A Parameter Estimation Approach to Time-Delay Estimation and Signal Detection

Y. T. CHAN, JACK M. RILEY, MEMBER, IEEE, AND J. B. PLANT, SENIOR MEMBER, IEEE

Abstract—**Present techniques that estimate the difference in arrival time between two signals corrupted by noise, received at two separate sensors, are based on the determination of the peak of the generalized cross correlation between the signals. To achieve good resolution and stability in the estimates, the input sequences are first weighted. Invariably, the weights are dependent on input spectra which are generally unknown and hence have to be estimated. By approximating the time shift as a finite impulse response filter, estimation of time delay becomes one of determination of the filter coefficients. With this formulation, a host of techniques in the well-developed area of parameter estimation is available to the time-delay estimation problem—with the possibilities of reduced computation time as compared with present methods. In particular, it is shown that the least squares estimation of the filter coefficients is equivalent to estimating the Roth processor. However, the parameter estimation approach is expected to have a smaller variance since it avoids the need for spectra estimation. Indeed, experimental results from two examples show that the Roth processor, found by least squares parameter estimation, has a smaller variance than the approximate maximum likelihood estimator of Hannan-Thomson where spectral estimation is required. A detector that uses the sum of the estimated parameters as a test statistic is also given, together with its receiver operating characteristics.**

I. Introduction

The estimation of time difference (or delay) between signals corrupted by noise, received at two sensors located at a known distance apart, have applications in many fields [1]. A familiar use is in passive sonar where the bearing of a signal source is related to the time delay [2], [3]. Other not so well-known examples make use of the time delay and known sensor separation to compute the speed of a ship, or the rolling speed of hot steel, or the flow rate of solids through a pneumatic conveyor [4].

The generalized correlation method [2], which unifies many of the existing time-delay estimators into a two-prefilter cross-correlator configuration, performs a weighting on the inputs through the prefilters. To compute the weights, the input spectra at the two sensors should ideally be known. The same is true of the method in [3]. Since the input spectra are not known in many problems, they must first be estimated in order to establish the weights prior to the estimation of time delay. Due to the inaccuracies associated with estimating spectrum and coherence [5], and hence the weights, these time-delay estimators fail to achieve in practice, where data length is finite, their theoretical performance (based on known spectra)—a difficulty recognized in [2].

Manuscript received December 18, 1978; revised April 30, 1979 and August 16, 1979.

The authors are with the Department of Electrical Engineering, Royal Military College of Canada, Kingston, Ont., Canada.

This paper presents an alternative approach to time-delay estimation by modeling the time delay as a finite impulse response filter (FIR). With this formulation, time-delay estimation becomes a parameter estimation problem, that of estimating the coefficients of the FIR filter. The literature on parameter estimation is extensive in the areas of control, economics, and speech [6], [7]. Many existing techniques are directly applicable to the present problem and it will be shown that the least squares estimation of the parameters is equivalent to the Roth processor in [2]. However, since spectra estimation is avoided in the least squares estimator, it should, in practice, have a lower variance than those realized via the method in [2]. While this claim is not proven in theory, it is at least substantiated by experimental results which show in two examples that for equal data length (1024 points) the Roth processor, realized by parameter estimation, has a smaller variance than the Hannan-Thomson processor [2]. The latter processor is a maximum-likelihood estimator, if the input spectra are known.

Section II contains the derivation that relates the time delay to the coefficients of the FIR filter. The delay is, in general, assumed to be a nonintegral multiple of the sampling period, otherwise the formulation is trivial. The least squares solution that gives the Roth processor is described in Section III, together with the experimental results. Section IV shows that the estimates themselves can be used in a detection scheme and goes on to present the development for a receiver operating characteristics (ROC) curve. The conclusions are in Section V.

II. Problem Formulation

Let $x(t)$ and $y(t) = x(t + \tau)$ represent, respectively, the signal and its delayed version, the delay being τ. Their corresponding sampled values are $x(iT)$ and $y(iT)$, and if the sampling interval T is adequately small for the bandwidth of $x(t)$, and assuming $x(t)$ is band limited, then [8]

$$x(t) = \sum_{i=-\infty}^{\infty} x(iT) \operatorname{sinc}(t - iT) \tag{1}$$

where

$$\operatorname{sinc}(\cdot) \triangleq \frac{\sin\left\{\dfrac{\pi(\cdot)}{T}\right\}}{\dfrac{\pi(\cdot)}{T}}. \tag{2}$$

Without loss of generality, let $T = 1$ and $\tau = (l + f)T$ where l is any integer and $0 < f < 1$, i.e., τ is a nonintegral multiple of T.

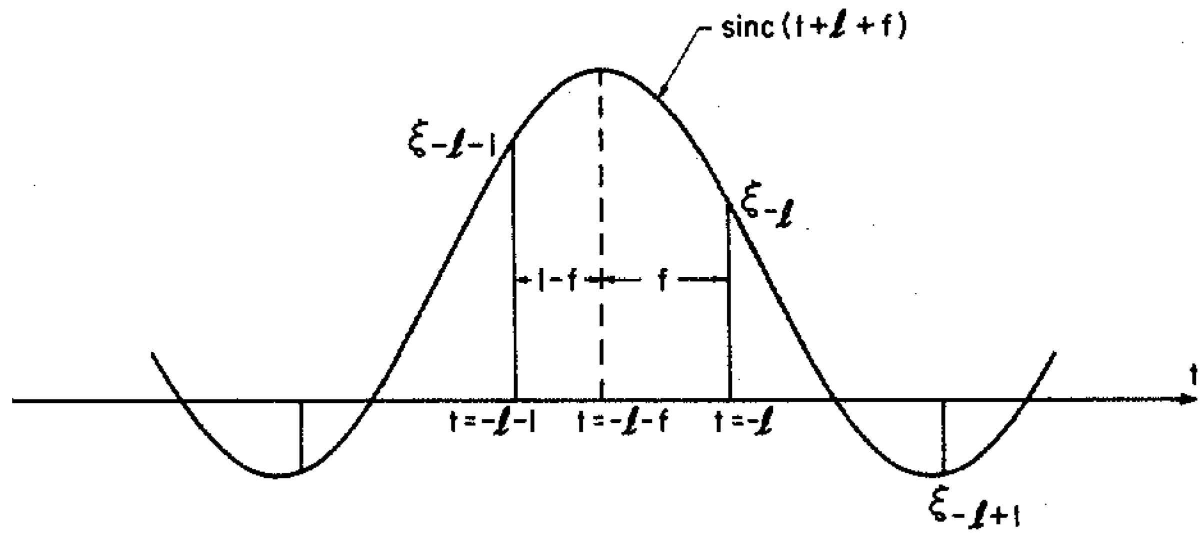


Fig. 1. The sinc $(t + e + f)$ function.

