[3] B. Picinbono, "Spherically invariant and compound Gaussian stochastic processes," *IEEE Trans. Inform. Theory* (Corresp.), vol. IT-16, pp. 77–79, Jan. 1970.
[4] S. Bochner, *Lectures on Fourier Integrals*. Princeton, N.J.: Princeton Univ. Press, 1959.
[5] E. Lukacs, *Characteristic Functions*. New York: Hafner, 1960.
[6] I. J. Schoenberg, "Metric spaces and completely monotone functions," *Ann. Math.*, vol. 39, pp. 811–841, 1938.
[7] J. L. Doob, *Stochastic Processes*. New York: Wiley, 1953.
[8] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*. New York: Academic, 1965.
[9] P. R. Halmos, *Measure Theory*. New York: Van Nostrand Reinhold, 1950.
[10] T. Apostel, *Mathematical Analysis*. Reading, Mass.: Addison-Wesley, 1957.
[11] H. Cramer, *Mathematical Methods of Statistics*. Princeton, N.J.: Princeton Univ. Press, 1946.
[12] K. S. Miller, *Multidimensional Gaussian Distributions*. New York: Wiley, 1964.
[13] D. K. McGraw and J. F. Wagner, "Elliptically symmetric distributions," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 110–120, Jan. 1968.
[14] P. W. Cooper, "The hyperplane in pattern recognition," *Cybernetica* (Brussels), vol. 5, pp. 215–238, 1962.
[15] ——, "The hypersphere in pattern recognition," *Inform. Contr.*, vol. 5, pp. 324–346, Dec. 1962.
[16] H. M. Hall, "A new model for 'impulsive' phenomena: application to atmospheric-noise communication channels," Stanford Electronics Lab., Stanford, Calif., Tech. Rep. SU-SEL-66-052, 1966.
[17] B. Picinbono and G. Vezzosi, "Detection d'un signal certain dan un bruit non stationaire et non Gaussien," *Ann. Telecommun.*, vol. 25, pp. 433–439, 1970.
[18] W. Feller, *An Introduction to Probability Theory and Its Applications*. New York: Wiley, vol. 2, 1966.
[19] P. Medgyessy, *Decomposition of Superpositions of Distribution Functions*. Budapest, Hungary: Hungarian Academy of Sciences, 1961.
[20] T. Y. Young and G. Coraluppi, "Stochastic estimation of a mixture of normal density functions using an information criterion," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 258–263, May 1970.
[21] E. Parzen, *Stochastic Processes*. San Francisco: Holden-Day, 1962.
[22] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1965.
[23] P. Hartman and A. Wintner, "On the spherical approach to the normal distribution law," *Amer. J. Math.*, vol. 62, pp. 95–110, 1972.
[24] G. Vezzosi and B. Picinbono, "Detection du'n signal certain dans un bruit spheriquement invariant, structure et caracteriques des receptions," *Ann. Telecommun.*, vol. 27, pp. 95–110, 1972.
[25] J. Goldman, "Moments of the sum of circularly symmetric random variables," *IEEE Trans. Inform. Theory* (Corresp.), vol. IT-18, pp. 297–298, Mar. 1972.
[26] J. F. C. Kingman, "On random sequences with spherical symmetry," *Biometrika*, vol. 59, pp. 492–494, 1972.
[27] D. Kelker, "Distribution theory of spherical distributions and a location-scale parameter generalization," *Sankhya: Ind. J. Stat.*, vol. 32, pp. 419–430, 1970.

Optimum Processing for Delay-Vector Estimation in Passive Signal Arrays

WILLIAM R. HAHN AND STEVEN A. TRETTER

Abstract—**For the purpose of localizing a distant noisy target, or, conversely, calibrating a receiving array, the time delays defined by the propagation across the array of the target-generated signal wavefronts are estimated in the presence of sensor-to-sensor-independent array self-noise. The Cramér–Rao matrix bound for the vector delay estimate is derived, and used to show that either properly filtered beamformers or properly filtered systems of multiplier-correlators can be used to provide efficient estimates. The effect of suboptimally filtering the array outputs is discussed.**

I. Introduction

In many physical problems of interest, with sonar, radar, and seismology as examples, the time records of the outputs of an array of sensors are observed over some time interval and used to estimate the position of a distant noise source. Conversely, the position of the distant noise source may be known, and the intent may be to estimate the positions of the sensors comprising the array. Typically, the sensor outputs are amplitude-scaled and delayed replicas of the waveform from the distant noise source, corrupted by additive noises which are usually local in origin. If the amplitude gradient across the array of the waveforms from the distant source is negligible, essentially all of the geometric information is encoded in the set of delays associated with the propagation across the array of the wavefronts from the distant source. This paper discusses the theoretical bounds on the precision with which the set of delays can be measured, and shows that either properly filtered beamformers or properly filtered systems of correlators can be used to obtain estimates that achieve the theoretical bound.

The theoretical bound discussed is the Cramér–Rao matrix bound (CRMB), which is the appropriate bound to use when large numbers of samples, or equivalently, long observation times, are used [1], [2]. For the purposes of this paper it is more convenient to use its inverse, the Fisher information matrix (FIM), and to compare the inverses of the measurement error covariance matrices for the beamformer and multiple-correlator delay measurement schemes to the FIM.

