The Joint Estimation of Differential Delay, Doppler, and Phase

MATI WAX

Abstract—**In radio and sonar applications it sometimes happens that narrow-band signals, originated from a remote source and observed at a pair of receivers, differ by unknown differential phase and Doppler shift in addition to the differential delay corresponding to the range difference. The correspondence presents the joint maximum likelihood (ML) estimate of the differential delay, Doppler, and phase and examines their accuracy by deriving the Cramér–Rao bound. It is shown that the joint ML estimators are the values of the delay and Doppler that maximize the magnitude of a generalized ambiguity function analogous to the one used in radar. It is also shown that for long observation time and high enough signal-to-noise ratio there is no degradation in the accuracy of the time-delay estimator due to the additional phase and Doppler uncertainty and that the differential Doppler is uncorrelated with the differential delay and phase estimators.**

I. Introduction

The location of a radiating source can be determined by observing its signal in an array of sensors. In the absence of detailed knowledge concerning the signal waveshape, all information about source location is contained in the differential delays between the signal components received by various sensor pairs. Differential delay estimation has therefore attracted a great deal of interest in recent years [1]–[13].

The signals received in a pair of sensors may differ by differential phase and frequency in addition to the differential delay corresponding to the range difference. Differential phase arises when one or both of the signals are reflected from some boundary, e.g., mountains, ocean bottom, etc. Differential frequency is present when there is a relative motion between the source and the sensors so as to cause a differential Doppler shift between the received signals. Differential phase and frequency may also be present when the sensors are widely separated and the received signals are preprocessed by noncoherent local oscillators which are offset in both phase and frequency.

Previous investigations confined themselves to partial problems. Knapp and Carter [4] derived the maximum likelihood estimator of the differential delay and Doppler, whereas Weinstein and Schultheiss [10], [11] investigated accuracy of this joint estimation by deriving the Cramér–Rao bound. The effect of differential phase on the maximum likelihood differential delay estimator structure and its accuracy was investigated by Wax [12].

This work extends the previous works to the situation where both differential phase and frequency are present. It presents the joint maximum likelihood (ML) estimate of the differential delay frequency (Doppler), and phase and examines the variance of these estimates by evaluating the Cramér–Rao bound. The development and notation parallel those of Weinstein and Schultheiss [10] and Knapp and Carter [2], [4] to ease the comparison of the analogous derivations.

II. Statistical Model

Consider two narrow-band signals, $r_1(t)$ and $r_2(t)$ received simultaneously in two separate sensors during the interval $(0, T)$. We assume that the two signals contain delayed replicas of the same signal, with relative phase and frequency offset embedded in additive uncorrelated noises.

A compact representation of the received signals is obtained by employing the pre-envelope (analytic signal) notation [14]. Denoting by the subscript p the pre-envelopes we can express $r_{1p}(t)$ and $r_{2p}(t)$, the pre-envelopes of the received signals $r_1(t)$ and $r_2(t)$, respectively, as

$$\begin{aligned} r_{1_p}(t) &= s_p(t) + n_{1_p}(t), \qquad 0 \le t \le T, \\ r_{2_p}(t) &= s_p(t-D)e^{-j[\omega_d(t-D)+\varphi]} + n_{2_p}(t), \end{aligned} \tag{1}$$

where $s_p(t)$, $n_{1_p}(t)$, and $n_{2_p}(t)$ are the pre-envelopes of the signal $s(t)$ and the additive noises, $n_1(t)$ and $n_2(t)$, respectively, D is the differential delay and φ and ω_d are differential phase and frequency, respectively.

We assume that $s(t)$, $n_1(t)$, and $n_2(t)$ and hence $s_p(t)$, $n_{1_p}(t)$, and $n_{2_p}(t)$ are narrow-band independent zero-mean Gaussian processes and that the observation time T is large compared to the correlation time $1/W$ of the processes, i.e., $TW \gg 1$. We also assume that the differential frequency is small compared to the signal's bandwidth, but large compared to $1/T$, i.e., $W \gg \omega_d \gg 1/T$.