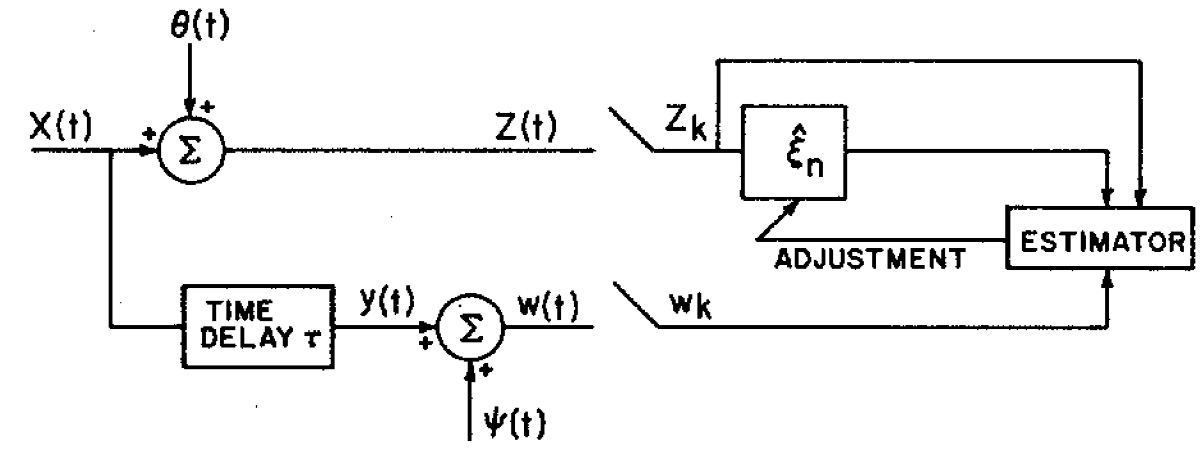


Fig. 2. Parameter estimation of time delay.

The delayed signal $y(t)$ can also be reconstructed from the samples $x(iT)$ by

$$y(t) = x[t + (l+f)] = \sum_{i=-\infty}^{\infty} x(i) \operatorname{sinc}(t + l + f - i), \tag{3}$$

so that for any integer k

$$y(k) = \sum_{i=-\infty}^{\infty} x(i) \operatorname{sinc}(k + l + f - i). \tag{4}$$

With the change of variable $k - n = i$, (4) becomes

$$y(k) = \sum_{k-n=-\infty}^{\infty} x(i) \operatorname{sinc}(n + l + f). \tag{5}$$

On defining

$$\zeta_n \triangleq \operatorname{sinc}(n + l + f) \tag{6}$$

and letting $y_k = y(k)$ and $x_k = x(k)$, (5) can be written as

$$y_k = \sum_{k-n=-\infty}^{\infty} \zeta_n x_{k-n}. \tag{7}$$

Since k is finite in (7) and the summation is from $-\infty$ to $+\infty$, (7) simplifies to

$$y_k = \sum_{n=-\infty}^{\infty} \zeta_n x_{k-n}. \tag{8}$$

Thus, the time series x_k is related to its delayed version y_k through a filter whose coefficients are ζ_n, the values of which are dependent on τ.

From (6), it is clear that ζ_n are the samples of the function sinc $(t + l + f)$, with the maximum at $t + l + f = 0$, as shown in Fig. 1. Hence, given the coefficients ζ_n, the delay τ is the value of t at which the maximum of $\zeta(t)$ given by

$$\zeta(t) = \sum_{n=-\infty}^{\infty} \zeta_n \operatorname{sinc}(t - n) \tag{9}$$

occurs. Alternatively, since

$$\zeta_n = \frac{(-1)^{n+l} \sin \pi f}{\pi(n + l + f)}, \tag{10}$$

the maximum value of ζ_n occurs at either ζ_{-l} (if $f < 0.5$) or ζ_{-l-1} (if $f > 0.5$) (see Fig. 1). Let ζ_j be the maximum value of ζ_n (i.e., $j = -l$ or $-l - 1$); then it is easy to verify, by direct substitution from (10), that

$$\tau = l + f = -j + \frac{\zeta_{j-1}}{\zeta_j + \zeta_{j-1}}. \tag{11}$$

Now with $x(t)$ as the signal, let $\theta(t)$ and $\psi(t)$ be the corrupting noise sources at the two sensors with the usual assumptions that the signal and noise sources are real, jointly stationary and independent random processes. The time-delay problem is as depicted in Fig. 2, where

$$z(t) = x(t) + \theta(t) \tag{12}$$

$$w(t) = y(t) + \psi(t) \tag{13}$$

and

$$z_k = x_k + \theta_k \tag{14}$$

$$w_k = y_k + \psi_k \tag{15}$$

are the samples (assuming an adequate sampling frequency) of $z(t)$ and $w(t)$, respectively. The estimator computes the estimates $\hat{\zeta}_n$, $n = -\infty$ to $+\infty$, using z_k and w_k, so that the delay τ can be obtained from either (9) or (11).

Clearly, it is not practical to estimate an infinite number of coefficients. However, since the function sinc $(n + l + f)$ approaches zero for large values of $n + l$, and since the maximum delay to be estimated (hence the largest possible value for l), is normally known, the series summation in (8) can be truncated at some predetermined, finite number p. Thus (8) changes to

$$y_k = \sum_{n=-p}^{p} \zeta_n x_{k-n}, \tag{16}$$

which models the time delay as an FIR filter. The modeling error introduced by (16) is examined in [9], which shows that, for an ideal low-pass [10] x_k and for $p = 4$, $l = 0$, the difference between the true delay and that produced by (16) varies, depending on the value of f, from a maximum of $-0.0231T$ at $f = 0.25T$ (about a 10 percent error) to a minimum of $-0.0004T$ at $f = 0.5T$. Similar results can be expected for sequences whose spectra are not ideal low pass. As an example in choosing p, suppose the desired accuracy is 10 percent and the maximum time difference that the estimator may encounter is $\pm 3T$. Then $p = 4 + 3 = 7$. It should be emphasized that the 10 percent error is the worst case, occurring at $\tau = 0.25T$. If $\tau = 1.25T$, the error is only $0.0231T/1.25T \times 100$ percent $=$ 1.85 percent. Quite often, the errors due to the approximation in (16) are insignificant compared with the variances of the estimator.