Manuscript received December 20, 1972; revised March 17, 1973.
W. R. Hahn is with the Naval Ordnance Laboratory, Silver Spring, Md. 20910.
S. A. Tretter is with the University of Maryland, College Park, Md. 20740.

II. The Fisher Information Matrix

Assume that the signal wavefronts from a distant noise source propagate across an M-element array of sensors and that the signal amplitude gradient across the array is negligible. The signal at the ith sensor is $s(t - d_i)$, where $s(t)$ is the signal at a reference point near the array, and d_i is the delay at the ith sensor. Without loss of generality, the reference point is assumed to be at the location of the first element in the array, so that $d_1 = 0$. The output of the ith sensor is

$$x_i(t) = s(t - d_i) + n_i(t) \tag{1}$$

where $n_i(t)$ is the additive sensor noise. The M sensors are observed for T s, $-T/2 \leq t \leq T/2$, and the M time records are represented by Fourier coefficients

$$X_i(\omega_k) = \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} x_i(t) \exp\{-jk\omega_0 t\}\, dt \tag{2}$$

where $\omega_0 = 2\pi/T$, and $\omega_k = k\omega_0$. The following assumptions will determine the joint density function for the random Fourier coefficients.

a) The random signal and each of the M additive sensor noises are all stationary zero-mean Gaussian random processes.

b) All of the random processes are independent.

c) T is large compared to the correlation times of the random processes and also to the time needed for the signal wavefronts to traverse the array.

Define a vector X containing the Fourier coefficients as elements, so that if $S(\omega)$ and $N_i(\omega)$ are the signal and noise power spectra at the ith sensor, and if only Fourier coefficients up to frequency $N\omega_0$ are to be processed, the density for X can be written as

$$p(X) = \left[\pi^{MN} \prod_{k=1}^{N} \det R(k)\right]^{-1} \cdot \exp\left\{-\sum_{k=1}^{N} X^T(k) R^{-1}(k) X^*(k)\right\} \tag{3}$$

where

$$\begin{aligned} X(k) &= [X_1(\omega_k), X_2(\omega_k), \cdots, X_M(\omega_k)]^T \\ X &= [X^T(1) \colon X^T(2) \colon \cdots \colon X^T(N)]^T \\ V(k) &= [1, \exp(j\omega_k d_2, \cdots, \exp(j\omega_k d_M)]^T \\ N(k) &= \operatorname{diag}[N_1(\omega_k), N_2(\omega_k), \cdots, N_M(\omega_k)] \\ R(k) &= N(k) + S(\omega_k) V^*(k) V^T(k). \end{aligned} \tag{4}$$

Superscripts * and T are used to denote complex conjugation and transposition, respectively, and diag $(N_1, N_2, \cdots, N_M)$ is a diagonal matrix whose ith diagonal element is N_i.

To avoid cumbersome notation, the frequency arguments of the functions discussed will generally be suppressed, $\sum_{B^+}$ will indicate a sum over the positive Fourier frequencies being considered, $\prod_{B^+}$ will indicate the corresponding product, and $\sum_i$ will be used to denote the array sum $\sum_{i=1}^{M}$. When ω appears following a sum $\sum_{B^+}$, it will be understood to stand for $k\omega_0$.

Since $\det R(k) = (1 + \sum_i S/N_i) \det N(k)$, only the exponential part of the density function will depend on the delays. Let the signal delay vector D be defined as

$$D = (d_2, d_3, \cdots, d_M)^T. \tag{5}$$

The likelihood function for D is

$$L(D) = \left[\pi^{MN} \prod_{B^+} \det R\right]^{-1} \exp\left\{-\sum_{B^+} X^T R^{-1} X^*\right\}. \tag{6}$$

The CRMB for unbiased estimators of the vector argument of the likelihood function is the inverse of the FIM, denoted by (FIM), where

$$(\text{FIM}) = -\langle \text{grad}\,(\text{grad} \ln L(D))^T \rangle. \tag{7}$$

$\langle \cdot \rangle$ is the expectation operator, grad f is the row vector which is the gradient of the scalar f, and the gradient of a vector is a matrix whose ith row is the gradient of the ith component of the vector. The gradients in (7) are taken with respect to the components of the vector D.

If G is defined by

$$G = \frac{S}{1 + \sum_i S/N_i} \tag{8}$$

the inverse of the matrix R is

$$R^{-1} = N^{-1} - G N^{-1} V^* V^T N^{-1}. \tag{9}$$

Provided that the elements of D depend only on parameters a and b, the typical element of the FIM has the form

$$\begin{aligned} &-\left\langle \frac{\partial}{\partial a} \frac{\partial}{\partial b} \ln L(D) \right\rangle \\ &= -\sum_{B^+} G \left\langle X^T N^{-1} \frac{\partial}{\partial a} \left(\frac{\partial}{\partial b} V^* V^T\right) N^{-1} X^* \right\rangle \\ &= \sum_{B^+} \omega^2 G \sum_k \sum_m \frac{S}{N_k N_m} \frac{\partial(d_k - d_m)}{\partial a} \frac{\partial(d_k - d_m)}{\partial b}. \end{aligned} \tag{10}$$

From (10) it follows that the FIM pertinent to the estimation of the vector D is

$$(\text{FIM}) = \sum_{B^+} 2\omega^2 \frac{S^2}{1 + \sum_i S/N_i} \cdot [(\operatorname{tr} N^{-1}) N_p^{-1} - N_p^{-1} \mathbf{1}\, \mathbf{1}^T N_p^{-1}]. \tag{11}$$

