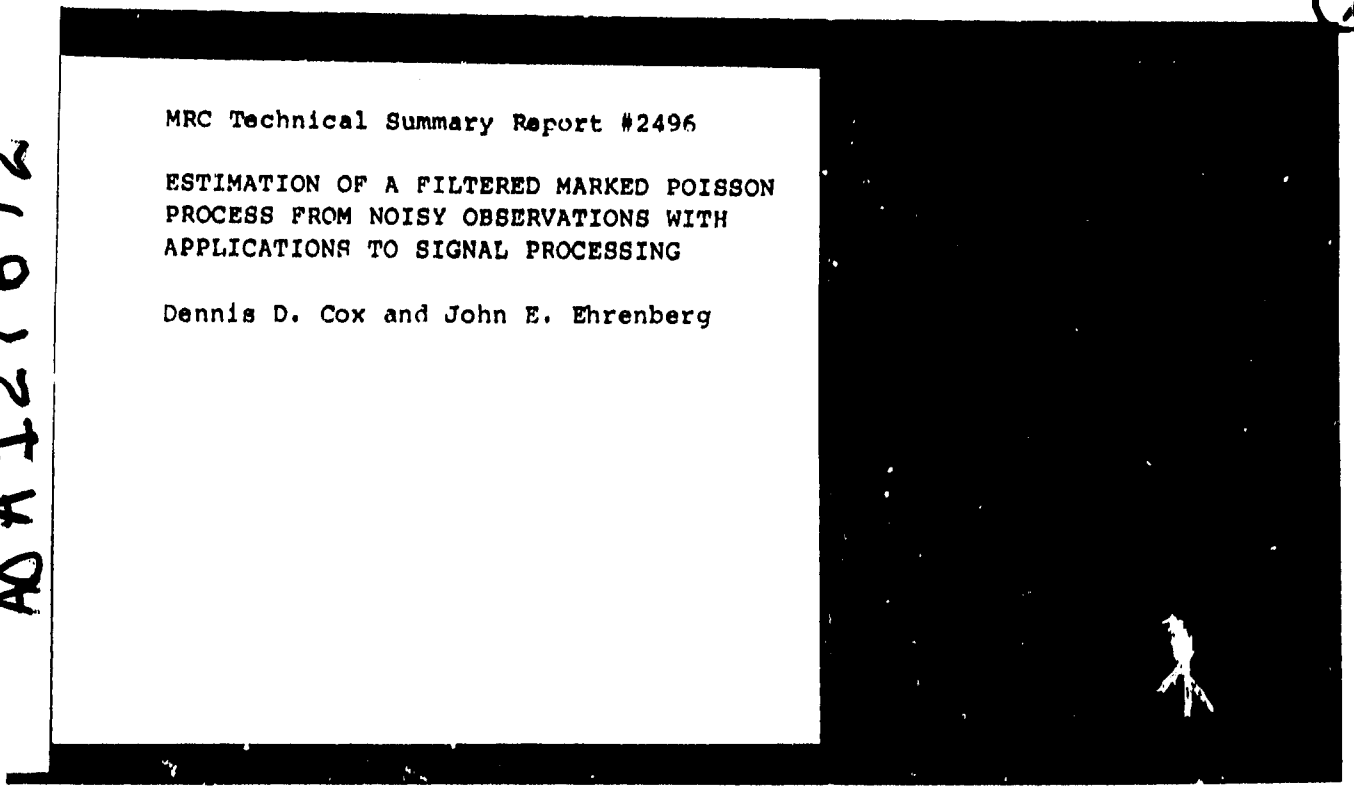
MRC Technical Summary Report #2496

ESTIMATION OF A FILTERED MARKED POISSON PROCESS FROM NOISY OBSERVATIONS WITH APPLICATIONS TO SIGNAL PROCESSING

Dennis D. Cox and John E. Ehrenberg



Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

March 1983

(Received June 11, 1982)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D.C. 20550

Office of Naval Research
Arlington, VA 22217

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

ESTIMATION OF A FILTERED MARKED POISSON PROCESS FROM NOISY OBSERVATIONS WITH APPLICATIONS TO SIGNAL PROCESSING

Dennis D. Cox*,[1,2,3] and John E. Ehrenberg**,[1]

Technical Summary Report #2496
March 1983

ABSTRACT

The problem of estimating the arrival times, amplitudes, and phases of an unknown number of signals in noise is treated. The signals are assumed to have a common, known waveform. A Bayesian model using a Poisson prior for the arrival times is specified, and a real time algorithm for computing the posterior mode is developed. Alternatively, the procedure may be looked upon as a penalized likelihood estimator with a penalty term which is a generalized form of Akaike's Information Criterion. Simulation results are presented which show that this approach can improve over classical, linear methods.

AMS(MOS) Subject Classification: Primary 62M99, Secondary 65U05, 60G55

Key Words: Filtered marked Poisson process, Bayesian estimation, penalized maximum likelihood estimation, Akaike's information criterion, real time signal processing

Work Unit Number 4 - Statistics and Probability

* Department of Statistics, University of Wisconsin-Madison.
** Applied Physics Laboratory and Department of Electrical Engineering, University of Washington-Seattle.

[1] Supported by the Office of Naval Research, Grant N0014-81-K0095.
[2] Sponsored by the U.S. Army under Contract No. DAAG29-80-C-0041.
[3] This material is based upon work supported by the National Science Foundation under Grant No. MCS-7927062, Mod.1.

SIGNIFICANCE AND EXPLANATION

The problem considered here is the extraction of an unknown number of signals from noise when the arrival times, amplitudes, and phases are unknown, but that the signals have a common, known waveform. This situation arises in sonar and radar signal processing. The signals are modelled as a filtered, marked, Poisson process and it is proposed that the maximum *a posteriori* estimate be used. A combinatorial problem arises involving the sets of possible arrival times, which has plagued many similar time series models. An algorithm is proposed for computing the estimate which circumvents this difficulty. Simulation results show the procedure can improve significantly on classical linear procedures.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

ESTIMATION OF A FILTERED MARKED POISSON PROCESS FROM NOISY OBSERVATIONS WITH APPLICATIONS TO SIGNAL PROCESSING

Dennis D. Cox*[1,2,3] and John E. Ehrenberg**[1]

1. Introduction

Consider the problem of estimating the arrival times, amplitudes, and phases of an unknown number of possibly overlapping signals with a known form of finite duration. A specific example of such a problem was treated by Ehrenberg, Ewart, and Morris (1978), denoted hereafter as EEM. Briefly, an acoustic pulse was transmitted underwater to a receiver some distance away. The received signal actually consisted of different replicas of the transmitted pulse, each arriving at a different time. The reason for this is the so called multipath phenomenon wherein the sound is transmitted over several paths of different lengths. The observations were contaminated by the usual ambient noise in the ocean and the internal noise of the receiving apparatus.

Similar problems arise in active sonar and radar wherein one sends out a pulse of energy and then receives the echoes, which are reflections of the pulse off of objects in the surrounding environment. The echoes have a known wave form, but the number of such echoes, their arrival times, amplitudes, and phases must be estimated. A more complete description of this problem is

* Department of Statistics, University of Wisconsin-Madison.
** Applied Physics Laboratory and Department of Electrical Engineering, University of Washington-Seattle.

[1] Supported by the Office of Naval Research, Grant N0014-81-K0095.
[2] Sponsored by the U.S. Army under Contract No. DAAG29-80-C-0041.
[3] This material is based upon work supported by the National Science Foundation under Grant No. MCS-7927062, Mod. 1.

given in section 3. A mathematically similar problem described by Barnard (1959) concerned a production process which undergoes level changes of unknown magnitudes and occurring at unknown times. The problem described by that author was different from the one considered here only in that the "signal" (a level change or step) is of infinite duration. The model we shall propose below is a simple case of a general time series model of Abraham and Box (1979), the so called "aberrent innovations" model. Those authors utilized an approximation wherein the number of signals is assumed small (≤ 2). We do not allow ourselves this luxury, so our analysis proceeds somewhat differently.

In EEM, a maximum likelihood estimation procedure was developed under the assumption that the noise is white, stationary, and Gaussian. For a given set of arrival times, the complex amplitudes of the associated signals were estimated by least squares. (The magnitude of the complex amplitude corresponds to the usual amplitude, and the argument of the complex amplitude corresponds to the phase.) For a given number of signals, all possible combinations of arrival times were examined, and the one with highest likelihood (minimal residual sum of squares) was chosen. This of course involves a rather heavy computational cost for two or more signals and a data record of even moderate length. The estimation of the actual number of signals presented considerable difficulty, since an increase in the number of assumed signals results in an increase in the likelihood, no matter what the true number of signals. The method for selecting the number of signals used in EEM was to match the residual sum of squares to the noise variance. This method seemed to work quite well.

