

Estimation of Pulse Heights and Arrival Times*

HUIBERT KWAKERNAAK†

The heights and arrival times of randomly arriving pulses observed with white measurement noise may be estimated using Rissanen's modification of the maximum likelihood principle.

Key Words—Identification; signal processing; parameter estimation; signal detection; maximum likelihood methods.

Abstract—The problem is studied of estimating the arrival times and heights of pulses of known shape observed with white additive noise. The main difficulty is estimating the number of pulses. When a maximum likelihood formulation is employed for the estimation problem, difficulties similar to the problem of estimating the order of an unknown system arise. The problem may be overcome using Rissanen's shortest data description approach. An estimation algorithm is described, and its consistency is proved. The results are illustrated by a simulation study using an example from seismic data processing also studied by Mendel.

1. INTRODUCTION

A WELL-KNOWN and well documented problem in statistical communication theory is the problem of estimating the size and arrival time of a pulse that is observed in the presence of noise (Van Trees, 1971). Problems of this kind arise in radar and sonar. A problem which seems to have been less extensively studied is that of estimating the sizes and arrival times of a sequence of pulses arriving with random intervals.

The mathematical model to be used for this problem is the following. It is assumed that a discrete-time random signal is observed, given by

$$Y_j = \sum_{i=1}^{\infty} H_i g(t_j - T_i) + W_j, \quad (1)$$

$j=1, 2, \dots, n$. Here g is a given pulse shape, assumed to be known and fixed from one pulse to the next. The observation instants are given by t_j

$=j\Delta$, $j=1, 2, \dots, n$, with Δ a fixed sampling interval. T_i , $i=1, 2, \dots$ are the (random) arrival times of the pulses, and H_i , $i=1, 2, \dots$, the (random) pulse heights. Finally, W_j , $j=1, 2, \dots, n$, are the observation errors. More detailed statistical assumptions are introduced in the next section. We shall consider the problem of estimating the arrival times and pulse heights from a single observed record.

I first learned about this problem in the context of ultrasound medical diagnosis (Mars, 1974). Mendel (1977) has studied the problem as it originates from seismic signal processing. In both cases, the delayed pulses $g(t_j - T_i)$ represent echos, with T_i the travelling time of the pulse, while the size H_i of the i th pulse represents the strength of the echo.

In earlier work on this problem (Kwakernaak, 1975; Nauta, 1977; Kwakernaak and Nauta, 1977) an attempt was made to solve different versions of it using martingale estimation theory.

To this end, it was assumed that the observations were continuous, and represented as

$$dY_t = \sum_{i=1}^{\infty} H_i g(t - T_i) dt + dW_t, \quad 0 \leq t \leq I,$$

with I the duration of the observation interval, and W_t , $t \geq 0$, Brownian motion. Using martingale estimation theory it is perfectly well possible to obtain recursive equations for the conditional expectations (given the observations) of processes such as the time elapsed since the arrival of the i th pulse, and the height of the i th pulse. These recursive equations, however, belong to a countably infinite set of coupled conditional moment equations. As in nonlinear filtering, the numerical solution of these equations forms a considerable problem. Naive approaches involv-

*Received 11 September 1979; revised 25 February 1980. The original version of this paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor H. Sorenson.

†Department of Applied Mathematics, Twente University of Technology, P.O. Box 217, 7500 AE Enschede, The Netherlands.

ing simple truncation of the moment equations produced very poor results in simulated problems (Nauta, 1977). During this work, it became clear that the difficulty of establishing whether or not a pulse has occurred at any given time is a central problem.

As it appears that in the intended applications (seismic and ultrasound signal processing) there is no pressing need for recursive estimation, but that batch processing of observed signal records is also acceptable, it was decided to look into maximum likelihood estimation methods. In Section 2 of this paper an approximate method of estimating the arrival times and pulse heights based on a matched filter type of approach is developed under the simplifying assumption that the number of pulses occurring during the observation interval is known. In Section 3 the problem of estimating the number of pulses is attacked using Rissanen's shortest data description approach. The consistency properties of the proposed method and the estimation of the statistical parameters are treated in Section 4. Section 5 is devoted to the presentation of some simulation results, while the conclusions of the work are summarized in Section 6.