In (11) N_p^{-1} is the submatrix formed by removing the first row and column of N^{-1}, and $\mathbf{1}$ is a column vector of ones. Because of the assumed smoothness of all of the spectra relative to the frequency increment $\omega_0 = 2\pi/T$, the FIM can also be written as

$$(\text{FIM}) = \frac{T}{2\pi} \int_B \omega^2 \frac{S^2}{1 + \sum_i S/N_i} \cdot [(\operatorname{tr} N^{-1}) N_p^{-1} - N_p^{-1} \mathbf{1}\, \mathbf{1}^T N_p^{-1}]\, d\omega \tag{12}$$

where B is the two-sided frequency interval $B = \{\omega \mid -\omega_N \leq \omega \leq \omega_N\}$.

III. The Maximum-Likelihood Estimate

It is well known that when the maximum-likelihood estimate (MLE) is based on a large number of independent samples, it is consistent, asymptotically normal, and asymptotically efficient [3]. Since the observation time T is large compared to the process correlation times, there should be, in some sense, a large number of independent samples. The covariance matrix for the error in the MLE for D should be the CRMB, at least to first order.

The results that follow are independent of the true delay vector, and the equations for the likelihood function and MLE are considerably simplified by assuming the true delay vector to be the zero vector, $\mathbf{0}$. The vector D of this section is the MLE and is, therefore, the measurement error. The steering vector corresponding to the error D is

$$V^T = (1, \exp\{j\omega d_2\}, \cdots, \exp\{j\omega d_M\}). \tag{13}$$

The MLE vectors D and V satisfy

$$\begin{aligned} 0 &= \operatorname{grad} \ln L(D) \\ &= \operatorname{grad} \sum_{B+} G \sum_i \sum_n \frac{X_i X_n^*}{N_i N_n} \exp\{j\omega(d_n - d_i)\} \\ &= \operatorname{grad}\,(A + BD + \tfrac{1}{2} D^T C D + \cdots) \end{aligned} \tag{14}$$

where by expanding $\exp\{j\omega(d_n - d_i)\}$ as a power series, the vector B is seen to be

$$B = \sum_{B+} j\omega G \mathbf{1}^T N^{-1}[X X_p^{*T} - X^* X_p^T] N_p^{-1} \tag{15}$$

while the matrix C is determined by

$$\frac{1}{2} D^T C D = \sum_{B+} \frac{1}{2} (j\omega)^2 G \sum_i \sum_k \frac{X_i X_k^*}{N_i N_k} (d_k - d_i)^2. \tag{16}$$

In (15), X_p is that part of the single-frequency data vector left after the first element is removed. The terms $X_i X_k^*$ in (16) are elements of the sample covariance matrix at a single frequency based on T s of data. These sample covariance elements do not converge, even if T is arbitrarily long [4]. However, since T is large compared to the process correlation times, the spectra are smooth enough so that the sample covariance can be averaged with samples from nearby frequencies to provide statistical convergence. The $\sum_{B+}$ summation in (16) provides such an averaging of the $X_i X_k^*$. Thus it is assumed that $X_i X_k^*$ can be replaced by $R_{ik} = \langle X_i X_k^* \rangle$ in (16), from which it follows that $C = \langle C \rangle$, or

$$\begin{aligned} C &= -\sum_{B+} 2\omega^2 \frac{S^2}{1 + \sum_i S/N_i} \\ &\quad \cdot [(\operatorname{tr} N^{-1}) N_p^{-1} - N_p^{-1} \mathbf{1}\,\mathbf{1}^T N_p^{-1}] \\ &= -(\mathrm{FIM}). \end{aligned} \tag{17}$$

From (15), it follows immediately that $\langle B \rangle = 0$, and not so immediately that

$$\begin{aligned} \langle B^T B^* \rangle &= \sum_{B+} 2\omega^2 \frac{S^2}{1 + \sum_i S/N_i} \\ &\quad \cdot [(\operatorname{tr} N^{-1}) N_p^{-1} - N_p^{-1} \mathbf{1}\,\mathbf{1}^T N_p^{-1}] \\ &= (\mathrm{FIM}). \end{aligned} \tag{18}$$

Neglecting the higher order terms, and assuming $C = \langle C \rangle$, the vector D is given by

$$D = -\langle C \rangle^{-1} B^T \tag{19}$$

so that

$$\langle D \rangle = \mathbf{0} \tag{20}$$

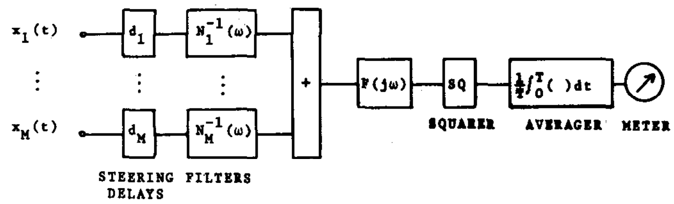

Fig. 1. Beamformer implementation of MLE processor.

and

$$\begin{aligned} \langle D^* D^T \rangle &= \langle C \rangle^{-1} \langle B^T B^* \rangle^* \langle C \rangle^{-1} \\ &= (\mathrm{FIM})^{-1} \\ &= (\mathrm{CRMB}) \end{aligned} \tag{21}$$

the CRMB.