To proceed with the estimation of ζ_n using z_k and w_k, their relationship is first developed via (15) and (8) to yield

$$w_k = \sum_{n=-\infty}^{\infty} \zeta_n x_{k-n} + \psi_k. \tag{17}$$

Substituting (14) into (17) gives

$$w_k = \sum_{n=-\infty}^{\infty} \zeta_n z_{k-n} - \sum_{n=-\infty}^{\infty} \theta_{k-n} + \psi_k \tag{18}$$

$$= \sum_{n=-p}^{p} \zeta_n z_{k-n} + d_k \tag{19}$$

where

$$d_k = \sum_{n=-\infty}^{-p-1} \zeta_n z_{k-n} + \sum_{n=p+1}^{\infty} \zeta_n z_{k-n} - \sum_{n=-\infty}^{\infty} \zeta_n \theta_{k-n} + \psi_k \tag{20}$$

is the disturbance that contains the noise terms and the modeling errors. For $k = p$ to N, (19) expands to the matrix equation

$$w = Z\zeta + d \tag{21}$$

where

$$w = \begin{bmatrix} w_p \\ w_{p+1} \\ \vdots \\ w_{p+N} \end{bmatrix}, \quad \zeta = \begin{bmatrix} \zeta_{-p} \\ \vdots \\ \zeta_0 \\ \vdots \\ \zeta_p \end{bmatrix}, \quad Z = \begin{bmatrix} z_{2p} & z_{2p-1} \cdots z_p \cdots z_1 z_0 \\ z_{2p+1} & z_{2p} \quad \cdots \quad z_1 \\ \vdots & \vdots \\ z_{2p+N} & \cdots \quad z_N \end{bmatrix}, \quad d = \begin{bmatrix} d_p \\ d_{p+1} \\ \vdots \\ d_{p+N} \end{bmatrix}. \tag{22}$$

The estimation problem is: given the measurements vector w and matrix Z and (21), find the estimates $\hat{\zeta}$ such that $\hat{\zeta}$ is close to ζ in some sense.

While (21) is a familiar equation in parameter estimation, it possesses some properties that are not normally present. The first one is that the measurements z_k are related with the disturbances d_k, so that $E\{Z^T d\} \neq 0$. Secondly, the disturbance sequence is not white, so that $E\{dd^T\}$ is not a diagonal matrix. Literature on parameter estimation refers to the first property as correlation between observation and disturbance and the second as nonspherical disturbances [11]. Together, they add complexities to the estimation of ζ. But, if the final objective is estimating τ and not ζ, analysis in the next section reveals that the standard parameter estimation procedures are still applicable.

III. The Time-Delay Parameter Estimator

As mentioned earlier, many techniques are available to solve the problem posed in Section II. Among them, the simplest one chooses $\hat{\zeta}_n$ to minimize $\{\sum_{k=p}^{N} (w_k - \sum_{n=-p}^{p} \hat{\zeta}_n z_{k-n})\}^2$, i.e., it minimizes the square of the differences between the output of the FIR filter and the observations w_k. This well-known solution [11] is given by

$$\hat{\zeta} = (Z^T Z)^{-1} Z^T w. \tag{23}$$

Now because $E\{Z^T d\} \neq 0$, the estimate $\hat{\zeta}$ is asymptotically biased [11], i.e., $\hat{\zeta}$ is not a consistent estimator of ζ. To confirm this, take the probability limit, as N approaches infinity, denoted by $p \lim_{N\to\infty}$, of (23). On using a corollary of Slutsky's theorem [11], the result is

$$\bar{\zeta} \triangleq \underset{N\to\infty}{p \lim}\, \hat{\zeta} = \left(\underset{N\to\infty}{p \lim} \frac{Z^T Z}{N}\right)^{-1} \underset{N\to\infty}{p \lim} \frac{Z^T w}{N}. \tag{24}$$

Let

$$R_{zz}(i) \triangleq E\{z_k z_{k+i}\}, \qquad R_{\theta\theta}(i) \triangleq E\{\theta_k \theta_{k+i}\} \tag{25}$$

and

$$R_{zw}(i) \triangleq E\{z_k w_{k+i}\}, \qquad R_{xx}(i) \triangleq E\{x_k x_{k+i}\}; \tag{26}$$

then,

$$\underset{N\to\infty}{p \lim} \frac{Z^T Z}{N} = \begin{bmatrix} R_{zz}(0) & R_{zz}(1) \cdots R_{zz}(-2p) \\ R_{zz}(1) & R_{zz}(0) \cdots \\ \vdots & \\ R_{zz}(2p) & \cdots \end{bmatrix} \triangleq \Sigma_{zz} \tag{27}$$

which is the covariance matrix of the sequence z_k. Similarly,

$$\underset{N\to\infty}{p \lim} \frac{Z^T w}{N} = [R_{zw}(-p) \cdots R_{zw}(p)]^T. \tag{28}$$

Thus, from (24),

$$\bar{\zeta} = \Sigma_{zz}^{-1} [R_{zw}(-p) \cdots R_{zw}(p)]^T. \tag{29}$$

Next, from (26), (17), and (14), and the fact that the signal x_k and the noise sources are uncorrelated with each other, one obtains

$$R_{zw}(i) = \sum_{n=-\infty}^{\infty} \zeta_n R_{xx}(i-n) \tag{30}$$

and, for sufficiently large p,

$$R_{zw}(i) \approx \sum_{n=-p}^{p} \zeta_n R_{xx}(i-n). \tag{31}$$

Therefore,

$$[R_{zw}(-p) \cdots R_{zw}(p)]^T = \Sigma_{xx}\zeta \tag{32}$$

where

$$\Sigma_{xx} \triangleq \begin{bmatrix} R_{xx}(0) & R_{xx}(-1) \cdots R_{xx}(-2p) \\ R_{xx}(1) & \vdots \\ \vdots & \vdots \\ R_{xx}(2p) & \cdots \quad R_{xx}(0) \end{bmatrix} \tag{33}$$

is the covariance matrix of x_k. From (14), (25), and (26), one easily deduces

$$\Sigma_{zz} = \Sigma_{xx} + \Sigma_{\theta\theta} \tag{34}$$

where $\Sigma_{\theta\theta}$ is the covariance matrix of θ_k. Finally, the combination of (29), (34), and (32) yields

$$\bar{\zeta} = (\Sigma_{xx} + \Sigma_{\theta\theta})^{-1} \Sigma_{xx}\zeta \tag{35}$$

showing that, in general, $\bar{\zeta} \neq \zeta$ unless $\Sigma_{\theta\theta} = 0$.

If x_k and θ_k are samples from ideal low-pass processes, Σ_{xx} and $\Sigma_{\theta\theta}$ become diagonal matrices with diagonal elements equal to σ_x^2, the signal power, and σ_θ^2, the noise power, respectively. Then, (35) simplifies to

$$\bar{\zeta} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2}\zeta. \tag{36}$$

Thus, although $\bar{\zeta} \neq \zeta$, the use of $\hat{\zeta}_n$ in (11) to estimate τ will be unbiased in the asymptotic sense because the factors $\sigma_x^2/(\sigma_x^2 + \sigma_\theta^2)$ cancel each other in (11). Of course, the assumption of x_k being an ideal low-pass process is not always valid. To use (11), unbiased estimates of ζ must be obtained by some of the more complex estimation procedures (the instrumental variable method [11], for example). However, for purposes of time-delay estimation, it will next be shown that using the estimates from (23) in (9) will suffice.