In this paper, we propose a Bayesian modification of the maximum likelihood procedure of EEM, and an efficient algorithm for computing the

estimate. In Section 2, a stochastic model for the arrival times and signals is specified. The arrival times are assumed to be a realization of a Poisson point process (more precisely, a discrete time approximation thereto), and the amplitudes of the signals are assumed to be independent complex normal. It is then proposed that we take as estimate that collection of arrival times and associated signal parameters which maximize the posterior likelihood. The resulting estimate is called the maximum a posteriori (MAP) estimate. It can also be described as a penalized maximum likelihood procedure. The likelihood of EEM is penalized in two ways. Firstly, there is a penalty on the signal amplitudes from the independent normal assumption. This "shrinkage" penalty appears in many Bayesian analyses. A second penalty results from the Poisson process prior on the event times. It is this penalty which determines the estimated number of events. In the log likelihood, this penalty leads to a term proportional to the number of events or signals (see equation (2.6) below), and is thus a generalized information criterion for selecting the number of events (see Akaike, 1974). Such a penalty for the likelihood has been proposed for a different problem by Bhansali and Downham (1977). Bayesian derivations of information criterion type penalties have been given in other contexts by Schwarz (1978) and Akaike (1978).

A main contribution of this article is the description of a relatively simple algorithm for the approximate computation of the MAP estimate. This is described in Section 4. The simplicity of the algorithm relies heavily on the assumption that the signals are of duration which is short compared to the overall observation record. Hence, the change in likelihood resulting from inclusion or exclusion of a particular time from a proposed set of event times can be calculated using only the observations near that particular time. Based on this remark, we develop a hill climbing procedure which increases the

likelihood with each iteration. Furthermore, because the estimate depends on the observations only locally, it is possible to do the computations in real time, with a delay. This is especially important for signal processing applications.

One of the main difficulties with this problem comes from the overlapping of signals. Indeed, if each signal lasts only one discrete time unit, then the information concerning whether a particular point is an event time (and the information about the associated signal amplitude) is contained in the observation itself at that point. However, when the signals are of a duration of two or more time units, then the information is spread out over the whole observation record, because it is necessary to account for signals which may overlap with a possible signal at the time of interest, and to account for signals which may overlap with those which may overlap with the signal of interest, etc. In terms of the problem of EEM, when the acoustic pulses overlap, the resultant may have greater or less amplitude than the components, depending on their relative arrival times, amplitudes, and phases. This is accounted for in the likelihood equations, but determination of the signal parameters requires solution of a nondiagonal linear system which is different for each possible set of event times. Our method for dealing with this difficulty is given in Section 4.

In Section 5, we present performance results for the MAP estimation procedure, relying primarily on simulations. The simulations consisted of a signal process sample path generated according to the stochastic model for the arrival times, including complex Gaussian amplitudes, with Gaussian white noise added in. The algorithm, and a couple of competitors, was applied to estimate the signal process. These results show that for sufficiently small values of the event intensity parameter (expected number of signal arrivals

per unit time), the MAP estimate will significantly outperform the classical linear estimates. For larger values of the intensity, there will be sufficient overlapping of signals that the underlying process will lose its impulsive character and begin to look like an ordinary Gaussian process (see e.g. Snyder, 1975) for which a linear estimate is optimal. The most important results are presented in Figures 1 and 2, where we compare the predictive root mean squared errors of our procedure and the best linear estimate. Figure 3 is also important in that it shows that the error is relatively insensitive to misspecification of the constant in the generalized information criterion (by up to a factor of 8). Other results presented in Section 5 indicate how well the MAP procedure estimates the total number of events (Fig. 4), the number of signals whose amplitude exceeds a certain threshold (Figs. 5 and 6), and the arrival time of the largest amplitude signal (Fig.7). It is only with respect to this latter criterion that the classical methods appear to offer any serious competition for the MAP procedure.

2. Model Description and Likelihoods

In this section we specify a discrete time stochastic model for the signal process described in the introduction. This model will be related to a continuous time model for the active sonar problem in the next section. Because this application involves sinusoidal signals, it will be expedient to use complex valued stochastic processes. An asterisk * will denote the operation (on vectors or matrices) of taking the complex conjugate transpose. Given a complex μ and $\sigma^2 > 0$, by a complex Gaussian random variable Z with mean μ and variance σ^2 (written $Z \sim CN(\mu,\sigma^2)$), we mean that the real and imaginary parts of Z (written Re Z and Im Z) are independent real valued random variables with distributions $N(\text{Re } \mu,\sigma^2)$ and $N(\text{Im } \mu,\sigma^2)$, respectively.

The observed vector $\underset{\sim}{Y}$ is a (complex) N vector which is assumed to satisfy the regression equation

$$\underset{\sim}{Y} = W\underset{\sim}{a} + \underset{\sim}{\varepsilon}, \tag{2.1}$$

where the noise vector $\underset{\sim}{\varepsilon}$ has independent and identically distributed (abbreviated i.i.d.) components with the $CN(0,\sigma^2)$ distribution, W denotes an $N \times N$ Toeplitz matrix with

$$W_{jk} = w(j-k) \tag{2.2}$$

for some complex function w defined on the integers (referred to as a waveform), and $\underset{\sim}{a}$ is an N-vector of complex amplitudes whose components are i.i.d. with the mixture distribution

$$a_k \sim p\, CN(0,\rho^2) + (1-p)\, \delta_0 \,. \tag{2.3}$$

Here, $0 < p < 1$, $\rho^2 > 0$, and, δ_0 denotes a unit point mass at 0. Thus, with probability p, a_k is chosen according to the $CN(o,\rho^2)$ distribution, while with probability $1 - p$, a_k is 0. We further assume that for some positive integer m, the waveform w satisfies

$$w(k) = 0 \quad \text{if} \quad k < 0 \quad \text{or} \quad k > m, \tag{2.4}$$

so that W is an $(m+1)$ - banded matrix. We assume that the parameters σ^2, p, ρ^2, and $w(\cdot)$ are known, and consider the problem of estimating $\underset{\sim}{a}$ given the observation $\underset{\sim}{y}$.

Under the assumptions set forth above, we have a Bayesian regression problem, and the posterior distribution of $\underset{\sim}{a}$ given $\underset{\sim}{y}$ can be obtained in the usual way. In order to give a simple description, define the (random) set of integers

$$K = \{k: 1 \leq k \leq N, \ a_k \neq 0\}.$$

In other words, K is the set of event times, i.e. the times k when a_k is chosen from the $CN(0,\rho^2)$ distribution, at least with probability one. We also put

$$|K| = \text{number of elements (cardinality) of } K$$

and write K in the form

$$K = \{k(1), k(2), \ldots, k(|K|)\}$$

with

$$k(1) < k(2) < \ldots < k(|K|),$$

and put

$$\underset{\sim}{a}_K = \big(a(k(1)), \ldots, a(k(|K|))\big)'.$$

We shall frequently pass from subscript notation to functional notation as in this last equation wherein we wrote $a(k)$ for a_k. We also define the $N \times |K|$ matrix W_K, which is the submatrix of W obtained by deleting the

columns whose indices are not in K. Hence, the equation

$$W\underset{\sim}{a} = W_K \underset{\sim}{a}_K$$

holds.

The posterior density for K, $\underset{\sim}{a}_K$ has the form

$$f(K, \underset{\sim}{a}_K | \underset{\sim}{y}) = c(\underset{\sim}{y}) \cdot [p/(1-p)]^{|K|} \cdot [2\pi\rho^2]^{-|K|} \cdot \tag{2.5}$$

$$\exp[\sigma^{-2} \mathrm{Re}\ \underset{\sim}{a}_K^* W_K^* \underset{\sim}{y} - \tfrac{1}{2} a_K^* (\rho^{-2} I_{|K|} + \sigma^{-2} W_K^* W_K) a_K].$$

Here, $c(\underset{\sim}{y})$ depends only on $\underset{\sim}{y}$ (and σ^2, p, ρ^2, and w) but not on K and $\underset{\sim}{a}_K$, and I_n is an $|n| \times |n|$ identity.