Thus in the limit of large T the MLE is unbiased and achieves the CRMB. The MLE processor can be readily implemented as indicated in Fig. 1. The MLE processor is just a steered and filtered beamformer followed by a square-law averager. The individual inputs are each steered and then filtered with filters whose frequency response is the inverse of the additive noise spectrum for that particular sensor. The beam sum is then formed and fed to a filter whose squared magnitude response is $|F(j\omega)|^2 = G(\omega)$. The MLE is determined as that set of steering delays that gives the maximum deflection of the output meter.

If the MLE processor of Fig. 1 is to provide unbiased estimates of the signal delays it is necessary that the signal components at any specific frequency be coherently processed. Thus the beamformer input filters, specified to be scalar functions of frequency, can have arbitrary phase responses, provided that all phase responses are identical. With conventional filters this means that except for scalar gain constants the filters must be identical. Thus the MLE processor of Fig. 1 is practical only if all the additive noise spectra have substantially the same shape. If not, then an estimator based on processing the Fourier coefficients, as in a digital computer, can be devised.

IV. Correlator Delay Measurement Systems

The following scheme can be used to estimate the unknown delay vector D. Let a system of correlators be used to form all the $M(M-1)/2$ cross correlations corresponding to processing all of the M input waveforms taken two at a time. The individual correlators are assumed to have input filters for each channel, and the position of the correlogram peak is used as a signal delay estimate for that sensor pair. If each correlator is to provide an unbiased estimate of the corresponding signal delay, the two input filters for each correlator must have the same phase response, and hence can be taken to be identical filters. A typical correlator is shown in Fig. 2. The steering delay is adjusted to give the maximum deflection of the meter, and this defines the delay estimate for the correlator.

Let d_{ij} be the correlator estimate for the signal delay from the ith to the jth sensor, based on the correlation of the $x_i(t)$ and $x_j(t)$ time records. Let e_{ij} and F_{ij} be, respectively, the error in the estimate d_{ij}, and the filter used on the inputs to the correlator. Let δ_{ij} be the Kronecker delta, and define

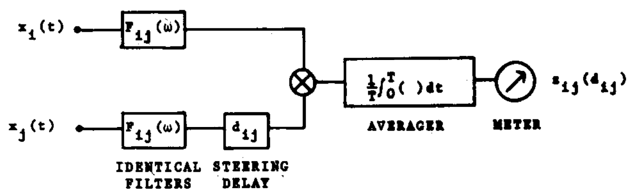

Fig. 2. Typical multiplier-correlator.

the following scalars, vectors, and matrices:

$$\begin{aligned}
G(ij;kl) &= [(S + N_i\delta_{ik})(S + N_j\delta_{jl}) \\
&\quad - (S + N_i\delta_{il})(S + N_j\delta_{jk})] \\
D_C &= (d_{12}, d_{13}, \cdots, d_{1M}, d_{23}, \cdots, d_{(M-1)M})^T \\
E &= (e_{12}, e_{13}, \cdots, e_{(M-1)M})^T \\
F &= \operatorname{diag}(|F_{12}|^2, |F_{13}|^2, \cdots, |F_{(M-1)M}|^2) \\
K &= \int_B \omega^2 SF\, d\omega \\
&= \operatorname{diag}(K_{12}, K_{13}, \cdots, K_{(M-1)M}) \\
G &= [G(ij;kl)].
\end{aligned} \tag{22}$$

The square matrix G has for its elements the scalars $G(ij; kl)$ positioned according to the scheme determined by the order of the subscripts in EE^T, where ij is the row designation, and kl the column designation.

Using (22), the covariance matrix for the correlator scheme measurement error vector can be compactly written [2]

$$\begin{aligned}
P_E &= \langle EE^T \rangle \\
&= \frac{2\pi}{T} K^{-1} \left(\int_B \omega^2 FGF\, d\omega \right) K^{-1}.
\end{aligned} \tag{23}$$

The correlator delay measurement vector is related to the vector to be estimated D by the equation

$$D_C = AD + E \tag{24}$$

where the matrix A, with its rows and columns labeled with the same sets of ordered subscripts used for the elements of D_C (for the rows) and D (for the columns), has for its element in the ij,k position

$$A(ij; k) = \delta_{jk} - \delta_{ik}. \tag{25}$$

Since $\langle E \rangle = \mathbf{0}$, the Gauss–Markov estimate [5] for the vector D based on the correlator measurements is

$$\hat{D} = [A^T P_E^{-1} A]^{-1} A^T D_C \tag{26}$$

and the covariance matrix for the Gauss–Markov estimate is

$$\langle (\hat{D} - D)(\hat{D} - D)^T \rangle = [A^T P_E^{-1} A]^{-1}. \tag{27}$$

