beamformer to presume that the source wavefront is curved and not planar. The individual sensor-to-sensor delays inserted are directly related to the hypothesized source and sensor locations.

For a particular array type A, Carter [15] has shown that, at high output SNR, the minimum variance of bearing and range estimates is given by

$$\sigma_A^2(\hat{B}) \cong \frac{K_B^{(A)}}{TMVL_e^2} \tag{62}$$

and

$$\sigma_A^2(\hat{R}) \cong \frac{K_R^{(A)} R^4}{TMVL_e^4} \tag{63}$$

where $\sigma^2(\hat{B})$ is measured in radians squared, the effective array length $L_e = L \sin B$, and the constants K_R and K_B for four array types are given in Table 3. M is the minimum of the number of sensors and the array length divided by the design half-wavelength; also at high-output SNR

$$V = \frac{1}{2\pi C^2} \int_0^{\Omega} \frac{S(\omega)}{N(\omega)} \omega^2 \, d\omega \tag{64}$$

where $S(\omega)$ is the signal power spectrum, $N(\omega)$ is the noise power spectrum, C is the speed of sound in the medium, Ω is the highest source (or receiver) frequency, and, as earlier, T is the observation time, R is the range, and B is the bearing.

Table 3 Constants for Four Arrays of Interest

Array Type	$\sqrt{K_R/K_B}$	K_R	K_B
Equispaced line	7.75	360	6
$M/2$, 0, $M/2$	∞	∞	2
$M/3$, $M/3$, $M/3$	6.9	144	3
$M/4$, $M/2$, $M/4$	5.7	128	4

D. Discussion

Doubling the number of sensors M or the observation time T will reduce the standard deviation of either the bearing estimate or range estimate by 1.4. In bearing estimation, we desire to make L_e large and the constant K_B small in order to reduce variance. Note that doubling the array length reduces the variance by four. Thus array length is a more important factor in bearing estimation than either integration time or the number of hydrophones when operating at high-output SNR.

The four different array types studied are an equispaced line array and three line arrays with M elements grouped at the two ends and the middle of the array. MacDonald and Schultheiss [53] have shown that, by placing half of the M elements at each end of a line array in an M over two, zero, M over two grouping, a bound on bearing variance is obtained. This bound, of course, is for a hypothetical array where the elements are collocated but still sense independent noises. The practical implications of MacDonald and Schultheiss' result are both to provide a bound on how well bearing can be estimated under ideal conditions and to suggest how to place a limited number of hydrophones over a large aperture. Namely, for bearing estimation half of the hydrophones should be positioned at each end of the array, placed at half-wavelength spacing for the design frequency, and none should be placed in the middle of the available aperture.

It is noteworthy that the variance of the range estimate depends on the fourth power of the range relative to the effective baseline. This fourth-order dependence is a fundamental physical limit and makes the task of passive ranging at long range with short arrays extremely difficult. The variance of the range estimate is reduced by making the effective array length L_e large. This can be done by making the array length L large or to a lesser degree by physically steering the array broadside to the source. The variance can also be reduced by decreasing the range to the source. Of course, reducing the range to the source can also increase SNR depending upon propagation conditions.

The constant K_R depends on the array type. For an equi-

spaced line array, K_R is 360. A bound on passive ranging performance is provided by an array configured with a quarter of the hydrophones at each end and half in the middle. Thus we see that the hydrophone configuration in a line array desired for bearing estimation and the one for range estimation differ. The bearing array should have its elements toward the array ends, while the ranging array should have half of its elements in the central portion. However, a line array with a third of its elements at each end and the middle will minimize the uncertainty region, that is the product of the cross range error and ranging error (see Carter [15]). Thus in this sense, a line array physically segmented into three equal groups of elements will outperform all other arrays for passively locating an acoustic source. This result is summarized in Table 4. The optimum processor coherently combines all M hydrophone outputs. If, however, only the beamformer output from each subarray is used for coherent processing, a nearly optimum technique is believed to result.

