minor changes it is also possible to obtain an optimum conditional estimate of the source waveform itself.

IV. Estimation Algorithm

The conditional log likelihood ratio, under the hypotheses of signal and no signal, is

$$\Lambda(R|X) = p_{r|x}(R|X)/p_n(R). \tag{2}$$

To aid in solution for the two probability densities which make up the likelihood ratio, we first define the reversible preprocessing operation, as described previously, shown in the upper left part of Fig. 2. It is shown by Grindon [6] that estimation of x given the preprocessed M-vector z is an equivalent problem. To solve for the probability densities, a Karhunen-Loeve expansion is used to represent the stochastic observation vector by a vector random variable of coefficients; the consequent infinite series are then isolated, found as expected to represent filtering functions, and shown to be solutions of integral equations of the Wiener-Hopf type. This sequence follows [6], so the detailed development is not repeated here.

The resultant formulation for the likelihood ratio is then manipulated to obtain the processing algorithm of Fig. 2. In Fig. 2, all terms are defined thereon or have been previously defined, except for the $M \times M$ matrix $H(t, x)$; this term arises because of system motion, and for slowly varying or stationary cases it may be approximated by the identity matrix. The interested reader is referred to [6].

From [4], the scalar impulse response $h_o(t, u)$ of the signal filter in Fig. 2 is the solution to the Wiener-Hopf equation,

$$\begin{aligned}(N_o/2)\,h_o(t,u) + \int_0^T & h_o(t,z)|H^{-1}(z,x)\,f(z,x)| \\ &\cdot |H^{-1}(u,x)\,f(u,x)|K_s(z,u)\,dz \\ = |H^{-1}&(t,x)\,f(t,x)||H^{-1}(u,x)\,f(u,x)| \\ &\cdot K_s(t,u), \qquad 0 \leqslant t,\ u \leqslant T\end{aligned} \tag{3}$$

where the M-vector $f(\cdot,\cdot)$ is defined on Fig. 2 and $N_o/2$ is the two-sided power spectral density of $n(t)$. While the general form for the optimal filter is given here, simple approximations to this filter are found satisfactory in practice.

The form illustrated separates into vector and scalar portions where the filter $h_o(t, u)$ can be moved into the scalar part. Note that $h_o(t, u)$ is the only element in the processor which depends upon the source covariance $K_s(t, u)$. In the form shown, then, only a single scalar filter is needed, regardless of the number of sensors. The vector portion of the processor is in the form of a generalized beam former which focuses the sensor array, jointly in both space and velocity coordinates, in response to X; the processor then scans the vector X in search of the MAP estimate $\hat{x}$, which is jointly optimal in both position and velocity.

V. Conclusion

An algorithm of broad applicability has been derived for jointly estimating the position, course, and speed of an underwater acoustic source from noisy signals received at a multisensor, possibly moving array. The algorithm, shown in the form of a focused beam former in combination with a filter-correlator, is a *globally optimal estimator* of the source state under the maximum *a posteriori* probability (MAP) criterion, which permits exploitation of the *a priori* statistics of the source state. In addition to its use in system implementations, the algorithm may be used as a reference in Monte Carlo tests for assessing the threshold performance of suboptimal or asymptotically optimal algorithms.

References

[1] G. C. Carter, "Time delay estimation for passive sonar signal processing," this issue, pp. 463-470.
[2] 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.
[3] W. B. Adams, J. P. Kuhn, and W. P. Whyland, "Correlator compensation requirements for passive time-delay estimation with moving source or receivers," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 158-168, Apr. 1980.
[4] H. L. Van Trees, *Detection, Estimation, and Modulation Theory—Part I.* New York: Wiley, 1968.
[5] I. N. El-Behery and R. H. MacPhie, "Maximum likelihood estimation of the number, directions, and strengths of point radio sources from variable baseline interferometer data," *IEEE Trans. Antennas Propagat.*, vol. AP-26, pp. 294-301, Mar. 1978.
[6] J. R. Grindon, "Estimation theory for a class of nonlinear random vector channels," Ph.D. dissertation, Washington Univ., St. Louis, MO, 1970; Ann Arbor, MI: University Microfilms no. 70-18, 948, 1970.
[7] W. R. Hahn, "Optimum signal processing for passive sonar range and bearing estimation," *J. Acoust. Soc. Amer.*, vol. 58, pp. 201-207, July 1975.
[8] V. H. MacDonald and P. M. Schultheiss, "Optimum passive bearing estimation in a spatially incoherent noise environment," *J. Acoust. Soc. Amer.*, vol. 46, pp. 37-43, July 1969.