From the definitions in (22), and with the filters defined by

$$|F_{ij}|^2 = \frac{S/N_i N_j}{1 + \sum_k S/N_k} \tag{28}$$

the FIM of (11) can be written as

$$(\mathrm{FIM}) = \frac{T}{2\pi} A^T K A \tag{29}$$

which suggests investigating the possibility that

$$A^T P_E^{-1} A \overset{?}{=} \frac{T}{2\pi} A^T K A. \tag{30}$$

With the filters given by (28), the matrix FGF can be written as

$$\begin{aligned}
FGF = {} & \frac{S^2}{1 + \sum_i S/N_i} \operatorname{diag}\left(\frac{1}{N_1 N_2}, \frac{1}{N_1 N_3}, \cdots, \frac{1}{N_{M-1} N_M} \right) \\
& - \sum_{1 \le \alpha < \beta < \gamma \le M} \frac{S^3/N_\alpha N_\beta N_\gamma}{\left(1 + \sum_i S/N_i\right)^2} U_{\alpha\beta\gamma} U_{\alpha\beta\gamma}^T
\end{aligned} \tag{31}$$

where $U_{\alpha\beta\gamma}$ is a column vector whose rows are labeled with the same scheme as is used for the elements of D_C, and whose element in the ij row is

$$U_{\alpha\beta\gamma}(ij) = \delta_{\alpha i}\delta_{\beta j} - \delta_{\alpha i}\delta_{\gamma j} + \delta_{\beta i}\delta_{\gamma j}. \tag{32}$$

Then the matrix

$$\int_B \omega^2 FGF\, d\omega = K - \sum_{1 \le \alpha < \beta < \gamma \le M} H_{\alpha\beta\gamma} U_{\alpha\beta\gamma} U_{\alpha\beta\gamma}^T \tag{33}$$

where

$$H_{\alpha\beta\gamma} = \int_B \omega^2 \frac{S^3/N_\alpha N_\beta N_\gamma}{\left(1 + \sum_i S/N_i\right)^2} d\omega. \tag{34}$$

Recursively applying one of Woodbury's modified matrix inversion formulas [6] to the right side of (33), the inverse can be written as

$$\left(\int_B \omega^2 FGF\, d\omega \right)^{-1} = K^{-1} + \sum_{1 \le \alpha < \beta < \gamma \le M} K^{-1} U_{\alpha\beta\gamma} M_{\alpha\beta\gamma} \tag{35}$$

where for the purposes of this paper it is not necessary to specify further the matrix $M_{\alpha\beta\gamma}$. From the relations defining the matrix A and the vector $U_{\alpha\beta\gamma}$, it follows that

$$A^T U_{\alpha\beta\gamma} = \mathbf{0} \tag{36}$$

where $\mathbf{0}$ is a matrix of zeros.

Thus the inverse of the covariance matrix for the Gauss–Markov estimate for D, with the correlator inputs filtered according to (28), is

$$\begin{aligned}
A^T P_E^{-1} A &= \frac{T}{2\pi} A^T K \left(\int_B \omega^2 FGF\, d\omega \right)^{-1} KA \\
&= \frac{T}{2\pi} A^T K A \\
&= (\mathrm{FIM}).
\end{aligned} \tag{37}$$

The correlator system optimally filtered according to (28) provides an efficient estimate.

V. Suboptimally Filtered Correlator Systems

For diverse reasons the decision may be made not to use the optimal filters of (28) at the input to a correlator delay measurement. It is then relevant to investigate the degree to which the delay estimate is degraded. The question can be answered in a simple way under the following hypotheses.

a) The ratio S/N_i is the same at each sensor.

b) Identical filters F are used on each input channel.

c) The suboptimally filtered Gauss–Markov delay estimate is used.

Under these hypotheses, the matrix FGF becomes

$$FGF = |F|^4(N^2 + MNS)I - \sum_{1 \le \alpha < \beta < \gamma \le M} |F|^4 SNU_{\alpha\beta\gamma}U_{\alpha\beta\gamma}^T \tag{38}$$

where I is the identity matrix and $U_{\alpha\beta\gamma}$ is defined in the same way as in the preceding section. If the Gauss–Markov estimate is formed from the suboptimally filtered delay estimates, the inverse of the covariance matrix for the Gauss–Markov estimate is

$$\begin{aligned} A^T P_E^{-1} A &= \frac{T}{2\pi} A^T K \left(\int_B \omega^2 FGF \, d\omega \right)^{-1} KA \\ &= \frac{T}{2\pi} \left(\int_B \omega^2 |F|^4 (N^2 + MNS) \, d\omega \right)^{-1} A^T A \\ &= \frac{T}{2\pi} \frac{(\int_B \omega^2 S |F|^2 \, d\omega)^2}{\int_B \omega^2 |F|^4 (N^2 + MNS) \, d\omega} A^T A. \end{aligned} \tag{39}$$

The FIM for this case is

$$(\mathrm{FIM}) = \frac{T}{2\pi} \left(\int_B \omega^2 \frac{S^2/N^2}{1 + M(S/N)} \, d\omega \right) A^T A. \tag{40}$$

Thus the covariance matrices for the optimally and suboptimally filtered estimates differ by a constant factor, and it is easy to take account of the effects of suboptimally filtering the inputs. Equations (39) and (40) can also be used to determine the degradation of the delay estimate due to an imprecise knowledge of either $S(\omega)$ or $N(\omega)$.