First, recall from (29) that

$$\sum_{n=-p}^{p} \bar{\zeta}_n R_{zz}(i-n) = R_{zw}(i) \tag{37}$$

where $\bar{\zeta}_n$ is the nth element of $\bar{\zeta}$. But (37) is also an approximation to the discrete solution of the unrealizable Wiener-Hopf equation [12]

$$\sum_{n=-\infty}^{\infty} h_n R_{zz}(i-n) = R_{zw}(i) \qquad -\infty \leqslant i \leqslant \infty \tag{38}$$

where h_n are the coefficients of the unrealizable filter that minimizes $E\{(w_k - \sum_{n=-\infty}^{\infty} h_n z_{k-n})^2\}$. Further, since [12]

$$\sum_{n=-\infty}^{\infty} h_n e^{-j\omega n} = \frac{G_{zw}(\omega)}{G_{zz}(\omega)} \tag{39}$$

with $G_{zw}(\omega)$, $G_{zz}(\omega)$, and $G_{ww}(\omega)$ denoting the cross and auto spectra of z_k and w_k, the coefficients h_n are samples, at $t = nT$, of the function

$$h(t) = \int_{-\infty}^{\infty} \frac{G_{zw}(\omega)}{G_{zz}(\omega)} e^{j\omega t}\, d\omega. \tag{40}$$

The function $h(t)$ is recognized as the output of the Roth processor [2]. Noting the similarity between $\bar{\zeta}_n$ in (37) and h_n in (38), we conclude that the $\bar{\zeta}_n$ are samples of the Roth processor and $\hat{\zeta}_n$ from (23) are estimates of h_n. The time delay $\hat{\tau}$ is now found from (9) by substituting $\hat{\zeta}_n$ for ζ_n, $n = -p$ to p, and computing the value of t that maximizes (9). The Roth processor is unbiased [2], but this method of realization has modeling errors as discussed in the paragraph following (16). The theoretical variance is [2]

$$\operatorname{var}\{\hat{\tau}\} = \frac{2\pi \displaystyle\int_{-\infty}^{\infty} \frac{\omega^2\{G_{zz}(\omega)G_{ww}(\omega) - |G_{zw}(\omega)|^2\}}{G_{zz}^2(\omega)}\, d\omega}{T_l\left[\displaystyle\int_{-\infty}^{\infty} \frac{\omega^2|G_{zw}(\omega)|}{G_{zz}(\omega)}\, d\omega\right]^2} \tag{41}$$

where T_l is the record length.

The computational requirement for the time-delay parameter estimator (TDPE) is rather modest. A recursive algorithm [13] is available for the implementation of (23). At the Nth sample, let $\hat{\zeta}^{(N)}$ be the estimate of ζ and $Z^{(N)}$ and $w^{(N)}$ be the measurement matrix and vector. Then at $(N+1)$, with

$$\underline{z}^{(N+1)} = [z_{2p+N+1} \cdots z_{N+1}], \tag{42}$$

the new estimate is

$$\hat{\zeta}^{(N+1)} = \hat{\zeta}^{(N)} + \frac{\Sigma^{(N)}\underline{z}^{(N)^T}(w_{N+1} - \underline{z}^{(N+1)}\hat{\zeta}^{(N)})}{D^{(N)}} \tag{43}$$

where

$$D^{(N)} = 1 + \underline{z}^{(N+1)}\Sigma^{(N)}\underline{z}^{(N+1)^T} \tag{44}$$

$$\Sigma^{(N+1)} = \Sigma^{(N)} - \frac{\Sigma^{(N)}\underline{z}^{(N+1)^T}\underline{z}^{(N+1)}\Sigma^{(N)}}{D^{(N)}}. \tag{45}$$

At the start of the algorithm, $\Sigma^{(0)}$ is a diagonal matrix with elements equal to some large numbers, 10^6 for example, and $\hat{\zeta}^{(0)} = 0$. This algorithm eliminates the need for matrix inversion and the requirement for storing a large quantity of data. In addition, because of its recursive nature, it is well suited for applications where data arrive sequentially, such as in sonar.

To determine the value of t that maximizes (9), a Newton-Raphson iterative scheme was used. In all the experiments below, the search converged to within a 0.0001 resolution in, at most, six iterations.

As an assessment of the TDPE, experiments were performed on a PDP 11/34 computer using Fortran IV and single precision arithmetic. Three random number generators produced the independent ideal low-pass sequences x_k, θ_k, and ψ_k. Then x_k was shifted via (16) to give y_k, with $p = 16$, so that the error due to truncation of series is negligible. The ζ_n in (16) were computed according to (6) to realize the desired delay. The estimator used an FIR filter of $p = 4$ in (21) to model the time delay which had a value of either $0.25T$ or $0.50T$. The mean and mean-squared errors (MSE), with respect to the time delay quoted below, were based on 50 independent runs.

Experiment 1: Since x_k is ideal low pass, after obtaining $\hat{\zeta}_n$, the time delay $\hat{\tau}$ can be estimated from either (9) or (11). Table I summarizes the results for $\tau = 0.25T$ and $\tau = 0.50T$. Comparisons between the mean values in the tables confirm that the error in using $p = 4$ is largest for $\tau = 0.25T$ at $-0.0231T$ and smallest for $\tau = 0.50T$ at $-0.0004T$. It was this error at $\tau = 0.25T$ that caused large MSE, compared with the theoretical values, at a signal-to-noise (S/N) ratio of 4. At lower S/N ratios, when the estimator variances dominate, the experimental MSE agrees well with the predicted values from (41).