Since the observation time is finite it is possible to represent (mean-square sense) the pre-envelopes of the received signals by Fourier series. Recalling that the pre-envelopes contain only positive frequencies [14] and assuming that the truncation error is negligible we can write

$$r_{i_p}(t) = \frac{1}{\sqrt{T}} \sum_{k=C_i-N}^{C_i+N} R_{i_p}(\omega_k)e^{j\omega_k t}, \qquad i = 1,2, \tag{2}$$

where C_i and N are integers corresponding to the center frequency and signal bandwidth, respectively, ω_k is defined by

$$\omega_k = k(2\pi/T), \tag{3}$$

and $R_{i_p}(\omega_k)$ are the Fourier coefficients of $r_{i_p}(t)$ given by

$$R_{i_p}(\omega k) = \frac{1}{\sqrt{T}} \int_0^T r_{i_p}(t)e^{-j\omega_k t}\,dt, \qquad i = 1,2. \tag{4}$$

Let us assume that the differential frequency is confined to the discrete set

$$\omega_d = \frac{2\pi}{T}d, \qquad d \text{ integer}. \tag{5}$$

Since we have assumed that $\omega_d \gg 1/T$ it follows that the allowed differential frequencies form a rather dense set and thus the assumption made is not restrictive and the results obtained should hold for any $W \gg \omega_d \gg 1/T$.

We may thus represent the received signals on the discrete set $\{k(2\pi/T)\}$ in the frequency domain. Recalling the assumption that the differential frequency is small compared with the signal bandwidth, i.e., $N \gg d$, we can neglect the nonoverlapping frequencies of the two offset signalbands and represent our processes by the random vector

$$\begin{aligned} \boldsymbol{R}' = \Big[& R_{1_p}(\omega_{c_1-N}), R_{2_p}(\omega_{c_1-N+d}), R_{1_p}(\omega_{c_1-N+1}) \\ & \cdot R_{2_p}(\omega_{c_1-N+d+1}), \cdots, R_{1_p}(\omega_{c_1+N-d}), R_{2_p}(\omega_{c_1+N})\Big]. \end{aligned} \tag{6}$$

Since the Fourier coefficients are obtained by linear operations on the Gaussian random-process $r_{i_p}(t)$ $(i = 1,2)$ it follows that they are zero-mean complex Gaussian random variables and thus their joint density, conditioned on the parameter vector $\boldsymbol{\Theta} =$

Manuscript received October 8, 1980; revised March 11, 1981. This work was supported by the Air Force Office of Scientific Command under Contract AF44-620-79-C-0058.

The author is with the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305.

0018-9448/82/0900-0817$00.75

(φ, D, ω_d), is given by [14]

$$P(\boldsymbol{R} \mid \boldsymbol{\Theta}) = \frac{1}{\det[\pi K]} \exp[-\boldsymbol{R}^{\dagger} K^{-1} \boldsymbol{R}], \tag{7}$$

where $R^{\dagger}$ denotes the complex-conjugate transpose and K is the covariance matrix of R. This covariance matrix has a block-diagonal structure since the Fourier coefficients, of a process for which $TW \gg 1$, that correspond to different frequencies are uncorrelated [6], [4]. Thus, as one can verify by straightforward computation.

$$K = E[\boldsymbol{R}\boldsymbol{R}^{\dagger}] = \begin{bmatrix} Q_{C_1-N} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & Q_{C_1+N-d} \end{bmatrix}, \tag{8}$$

where Q_k denotes the covariance matrix of $[R_1(\omega_k) R_2(\omega_{k+d})]$ given by

$$Q_k = \begin{bmatrix} S_p(\omega_k) + N_{1_p}(\omega_k) & S_p(\omega_k) e^{j[(\omega_k+\omega_d)D+\varphi]} \\ S_p(\omega_k) e^{-j[(\omega_k+\omega_d)D+\varphi]} & S_p(\omega_k) + N_{2_p}(\omega_k) \end{bmatrix}. \tag{9}$$