VI. Suboptimal Filtering, $M = 2$

The degradation in the correlator estimator performance due to suboptimal filtering is studied here for $M = 2$. It is assumed that both noises have the same spectrum. The signal and noise spectra $S(\omega)$ and $N(\omega)$, respectively, are both taken to be band limited with constant slopes of 0, -3, or -6 dB/octave.

$$S(\omega) = \begin{cases} 2A(\omega/\omega_1)^{-a}, & 1 \le \omega/\omega_1 \le 1 + W \\ 0, & \text{elsewhere.} \end{cases} \tag{41}$$

$$N(\omega) = \begin{cases} 2B(\omega/\omega_1)^{-b}, & 1 \le \omega/\omega_1 \le 1 + W \\ 0, & \text{elsewhere.} \end{cases} \tag{42}$$

In (41) and (42), $S(\omega)$ and $N(\omega)$ are the one-sided spectra, and W is a bandwidth variable. For $W = 1$, 3, 7, and 15, the bandwidths of the spectra are 1, 2, 3, and 4 octaves, respectively. The spectra slopes are determined by a and b, each of which will be equal to 0, 1, or 2. The filters will be defined by

$$|F_{\mathrm{NF}}|^2 = 1 \tag{43}$$

$$|F_{\mathrm{ECK}}|^2 = S/N^2 \tag{44}$$

$$|F_{\mathrm{WH}}|^2 = 1/N \tag{45}$$

and

$$|F_{\mathrm{OPT}}|^2 = \frac{S/N^2}{1 + 2(S/N)}. \tag{46}$$

F_{ECK} is conventionally called the Eckart filter. It optimizes the cross correlator for detection [7]. F_{WH} is the filter that whitens the input noise. F_{OPT} is the optimal filter for delay estimation. Notice that for large signal-to-noise ratios (SNR) the optimal filter essentially whitens the noise. (The system performance is unaffected by filter gain constants.) For small SNR the optimal filter is essentially an Eckart filter.

For the assumed spectra and filters, the integrations defined by (22) and (23) are readily performed, with the results involving, at most, elementary functions. For economy of presentation the results are not given in equation form, but rather only as curves, with processed bandwidth, signal-to-noise power, and filter choice as parameters.

The system dependence on the input SNR is studied by plotting curves for five SNR determined by

$$\left(\frac{A}{B}\right) = 2^{-k}, \qquad k = 0,1,2,3,4.$$

The ordinates of the curves will be the measurement standard deviations for the optimally filtered case, or the degradation in dB when suboptimal filters are used. The variable W is used for the abscissa. W determines the processed bandwidth.

Whelchel [8] used the spectra of (41) and (42) in an analysis of the effects of suboptimal filtering on correlator performance. Whelchel was primarily interested in signal detection, and used the output SNR of the correlator as a measure of detection capability. The study of suboptimally filtered correlator estimators in this section in part parallels Whelchel's study of suboptimally filtered correlator detectors.

Denote the measurement variances by D_{NF}^2, D_{ECK}^2, D_{WH}^2, and D_{OPT}^2, to correspond to the filters F_{NF}, F_{ECK}, F_{WH}, and F_{OPT}.

Let $d = a - b$. When the spectra of (41) and (42) are used with the filters defined by (43)–(46), and the estimation variance is calculated using (22) and (23), the result in all cases except one depends on d, and not specifically on a and b. Only when the filters F_{WH} are used does the corresponding variance depend specifically on both a and b.

For $d = 0$ and $d = 1$, $D_{\mathrm{OPT}}(W)$ can be made arbitrarily small by letting W be sufficiently large. This is not true for $d = 2$. Fig. 3 shows how the standard deviations of the optimally filtered delay estimates depend on d, W, and A/B. The curves are normalized by the factor $\sqrt{\omega_1^3 T}$, so that $\sqrt{\omega_1^3 T}\, D_{\mathrm{OPT}}(W)$ is dimensionless. For example, if ω_1 and T are chosen so that $\sqrt{\omega_1^3 T} = 10^3\ \mathrm{s}^{-1}$, then the ordinates in Fig. 3 read directly in milliseconds. The curves in Fig. 3 for $d = 0$ also apply to D_{WH} and D_{ECK}, since for $d = 0$, S/N is simply a constant. The $d = 0$ curves in Fig. 3 also apply to D_{NF} when the noise and signal spectra are both flat ($a = b = 0$).

The asymptotic nature of the curves for small values of W may not truly represent the behavior of the system