Related work on the estimation of pulse heights and arrival times has been done by several other authors. Liechtenstein and Young (1968) formulated a maximum-likelihood hypothesis test for a general problem of parameter estimation, which is specialized to the problem of range resolution. Bounds on the interarrival times are imposed and sub-optimal solutions are obtained for special pulse shapes in a matched-filter setting. Young (1969) develops a recursive likelihood test for pulse arrival detection and pulse height estimation based on simplifying assumptions. Au and Haddad (1978) consider the estimation of the state of a linear system driven by Poisson white noise and employ a sequential maximum *a posteriori* likelihood schema to detect and estimate the arrival of pulses.

By assuming that the pulse shape g is the impulse response of a finite-dimensional system, Mendel (1977) represents the observed signal as the noisy output of a linear system excited by Poisson white noise. By applying optimal smoothing theory the best linear filter to estimate the pulse heights is obtained. The arrival time estimation problem is not really tackled in this work. Kormylo (1979) uses a maximum-likelihood approach related but not identical to the present approach, and resolves it using Kalman smoothing techniques. Various approximate solutions are proposed.

In none of the works referenced is a consistency analysis attempted, nor is the problem of

estimating the number of pulses explicitly discussed. A comment on the consistency of Kormylo's method is made in Section 4.

2. ESTIMATION OF THE PULSE HEIGHTS AND ARRIVAL TIMES

In this section, we shall consider the problem of estimating the pulse heights and arrival times, where we assume that the number of pulses is given. In the next section we shall consider the more difficult situation where this number is not known.

First we state more detailed statistical assumptions for the model (1) introduced in the previous section. The random variables W_j , $j=1, 2, \dots, n$, H_i , $i=1, 2, \dots$ and $T_{i+1} - T_i$, $i=1, 2, \dots$, will all be assumed to be independent. For each j , the observation error W_j is assumed to be Gaussian with zero expectation and variance σ^2 . For each i , the pulse height H_i is also assumed to be Gaussian, with expectation zero and variance v^2 , while the interarrival time $T_{i+1} - T_i$ will be taken to be exponentially distributed with parameter λ .

We next define the random variable $Q = \max\{i | T_i \leq I\}$, where $I = n\Delta$ is the duration of the observation interval. Q is the number of pulses occurring during the observation interval and has the probability distribution

$$P(Q=q) = \frac{(\lambda I)^q}{q!} e^{-\lambda I}, \quad q=0, 1, 2, \dots \quad (2)$$

This distribution is determined by the arrival time rate λ , which is a deterministic statistical parameter, and the observation period I .

Since the fact that the number of pulses occurring during the observation interval is a random variable, and thus is not known, considerably complicates the estimation problem, we shall first consider the situation that the number of pulses is given and equal to q , with q a natural number. The log-likelihood function to be used for the estimation of the pulse heights and arrival times will then consist of three terms. The first term is the log-likelihood of the observations Y_1, Y_2, \dots, Y_n , given the pulse heights and arrival times, and can be written as

$$\begin{aligned} L_1 &= \log[p_{Y_1, \dots, Y_n | H_1, H_2, \dots, H_q, T_1, T_2, \dots, T_q}(y | h, \theta)] \\ &= \sum_{j=1}^n \log \left[p_w \left(y_j - \sum_{i=1}^q h_i g(t_j - \theta_i) \right) \right]. \end{aligned} \quad (3)$$