III. The Maximum Likelihood Estimator

The maximum likelihood estimator is the value $\hat{\Theta}_{\mathrm{ML}} = (\hat{\varphi}_{\mathrm{ML}}, \hat{D}_{\mathrm{ML}}, \hat{\omega}_{d_{\mathrm{ML}}})$ that maximizes the conditional density $P(\boldsymbol{R} \mid \boldsymbol{\Theta})$. Noting that the determinant of K is independent of the parameters vector $\boldsymbol{\Theta}$ it follows that the maximization should be done over the exponent

$$I_1(\boldsymbol{\Theta}) = -\boldsymbol{R}^{\dagger} K^{-1} \boldsymbol{R}. \tag{10}$$

Substituting (6), (8), and (9) in (10) and neglecting terms independent of $\boldsymbol{\Theta}$, one obtains that $\hat{\Theta}_{\mathrm{ML}}$ should maximize the expression

$$\begin{aligned} I_1(\varphi, D, \omega_d) = &\sum_{k=C_1-N}^{C_1+N-d} \frac{R_{1p}^*(\omega_k) R_{2_p}(\omega_k + \omega_d)}{|G_{12_p}(\omega_k)|} \\ &\cdot \frac{C_p(\omega_k)}{1 - C_p(\omega_k)} e^{+j[(\omega_k+\omega_d)D+\varphi]} \\ &+ \sum_{k=C_1-N}^{C_1+N-d} \frac{R_{1_p}(\omega_k) R_{2_p}^*(\omega_k + \omega_d)}{|G_{12_p}(\omega_k)|} \\ &\cdot \frac{C_p(\omega_k)}{1 - C_p(\omega_k)} e^{-j[(\omega_k+\omega_d)D+\varphi]}, \end{aligned} \tag{11}$$

where $C_p(\omega_k)$ is the magnitude-squared coherence at the frequency ω_k [11]

$$C(\omega_k) = \frac{|G_{12_p}(\omega_p)|^2}{G_{11_p}(\omega_k) G_{22_p}\omega_k)} \tag{12}$$

and $G_{11_p}(\omega_k)$, $G_{22_p}(\omega_k)$, and $G_{12_p}(\omega_k)$ are the auto- and cross-power spectra of $r_{1_p}(t)$ and $r_{2_p}(t)$, evaluated at the frequency ω_k.

Recalling the relation between the Fourier transform of the signal and its pre-envelope [14] we obtain

$$R_{i_p}(\omega) = \begin{cases} 2R_i(\omega), & \omega \geq 0, \\ 0, & \omega < 0, \end{cases} \tag{13}$$

where $R_i(\omega)$ is the Fourier transform of $r_i(t)$. Thus, substituting (13) in (11) and rewriting the result in a integral form, yield

$$I_1(\varphi, D, \omega_d) = \frac{T}{\pi} \left| \int_0^{\infty} R_1(\omega) R_2^*(\omega + \omega_d) W_g(\omega) e^{-j(\omega+\omega_d)D} \, d\omega \right| \cdot \cos(\alpha - \varphi), \tag{14}$$

where

$$W_g(\omega) = \frac{1}{|G_{12}(\omega)|} \frac{C(\omega)}{1 - C(\omega)}, \tag{15}$$

$$C(\omega) = \frac{|G_{12}(\omega)|^2}{G_{11}(\omega) G_{22}(\omega)}, \tag{16}$$

and

$$\alpha = \arg\left[\int_0^{\infty} R_1(\omega) R_2^*(\omega + \omega_d) W_g(\omega) e^{-j(\omega+\omega_d)D} \, d\omega \right], \tag{17}$$

where $G_{11}(\omega)$, $G_{22}(\omega)$, and $G_{12}(\omega)$ are the auto- and cross-power spectra of the received signals $r_1(t)$ and $r_2(t)$, respectively.