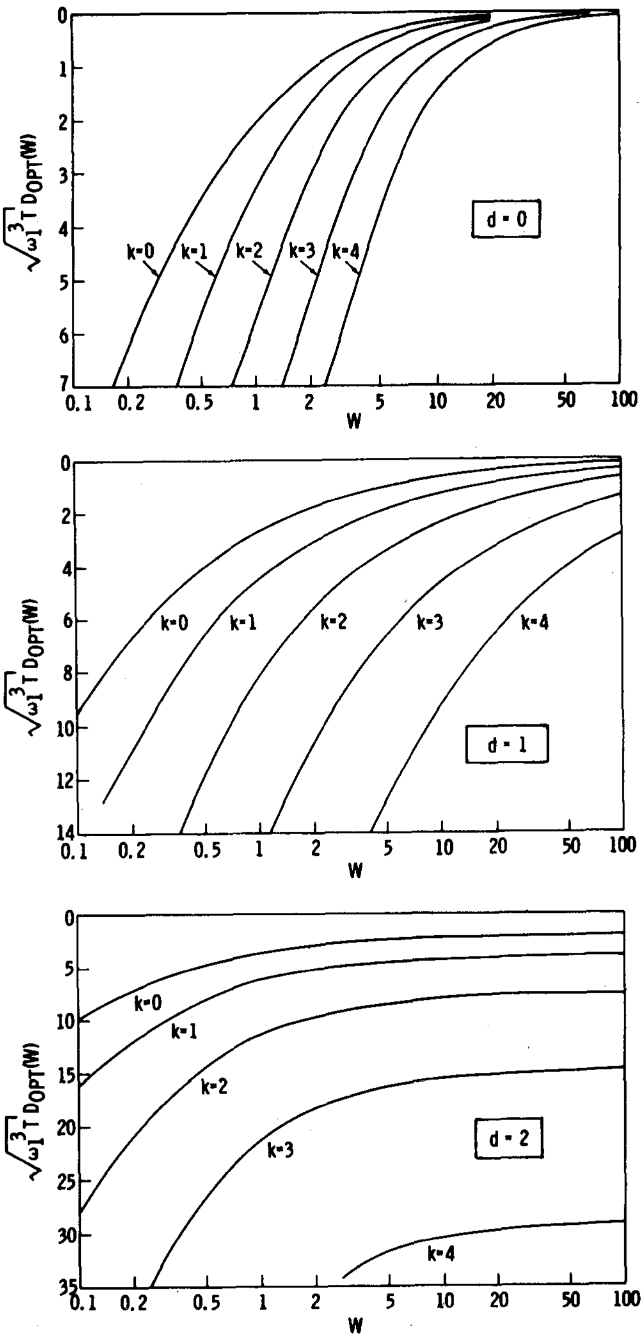

Fig. 3. Normalized standard deviation, optimum.

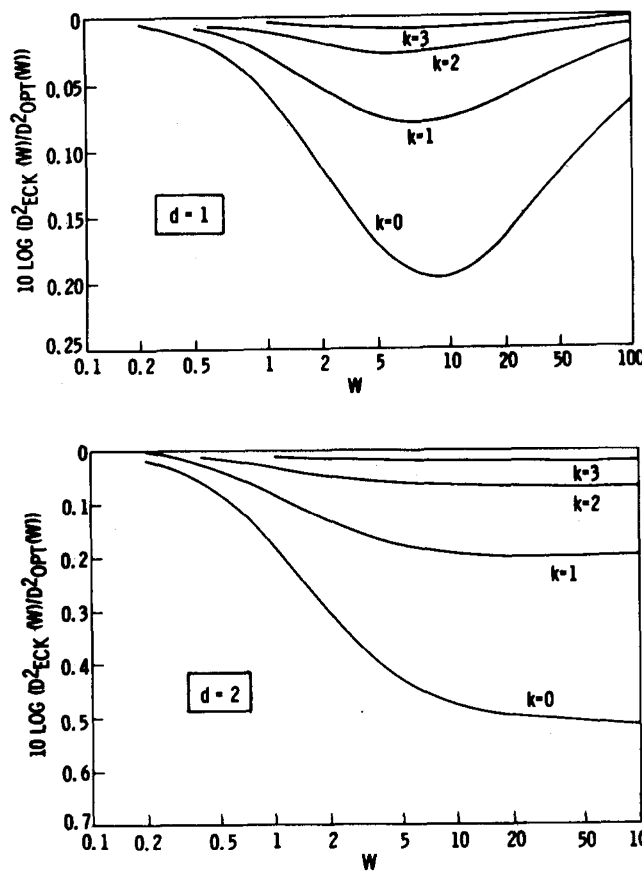

Fig. 4. System degradation using $|F_{\mathrm{ECK}}|^2$.

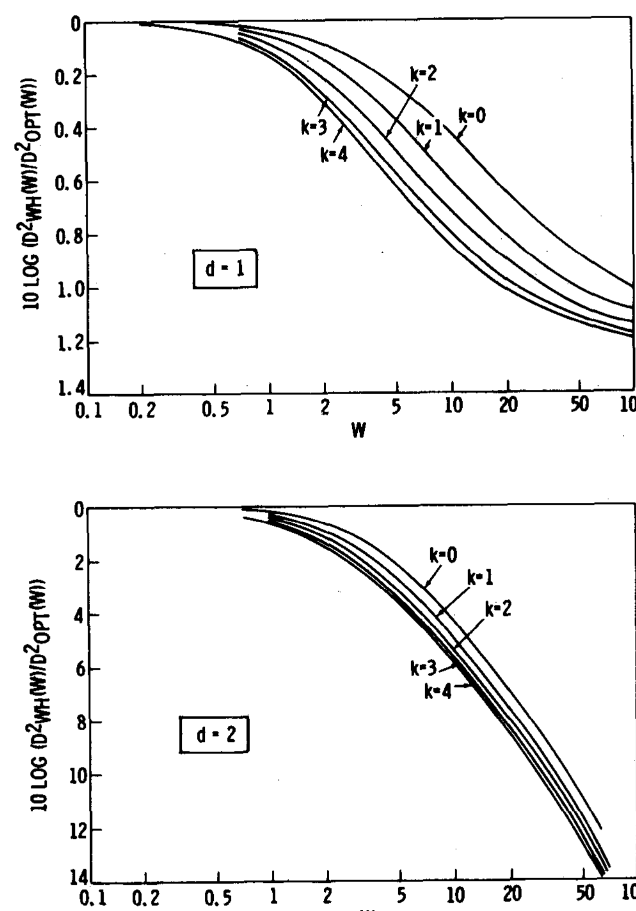

Fig. 5. System degradation using $|F_{\mathrm{WH}}|^2$.

measurement error. This is because the derivation leading to (23) assumes a sufficiently large time-bandwidth product to yield measurements with small errors.