Given the posterior density (2.5), it is possible at least in principle to obtain an optimal Bayes estimate of the amplitude vector $\underset{\sim}{a}$, such as the posterior mean. This will involve a complicated integration. For each $K \subseteq \{1,2,\ldots,N\}$, it is possible (in principle) to integrate out $\underset{\sim}{a}_K$. This involves computing the inverse and determinant of a matrix which depends on K. However, there are 2^N values of K, so that an exact procedure is impractical for even moderate sample sizes N. Under certain circumstances, the problem is tractable (e.g. when signal length is 1, i.e. $m = 0$), and there are various approximations that would work sometimes. For example, if p is sufficiently small, then only values of K for which $|K| \leq 1$ would contribute to the integral. However, we are interested in the general case where such approximate methods would not be accurate.

One estimate which is usually easier to compute than the posterior mean is the posterior mode, also known as the maximum a posteriori (MAP) estimate. In this situation, the MAP estimate $\hat{K}$, $\hat{\underset{\sim}{a}}_{\hat{K}}$ are those values of

$K, \underset{\sim}{a}_K$ which maximize the posterior likelihood (2.5), or equivalently, minimize

$$L(K, \underset{\sim}{a}_K | \underset{\sim}{\chi}) = -\mathrm{Re}\ \underset{\sim}{a}_K^* \underset{\sim}{\mu}_K + \tfrac{1}{2} \underset{\sim}{a}_K^* Q_K \underset{\sim}{a}_K + |K|\alpha \tag{2.6}$$

$$= -\mathrm{Re}\ \underset{\sim}{a}(K)^* \underset{\sim}{\mu} + \tfrac{1}{2} \underset{\sim}{a}(K)^* Q\ \underset{\sim}{a}(K) + |K|\alpha.$$

Here we have introduced the notations

$$\begin{aligned} \underset{\sim}{\mu} &= W^* \underset{\sim}{\chi} \in C^N \\ \underset{\sim}{\mu}_K &= W_K^* \underset{\sim}{\chi} \in C^{|K|} \\ Q &= (\sigma/\rho)^2 I_N + W^* W \\ Q_K &= (\sigma/\rho)^2 I_{|K|} + W_K^* W_K \\ \alpha &= -\sigma^2 \log \{p/(2\pi\rho^2[1-p])\}. \end{aligned} \tag{2.7}$$

We have also written $\underset{\sim}{a}(K)$ for $\underset{\sim}{a} \in C^N$ to make explicit the set K of indices corresponding to nonzero entries.

Now, for a given set $K \subseteq \{1,2,\ldots,N\}$, we may obtain the MAP estimate of $\underset{\sim}{a}$, denoted $\hat{\underset{\sim}{a}}(K) \in C^N$ (or $\hat{\underset{\sim}{a}}_K \in C^{|K|}$ if the zero components are deleted) by solving a $|K| \times |K|$ linear system. The result is

$$\hat{\underset{\sim}{a}}_K = Q_K^{-1} \underset{\sim}{\mu}_K \tag{2.8}$$

and then the negative log likelihood becomes

$$L(K, \hat{\underset{\sim}{a}}_K | \underset{\sim}{\chi}) = -\tfrac{1}{2} \hat{\underset{\sim}{a}}_K^* \underset{\sim}{\mu}_K + |K|\alpha.$$

We are then obliged to search over all 2^N values of K to find a $\hat{K}$ which minimizes the last expression, leading us back to the same dilemma as presented by the posterior mean computation. Since $p \approx 0$, it is reasonable to suppose that $\alpha > 0$, whereas if $\alpha \leq 0$, we simply have $\hat{K} = \{1,2,\ldots,N\}$.

In Section 4 below, we present a hill climbing algorithm which attempts to find $\hat{K}$, or an approximation thereto.

We would like to now briefly point out some of the salient features of the MAP estimation procedure. Firstly, note that the problem of determining $\hat{a}_K$ for a given K (the solution given in (2.8)) is a standard regression problem, although our use of a Gaussian prior introduces some shrinkage (the $(\sigma/\rho)^2 I_K$ terms in (2.7)). A non-Bayesian may be content to consider this as a ridge regression type estimate. There is now a considerable body of literature on how to estimate the ridge parameter (or the hyperparameters in the Bayesian model), and such estimates could be transformed into estimates of ρ^2.

Within this regression context, the estimation of the set K of event times translates into a model selection problem, wherein K corresponds to the set of variables in the model. One then recognizes (2.6) as a penalized (posterior) log likelihood, where the penalty term $|K|\alpha$ corresponds to a generalized version of Akaike's Information Criterion (AIC), as described in the introduction.

We will now describe two other procedures which could be applied to this problem, which we refer to as the matched filter estimate (abbreviated MFE) and best linear estimate (BLE). Still other methods are described in EEM. If one believed that there was an event at time k whose effect did not overlap with any others, then an unbiased least squares estimate of the amplitude is

$$\bar{a}_k = \mu_k/[Q_{kk} - (\sigma/\rho)^2].$$

In one form or another, such estimates are widely used in signal processing, and will be referred to as matched filter estimates. Our MAP procedure is a modification of this matched filter in that the shrinkage term $(\sigma/\rho)^2$ is

included in the denominator, and an effort is made to take into account the overlaps of the various effects (which leads to the system of equations whose solution is given in (2.8)). Indeed, if there is no overlap (i.e. if $m = 0$). then the MAP estimate is essentially the MF estimate for those times where the MFE is above a suitable threshold determined by the information penalty. Under these circumstances, the MAP procedure just amounts to simultaneous signal detection and estimation. However, when overlapping occurs, the MFE becomes quite confused and will indicate the presence of large amplitude signals when sufficiently many small signals overlap and add more or less in phase. The MAP procedure provides a means for correcting this defect.

Another approach would be to use the linear estimate which is optimal for the Bayesian model (over all linear estimates) in the sense of minimizing mean squared error. This estimate is the minimizer over $\underset{\sim}{a} \in C^N$ of the quadratic function

$$(2.9) \qquad 1/(2p\rho^2)\ \underset{\sim}{a}^*\underset{\sim}{a} + 1/(2\sigma^2)(\underset{\sim}{y} - W\underset{\sim}{a})^*(\underset{\sim}{y} - W\underset{\sim}{a}).$$

We will refer to this estimate as the BLE.

Both the MF and BL estimates have the difficulty that they don't take into account the impulsive nature of the stochastic model for $\underset{\sim}{a}$. Both procedures will almost never estimate $a_j = 0$ for some j, and hence do not immediately give estimates for the event times. Also, both methods have difficulty separating a large number of small overlapping signals from a single large signal, and they will give large amplitude estimates at times near a large signal rather than just at its single event time. The MAP estimate, however, does not suffer from these problems.

3. Applications to Sonar Signal Processing

In this section we briefly describe a class of real problems which motivated th.; current study, and show how to link them up with the model of the previous section. The use of sonar and radar models involving filtered Poisson processes is not new. Middleton (1979) has developed a very general model incorporating such a noise component and has written at length on the various estimation problems for such models. See also Ehrenberg (1976). In active sonar, a narrowband sonic pulse of known frequency f_o sec.$^{-1}$ and with a known envelope function $w_o(t)$ is transmitted, and the echoes formed by reflections of the sonic pulse off of solid objects are received. The signal process which results from many such echoes is commonly called reverberation in the sonar literature and clutter in the radar literature. The echoes have unknown return times (depending on the distance of the reflecting object) and amplitudes (depending on the size and distance of the object), but, if the reflecting objects are stationary so that there is no Doppler shift, the echoes are of the same frequency and shape as the transmitted pulse. Therefore, the received echoes as a function of time constitute a "signal" of the form

$$S(t) = \mathrm{Re} \sum_j A(j)\, w_0(t - \tau(j))\, \exp[2\pi i f_0 (t - \tau(j))]$$

where $A(j)$ and $\tau(j)$ are the complex amplitude and arrival time of the j^{th} echo. In practice, there is always noise contaminating the signal, so that the observation model is

$$Y(t) = S(t) + N(t), \quad 0 \leq t \leq T$$

where $N(\cdot)$ denotes the noise, which we assume is white Gaussian noise with two sided power spectral density $\gamma > 0$.