Time Delay Estimation in Nonlinear Systems

N. J. I. MARS AND G. W. VAN ARRAGON

***Abstract*—Methods for the estimation of time delay usually require the assumption of a linear channel between pairs of signals. In this paper we study an estimator (using the concept of average mutual amount of information) which is usable in the case of nonlinear channels. Simulation experiments for a linear system, a squaring system, and a rectifying system show encouraging results.**

I. Introduction

In many methods for *time delay estimation*, a linear channel between pairs of signals is assumed. In this paper, however, we study systems described by

$$s_2(t) = F(s_1(t+D)) + n(t) \tag{1}$$

in which $s_1(t)$ is the input signal, $s_2(t)$ is the output signal, and $n(t)$ is Gaussian noise, uncorrelated with s_1 or s_2; F is a (possibly nonlinear) functional and D is the time delay to be determined from observations of $s_1(t)$ and $s_2(t)$. As examples of the functional F we will discuss in some detail the linear system, the squaring system and the rectifying system. Our method is not restricted to these, however.

A common method for determining the time delay D in linear systems is to estimate the *cross correlation function*

$$R_{s_1 s_2}(\tau) = \frac{1}{T-\tau}\int_\tau^T s_1(t)\cdot s_2(t-\tau)\,dt. \tag{2}$$

The value of τ where (2) is maximum provides an estimate of D. Variations on this method have been proposed [1] in which $s_1(t)$ and $s_2(t)$ are prefiltered to improve the estimate of D.

Manuscript received June 4, 1980; revised September 12, 1980.
The authors are with the Department of Electrical Engineering, Twente University of Technology, P.O. Box 217, 7500 AE Enschede, The Netherlands.

0096-3518/81/0600-0619$00.75 © 1981 IEEE

It can be shown that in such a way an optimal estimate of D can be achieved. In case the functional F in (1) is nonlinear, however, the optimality of that estimator is no longer assured. We have developed an estimator of D for nonlinear functionals. Although we cannot prove its optimality either for all or even for some nonlinear functionals, empirical results with it are encouraging.

Our motivation for the development of this estimator comes from a medical problem, namely, the study of the spread of electrical activity through the human brain in epileptic patients. The pathways of propagation in the brain are highly nonlinear, necessitating other analysis methods than the cross correlation function for the determination of (propagation) time delay.

II. Average Mutual Amount of Information

By writing the cross correlation function in the form

$$R_{s_1 s_2}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_1(t) \cdot s_2(t+\tau) \cdot f(s_1, s_2, \tau)\, ds_1 ds_2 \tag{3}$$

in which f is the joint probability density function of $s_1(t)$ and $s_2(t+\tau)$, we obsêrve that only the second-order moments of the probability density function of s_1 and s_2 contribute to the cross correlation function. It is therefore not surprising that for nonlinear functionals an improved estimator of D can be found which makes use of the higher order moments of the distributions of s_1 and s_2. We have used the concept of *average mutual amount of information* for this purpose.

Average mutual amount of information (AMAI) was defined in 1959 by Gel'fand and Yaglom [2] as a measure for the predictability of one signal, given another. The AMAI between two sequences X and Y is defined as

$$\text{AMAI}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) \cdot \log \frac{f(x,y)}{f(x) \cdot f(y)}\, dx\, dy \tag{4}$$

in which $f(x, y)$ is the simultaneous probability density function of (X, Y), $f(x)$ is the marginal probability density function of X, and $f(y)$ is the marginal probability density function of Y. AMAI ranges from 0 (no predictability) to ∞ (perfect predictability) and is symmetrical in its arguments (AMAI(X, Y) = AMAI(Y, X)).

In our application to time delay estimation in *nonlinear systems*, we use AMAI in a manner analogous to the cross correlation function by computing (4) for a range of lag values τ between s_1 and s_2:

$$\text{AMAI}(s_1, s_2, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s_1(t), s_2(t+\tau)) \cdot \log \left\{ \frac{f(s_1(t) \cdot s_2(t+\tau))}{f(s_1) \cdot f(s_2)} \right\} ds_1 ds_2. \tag{5}$$

The value of τ where AMAI(τ) has a maximum provides an estimate of the delay D.