Figs. 4 and 5 show the processor performance (in dB, relative to the optimal) when suboptimum filters D_{WH} and D_{ECK} are used. In these curves both the optimum and the suboptimum systems process the same band of frequencies, determined by the argument W. In Figs. 4 and 5 the system degradation is given by

$$\text{dB loss} = 10 \log_{10} (D_{\mathrm{ECK}}^2 (W)/D_{\mathrm{OPT}}^2(W)) \tag{48}$$

for $d = 1$ and $d = 2$. As indicated by Fig. 4, for the assumed spectra, the Eckart-filtered system sustains a modest processing loss relative to the optimum of at most about

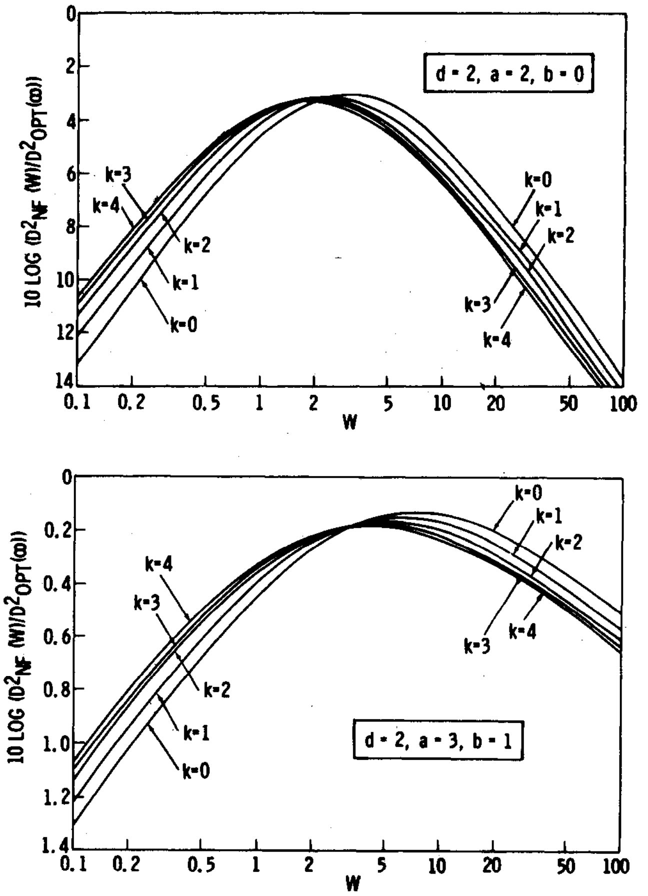

Fig. 6. System degradation using $|F_{NF}|^2$.

0.5 dB. The processing loss decreases as the input SNR decreases, as expected. In the $d = 1$ case it perhaps appears that the Eckart-filtered estimate is using the signal energy at higher frequencies more efficiently than the optimum estimate. This is not so. It is only that the Eckart-filtered estimate achieves a greater percentage reduction in its larger variance by processing a larger bandwidth when $W \geq 10$, approximately. Over any common frequency band an optimally filtered estimate will always have a smaller variance than the Eckart-filtered estimate.

Fig. 6 depicts the system performance if $d = 2$ and only flat bandpass filters, F_{NF}, are used. The separate sets of curves are for $d = 2$, with $a = 2$ and $b = 0$ (upper set), and $a = 3$ and $b = 1$ (lower set). In this figure the system degradation is measured relative to the infinite-bandwidth ($W = \infty$) optimally filtered system. Thus the

$$\text{dB loss} = 10 \log_{10} (D_{NF}{}^2(W)/D^2_{OPT}(\infty)) \tag{49}$$

in Fig. 6. Note that the upper set of curves also gives the dB loss of $D^2_{WH}(W)$ relative to $D^2_{OPT}(\infty)$ for $d = 2$. This is because $b = 0$. In both sets of curves the processing loss at first decreases to a minimum, and then, as too much high-frequency noise is processed, the variance then increases. Since the signal spectrum is falling off 6 dB/octave faster than the noise spectrum, the processor will behave ever more poorly as the processor bandwidth is made larger and larger.

References

[1] A. V. Balakrishnan, "Filtering and prediction theory," in *Communication Theory.* New York: McGraw-Hill, 1968.
[2] W. R. Hahn, "Optimum estimation of a delay vector caused by a random field propagating across an array of noisy sensors," Ph.D. dissertation, Dep. Elec. Eng., Univ. Maryland, College Park, Md., 1972. Reprinted by the Naval Ordnance Lab., Silver Spring, Md., as NOLTR 72-120, June 16, 1972.
[3] R. Deutsch, *Estimation Theory.* Englewood Cliffs, N.J.: Prentice-Hall, 1965.
[4] G. M. Jenkins and D. G. Watts, *Spectral Analysis and its Applications.* San Francisco: Holden-Day, 1968.
[5] D. G. Luenberger, *Optimization by Vector Space Methods.* New York: Wiley, 1969.
[6] M. A. Woodbury, "Inverting modified matrices," Stat. Res. Group, Princeton, N.J., Memo. Rep. 42, 1950; see also *Math. Rev.*, vol. 21, p. 361, 1951.
[7] C. Eckart, Marine Physical Lab., Scripps Inst. Oceanography, Univ. California, San Diego, Calif., SIO Ref. 52-11, 1952.
[8] R. J. Whelchel, "Optimization of conventional passive sonar detection systems," Naval Avionics Facility, Indianapolis, Ind., NAFI TR 1440, June 1969.