The maximization of (14) with respect to (φ, D, ω_d) is straightforward. Clearly, $\hat{D}_{\mathrm{ML}}$ and $\hat{\omega}_{d_{\mathrm{ML}}}$ are the values of D and ω_d that maximize the expression

$$I(D, \omega_d) = \left| \int_0^{\infty} R_1(\omega) R_2^*(\omega + \omega_d) W_g(\omega) e^{-j(\omega+\omega_d)D} \, d\omega \right|, \tag{18}$$

while $\hat{\varphi}_{\mathrm{ML}}$ is given by

$$\hat{\varphi}_{\mathrm{ML}} = \arg\left[\int_0^{\infty} R_1(\omega) R_2^*(\omega + \hat{\omega}_{d_{\mathrm{ML}}}) W_g(\omega) e^{-j(\omega+\hat{\omega}_{d_{\mathrm{ML}}})\hat{D}_{\mathrm{ML}}} \, d\omega \right]. \tag{19}$$

In case $W_g(\omega)$ does not vary rapidly with ω, recalling that $\omega_d \ll W$, it follows that $W_g(\omega) \simeq W_g(\omega + \omega_d)$. Now, using this relation and Parseral's theorem we can rewrite $I(D, \omega_d)$ as

$$I(D, \omega_d) = \left| \int_0^T \tilde{r}_1(t) \tilde{r}_2^*(t + D) e^{-j\omega_d t} \, dt \right|, \tag{20}$$

where $\tilde{r}_1(t)$ and $\tilde{r}_2(t)$ are the filtered versions of $r_1(t)$ and $r_2(t)$, respectively, obtained by passing the signals through a filter whose transfer function is given by

$$H(\omega) = [W_g(\omega)]^{1/2}. \tag{21}$$

Observing (20) it is clear that

$$A(D, \omega_d) = \int_0^T \tilde{r}_1(t) \tilde{r}_2^*(t + D) e^{-j\omega_d t} \, dt. \tag{22}$$

should be regarded as the generalization of the ambiguity function used extensively in radar [16]. Obviously, the expressions suggested on intuitive grounds in [9] and [13], for the generalized ambiguity function are correct only when $W_g(\omega)$ is constant, i.e., when the signal and noises are white.

A possible realization of the maximum likelihood estimator, readily obtained from (20), is shown in Fig. 1. Looking at Fig. 1 we see that the scheme resembles the generalized cross-correlator [2] only that now one of the signals is frequency shifted prior to

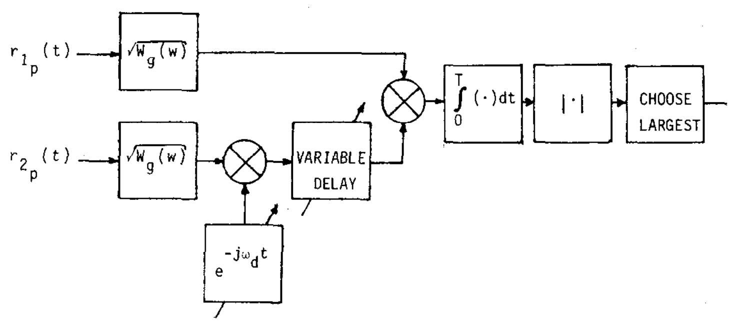

Fig. 1. Maximum likelihood (ML) time-delay estimator.

the correlation as one would expect intuitively. The maximization is now done in two dimensions and one should vary both the time-delay D and the differential frequency ω_d until the overall maximum of the magnitude of the generalized correlator output is found.