Experiment 2: The signals x_k and y_k used in this experiment were nonwhite sequences obtained from their ideal low-

TABLE I
TDPE, x_k WHITE

	S/N Ratio	$\hat{\tau}$ from (11)						$\hat{\tau}$ from maximization of (9)					
		N = 500			N = 1000			N = 500			N = 1000		
		Mean	MSE Theo.	MSE Exp.	Mean	MSE Theo.	MSE Exp.	Mean	MSE Theo.	MSE Exp.	Mean	MSE Theo.	MSE Exp.
$\tau = 0.25$	4	.249	3.4×10^{-4}	4.1×10^{-4}	.249	1.7×10^{-4}	2.2×10^{-4}	.225	3.4×10^{-4}	1.1×10^{-3}	.226	1.7×10^{-4}	8.1×10^{-4}
	1	.262	1.8×10^{-3}	2.2×10^{-3}	.263	9.0×10^{-4}	1.6×10^{-3}	.236	1.8×10^{-3}	1.8×10^{-3}	.239	9.0×10^{-4}	1.2×10^{-3}
	0.25	.242	1.5×10^{-2}	1.5×10^{-2}	.245	7.3×10^{-2}	9.5×10^{-3}	.219	1.5×10^{-2}	1.4×10^{-2}	.215	7.3×10^{-3}	8.0×10^{-3}
$\tau = 0.50$	4	.500	3.4×10^{-4}	3.9×10^{-4}	.500	1.7×10^{-4}	1.7×10^{-4}	.498	3.4×10^{-4}	5.2×10^{-4}	.500	1.7×10^{-4}	2.2×10^{-4}
	1	.509	1.8×10^{-3}	1.7×10^{-3}	.509	9.0×10^{-4}	1.9×10^{-3}	.509	1.8×10^{-3}	2.3×10^{-3}	.510	9.0×10^{-4}	1.4×10^{-3}
	0.25	.505	1.5×10^{-2}	1.0×10^{-2}	.500	7.3×10^{-3}	7.3×10^{-3}	.504	1.5×10^{-2}	1.7×10^{-2}	.505	7.3×10^{-3}	9.6×10^{-3}

TABLE II
TDPE, x_k NONWHITE

	S/N Ratio	$\hat{\tau}$ from maximization of (9)					
		N = 500			N = 1000		
		Mean	MSE Theo.	MSE Exp.	Mean	MSE Theo.	MSE Exp.
$\tau = 0.25$	4	.262	6.4×10^{-4}	1.3×10^{-2}	.247	3.2×10^{-4}	9.1×10^{-4}
	1	.255	4.9×10^{-3}	9.6×10^{-3}	.266	2.5×10^{-3}	6.4×10^{-3}
	0.25	.258	3.1×10^{-2}	5.2×10^{-2}	.223	1.5×10^{-2}	5.0×10^{-2}
	0.125	.329	9.2×10^{-2}	2.0×10^{-1}	.347	4.6×10^{-2}	1.2×10^{-1}
$\tau = 0.50$	4	.504	6.4×10^{-4}	4.3×10^{-3}	.497	3.2×10^{-4}	1.0×10^{-3}
	1	.492	4.9×10^{-3}	1.1×10^{-2}	.507	2.5×10^{-3}	5.6×10^{-3}
	0.25	.495	3.1×10^{-2}	4.9×10^{-2}	.469	1.5×10^{-2}	4.0×10^{-2}
	0.125	.515	9.2×10^{-2}	1.2×10^{-1}	.560	4.6×10^{-2}	1.1×10^{-1}

TABLE III
TDPE VERSUS *HT* ESTIMATOR $\tau = 0.25$

	S/N RATIO	HT estimator N=1024			TDPE N=1000		
		Mean	MSE Theo.	MSE Exp.	Mean	MSE Theo.	MSE Exp.
White	1	.222	9.0×10^{-4}	2.3×10^{-3}	.239	9.0×10^{-4}	1.2×10^{-3}
	0.25	.063	7.3×10^{-3}	4.3×10^{-2}	.215	7.3×10^{-3}	8.0×10^{-3}
Nonwhite	1	.159	1.5×10^{-3}	1.2×10^{-2}	.266	2.5×10^{-3}	6.4×10^{-3}
	0.25	.036	1.2×10^{-2}	5.3×10^{-2}	.223	1.5×10^{-2}	5.0×10^{-2}

pass counter parts by passing each of them through the filter

$$a_k = b_k + b_{k-1} \tag{46}$$

where b_k is the input to the filter and a_k is its output. The resultant x_k and y_k sequences have the nonwhite spectrum [14] $[\sin^2 \omega]/[\sin^2 (\omega/2)]$. Since θ_k and ψ_k remained white, the S/N ratio was no longer constant across the spectrum and was taken as the ratio of total signal power-to-noise power. The results are in Table II. In comparison with Table I, the nonwhite case has higher MSE, as predicted.

Experiment 3: This experiment compares the performance between the TDPE and the Hannan-Thomson (HT) estimator, which was computed according to [1]. For spectral estimation, the program divided 1024 points into 7 segments of 256 points each (hence a 50 percent overlap). It then applied a Hanning weighting to each segment before taking the fast Fourier transform and producing 128 Fourier coefficients. The nonwhite signals were produced in the same manner as described in experiment 2. Results in Table III show that, although the HT estimator has theoretical variances equal to (when x_k white) or less than (when x_k nonwhite) the TDPE, the latter has smaller experimental MSE. This, as conjectured in Section I, is because the TDPE avoids the need for spectral estimation. Mean values in Table III also indicate that the HT estimator exhibits a larger bias.

Several comments are in order on the experimental results. First, it is well-known [11] that the sample, or experimental, MSE are unbiased and have a variance equal to $\sigma^4/25$ for Gaussian distribution and 50 independent samples, and a theoretical (or population) MSE of σ^2. Thus, in 68 percent of the time (for one standard deviation $= \sigma^2/5$) the experimental MSE is within ± 20 percent of the theoretical value. Second, in experiment 3, the choice of 128 Fourier coefficients in the HT estimator is ad hoc. This choice is taken to cover a reasonable variation in the signal spectrum. If it were known *a priori* that the signal spectrum was white, for example, a much smaller number of coefficients, say eight, could have been used. The resultant experimental MSE would be much closer to the theoretical value because more segments would then be available for averaging. Of course, in practice, the shape of signal spectrum is unknown and a reasonable number of frequency points have to be used to ensure a sufficient resolution in frequencies. Finally, it should be noted that the experimental MSE in the HT estimator is in part caused by a large bias. This bias could possibly be due to the bias associated with coherence estimation [16], which would then bias the location of the maximum of the cost function [1] used in the HT estimator.

IV. The Detector

The development of the signal detection scheme is rather straightforward. This scheme is nonparametric with the test statistic

$$\Lambda = \sum_{n=-p}^{p} \hat{\zeta}_n. \tag{47}$$

Assuming that x_k and the noise sources are ideal low pass (an assumption necessary for the derivation of the ROC curves) so that the S/N ratio is constant over the band, then it follows from (36), for large N,

$$E\{\Lambda\} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2} \sum_{n=-p}^{p} \zeta_n. \tag{48}$$

In addition, for large N, the estimate $\hat{\zeta}$ is Gaussian [11] with the covariance (see Appendix I) given by

$$\frac{I}{N} \frac{1 + 2r}{(1 + r)^2},$$

where I is an identity matrix. Thus the $\hat{\zeta}_n$ are independent of each other and the variance of Λ is

$$\operatorname{var}\{\Lambda\} = \frac{(2p + 1)}{N} \frac{(1 + 2r)}{(1 + r)^2}. \tag{49}$$

It is shown in [15] that $\Sigma_{n=-\infty}^{\infty} \zeta_n = 1$ and for ζ_n as given by (6) with $p = 4$, $l = 0$, direct evaluation of the series $\Sigma_{n=-p}^{p} \zeta_n$ gives a sum of 0.992 for $f = 0.5$ and 0.997 for $f = 0.25$. Hence, with negligible error, (48) becomes

$$E\{\Lambda\} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2} = \frac{r}{1 + r} \tag{50}$$

where $\sigma_x^2/\sigma_\theta^2 = r$ is the S/N ratio and is equal to zero when no signal is present.