The initial stage of the signal processing involves reduction of the continuous time data to the discrete time quadrature demodulator statistics given by

$$y_k = \frac{N}{T} \int_{(k-1)T/N}^{k\ T/N} y(t) \exp\ [-2\pi i f_0 t]\ dt, \quad \text{for } 1 \leq k \leq N.$$

It is assumed that $f_0 T/N$ is an integer, and also that w_0 is nearly constant over a time interval of length T/N. Under the white noise assumption, the noise vector $\underset{\sim}{\varepsilon}$ is given by

$$\varepsilon_k = \frac{N}{T} \int_{(k-1)T/N}^{kT/N} N(t) \exp\ [-2\pi i f_0 t]\ dt$$

where these components are i.i.d. $CN(0, \sigma^2)$ distribution, with

$$\sigma^2 = \gamma N/T. \tag{3.1}$$

Now suppose that the echo arrival times $\tau(1), \tau(2), \ldots$ are the event times of a homogeneous Poisson process with constant intensity λ, and that

$$p = \lambda\ T/N \tag{3.2}$$

is sufficiently small that we may assume that at most one echo arrives during one of the Fourier processing intervals of duration T/N. Further assume that $w_0(\cdot)$ is sufficiently smooth so as to be nearly constant over each Fourier processing interval and then set

(3.3) $$w(k) = w_0(k\ N/T).$$

Finally, suppose that $f_0\ T/N >> 1$ so that the phase of the arriving signal is essentially random, and that the magnitude $|A(j)|$ of $A(j)$ has the Rayleigh distribution with parameter ρ^2 (see Middleton, 1960 for definitions), and hence the complex amplitude of the arriving signal will have the $CN(0,\rho^2)$ distribution. This discussion, together with equations (3.1) through (3.3), specifies how to fit this sonar problem into framework of the model defined in Section 2. We note that the problem treated in EEM may also be fit in this framework in an obvious way, if the amplitudes are Rayleigh.

A possibly better model of the active sonar problem would be obtained by assuming that the solid scatterers which produce the echoes have a homogeneous spatial Poisson distribution with parameter α per unit of volume and that at a unit distance from the transmitter, their echo amplitude would be $CN(0,\rho_1^2)$. Also suppose, for the sake of simplicity, that the sonar transmit beam is conical, subtending a solid angle of measure ω. Then the echo arrival time process is a nonhomogeneous Poisson process with rate function

$$\lambda(t) = \alpha\omega c^3 t^2,$$

c being sound velocity, and the complex amplitude of an echo arriving at time τ is $CN(0,\rho^2(\tau))$, where

$$\rho(\tau) = \rho_1/(c\tau)\ .$$

These complications can be easily incorporated in the model of section 2, and will change the log likelihood (2.6). For example, the term $|K|\alpha$ will be replaced by

$$\sum_{k \in K} \alpha(k)$$

where

$$\alpha(k) = -\log \left[\{\lambda(kT/N)T/N\} / \{2\pi\rho^2(kT/N)(1-\lambda(kT/N)T/N)\} \right].$$

We maintain that the necessary modifications are more notational than substantial, and also that an understanding of the MAP estimation problem posed in Section 2 will immediately lead to an understanding of the corresponding MAP estimation procedure for this more complicated problem.

4. Recursive Computation of the MAP Estimate

In this section, we describe an algorithm for approximate computation of the MAP estimate described in Section 2. This algorithm starts with an initial estimate of K and the corresponding $\hat{\underset{\sim}{a}}_K$ given by (2.8) (e.g. initially $K = \phi$, in which case no $\hat{\underset{\sim}{a}}_K$ is required), and produces a final estimate of K, together with the corresponding $\hat{\underset{\sim}{a}}_K$, which has higher posterior likelihood. One of the main features of the procedure is that it can be implemented in such a way that only one pass through the data is required, and such that at any time only a small span of the data is being processed. This is significant to statisticians with large data sets, and to engineers who are interested in real time processing.

We now give an overview of the algorithm via an hierarchal description, proceeding from macro to micro. The interested reader may consult the Fortran program segment listed in Appendix A.

(i) There are several (say n) applications of the posterior likelihood increasing procedure described in stage (ii). These can be working simultaneously on adjacent sections of the data stream.

(ii) Local perturbations in the set K of estimated event times are successively examined. With each such perturbation, the updated amplitude vector estimate $\hat{\underset{\sim}{a}}_K$ and the log likelihood increment are computed as described in (iii) and (iv). The simplest example of such a perturbation is, for a particular time epoch j, to augment the current K with j if $j \notin K$, or to delete j from K if $j \in K$. Afterwards, the negative log likelihood is checked for a decrease. If there is one, the K is updated, and otherwise it is left unchanged. The processor would then proceed to the $(j + 1)^{st}$ epoch to repeat the procedure there, and so forth.

(iii) Each time K is perturbed, say from K_0 to K_1, it is necessary to update $\hat{\underset{\sim}{a}}_K$, from $\hat{\underset{\sim}{a}}(K_0)$ to $\hat{\underset{\sim}{a}}(K_1)$, which involves solving a new linear system (see (2.8)). A method for doing this is described in (iv). The increment in the negative log likelihood is evaluated by

$$\Delta L = -\tfrac{1}{2}\left(\hat{\underset{\sim}{a}}(K_1) - \hat{\underset{\sim}{a}}(K_0)\right)^* \underset{\sim}{\mu} + (|K_1| - |K_0|)\ \alpha.$$

If $\Delta L < 0$, then K_0 is replaced by K_1, and otherwise K_0 is left unmodified.

(iv) Given K_0, $\hat{\underset{\sim}{a}}(K_0)$, and K_1, it is required to compute $\hat{\underset{\sim}{a}}(K_1)$. Since K_1 is obtained by a local perturbation on K_0, $\hat{\underset{\sim}{a}}(K_1)$ should be obtainable by locally updating $\hat{\underset{\sim}{a}}(K_0)$. This can be accomplished by a few applications of a Gauss-Seidel recursion. A single iteration involves replacing the k^{th} component of $\hat{\underset{\sim}{a}}_K$ by the updated value

$$\hat{a}_{K,k} = \left(\mu_{K,k} - \sum_{j \neq k} Q_{K,kj}\ \hat{a}_{K,j}\right)/Q_{K,kk}$$

Since Q_K is $2m + 1$ banded, there are really only $2m$ terms in the summation.

We now indicate the span of the data stream which is involved in the updating at any one time. Each Gauss-Seidel iteration involves $(2m + 1)$ adjacent data points, and if M such iterations are performed, then $2mM + 1$ adjacent data points are included. Note that if we are updating $\hat{\underset{\sim}{a}}$ as a result of a perturbation of K at j, this entire updating procedure must be completed before moving on to the perturbation at $j + 1$. However, the Gauss-Seidel iterations can be condensed into a single matrix multiplication of the $2mM + 1$ data points by a square matrix. Now if n of the overall

likelihood increasing steps in (i) are applied, then a span of $n(2mM + 1)$ data points are actually being processed at one time. Our simulation experience (results of which are reported in the next section) indicate $n = 2$ and $M = 3$ are usually sufficient.

We now indicate two modifications of the basic algorithm which were used for the simulations reported in the next section. One can expand the local perturbation in the set K as follows. For a given time epoch j of interest, one can examine all 8 combinations of inclusion/exclusion of the epochs j, $j + 1$, and $j + 2$ from the set K, compute the likelihood change for each, and modify K accordingly. Then at the next step, look at $j + 1$, $j + 2$, and $j + 3$, etc. This procedure can be extended to looking at blocks of times of length 4, 5, etc. Also, it is not necessary to use successive blocks of epochs, one could say skip from j to $j + 2$, as long as there was substantial overlap of the blocks. However, we stayed with the simple version of successive blocks of 3. Another modification is to run through the data with the Gauss Seidel smoothing procedure, but keeping the current K fixed. This serves to give better accuracy in solving the linear system of equations corresponding to K.

5. Simulation Results

In this section, we present simulation results for the MAP estimate. The purpose of these simulations is to show that the MAP procedure can be implemented with a small number of each of the basic steps and still perform quite well in comparison with other procedures. The good performance should come as no suprise, given the results presented in EEM. What is significant is that our algorithm achieves this level of performance. The MAP estimate was compared wherever possible with the MF and BL estimates, and only for one performance measure (estimating the maximum amplitude event time) were they serious competitors. There are four different performance criteria which we consider, which are estimation of the total signal (the filtered point process, or the superposition of all the individual signals). the total number of signals, the number of signals with amplitudes larger than a certain threshold, and the event time of the maximum amplitude signal. We also considered the performance of the MAP estimate when the information penalty factor α (see (2.7)) is misspecified, and found that it was surprisingly insensitive to this misspecification. This is also apparent in the simulation results of Bhansali and Downham (1977). Perhaps this is one of the reasons AIC has worked so well in practice.