Table 4 Optimum Sensor Configuration

For Estimating	Best Array Configuration
Bearing	● ● ● ● ● ● (M/2) ● ● ● ● ● ● (M/2)
Range	● ● ● (M/4) ● ● ● ● ● ● (M/2) ● ● ● (M/4)
Position	● ● ● ● (M/3) ● ● ● ● (M/3) ● ● ● ● (M/3)

Of considerable concern when attempting to predict the performance of a passive localization technique are values such as SNR, number of sensors, and integration time. It is interesting that these terms, together with constant such as 2π, can all be attributed to the standard deviation of the bearing estimates (measured in radians). Then the relative range error given by the standard deviation of the range estimate divided by the true range is given by a constant times standard deviation of the bearing estimate times a term that depends linearly on the range to the source relative to the effective array length. In particular

$$\frac{\sigma_A(\hat{R})}{R} = \sqrt{\frac{K_R}{K_B}}\frac{R}{L_e}\sigma_A(\hat{B}). \tag{65}$$

For example, suppose an equispaced line array had an inherent standard deviation of $\frac{1}{10}$ rad (5.7°) and was to estimate the range to a source ten times as far away as the effective array length. In that case, the relative range error is 7.75, or more than 700 percent. Hence, we see that it is extremely difficult to passively estimate range of a distant source even under ideal conditions with high-output SNR.

One of the advantages of expressing relative range errors in this form is that the standard deviation of bearing estimates is a term familiar to sonar engineers and signal processors. Moreover, the ocean medium may inherently limit the practical ability to estimate bearing even though theory predicts that with enough SNR or integration time, the bearing can be measured arbitrarily well. The expression given here clearly points out the need to make the array length large when the source range cannot be reduced. Of interest is that this conclusion is extremely insensitive to the type of array, provided the array has some ranging capability. This can be seen from the similarity of the constants given in Table 3.

To summarize, we desire to know how to place a limited number of hydrophones over a baseline of fixed length. The hydrophones should be placed in groups, with the hydrophones in each group placed at half-wavelength spacing for the design frequency. For passive bearing estimation, half of the M hydrophones should be placed at each end of the array. For passive range estimation, a quarter of the hydrophones should be placed at each end of the array and half placed in the middle. For simultaneously estimating range and bearing, both passively, the hydrophones are placed in three groups or subarrays, each with M over three hydrophones. If the baseline remained of fixed length and we had more hydrophones to add, we would add the hydrophones at half-wavelength spacing approaching an equispaced line array. On the other hand, if the number of hydrophones were limited but the baseline were not, we would keep the hydrophones at half-wavelength spacing and increase the distance between subarrays. By keeping the hydrophones at half-wavelength spacing in the subarrays ambiguities would be minimized.

Summary

This paper has presented a review of research on coherence estimation and time-delay estimation. References to much of the relevant work in these two fields are included.

Acknowledgment

As evident from the paper and list of references this work is a review of research by this author conducted alone and with the following colleagues: Dr. C. Knapp, Dr. K. Scarbrough, E. Scannell, Dr. P. Cable, Dr. A. Nuttall, R. Tremblay, and Dr. P. Abraham. The author has been honored by Prof. G. Tacconi, Prof. L. Bjørnø, Prof. Y. T. Chan, Prof. J. Plant, Dr. H. Urban, Prof. T. Durrani, and Prof. J. L. Lacoume to present this work at several meetings over the past decade in Italy, France, the United Kingdom, Denmark, West Germany, and Canada as well as in the United States.

References

[1] W. B. Adams, J. P. Kuhn, and W. P. Whyland, "Correlator compensation requirements for passive time delay estimation with moving source or receivers," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, no. 2, pp. 158–168, 1980.

[2] H. Akaike and Y. Yamanouchi, "On the statistical estimation of frequency response function," *Ann. Inst. Statist. Math.*, vol. 14, pp. 23–56, 1963.

[3] W. J. Bangs and P. M. Schultheiss, "Space time processing for optimal parameter estimation," in *Signal Processing*, J. W. R. Griffiths, P. L. Stocklin, and C. Van Schooneveld, Eds. New York: Academic, 1973, pp. 577–590.

[4] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures.* New York: Wiley, 1971.