Let $\Lambda/1$ denote the test statistic when a signal is present and $\Lambda/0$ when not; then (50) and (49) give

$$E\{\Lambda/0\} = 0, \qquad \operatorname{var}\{\Lambda/0\} = \frac{2p + 1}{N} \tag{51}$$

and

$$E\{\Lambda/1\} = \frac{r}{1 + r}, \qquad \operatorname{var}\{\Lambda/1\} = \frac{(2p + 1)}{N} \frac{(1 + 2r)}{(1 + r)^2}. \tag{52}$$

With the mean and variance of the Gaussian random variable Λ known, the probability of false alarm P_{FA} and detection P_D are given by

$$P_{FA} = \int_{TH}^{\infty} f(\Lambda/0)\, d\Lambda, \tag{53}$$

which is independent of the ambient noise levels and

$$P_D = \int_{TH}^{\infty} f(\Lambda/1)\, d\Lambda. \tag{54}$$

The detection threshold is TH, i.e., $\Lambda > TH$ is a detection and $f(\cdot)$ denotes the probability density function.

We have computed, for $p = 4$ and $N = 1000$, the values of P_D and P_{FA} as a function of r. The results are in Table IV and the ROC curves are plotted in Fig. 3.

This detection technique is easily extendable to a linear array of 2R receivers. The spacings between receivers can be arbitrary, but cannot be too close to invalidate the assumption of independent noise sources at the receivers. The test statistic for 2R receivers, γ, is simply the sum of the Λ_j statistic from each pair of receivers. Thus,

$$\gamma = \sum_{j=1}^{R} \Lambda_j. \tag{55}$$

Any two receivers can form a pair, but each is used only once to ensure the independence of the Λ_j. Since the signal is ideal low pass and the noise sources are independent, the Λ_j are independent Gaussian random variables so that, for constant S/N ratio in each receiver in the array, γ has mean

$$E\{\gamma\} = \sum_{j=1}^{R} E\{\Lambda_j\} = RE\{\Lambda\} \tag{56}$$

and variance

$$\operatorname{var}\{\gamma\} = \sum_{j=1}^{R} \operatorname{var}\{\Lambda_j\} = R \operatorname{var}\{\Lambda\}. \tag{57}$$

The ROC curves for the γ statistic can be established with the same procedures as before.

V. Conclusions

This paper has introduced a new formulation of the time-delay estimation problem by modeling the delay as an FIR filter. This formulation is most useful when the largest ex-

TABLE IV
P_D FOR $N = 1000$

r \ P_{FA}	.001	.005	.02	.035	.045
0.1	.02	.05	.15	.20	.23
0.25	.16	.32	.53	.61	.66
0.35	.36	.57	.76	.83	.86
0.5	.67	.84	.94	.96	.97

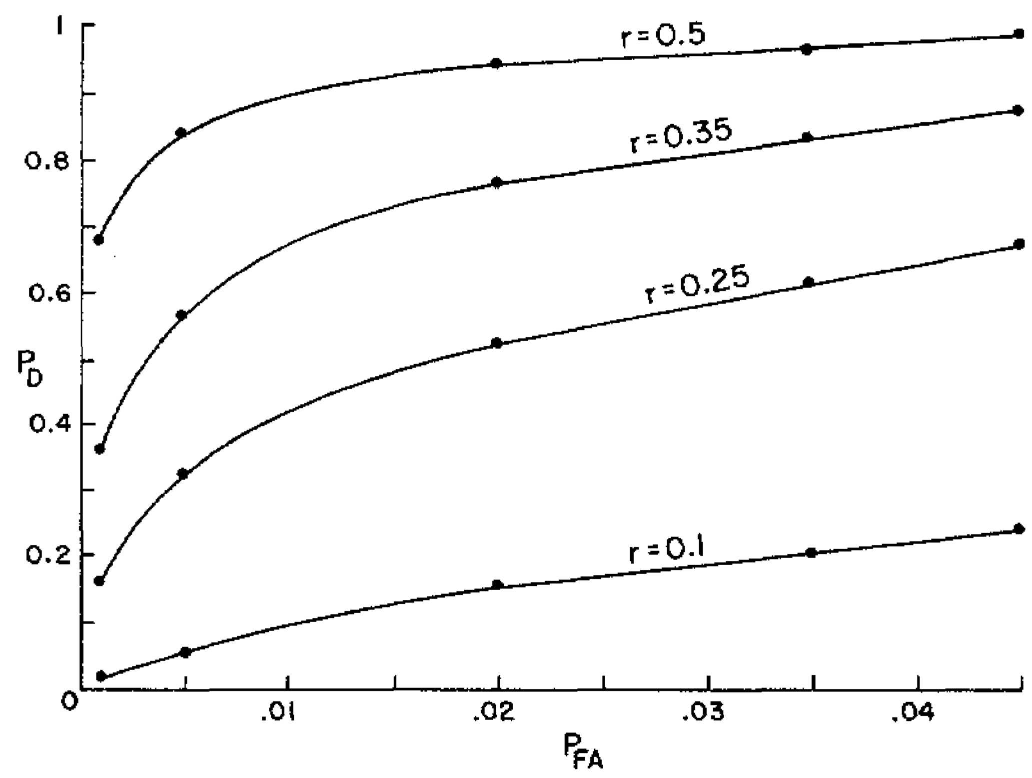


Fig. 3. ROC curves.

pected delay does not exceed several (3 or 4, say) times the sampling period, for then the order of the FIR filter can take on a reasonable number. While there are many methods available to estimate the coefficients of an FIR filter, only the least squares method is considered and is shown to be equivalent to the Roth processor. The effectiveness of this least squares algorithm has been demonstrated by experiments which also compare the TDPE with the HT estimator. The TDPE achieves better performance than the HT estimator in two examples in which one has a white input sequence and the other a nonwhite sequence. We attribute this outcome to the additional variances created by spectral estimations in the HT estimator. We have not given a comparison of computational times between the TDPE and those estimators in [2]. This is because the real time requirement can be very dependent on the special features of computing equipment. The choice of an estimator may well be decided by the particular application.