We now specify some of the parameters of the Monte Carlo experiment. All points plotted in the accompanying figures represent an average of 25 trials of $N = 100$ observations each. In all cases, the waveform w was unity, i.e.

$$w(j) = 1 \quad \text{for} \quad 0 \leq j \leq m. \tag{5.1}$$

The signal length $(m + 1)$ was either 4, 7, or 10. We use L to denote $(m + 1)$ for convenience. The noise variance σ^2 was fixed at 1. The event

probability p was always chosen as a power of $1/2$, either .25, .125, .0625, or .03125, corresponding to values of 2, 3, 4, or 5 for $-\log_2 p$. Note that the expected number of signals being received at any particular time is just Lp. The amplitude variance ρ^2 was chosen so that the quantity

$$c = p\rho^2$$

took on one of the values .25, .5, 1, or 2. Hence, as p decreases ($-\log_2 p$ increases), the arriving signals become fewer and larger in amplitude. Note that the process covariance is determined by c, L, and σ^2, so that specifying c was convenient for the BLE. This also means that the signal to noise ratio is the same for the simulations at the same value of c. In all cases, $n = 3$ passes with the basic perturbation algorithm were made (step (i) of the previous section), and $M = 3$ Gauss Seidel iterations (step (iv)) were performed at each step. Also, there was one final pass through the data with a Gauss-Seidel iteration, as described in the last section.

The first performance criterion we examine is the integrated mean squared error (IMSE) in the estimate of total signal process, where the total signal process is taken to be the sum of all the individual signals (received echoes). More formally

$$\text{IMSE} = E \, || \, W(\hat{\underset{\sim}{a}} - \underset{\sim}{a}) \, ||^2,$$

where $||\cdot||$ denotes the usual norm on C^N. The ordinates in Figures 1 through 3 are the square root of the IMSE (referred to as just the RMS error), divided by the total RMS signal power, which is given by

$$(\text{RMS signal power}) = [NpL(2\rho^2)]^{1/2} = [2NcL]^{1/2} .$$

Note that the total noise power is just $2N\sigma^2$. The estimates of the normalized RMS error, each based on a sample of 25 simulation trials, are reasonably accurate. In no case did the (estimated) standard error exceed .025, and was usually less than .015. Three points in Figure 1 were replicated, and one notes how close the values were in each case.

On Figures 1 and 2, we have also shown the normalized steady state RMS error for the BLE. Using standard spectral methods as in Koopmans (1974), the steady state MSE per observation is given by (see Appendix B)

$$\frac{\sigma^2}{\pi} \int_{-\pi}^{\pi} \frac{|\hat{w}(\lambda)|^2}{c^{-1}\sigma^2 + |\hat{w}(\lambda)|^2} \, d\lambda ,$$

where

$$|\hat{w}(\lambda)|^2 = \sin^2 (L\lambda/2)/\sin^2 (\lambda/2) .$$

This will in fact be a lower bound for the exact IMSE for the BLE for a finite data record because of boundary effects, and the accuracy will decrease with increasing L or c. However, we believe these numbers are reliable. The main conclusion to be drawn from Figures 1 and 2 is that for $p \lesssim .08$, the MAP estimate does better than the BLE, in terms of IMSE. For $L = 10$ and $p = .08$, one expects .8 signals being received at any one time (total of 8 expected events in the whole record), so there is considerable overlapping of signals. Also, on these two figures are shown some limiting IMSE values for the MAP estimate as $p \to 0$ (holding $p\rho^2 = c$ constant). Under these conditions, there will be essentially no overlapping effects and the event time set K will be perfectly estimated. It can then be shown (Appendix C) that

$$IMSE(p\to 0) = 2NpL\sigma^2/(\sigma^2/\rho^2 + L).$$

In any event, this is always a lower bound for the MAP IMSE, with decreasing accuracy for increasing L.

Figure 3 shows a plot for the normalized RMS error when the penalty α is mismatched. The "correct" value, denoted just α, is given in (2.7), but this is not necessarily optimal. Denoting by $\bar{\alpha}$ the value which was in fact used, we set $\bar{\alpha} = \alpha$ in all simulations except for those reported in Figures 3 and 6. It is clear from Figure 3 that α is very nearly the optimal value (for minimizing IMSE) and furthermore that the IMSE of the estimate is quite insensitive to the actual $\bar{\alpha}$ used.

Our second performance criterion is the error in estimating the total number of individual signals. In Figure 4, the ordinates are the RMS error in the estimate of the total number of signals in the simulated data record of 100 observations, normalized by dividing by $(Np(1 - p))^{1/2}$, the standard deviation of the number of signals. Note that this latter quantity would be the RMS error if one used the constant Np as the estimate of the number of signals. There is considerable sampling variability in the results of Figure 4, as is clearly seen by comparing the replicated points. However, for $p \lesssim .1$, it appears that the MAP estimate does do some good. In simulations with mismatched α, the performance with respect to this indicator degraded quite severely. The reason for this is of course the fact that there are many signals of small magnitude which are difficult to accurately detect, and if $\bar{\alpha}$ is set very low to detect them, then many false detections are made. Also note that there is no simple way to use the MFE or BLE to estimate the total number of signals.

Figures 5 and 6 present the RMS error in the MAP estimate of the number of signals for which the amplitude $|a_j|$ exceeded ρ. The estimate was obtained by counting the number of MAP amplitude estimates exceeding ρ in magnitude. For a given signal, the probability its amplitude will exceed ρ is just $P[\chi_2^2 > 1] = e^{-1/2} \doteq .6065$, so $.6065Np$ such signals are expected, and the standard deviation of the number of such signals is

$$[.6065Np\ (1 - .6065p)]^{1/2}.$$

The RMS errors of the MAP estimate were divided by this last quantity for normalization. In Figure 5, we see that the MAP estimate is not significantly better than the constant estimate of $.6065Np$ for $p = .25$, but exhibits sharp improvement as p decreases. Compare the slope in Figure 5 with that in Figure 4. We also note from Figure 6 that, at least for $p = .0625$, the performance is relatively stable under misspecification of the information penalty fact α by a factor of 2. In that figure, the point $p = .125$, $\bar{\alpha}/\alpha = .5$ was replicated because the first value (with error /st. dev. value of .9) indicated a discontinuity in going from $\bar{\alpha}/\alpha = .5$ to 1. However, the second value did not clear up matters very much. These results of Figure 6 are important to those who are interested in just estimating the number of signals, or the event probability p, since p only enters in the procedure through α. It is of course possible to use the BL and MF estimates to estimate the number of signals with amplitudes exceeding ρ, and we did in fact simulate these. The BLE was much better than the MFE in this regard, but was always much worse than the MAP estimate with RMS errors two or more times those of the MAP estimate.

Our final performance criterion concerns the estimation of the arrival time of the largest amplitude signal, abbreviated MAET for maximum amplitude event time. For each type of estimation procedure, the MAET was estimated by the time of the largest estimated amplitude. Some results are presented in Figure 7, where the ordinate represents the fraction of 25 simulation runs that the procedure picked the correct MAET. We should note that with very few exceptions, each of the procedures chose exactly the correct time, or was off by three or more time periods. Our expectation was that the MFE and BLE would be frequently "fooled" by overlapping signals which add to give a large amplitude, whereas the MAP would separate out the individual signals. However, the results seem to indicate rough equality of the three different estimates. We do note that there is an apparent trend for the MAP estimate to perform better than the other two for small values of p. Unfortunately, it would take a large number of simulation trials to validate this.