[5] D. R. Brillinger, *Time Series Data Analysis and Theory.* New York: Holt, Rinehart, and Winston, 1975.

[6] J. S. Bendat and A. G. Piersol, *Engineering Applications of Correlation and Spectral Analysis.* New York: Wiley, 1980.

[7] V. A. Benignus, "Estimation of coherence spectrum of non-Gaussian time series populations," *IEEE Trans. Audio Electroacoust.*, vol. AU-17, pp. 198–201, Sept. 1969; vol. AU-18, p. 320, Sept. 1970.

[8] J. Betz, Ph.D. dissertation, Northeastern Univ., Boston, MA, 1984.

[9] L. Bjørnø, Ed., *Proc. 1980 NATO Advanced Study Institute on Underwater Acoustics and Signal Processing.* Boston, MA: Reidel, 1981.
[10] R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra.* New York: Dover, 1958.
[11] G. C. Carter, C. H. Knapp, and A. H. Nuttall, "Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 337–344, Aug. 1973.
[12] G. C. Carter, A. H. Nuttall, and P. G. Cable, "The smoothed coherence transform," *Proc. IEEE*, vol. 61, pp. 1497–1498, 1973.
[13] G. C. Carter and C. H. Knapp, "Coherence and its estimation via the partitioned modified chirp-z transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp. 257–264, June 1975.
[14] G. C. Carter, "Time delay estimation," Ph.D. dissertation, Univ. of Connecticut, Storr, CT, 1976.
[15] ——, "Variance bounds for passively locating an acoustic source with a symmetric line array," *J. Acoust. Soc. Amer.*, vol. 64, pp. 922–926, 1977.
[16] ——, "Receiver operating characteristics for a linearly threshold coherence estimation detector," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 90–94, Feb. 1977.
[17] ——, "A brief description of the fundamental difficulties of passive ranging," *IEEE J. Oceanic Eng.*, vol. OE-3, no. 3, pp. 65–66, July 1978.
[18] G. C. Carter and P. B. Abraham, "Estimation of source motion from time delay and time compression measurement," *J. Acoust. Soc. Amer.*, vol. 67, no. 3, pp. 830–832, 1980.
[19] G. C. Carter, "Time delay estimation for passive sonar signal processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, no. 3, pp. 463–470, June 1981.
[20] Y. T. Chan, "Time delay estimation in the presence of multipath propagation," in *Proc. 1984 NATO Advanced Study Institute.*
[21] O. G. Childers, Ed., *Modern Spectrum Analysis.* New York: IEEE PRESS, 1978.
[22] S. K. Chow and P. M. Schultheiss, "Delay estimation using narrow-band processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 478–484, June 1981.
[23] W. S. Cleveland and E. Parzen, "The estimation of coherence, frequency response, and envelope delay," *Technometrics*, vol. 17, no. 2, pp. 167–172, 1975.
[24] H. Cramér, *Mathematical Methods of Statistics.* Princeton, NJ: Princeton Univ. Press, 1946.
[25] C. Eckart, "Optimal rectifier systems for the detection of steady signals," Univ. California, Scripps Inst. Oceanography, Marine Physical Lab. Rep. SIO Ref. 52-11, 1952.
[26] L. D. Enochson and N. R. Goodman, "Gaussian approximations to the distribution of sample coherence," Tech. Rep. AFFDL-TR-65-57, June 1965 (AD520987).
[27] A. S. Gevins, C. L. Yeager, S. L. Diamond, J. P. Spire, G. M. Zeitlin, and A. H. Gevins, "Automated analysis of the electrical activity of the human brain (EEG): A progress report," *Proc. IEEE*, vol. 63, pp. 1382–1399, Oct. 1975.
[28] J. J. Gosselin, "Comparative study of two-sensor (magnitude-squared coherence) and single-sensor (square-law) receiver operating characteristics," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 311–314, 1977.
[29] J. W. R. Griffiths, P. L. Stocklin, and C. Van Schooneveld, Eds., *Signal Processing.* New York: Academic Press, 1973 (including E. B. Lunde, "Wavefront stability in the ocean," W. J. Bangs and P. M. Schultheiss, "Space-time processing for optimal parameter estimation").
[30] L. Griffiths, "Time varying filtering and spectrum estimation," in *Proc. 1984 NATO Advanced Study Institute.*
[31] W. R. Hahn and S. A. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Trans. Inform. Theory*, vol. IT-19, no. 5, pp. 608–614, 1973.