Appendix
The Asymptotic Covariance of $\hat{\zeta}$

In [11] the asymptotic covariance (ACV) of $\hat{\zeta}$, as estimated from (23), is given for the case $E\{Z^T d\} = 0$. When elements of the measurement matrix Z are correlated with the disturbances d_k in (21), the derivation for the ACV is considerably more complex and is not available in the literature. The ACV for $\hat{\zeta}$ is needed for the construction of the ROC curves in Section IV, where it is assumed that x_k and the noise sources are all ideal low pass. Without this assumption, the development for the ACV of $\hat{\zeta}$ would be untractable.

From [11], the ACV of $\hat{\zeta}$ is defined as

$$\mathrm{ACV}(\hat{\zeta}) \triangleq \frac{1}{N} \lim_{N\to\infty} E\{\sqrt{N}\,(\hat{\zeta} - \bar{\zeta})\sqrt{N}\,(\hat{\zeta} - \bar{\zeta})^T\} \tag{A.1}$$

where $\bar{\zeta}$ is given by (36). Note that, since $\bar{\zeta} \neq \zeta$, this variance of $\hat{\zeta}$ is in respect to its own mean and not the true value. The evaluation of (A.1) is rather tedious and the reader is referred to [15] for complete details. The procedures involve considering the asymptotic distribution of the terms in $\sqrt{N}\,(\hat{\zeta} - \bar{\zeta})$ and frequent use of the assumptions that signal and noise sources are white and are uncorrelated with each other. The result is an ACV($\hat{\zeta}$) which has all diagonal elements equal to

$$\frac{1}{N}\frac{1+2r}{(1+r)^2} \tag{A.2}$$

and all off diagonal elements equal to

$$\frac{1}{N}\frac{r}{(1+r)^2}\sum_{n=-p}^{p-1} \zeta_n \zeta_{n+1} \tag{A.3}$$

where $r \triangleq \sigma_x^2/\sigma_\theta^2$ is the S/N ratio. It is further shown in [15] that $\Sigma_{n=-\infty}^{\infty} \zeta_n\zeta_{n+1} = 0$. For $p = 4$ and $f = 0.5$, direct evaluation of $\Sigma_{n=-4}^{4} \zeta_n\zeta_{n+1}$ gives 0.052. Hence, with negligible error, ACV($\hat{\zeta}$) is a diagonal matrix with diagonal elements equal to

$$\frac{1}{N}\frac{1+2r}{(1+r)^2}.$$

To verify this theoretical ACV($\hat{\zeta}$), we have computed in experiment 1 the covariance of $\hat{\zeta}$ for $N = 1000$, $f = 0.5$, and $r = 4$ and 1, and are reproduced below. Only the upper triangular elements are given because ACV($\hat{\zeta}$) is a symmetric matrix.

$$r=4,\ \frac{1}{N}\frac{1+2r}{(1+r)^2}=3.6\times 10^{-4},\qquad \frac{1}{N}\frac{r}{(1+r)^2}\sum_{n=-p}^{p}\zeta_n\zeta_{n+1}=8.3\times 10^{-6}$$

Experimental ACV($\hat{\zeta}$):

4.8×10^{-4}	1.9×10^{-5}	2.9×10^{-5}	5.6×10^{-6}	-7.7×10^{-5}	-1.9×10^{-5}	-2.3×10^{-5}	6.5×10^{-5}	1.9×10^{-5}
	4.2×10^{-4}	-2.3×10^{-5}	-5.3×10^{-5}	-6.0×10^{-5}	-3.4×10^{-5}	-2.9×10^{-5}	-2.2×10^{-5}	-3.5×10^{-5}
		2.5×10^{-4}	6.4×10^{-5}	4.0×10^{-5}	-1.1×10^{-5}	-3.4×10^{-5}	4.6×10^{-5}	-3.4×10^{-5}
			4.6×10^{-4}	1.1×10^{-4}	-4.3×10^{-5}	-4.0×10^{-3}	6.5×10^{-5}	3.3×10^{-5}
				3.4×10^{-4}	1.3×10^{-5}	9.5×10^{-6}	-2.4×10^{-5}	1.5×10^{-5}
					3.9×10^{-4}	-3.5×10^{-5}	6.1×10^{-5}	9.4×10^{-5}
						2.2×10^{-4}	-9.4×10^{-5}	3.7×10^{-5}
							3.1×10^{-4}	3.4×10^{-6}
								3.7×10^{-4}

$$r=1,\ \frac{1}{N}\frac{1+2r}{(1+r)^2}=7.5\times 10^{-4},\qquad \frac{1}{N}\frac{r}{(1+r)^2}\sum_{n=-p}^{p}\zeta_n\zeta_{n+1}=1.3\times 10^{-5}$$

Experimental ACV($\hat{\zeta}$):

8.7×10^{-4}	4.0×10^{-5}	-1.6×10^{-4}	-9.5×10^{-6}	-2.0×10^{-4}	1.7×10^{-4}	9.6×10^{-5}	-9.2×10^{-5}	-5.7×10^{-5}
	7.9×10^{-4}	-3.0×10^{-5}	7.9×10^{-5}	-3.4×10^{-5}	-1.2×10^{-4}	-8.8×10^{-5}	3.7×10^{-5}	7.5×10^{-5}
		9.6×10^{-4}	5.0×10^{-5}	1.1×10^{-4}	-7.5×10^{-5}	-3.7×10^{-5}	-6.2×10^{-5}	1.5×10^{-5}
			8.4×10^{-4}	-2.0×10^{-4}	-9.3×10^{-5}	1.3×10^{-4}	-1.1×10^{-4}	2.2×10^{-4}
				6.2×10^{-4}	-1.1×10^{-4}	-1.1×10^{-4}	1.3×10^{-4}	-9.3×10^{-5}
					7.9×10^{-4}	-3.5×10^{-5}	-7.1×10^{-5}	1.5×10^{-4}
						6.9×10^{-4}	9.9×10^{-5}	-1.8×10^{-4}
							7.3×10^{-4}	-3.9×10^{-5}
								8.7×10^{-4}

References

[1] B. V. Hamon and E. J. Hannan, "Spectral estimation of time delay for dispersive and non-dispersive systems," *Appl. Statist.*, vol. 23, no. 2, pp. 134–142, 1974.

[2] 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.

[3] 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. Acoust., Speech, Signal Processing*, vol. ASSP-26, June 1978.

[4] H. K. Whitsel and L. W. Griswold, "Speed sensors for high speed surface craft," presented at 5th Ship Control Systems Symposium, Annapolis, MD, Nov. 1978.

[5] E. H. Scannell and G. C. Carter, "Confidence bounds for coherence estimates," presented at IEEE Intl. Conf. Acoust., Speech, Signal Processing, Tulsa, OK, April 1978.