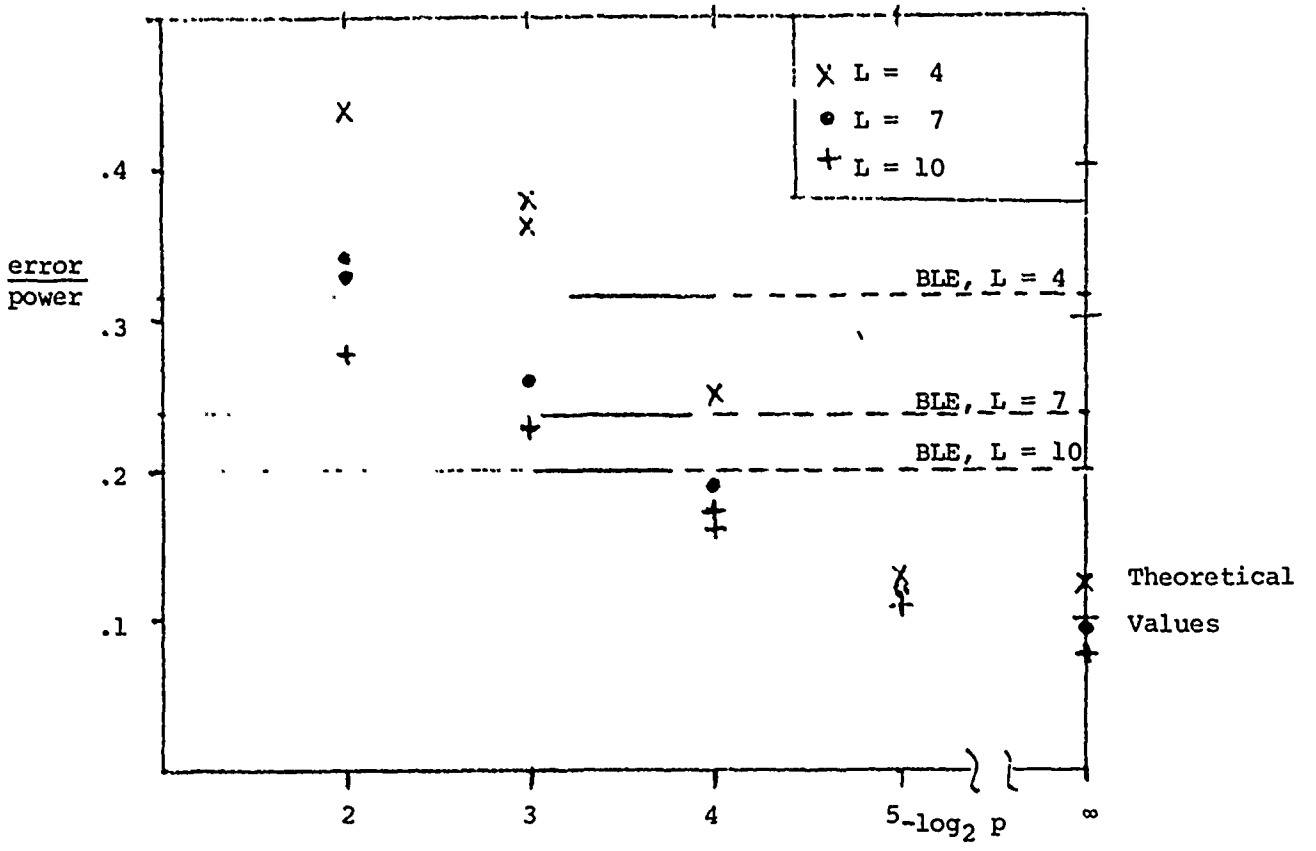


Figure 1. Normalized RMS signal estimation error for different signal lengths. (c = .5)

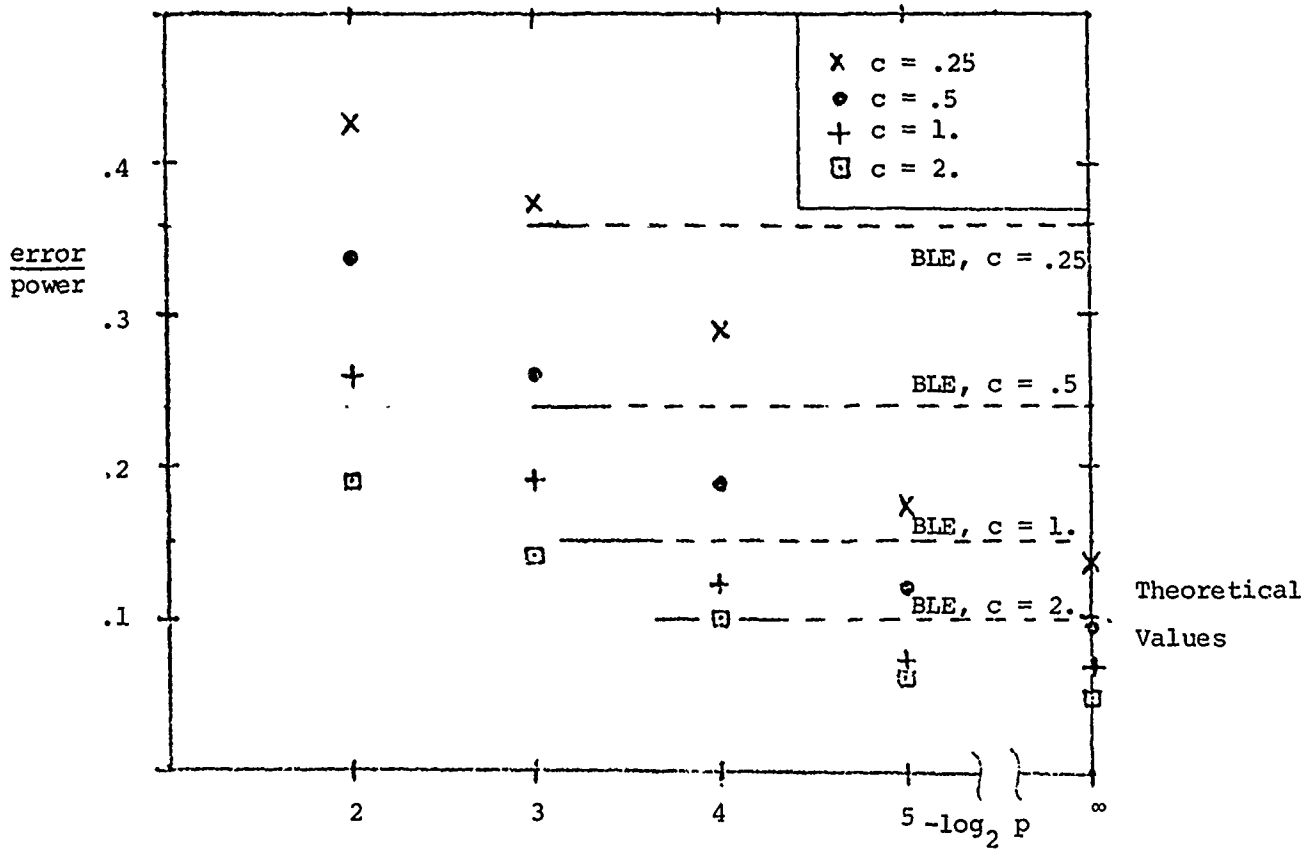


Figure 2. Normalized RMS signal estimation error for different covariances. (L = 7)

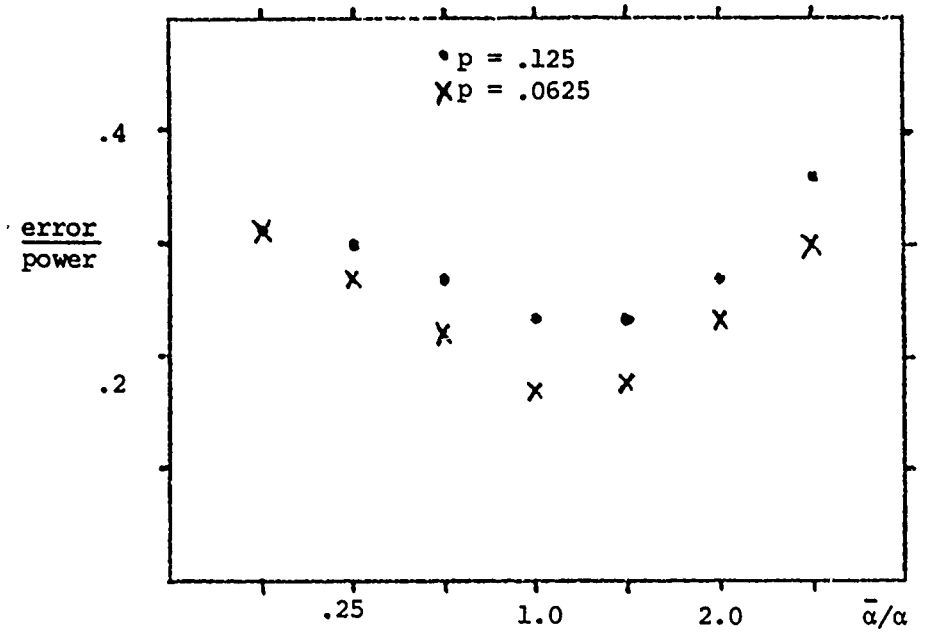


Figure 3. RMS Signal estimation error when the information penalty factor is misspecified. (c = .5)

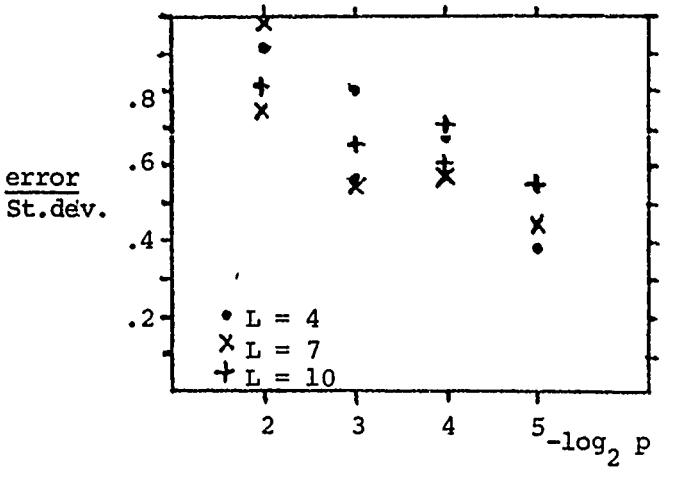


Figure 4. RMS error in the estimate of the total number of signals. (c = .5)