We observe from (5) that the computation of AMAI requires knowledge of the joint and marginal probability density functions of s_1 and s_2. In the next paragraph we will discuss methods to estimate these densities. We also observe that (5) is not only a function of the second-order moments of the probability density functions, but also of higher order moments, as required by our application to nonlinear systems.

III. Estimation of Probability Density Functions

In many practical situations the joint and marginal probability density functions in (5) will not be known and will have to be estimated.

Of the different approaches to *probability density function estimation* described in the literature (for reviews see [3]-[5]) we have chosen the well-known Parzen kernel estimator [6]

$$f_N(x_1, x_2, \cdots, x_M) = \frac{1}{N} \sum_{n=1}^{N} \prod_{m=1}^{M} \frac{1}{h_m(N)} K\left(\frac{x_m - X_{mn}}{h_m(N)}\right). \tag{6}$$

Epanechnikov [7] has derived the nonnegative kernel form and kernel width which minimize the relative global approximation error over all densities, in the case where the true probability density function has a Taylor expansion in all its arguments everywhere. As the optimal kernel form he has found

$$K(y) = \frac{3}{4\sqrt{5}} \left(1 - \frac{y^2}{5}\right) \quad \text{if } -\sqrt{5} \leqq y < +\sqrt{5}$$
$$= 0 \qquad \text{elsewhere.} \tag{7}$$

This kernel function has a simple form and a finite support and is independent of the true probability density function and of the sample size.

Within this class of optimal kernel functions the kernel width which minimizes the relative global approximation error is given by

$$h_0(N) \approx \left[\frac{ML^M}{ND}\right]^{1/(M+4)} \tag{8}$$

in which

$$L = \int_{-\infty}^{\infty} K^2(y)\, dy = \int_{-\sqrt{5}}^{\sqrt{5}} \left[\frac{3}{4\sqrt{5}} \left(1 - \frac{y^2}{5}\right)\right]^2 dy = \frac{3}{5\sqrt{5}} \tag{9}$$

and

$$D = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\sum_{m=1}^{M} \frac{\delta^2}{\delta x_m^2} f(x_1, x_2, \cdots, x_M)\right]^2 \cdot dx_1 dx_2 \cdots dx_M. \tag{10}$$

Thus, the optimal kernel width is a function both of the number of samples and of the density to be estimated. We have used this probability density function estimator in an iterative way, in which an initial estimate of the probability density functions is used in (8) to compute an estimate of the optimal kernel width, which in turn is used to estimate the probability density functions using (6), etc. Although we have not been able to prove the convergence of this procedure, we have not experienced cases of malconvergence in a great many experiments with samples from known probability density functions.

Both the joint and marginal probability density functions are used in (4) to compute the AMAI by numerically integrating the integrand of (4).

The whole procedure is repeated for every value of the lag value τ. The range over which τ is varied is determined *a priori* from the known physical properties of the problem. The value of τ where AMAI(τ) has a maximum is considered the best estimate of D.

IV. Experimental Results

To assess the practical usefulness of the AMAI estimator of time delay, we performed simulations in which the system described by (1) was used. Three functionals were studied: the linear system, the squaring system and the rectifying system. As input signal $s_1(t)$ we used low-pass filtered Gaussian noise, generated according to

$$s_1(t) = \text{rho} \cdot s_1(t-1) + \sqrt{1 - \text{rho}^2}\, n_1(t); \tag{11}$$

TABLE I
QUALITY OF TIME DELAY ESTIMATION [ACCORDING TO (13)] AS A FUNCTION OF S/N RATIO AND CORRELATION BETWEEN SAMPLES RHO FOR A LINEAR SYSTEM

rho \ S/N	0.111	0.250	1.000	2.333	9.000
0.00	38.827	59.075	236.218	479.983	1074.509
0.50	8.545	9.145	11.946	15.122	22.275
0.70	4.450	4.638	5.790	6.970	9.623
0.90	2.587	2.656	2.977	3.396	4.378
0.95	2.084	2.209	2.472	2.730	3.438

TABLE II
QUALITY OF TIME DELAY ESTIMATION [ACCORDING TO (13)] AS A FUNCTION OF S/N RATIO AND CORRELATION BETWEEN SAMPLES RHO FOR A SQUARING SYSTEM

rho \ S/N	0.111	0.250	1.000	2.333	9.000
0.00	52.266	97.724	363.089	271.439	353.439
0.50	24.695	31.245	45.146	49.276	74.977
0.70	9.680	10.278	13.008	15.120	19.576
0.90	3.391	3.647	4.135	4.695	6.028
0.95	2.522	2.739	2.983	3.395	4.054