IV. The Cramér–Rao Bound

It is well-known that the mean-square error of the maximum likelihood estimator is appropriately described by the Cramér–Rao bound if the observation time is large compared to the correlation times of the processes and if the signal-to-noise ratio is high enough [6], [15]. The Cramér–Rao bound asserts that the mean-square error of $\hat{\Theta}_k$, any estimate of the kth parameter of the vector $\boldsymbol{\Theta}$, is lower bounded by

$$\sigma^2(\hat{\Theta}_k) \geq (J^{-1})_{kk}, \tag{23}$$

where J is the Fisher information matrix whose elements are

$$J_{nl} = -E\frac{\partial^2 \ln P(R \mid \Theta)}{\partial\Theta_n \partial\Theta_l}. \tag{24}$$

From results available in the literature [6] one can readily write down the elements of the Fisher information matrix in terms of the covariance matrix K as

$$J_{nl}(\boldsymbol{\Theta}) = \operatorname{tr}\left\{K^{-1}\frac{\partial K}{\partial \Theta_n} K^{-1} \frac{\partial K}{\partial \Theta_l}\right\}, \tag{25}$$

where $\operatorname{tr}\{\cdot\}$ stands for the trace of the bracketed matrix.

The computation of the Fisher information matrix can be facilitated by the use of the method introduced in [10]. It was proved there that the Fisher information matrix is invariant to reversible operations on the received signals. Thus, we can phase shift, time shift, and frequency shift $r_{2_p}(t)$, as shown in Fig. 2, and compute the Fisher information matrix for the signals $\tilde{r}_{1_p}(t)$ and $\tilde{r}_{2_p}(t)$.

Evaluating the Fisher information matrix for the special value $\hat{\boldsymbol{\Theta}} = \boldsymbol{\Theta}$, namely the true parameters, simplifies (somewhat) the calculations. To start with, we first calculate the covariance matrix $\tilde{K}$ of the data vector $\tilde{R}$ whose components are now

$$\tilde{R}_{i_p}(\omega_k) = \frac{1}{\sqrt{T}} \int_0^T \tilde{r}_{i_p}(t) e^{-j\omega_k t}\, dt. \tag{26}$$

Straightforward computation yields that the elements of $\tilde{K}$ are given by

$$E\left\{\tilde{R}_{i_p}(\omega_n)\tilde{R}^*_{k_p}(\omega_l)\right\} \big|_{\hat{\boldsymbol{\Theta}}=\boldsymbol{\Theta}} = \left[S_p(\omega_n) + N_p(\omega_n)\delta_{ik}\right]\delta_{nl}, \tag{27}$$

where δ_{ij} is the Kroneker deltas. If one arranges the data vector in the order

$$\boldsymbol{R}' = \left[\tilde{R}_{1_p}(\omega_{C_1-N}), \tilde{R}_{2_p}(\omega_{C_1-N}), \cdots, \tilde{R}_{1_p}(\omega_{C_1+N}), \tilde{R}_{2_p}(\omega_{C_1+N})\right],$$

the covariance matrix $\tilde{K}$ becomes block diagonal and its analytical inversion becomes feasible.

There remains the computation of the derivatives $\partial\tilde{K}/\partial\Theta_i$. Following the methods of [10] one obtains after tedious, though straightforward computation, the following results:

$$\frac{\partial}{\partial\varphi} E\left[\tilde{R}_{i_p}(\omega_n)\tilde{R}^*_{k_p}(\omega_l)\right]\Big|_{\hat{\boldsymbol{\Theta}}=\boldsymbol{\Theta}} = jS_p(\omega_n)\delta_{nl}(1-\delta_{ik}), \tag{28}$$

$$\frac{\partial}{\partial D} E\left[\tilde{R}_{i_p}(\omega_n)\tilde{R}^*_{k_p}(\omega_l)\right]\Big|_{\hat{\boldsymbol{\Theta}}=\boldsymbol{\Theta}} = j\omega_n S_p(\omega_n)\delta_{nl}(1-\delta_{ik}), \tag{29}$$