[32] W. R. Hahn, "Optimum signal processing for passive sonar range and bearing estimation," *J. Acoust. Soc. Amer.*, vol. 58, pp. 201–207, 1975.
[33] W. G. Halvorsen and J. S. Bendat, "Noise source identification using coherent output power spectra," *J. Sound Vibration*, pp. 15–24, Aug. 1975.
[34] E. J. Hannan and P. J. Thomson, "The estimation of coherence and group delay," *Biometrika*, vol. 58, pp. 469–481, 1971.
[35] ——, "Estimating group delay," *Biometrika*, vol. 60, pp. 241–253, 1973.
[36] ——, "Delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 485–490, June 1981.
[37] J. C. Hassab and R. E. Boucher, "A quantitative study of optimum and suboptimum filters in the generalized correlator," in *Proc. 1979 IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 124–127.
[38] ——, "Performance of the generalized cross correlator in the presence of a strong spectral peak in the signal," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 549–555, June 1981.
[39] R. A. Haubrich, "Earch noise 5 to 500 millicycles per second, 1. Spectral stationarity, normality and nonlinearity," *J. Geophys. Res.*, vol. 70, pp. 1415–1427, Mar. 1965.
[40] J. P. Ianniello, "Time delay estimation via cross-correlation in the presence of large estimation errors," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 998–1003, Dec. 1982.
[41] J. P. Ianniello, E. Weinstein, and A. Weiss, "Comparison of the Ziv-Zakai lower bound on time delay estimation with correlator performance," in *Proc. 1983 IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 875–878.
[42] G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications.* San Francisco, CA: Holden-Day, 1968.
[43] G. W. Johnson, D. E. Ohlms, and M. L. Hampton, "Broadband correlation processing," in *Proc. 1983 IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 583–586.
[44] S. B. Kesler, Ed., *Modern Spectrum Analysis II.* New York: IEEE PRESS, 1986.
[45] R. L. Kirlin and J. N. Bradley, "Delay estimation simulations and a normalized comparison of published results," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 508–511, June 1982.
[46] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, Aug. 1976.
[47] ——, "Estimation of time delay in the presence of source or receiver motion," *J. Acoust. Soc. Amer.*, vol. 61, pp. 1545–1549, 1977.
[48] L. H. Koopmans, *The Spectral Analysis of Time Series.* New York: Academic Press, 1974.
[49] J. P. Kuhn, "Detection performance of the smoothed coherence transform (SCOT)," in *Proc. 1978 IEEE Int. Conf. Acoust., Speech, Signal Processing.*
[50] W. C. Lindsey and H. Meyr, "Complete statistical description of the phase-error process generated by correlative tracking system," *IEEE Trans. Inform. Theory*, vol. IT-23, no. 2, pp. 194–202, 1977.
[51] J. F. Lynch, "On the use of focused horizontal arrays as mode separation and source location devices in Ocean acoustics, Part I: Theory," *J. Acoust. Soc. Amer.*, vol. 74, pp. 1406–1416, 1983.
[52] J. F. Lynch, D. K. Schwartz, and K. Sivaprashad, "On the use of focused horizontal arrays as mode separation and source location devices in ocean acoustics. Part II: Theoretical and Modeling results," *J. Acoust. Soc. Amer.*, vol. 78, pp. 575–586, 1985.
[53] V. H. MacDonald and P. M. Schultheiss, "Optimum passive bearing estimating in a spatially incoherent noise environment," *J. Acoust. Soc. Amer.*, vol. 46, pp. 37–43, 1969.
[54] S. L. Marple, Jr., "Spectrum analysis; overview of classical and high resolution spectral estimation," *Proc. 1984 NATO Advanced Study Institute.*
[55] H. Meyr, "Delay lock tracking of stochastic signals," *IEEE Trans. Commun.*, vol. COM-24, no. 3, pp. 331–339, 1976.
[56] A. H. Nuttall, "Theory and application of the separable class of random processes," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, RLE Rep. 343, May 1958.
[57] ——, "Spectral estimation by means of overlapped FFT processing of windowed data," Naval Underwater Systems Center, New London, CT, Rep. 4169, Oct. 1971.
[58] A. H. Nuttall and G. C. Carter, "Bias of the estimate of magnitude squared coherence," *IEEE Trans. Acoust., Speech, Sig-*