[6] K. J. Astrom and P. Eykhoff, "System identification–A survey," *Automatica*, vol. 7, 1971.

[7] J. Makhoul, "Linear prediction–A tutorial review," *Proc. IEEE*, vol. 63, Apr. 1975.

[8] M. Schwartz and L. Shaw, *Signal Processing–Discrete Spectral Analysis, Detection, and Estimation.* New York: McGraw-Hill, 1975.

[9] Y. T. Chan, J. M. Riley, and J. B. Plant, "A method to shift a time series by a nonintegral multiple of sampling period," submitted to *IEEE Trans. Syst., Man, Cybern.*

[10] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 1965.

[11] A. S. Goldberger, *Econometric Theory.* New York: Wiley, 1964.

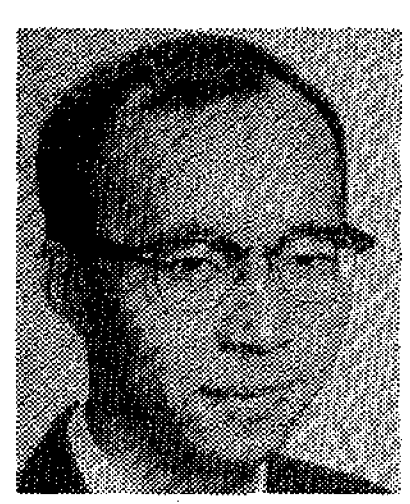
[12] H. Freeman, *Discrete-Time Systems.* New York: Wiley, 1965.

[13] J. M. Mendel, *Discrete Techniques of Parameter Estimation.* New York: Marcel Dekker Inc., 1973.

[14] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1975.

[15] J. M. Riley, "The least squares estimation of time delay," M.Eng. thesis, Dept. Elec. Eng., Royal Military College of Canada, Kingston, Ont., Canada, Oct. 1978.

[16] 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, June 1975.



Y. T. Chan was born in Hong Kong and received the B.Sc. and M.Sc. degrees from Queen's University, Kingston, Ont., Canada, and the Ph.D. degree from the University of New Brunswick, Fredericton, N.B., Canada, all in electrical engineering.

He has been with the Northern Electric Company, Ltd., London, Ont., Canada, the Royal Military College of Canada, Kingston, Ont., and Bell-Northern Research, Ottawa, Ont., Canada. Since 1973 he has been a Research Associate at the Royal Military College, Kingston, conducting research on weapons systems for the Directorate of Maritime Combat Systems, Department of National Defence, Canada.

Jack M. Riley (S'76-M'78) was born in Russell, Manitoba, Canada, on February 20, 1944. He received the B.Sc. in electrical engineering from the University of Manitoba, Canada in 1966, and the M.Eng. degree from The Royal Military College of Canada, Kingston, Ont., Canada, in 1979.

He is a member of the Canadian Armed Forces' Communications and Electronics Engineering Branch and has held engineering and project management positions at the Avionics Maintenance Development Unit, Trenton, Ont., Canada, National Defence Headquarters, Ottawa, Ont., Canada, and Air Defence Command/Air Command Headquarters, North Bay, Ont., Canada. He is presently an Assistant Professor with the Department of Electrical Engineering, The Royal Military College of Canada, Kingston, Ont., Canada.

J. B. Plant (S'63-M'65-SM'73) was born in Smiths Falls, Ont., Canada. He attended the Royal Military College of Canada, Kingston, Ont., and the Royal Naval Engineering College, Plymouth, England, graduating as a Lieutenant (E) in marine engineering in 1957. He received the Ph.D. degree from Massachusetts Institute of Technology, Cambridge, MA, in electrical engineering (specializing in optimal control) in 1965.

He has held an academic appointment in electrical engineering at the Royal Military College of Canada since that time as Head of the Department Electrical Engineering from 1967 to 1971, and as Dean of Graduate Studies and Research since 1971. During the academic year 1978 to 1979, he was on sabbatical leave at Thomson CSF, Paris, France.

Two-Dimensional Recursive Filter Design—A Spectral Factorization Approach

MICHAEL P. EKSTROM, SENIOR MEMBER, IEEE, RICHARD E. TWOGOOD, MEMBER, IEEE, AND JOHN W. WOODS, MEMBER, IEEE

Abstract—**This paper concerns development of an efficient method for the design of two-dimensional (2-D) recursive digital filters. The specific design problem addressed is that of obtaining half-plane recursive filters which satisfy prescribed frequency response characteristics. A novel design procedure is presented which incorporates a spectral factorization algorithm into a constrained, nonlinear optimization approach. A computational implementation of the design algorithm is described and its design capabilities demonstrated with several examples.**

I. Introduction

Recent technological advances admit the potential of using dedicated computer systems (or special-purpose hardware) to perform two-dimensional (2-D) signal processing tasks previously requiring large-scale scientific computers. In achieving this potential, such processing would be accessible to an enormously broad spectrum of applications. The principal characteristic of these applications is their need for "timely" data manipulation with affordable (small-scale) hardware.

In many aspects, recursive processors seem ideally suited to fulfill this need. Because of their reduced computational complexity over that of nonrecursive filter forms (fewer arithmetic operations, smaller memory requirements), recursive structures appear more compatible with small-scale hardware implementations, such as those involving mini- and microcomputers [1], [2]. That 2-D recursive structures have not enjoyed the success of their 1-D antecedents is attributable in large measure to difficulties in their design and analysis. Many of the basic techniques and tools used in analyzing 1-D filters rely rather fundamentally on the factorability of polynomials, and therefore cannot directly be applied to the 2-D problem. As a consequence, the issues of prime importance in recursive filtering, that of stability testing and filter design, are substantively less-tractable problems in the 2-D case [3].

While much progress has been made recently in developing stability theorems and practical tests based on these theorems [4]-[6], advances in recursive filter design have been less satisfying. The available literature features two prominent design approaches, those involving spectral transformations [7]-[9] and parameter optimization [10]-[12]. For the most part, these contributions have employed constrained filter forms (e.g., second-order, quarter plane) and addressed quite specific design problems (e.g., circularly-symmetric, lowpass). This appears to have been motivated by attendant simplifications derived from the special cases considered. These are simplifications involving stability testing and/or the availability of traditional 1-D filter design results. Despite this specificity (or perhaps because of it), their demonstrated design performance seems less than desirable, particularly for generic applications.

Manuscript received January 22, 1979; revised August 30, 1979. This work was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore Laboratory under Contract W-7405-ENG-48 and partially supported by the National Science Foundation under Grant ENG-78-04240.

M. P. Ekstrom and R. E. Twogood are with the Lawrence Livermore Laboratory, University of California, Livermore, CA 94550.

J. W. Woods is with the Department of Electrical and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12181.

0096-3518/80/0200-0016$00.75