$$\frac{\partial}{\partial\omega} E\left[\tilde{R}_{i_p}(\omega_n)\tilde{R}^*_{k_p}(\omega_l)\right]\Big|_{\hat{\boldsymbol{\Theta}}\text{-}\boldsymbol{\Theta}} = \begin{cases} \dfrac{1}{4\pi} S_p^{(1)}(\omega_n), & n = l, \\[2ex] \dfrac{(-1)^{l-n}}{4\pi(l-n)} T\left[S_p(\omega_n) + S_p(\omega_l)\right](1+\delta_{ik}), & n \neq l, \end{cases} \tag{30}$$

where $j = \sqrt{-1}$ and $S_p^{(1)}(\omega)$ is the first derivative of the spectrum of $s_p(t)$.

The evaluation of the Fisher information matrix by (25) is now straightforward. Again, following the methods of [10], one obtains, after tedious computation, that the Fisher information matrix is given by

$$J = \begin{bmatrix} 2T\int_0^\infty \psi(f)\, df & 2T\int_0^\infty (2\pi f)\psi(f)\, df & 0 \\ 2T\int_0^\infty (2\pi f)\psi(f)\, df & 2T\int_0^\infty (2\pi f)^2\psi(f)\, df & 0 \\ 0 & 0 & \dfrac{T^3}{6}\int_0^\infty \psi(f)\, dt \end{bmatrix}, \tag{31}$$

where $\psi(f)$ is defined by

$$\psi(f) = \frac{C(f)}{1 - C(f)} \tag{32}$$

and is related to the signal-to-noise power ratio [2].

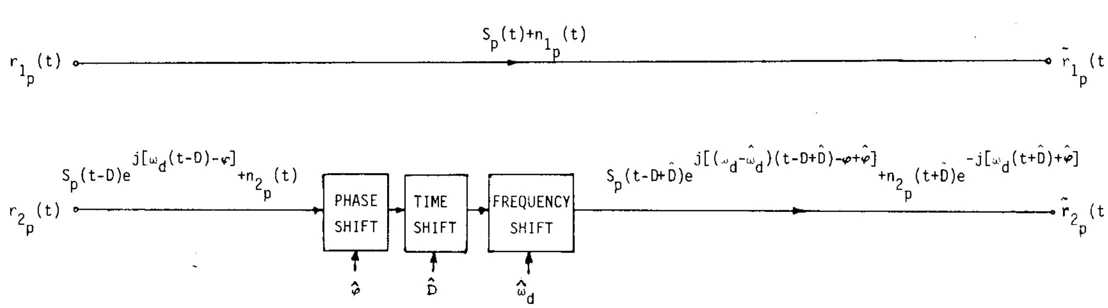

Fig. 2. Reversible operation on the received signals.

The block-diagonal structure of the Fisher information matrix indicates that the Doppler estimate is asymptotically uncorrelated with the differential delay and phase estimates. The Cramér–Rao bound on the accuracy of these estimates is obtained by the corresponding elements of J^{-1}. Following (23), we obtain after some algebraic manipulation

$$\sigma^2(\hat{\varphi}-\varphi) \geq \frac{\beta^2}{2T\int_0^\infty \left[2\pi(f-\bar{f})\right]^2 \psi(f)\, df}, \tag{33}$$

$$\sigma^2(\hat{D}-D) \geq \frac{1}{2T\int_0^\infty \left[2\pi(f-\bar{f})\right]^2 \psi(f)\, df}, \tag{34}$$

$$\sigma^2(\hat{\omega}_d-\omega_d) \geq \frac{1}{\frac{T^3}{6}\int_0^\infty \psi(f)\, df}, \tag{35}$$

where $\bar{f}$ denotes the center frequency of $\psi(f)$:

$$\bar{f} = \frac{\int_0^\infty f\psi(f)\, df}{\int_0^\infty \psi(f)\, df}, \tag{36}$$

and β denotes the radius of gyration of $\psi(f)$:

$$\beta^2 = \frac{\int_0^\infty (2\pi f)^2 \psi(f)}{\int_0^\infty \psi(f)\, df}. \tag{37}$$