$n(t)$ from (1) and $n_1(t)$ from (11) have the same bandwidth. The correlation between neighboring samples was thus rho. The output signal-to-noise ratio is defined as

$$S/N = \text{var}(s_1)/\text{var}(n). \tag{12}$$

Without loss of generality the time delay D in (1) was set to zero. We performed simulations in which both the S/N ratio and the correlation between successive samples of the input signal were varied systematically. To summarize the results of these experiments we define a measure of "peakedness" of the AMAI as a function of lag, as an approximate measure for the quality of our time delay estimator, according to

$$\text{Quality} = \frac{\text{AMAI}(k) - 0.125 \sum_{|i|=1}^{4} \text{AMAI}(k+i)}{\sigma_{\text{nonpeak}}} \tag{13}$$

in which k is the index of the theoretical value of the delay D and σ_{nonpeak} is the standard deviation of the eight AMAI values surrounding the peak. In other words, this provides an estimate of how well (in terms of standard deviations) the AMAI peak "stands out" from the background.

The results for the linear system, the squaring system, and the rectifying system are given in Tables I, II, and III, respectively. As can be seen from these tables, and as expected, the "peakedness" decreases with increasing correlation between the samples of the input signal and with decreasing signal-to-noise ratio. Both for the linear and the nonlinear systems the correct value for the time delay D is found.

V. CONCLUSIONS

An estimator for time delay using the concept of average mutual amount of information is described. Although analogous to the cross correlation function, this estimator is not restricted to time delay estimation in linear systems. Simulations show it to perform quite well in two nonlinear systems: the squaring system and the rectifying system.

TABLE III
QUALITY OF TIME DELAY ESTIMATION [ACCORDING TO (13)] AS A FUNCTION OF S/N RATIO AND CORRELATION BETWEEN SAMPLES RHO FOR A RECTIFYING SYSTEM

rho \ S/N	0.111	0.250	1.000	2.333	9.000
0.00	47.854	124.506	528.972	684.887	1054.363
0.50	30.290	37.418	51.559	66.928	105.042
0.70	10.475	11.290	14.679	18.565	24.252
0.90	3.474	3.715	4.615	5.480	7.275
0.95	2.555	2.738	3.388	3.778	4.804

REFERENCES

[1] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 24, pp. 320–327, 1976.
[2] I. M. Gel'fand and A. M. Yaglom, "Calculation of the amount of information about a random function contained in another such function," *Amer. Math. Soc. Trans.*, vol. 12, pp. 199–246, 1959.
[3] E. J. Wegman, "Nonparametric probability density estimation, I: A summary of available methods," *Technometrics*, vol. 14, pp. 533–546, 1972.
[4] —, "Nonparametric probability density estimation, II: A comparison of density estimation methods," *J. Statist. Comput. Sim.*, vol. 1, pp. 225–245, 1972.
[5] M. J. Fryer, "A review of some nonparametric methods of density estimation," *J. Inst. Math. Its Appl.*, vol. 20, pp. 335–354, 1977.
[6] E. Parzen, "On the estimation of a probability density function and mode," *Ann. Math. Statist.*, vol. 33, pp. 1065–1076, 1962.
[7] V. A. Epanechnikov, "Nonparametric estimation of a multivariable probability density," *Theory Prob. Appl.*, vol. 14, pp. 153–158, 1969.

Delay Estimation of Disturbances on the Basilar Membrane

MARY MORTON GIBSON

Abstract—**With current measurement techniques, it is impossible to directly observe the traveling wave on the basilar membrane over its entire length. Treating the mechanical propagation as a pure conduction delay, the travel time to a particular region can be inferred from the phase-frequency characteristics of the neural response.**

INTRODUCTION

Changes in sound pressure around us are transmitted to the fluids of the inner ear where they are transduced into electrical impulses which are sent to the brain. The external ear directs the impinging sound pressures down the auditory canal and against the eardrum, causing it to move. This motion causes the bones of the middle ear to move, transmitting the motion to the oval window membrane. The middle ear behaves as an impedance matching device between the outside air and the

Manuscript received May 26, 1980; revised October 24, 1980. This work was supported by NIH Grants NS06225, NS12732, NS0536, and NS30897.

The author is with the Department of Neurophysiology, University of Wisconsin, Madison, WI 53706.