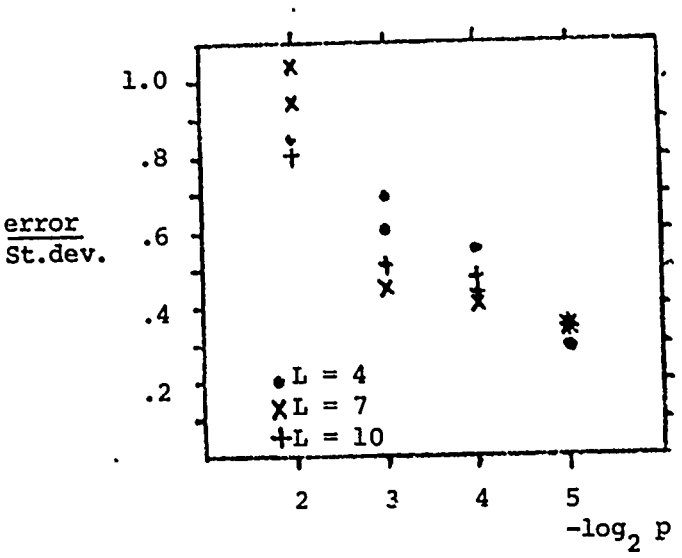


Figure 5. RMS error in the estimate of the number of signals with amplitude bigger than ρ. (c = .5)

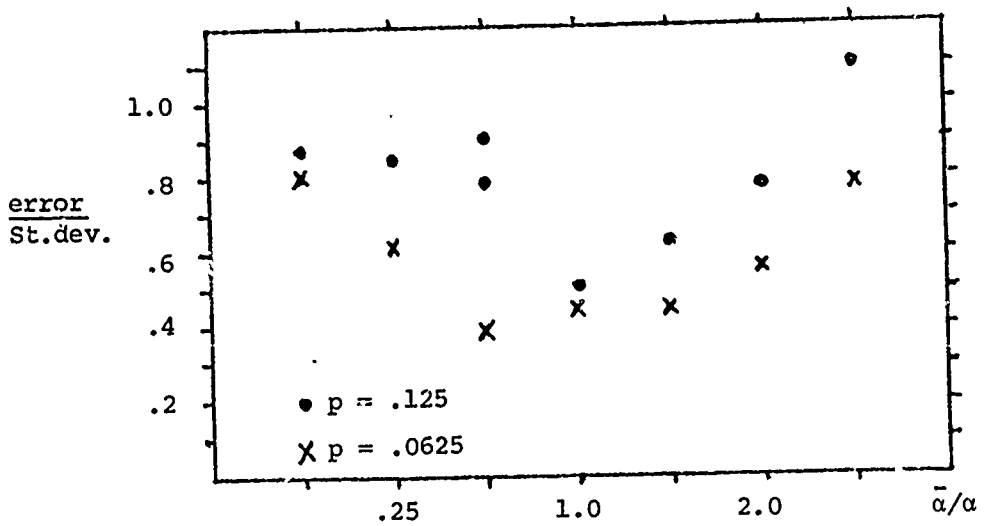


Figure 6. RMS error in the estimate of the number of signals with amplitude bigger than ρ when the information penalty factor is misspecified (c = .5)

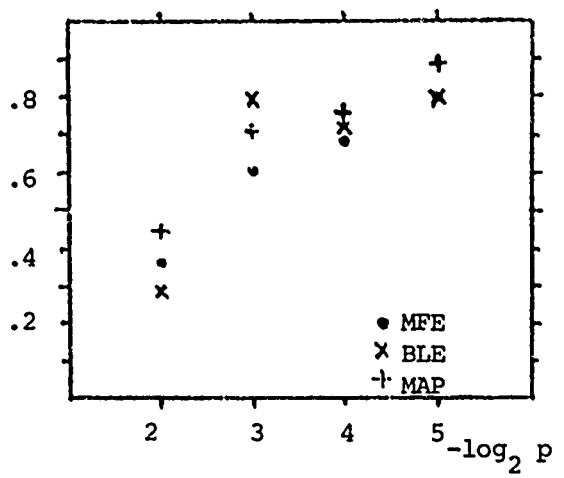


Figure 7. Proportion of correctly estimated maximum amplitude event times. (c = .5)

6. Summary

The statistical problem treated here has been investigated by other authors, but there has been relatively little progress for moderate values of the event time intensity parameter p. This is because of the apparent need to examine at least a large fraction of the collection of all sets of event times. We have proposed an estimation procedure which avoids this difficulty, and in fact can be implemented in a real time algorithm. Simulation results indicate that the procedure can do much better in many respects than classical methods. We hope that this work will stimulate others to improve upon the basic algorithm, and to develop a better understanding of its statistical properties.

REFERENCES

[1] Abraham, B., and Box, G.E.P. (1979), "Bayesian analysis of some outlier problems in time series", Biometrika 66, 229-36.

[2] Akaike, H. (1974), "A new look at the statistical model identification", IEEE Trans. Automat. Contr. AC 19, 716-23.

[3] Akaike, H. (1978), "A Bayesian analysis of the minimum AIC procedure", Ann. Inst. Statist. Math. 30, 9-14.

[4] Barnard, G. (1959), "Control charts and stochastic processes", Jrnl. Royal. Statist. Soc., series B, 21, 239-71.

[5] Bhansali, R.J. & Downham. D.Y. (1977) "Some properties of the order of an autoregressive model selected by a generalization of Akaike's FPE criterion", Biometrika 64, 547-51.

[6] Ehrenberg, J. (1976), "An approximate bivariate density function for orthogonal filtered Poisson processes with applications to volume reverberation problems", Jrnl. Acoust. Soc. Am. 59, 374-80.

[7] Ehrenberg, J., Ewart, T., and Morris, R. (1978), "Signal processing techniques for resolving individual pulses in a multipath signal, Jrnl. Acoust. Soc. Am. 63, 1861-65.

[8] Koopmans, L. (1974), Spectral Analysis of Time Series, Academic Press.

[9] Middleton, D. (1960) An Introduction to Statistical Communication Theory, McGraw-Hill, New York.

[10] Middleton, D. (1979), "Canonical non-Gaussian noise models: their implications for measurement and for prediction of receiver performance", IEEE Trans. Electromag. Compat., EMC 21, 209-20.

[11] Schwarz, G. (1978), "Estimating the dimension of a model", Ann. Statist 6, 461-464.

[12] Snyder, D.L. (1975) Random Point Processes, Wiley, New York.

Appendix A; An ASCII Fortran program segment for implementing the MAP estimation algorithm. The complex array A stores the current set of amplitude estimates, with many elements equal to 0, and the array B is used for scratch work to determine the effect of a perturbation in K. The internal subroutine MAPINC determines the effect of the change in K, using the Gauss-Seidel computations in CINC. The vector μ is in RMU, and the matrix Q is stored in a 1 dimensional array since it is Toeplitz and Hermitian. The outer loop (DO 20) is to implement (i), while the big IF...THEN...ELSE implements (ii). The change in negative log posterior likelihood (DLIK) is computed by LIKCHG. If the change is negative, the A is replaced by B with BTOA, and otherwise B is restored to be equal to A by ATOB

```
      DO 10   J = 1, N
      A(J) = 0.
10    B(J) = 0.

      DO 20   IND = 1, NØ
      DO 20   J = 1, N
      I1 = J
      J1 = J

      IF (CABS (A(J)).LT. 1.E-10) THEN
      B(J) = CINC (J)
      DO 40  I = 1,M
      CALL MAPINC
40    CONTINUE

      CALL LIKCHG

      DLIK = -ALPHA + DLIK *.5

      ELSE
      B(J) = 0.
      DO 50 I = 1,M
      CALL MAPINC
50    CØNTINUE
      CALL LIKCHG
      DLIK = ALPHA + DLIK *.5
      END IF

      IF (DLIK .GT.0) THEN
        CALL ATOB
      ELSE
        CALL BTOA
      END IF
      20 CONTINUE

      SUBROUTINE MAPINC
      I1 = MAX(1, I1 - L + 1)
      J1 = MIN(N, J1 + L - 1)
      DO 10  I = I1, J1
      IF (CABS(B(I)), GT. 1.E - 10) B(I) = CINC(I)
10    CONTINUE
      RETURN
```

```
      FUNCTION CINC (I)
      CØMPLEX CINC
      I3 = MAX (1, I - L + 1)
      J3 = MIN (N, I + L - 1)
      B(I) = 0.
      IP1 = I + 1
      IM1 = I - 1
      DO  10  JI = I3, IM1
10    B(I) = B(I) - CØNJG (Q(I - JI + 1)) * B(JI)
      DO  20  JI = IP1, J3
20    B(I) = B(I) - Q(JI - I + 1) * B(JI)
      CINC = (RMU(I) + B(I))/Q(1)
      RETURN

      SUBROUTINE LIKCHG
      I2 = MAX(1, I1 - L + 1)
      J2 = MIN(N, JI + L - 1)
      DLIK = 0
      DO  10 I = I2, J2
10    DLIK = DLIK + REAL (CONJG (B(J) - A(J)) * RMU(J))
      RETURN

      SUBROUTINE ATOB
      DO  10 I = 1, N
10    B(I) = A(I)
      RETURN

      SUBROUTINE BTOA
      DO  10 I = 1, N
10    A(I) = B(I)
      RETURN
```