The bound on the accuracy of the delay estimate (34) is identical to that obtained and discussed in [12]. Thus, as was also inferred from the block-diagonal structure of the Fisher information matrix, it follows that the additional frequency uncertainty does not degrade, asymptotically, the accuracy of the delay estimator. It should be noted that in [12] it was shown that, due to the ambiguity present in the output of the generalized correlator, the accuracy obtained in the pure delay case is actually equal to that obtained in the presence of phase uncertainty. Thus it follows that, asymptotically, there is no degradation in the attainable accuracy of the delay estimate in the presence of phase and frequency uncertainty.

The bound on the accuracy of the differential Doppler is identical to that obtained and discussed in [11]. Thus, as was also inferred from the block-diagonal structure of the Fisher information matrix, the attainable accuracy in the differential Doppler estimator is unaffected, asymptotically, by the presence of the phase and delay uncertainty.

References

[1] B. V. Hamon and E. J. Hannan, "Spectral estimation of time delay for dispersive and nondispersive systems," *J. Roy. Stat. Soc.*, Ser. C (Appl. Stat.), vol. 23, pp. 134–142, 1974.

[2] C. H. Knapp and G. C. Carter, "The generalized correlation method for the estimation of time delay," *IEEE Trans. Acous. Speech, Signal Processing*, vol. ASSP-24, pp. 320–327, Aug. 1976.

[3] 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.

[4] C. H. Knapp and G. C. Carter, "Estimation of time delay in the presence of source or receiver motion," *J. Acoust. Soc. Amer.*, vol. 61, pp. 1545–1549, June 1977.

[5] J. C. Hassab and R. E. Boucher, "Optimum estimation of time delay by a generalied correlator," *IEEE Trans. Acous., Speech, Signal Processing*, vol. ASSP-27, pp. 373–380, 1979.

[6] W. J. Bangs, "Arrays processing with generalized beamformers," Ph.D. dissertation, Yale University, New Haven, CT, 1971.

[7] S. K. Chow and P. M. Schultheiss, "Delay estimation using narrowband processes," *IEEE Trans. Acous., Speech, and Signal Processing,* (Special Issue on Time Delay Estimation, Part II), vol. 29, pp. 478–484, June 1981.

[8] Y. T. Chan, R. V. Hattin, and J. B. Plant, "The least squares estimation of time delay and its use in signal detection," *IEEE Trans. Acous., Speech, Signal Processing*, vol. ASSP-26, pp. 217–222, June 1978.

[9] J. C. Patzewitch, M. D. Srinath, and C. I. Black, "Near field performance of passive correlation sonars," *J. Acoust. Soc. Amer.*, vol. 64, pp. 1412–1423, Nov. 1978.

[10] E. Weinstein and P. M. Schultheiss, "Localization of a moving source using passive array data," Naval Ocean System Center Tech. Rep., Dec. 1978.

[11] P. M. Schultheiss and E. Weinstein, "Estimation of differential Doppler shifts," *J. Acoust. Soc. Amer.*, vol. 66, no. 5, pp. 1412–1419, Nov. 1979.

[12] M. Wax, "The estimation of the time delay between two signals with random relative phase," *IEEE Trans. Acous., Speech, Signal Processing*, (Special Issue on Time Delay Estimation, Part II), vol. 29, pp. 497–501, June 1981.

[13] S. Stein, "Algorithms for ambiguity function processing," *IEEE Trans. Acous., Speech, Signal Processing*, (Special Issue on Time Delay Estimation, Part II), vol. 29, pp. 588–599, June 1981.

[14] A. D. Whalen, Detection of Signals in Noise. New York: Academic, 1971.

[15] H. Van Trees, Detection, Estimation, and Modulation Theory, Part I. New York: Wiley and Sons, 1968.

[16] ——, Detection, Estimation, and Modulation Theory, Part III. New York: Wiley and Sons, 1971.