nal Processing, vol. ASSP-24, pp. 582–583, Dec. 1976.
[59] ——, "A generalized framework for power spectral estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, no. 3, pp. 334–335, June 1980.
[60] ——, "Spectral estimation using combined time and lag weighting," *Proc. IEEE*, vol. 70, no. 9, pp. 1115–1125, Sept. 1982.
[61] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.
[62] N. Owsley, "Overview of adaptive array processing techniques," in *Proc. 1984 NATO Advanced Study Institute*.
[63] B. Picinbono, "Adaptive, robust and non-parametric methods in signal detection," in *Proc. 1984 NATO Advanced Study Institute*.
[64] A. H. Quazi, "An overview on the time delay estimate in active and passive systems for target localization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 527–533, June 1981.
[65] E. H. Scannell, Jr., and G. C. Carter, "Confidence bounds for magnitude-squared coherence estimates," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1978.
[66] K. Scarbrough, N. Ahmed, and G. C. Carter, "On the simulation of a class of time delay estimation algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 534–540, June 1981.
[67] K. Scarbrough, R. J. Tremblay, and G. C. Carter, "Performance predictions for coherent and incoherent processing techniques of time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, no. 5, pp. 1191–1196, Oct. 1983.
[68] K. Scarbrough, Ph.D. dissertation, Kansas State Univ., Topeka, KS, 1984.
[69] P. M. Schultheiss and E. Weinstein, "Source tracking using passive array data," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 600–607, June 1981.
[70] G. Tacconi, Ed., *Aspects of Signal Processing*. Boston, MA: Reidel, 1977.
[71] H. L. Van Trees, *Detection, Estimation, and Modulation Theory Part I*. New York: Wiley, 1968.
[72] A. Weiss and E. Weinstein, "Composite bound on the attainable mean-square error in passive time-delay estimation from ambiguity prone signals," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 977–979, Nov. 1982.
[73] ——, "Fundamental limitations in passive time delay estimation—Part I: Narrow-band systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 472–485, Apr. 1983.
[74] P. D. Welch, "The use of fast Fourier transform for estimation of power spectra: A method based on time averaging over short modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, pp. 70–73, 1967.

G. Clifford Carter (Senior Member, IEEE) was born in Oak Park, IL, in the 1940s. He was graduated from the U.S. Coast Guard Academy and served as a deck watch officer and antisubmarine warfare (ASW) training officer. He received the M.S.E.E. and Ph.D.E.E. degrees from the University of Connecticut, Storrs, CT.

Following Coast Guard Service, he joined the Naval Underwater Systems Center (NUSC), then the Underwater Sound Laboratory, New London, CT. At the NUSC, he has successfully run a multiyear multimillion dollar state-of-the-art sonar signal processing R&D program, formulating, executing, and bringing the program to completion on time and within cost despite unusual difficulties. He currently leads a thirty person branch responsible for technical leadership in all aspects of the U.S. Navy's surface ship sonar signal processing. During more than five years in his position, he has significantly increased the size and improved the quality and influence of NUSC's surface ship sonar signal processing activities. He is the recipient of numerous cash awards and pay raises under the government's merit pay system for his technical contributions to the U.S. Navy and for his management accomplishments. He has lectured on technical aspects of sonar signal processing throughout the United States, Canada, and Europe, and is the author or coauthor of numerous journal articles and reports.

Dr. Carter has served in a variety of IEEE positions, including Associate Editor and officer in the Acoustics, Speech, and Signal Processing Society.