Appendix B: Derivation of the steady state error for the best linear estimate. The processes $\{\varepsilon_j : j \in \mathbb{Z}\}$ and $\{a_j : j \in \mathbb{Z}\}$ (here, $\mathbb{Z}$ denotes the integers, positive and negative) have spectral representations of the form

$$\varepsilon_j = \int_{-\pi}^{\pi} e^{i\lambda j}\, dZ(\lambda),$$

$$a_j = \int_{-\pi}^{\pi} e^{i\lambda j}\, dA(\lambda),$$

where dZ and dA are random spectral measures satisfying the operational conventions

$$E\, dZ(\lambda)\, dZ(\mu)^* = (\sigma^2/\pi)\, \delta(\lambda-\mu)\, d\lambda,$$

$$E\, dA(\lambda)\, dA(\mu)^* = (p\rho^2/\pi)\, \delta(\lambda-\mu)\, d\lambda,$$

$$E\, dA(\lambda)\, dZ(\mu) = 0$$

where $\delta(\cdot)$ denotes Dirac's delta function. See section 3.3 and equation (2.32) of Koopmans (1974), and recall that $\{\varepsilon_j\}$ and $\{a_j\}$ are complex valued white noises, which accounts for a factor of 2 in the latter equations. The BLE, which is the minimizer of (2.9), is given by

$$\hat{\underset{\sim}{a}} = [bI + W^*W]^{-1} W^* \underset{\sim}{y}$$

where

(B1) $$b = \sigma^2/(p\rho^2).$$

To write this in spectral form, we first define

$$\hat{w}(\lambda) = \sum_k w(k) e^{-i\lambda k},$$

where w is the waveform of (2.2). Note that operating with W is just convolution with w. For the particular wave for used in the simulations of Section 5,

$$\text{(B2)} \qquad \Psi(\lambda) = \sum_{k=0}^{L-1} e^{-i\lambda k}$$

$$= \exp[-i(L-1)\lambda/2] \sin[L\lambda/2]/\sin[\lambda/2].$$

Now it is easy to see that

$$[W\underset{\sim}{\hat{a}}]_j = [(bI + W^*W)^{-1} W^*W\underset{\sim}{y}]_j$$

$$= \int_{-\pi}^{\pi} \frac{|\Psi(\lambda)|^2}{b + |\Psi(\lambda)|^2} e^{i\lambda j} \{dZ(\lambda) + \Psi(\lambda)\, dA(\lambda)\}.$$

Here, we are using $[W\underset{\sim}{\hat{a}}]_j$ to denote the j^{th} component of $W\underset{\sim}{\hat{a}}$. The steady state error (per observation) in the estimate of the signal $W\underset{\sim}{a}$ is then

$$E|\,[W(\underset{\sim}{\hat{a}} - \underset{\sim}{a})]_0|^2 = E\Big|\int_{-\pi}^{\pi} \frac{|\Psi(\lambda)|^2}{b + |\Psi(\lambda)|^2} dZ(\lambda) + \int_{-\pi}^{\pi} \frac{-b\,\Psi(\lambda)}{b + |\Psi(\lambda)|^2} dA(\lambda)\Big|^2 =$$

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sigma^2|\Psi(\lambda)|^4}{(b + |\Psi(\lambda)|^2)^2} + \frac{p\rho^2 b^2|\Psi(\lambda)|^2}{(b + |\Psi(\lambda)|^2)^2}\, d\lambda = \frac{\sigma^2}{\pi}\int_{-\pi}^{\pi} \frac{|\Psi(\lambda)|^2}{b + |\Psi(\lambda)|^2}\, d\lambda,$$

where (B1) was used at the last step. Multiplication of this latter quantity by N gives a good approximation for the integrated mean squared signal estimation error for the BLE from a set of N observations. It is a lower bound since the boundary effects of a finite observation interval always increase the MSE.

Appendix C: Derivation of the limiting IMSE for the MAP estimate. The integrated mean squared error (IMSE) for the MAP estimate of the total signal process is given by

$$\text{IMSE} = E \sum_{k=1}^{N} |[W(\underset{\sim}{\hat{a}} - \underset{\sim}{a})]_k|^2 ,$$

where $\underset{\sim}{\hat{a}}$ denotes the MAP estimate of $\underset{\sim}{a}$, and $[W\underset{\sim}{u}]_k$ denotes the k^{th} component of the N-vector $W\underset{\sim}{u}$. Now, we will fix σ^2 and $c = p\rho^2$, and let $p \to 0$. Under these conditions, the events will occur less frequently and the individual signals become more pronounced (since ρ^2 must tend to ∞). Hence, the following two assumptions will become approximately valid:

(i) there is no overlapping of signals;

(ii) the event time set K will be estimated perfectly.

In any case, the IMSE derived under these two assumptions will be a lower bound to the true IMSE for any $p > 0$, since overlapping signals and uncertainty in K will only increase the error.

If the waveform w is given by (5.1), then it follows from (2.8) and (2.7) that for $j \in$,

$$\hat{a}_j = \sum_{k+j}^{j+L-1} y_k/[(\sigma^2/\rho^2) + L].$$

Hence, the IMSE resulting from this individual signal is

$$L\, E|\hat{a}_j - a_j|^2 = L\,[1 - 1/(\sigma^2/\rho^2 + L)]^2\, E|a_j|^2 +$$

$$L\,[1/(\sigma^2/\rho^2 + L)]\, E\Big|\sum_{k=j}^{j+L-1} \varepsilon_j\Big|^2 = \frac{L[(\sigma^4/\rho^4)2\rho^2 + 2\sigma^2 L]}{[\sigma^2/\rho^2 + L]^2}$$

$$= 2L\sigma^2/(\sigma^2/\rho^2 + L).$$

Since the expected number of events is Np, the total IMSE is given by

$$IMSE = 2NPL\sigma^2/(\sigma^2/\rho^2 + L).$$

DDC/JEE/jik

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 2496	2. GOVT ACCESSION NO.: AD-A127692
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): ESTIMATION OF A FILTERED MARKED POISSON PROCESS FROM NOISY OBSERVATIONS WITH APPLICATIONS TO SIGNAL PROCESSING	5. TYPE OF REPORT & PERIOD COVERED: Summary Report - no specific reporting period
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Dennis D. Cox and John E. Ehrenberg	8. CONTRACT OR GRANT NUMBER(s): DAAG29-80-C-0041, MCS-7927062, Mod. 1, and N0014-81-K0095.
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS: * see below	12. REPORT DATE: March 1983
	13. NUMBER OF PAGES: 37
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

* U.S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina, 27709; National Science Foundation, Washington, D.C. 20550; Office of Naval Research, Arlington, VA 22217

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Filtered marked Poisson process, Bayesian estimation, penalized maximum likelihood estimation, Akaike's information criterion, real time signal processing.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The problem of estimating the arrival times, amplitudes, and phases of an unknown number of signals in noise is treated. The signals are assumed to have a common, known waveform. A Bayesian model using a Poisson prior for the arrival times is specified, and a real time algorithm for computing the posterior mode is developed. Alternatively, the procedure may be looked upon as a penalized likelihood estimator with a penalty term which is a generalized form of Akaike's Information Criterion. Simulation results are presented which show that this approach can improve over classical, linear methods.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)