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

IV. ESTIMATION ALGORITHM

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

$$\Lambda(R|X) = p_{R|x} R(X) p_{R}(R).$$

(2)

To aid in solution for the two probability densities which make up the likelihood ratio, we first define the reversible pre-
processing operation, as described previously, shown in the upper left part of Fig. 2. It is shown by Grindon [6] that esti-
mation 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 con-
sequent infinite series are then isolated, found as expected to repre-
sent 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 de-
defined, except for the $M \times M$ matrix $H(t,x)$; this term arises be-
cause of system motion, and for slowly varying or stationary cases it may be approximated by the identity matrix. The in-
terested reader is referred to [6].

From [41], the scalar impulse response $h_0(t,u)$ of the signal filter in Fig. 2 is the solution to the Wiener-Hopf equation,

$$(N_0/2) h_0(t,u) + \int_0^T h_0(t,z) H^{-1}(z,x) f(z,x) \, dz$$

$$= |H^{-1}(u,x)f(u,x)| K_2(z,u) \, dz$$

$$|H^{-1}(t,x)f(t,x)| |H^{-1}(u,x)f(u,x)|$$

$$K_2(t,u), \quad 0 \leq t, \leq T$$

(3)

where the $M$-vector $f(\cdot , \cdot )$ is defined on Fig. 2 and $N_0/2$ 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 $h_0(t,u)$ is moved into the scalar part. Note that $h_0(t,u)$ is the only element in the processor which depends upon the source covariance $K_2(z,u)$. 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 coor-
dinates, 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 under-
water acoustic source from noisy signals received at a multi-
sensor, 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 un-
der the maximum a posteriori probability (MAP) criterion, which permits exploitation of the a posteriori 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 pro-
cessing," this issue, pp. 463-470.


ensation requirements for passive time-delay estimation with mov-
ing source or receivers," IEEE Trans. Acoust., Speech, Signal Pir-


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. Sim-
ulation experiments for a linear system, a squaring system, and a recti-
fying 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

$$s_2(t) = F(s_1(t + D)) + n(t)$$

(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 (possi-
bly nonlinear) functional and $D$ is the time delay to be deter-
mored from observations of $s_1(t)$ and $s_2(t)$. As examples of the functional $F$ we will discuss in some detail the linear sys-
tem, 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

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

(2)

The value of $\tau$ 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

$$R_{s_1s_2}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_1(t) \cdot s_2(t + \tau) \cdot f(s_1, s_2, \tau) \, ds_1 \, ds_2$$

(3)

in which $f$ is the joint probability density function of $s_1(t)$ and $s_2(t + \tau)$, we observe 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

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

(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 $\infty$ (perfect predictability) and is symmetrical in its arguments ($\text{AMAI}(X, Y) = \text{AMAI}(Y, X)$).

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 $\tau$ between $s_1$ and $s_2$:

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

(5)

The value of $\tau$ where $\text{AMAI}(\tau)$ 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]

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

(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

$$K(y) = \frac{3}{4\sqrt{5}} \left( 1 - \frac{y^2}{5} \right) \quad \text{if} \quad -\sqrt{5} \leq y < +\sqrt{5}$$

$$= 0 \quad \text{elsewhere}.$$  

(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

$$h_0(N) = \left[ \frac{ML^2}{ND} \right]^{1/(M+4)}$$

(8)

in which

$$L = \int_{-\sqrt{5}}^{\sqrt{5}} K(y) \, dy$$

$$= \frac{\sqrt{5}}{\sqrt{5}} \left[ \frac{3}{4\sqrt{5}} \left( 1 - \frac{y^2}{5} \right)^2 \right] \, dy = \frac{3}{5\sqrt{5}}$$

and

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

(9)

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 $\tau$. The range over which $\tau$ is varied is determined a priori from the known physical properties of the problem. The value of $\tau$ where $\text{AMAI}(\tau)$ 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 $x(t)$ we used low-pass filtered Gaussian noise, generated according to

$$s_1(t) = \rho_0 \cdot s_1(t - 1) + \sqrt{1 - \rho_0^2} n_1(t);$$

(10)
The correlation between neighboring samples was thus \( \rho \).

\[ n(t) \]

\( \text{var}(n) \) 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.

REFERENCES

[2] I. M. Gel'fand and A. M. Yaglom, "Calculation of the amount of
information about a random function contained in another such
[3] E. J. Wegman, "Nonparametric probability density estimation, I:
[4] ——, "Nonparametric probability density estimation, II: A comparision

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
total 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, NS012732, NS0536, and
NS30897.

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