Here $y = \text{col}(y_1, y_2, \dots, y_n)$ represents the observed signal, $h = \text{col}(h_1, h_2, \dots, h_q)$ the pulse heights, and $\theta = \text{col}(\theta_1, \theta_2, \dots, \theta_q)$ the pulse arrival times. Finally, p_w is the density of the observation error. The two other terms in the log-likelihood function represent the *a priori* log-likelihoods L_2 and L_3 on the pulse heights and arrival times, respectively, and are given by

$$L_2 = \sum_{i=1}^q \log[p_H(h_i)], \tag{4}$$

where p_H is the density of the pulse heights, and

$$L_3 = \log[p_{T_1, T_2, \dots, T_q|Q}(\theta|q)], \tag{5}$$

with $p_{T_1, T_2, \dots, T_q|Q}(\theta|q)$ the joint probability density of T_1, T_2, \dots, T_q given that the number of pulses during the observation interval $[0, I]$ equals q . This density may be found to be given by

$$p_{T_1, T_2, \dots, T_q|Q}(\theta|q) = \begin{cases} q!/I^q & \text{for } 0 < \theta_1 < \theta_2 < \dots < \theta_q < I, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Substituting Gaussian densities with appropriate variances for p_w and p_H we find that the neg-log-likelihood $N = -L_1 - L_2 - L_3$ is given by the expression

$$N(h, \theta, q) = \frac{1}{2\sigma^2} \sum_{j=1}^n \left[y_j - \sum_{i=1}^q h_i g(t_j - \theta_i) \right]^2 + \frac{1}{2v^2} \sum_{i=1}^q h_i^2 + \frac{1}{2}n \log(2\pi\sigma^2) + \frac{1}{2}q \log(2\pi v^2) - \log(q!/I^q). \tag{7}$$

To find maximum-likelihood estimates for the pulse heights and arrival times, this expression has to be minimized with respect to h and θ . For this purpose, only the first two terms of N are relevant. These can be written as

$$\frac{1}{2} \left(\frac{d}{v} \right)^2 \left[\frac{\Delta}{\alpha} \sum_{j=1}^n y_j^2 - 2 \sum_{i=1}^q h_i z(\theta_i) + \sum_{i=1}^q \sum_{k=1}^q h_i h_k r(\theta_i, \theta_k) + \frac{1}{d^2} \sum_{i=1}^q h_i^2 \right]. \tag{8}$$

Here

$$\alpha := \Delta \sum_{j=1}^n g^2(t_j) \simeq \int_0^I g^2(s) ds \tag{9}$$

is a quantity that is a measure for the 'strength' of the pulse,

$$d := \left(\frac{\alpha v^2}{\Delta \sigma^2} \right)^{1/2} \tag{10}$$

is a quantity that is called the detection index in communication theory and is a measure for the signal-to-noise ratio,

$$z(\theta_i) := \frac{\Delta}{\alpha} \sum_{j=1}^n y_j g(t_j - \theta_i), \tag{11}$$

$i = 1, 2, \dots, q,$

is a signal that is to be computed from the observed signal, while

$$r(\theta_i, \theta_k) := \frac{\Delta}{\alpha} \sum_{j=1}^n g(t_j - \theta_i) g(t_j - \theta_k), \tag{12}$$

$i, j = 1, 2, \dots, q,$

is determined by the pulse shape alone. If both θ_i and θ_k are integral multiples of the sampling period Δ and neither θ_i nor θ_k is close to 0 or I , we may approximate

$$r(\theta_i, \theta_k) \simeq \frac{\Delta}{\alpha} \sum_{j=-\infty}^{\infty} g(t_j - \theta_i) g(t_j - \theta_k) = \rho(\theta_i - \theta_k), \tag{13}$$

where

$$\rho(\tau) := \frac{\Delta}{\alpha} \sum_{j=-\infty}^{\infty} g(t_j) g(t_j - \tau). \tag{14}$$

The function ρ is of course the normalized auto-correlation of the pulse shape. We can now rewrite (8) in vector form as

$$\frac{1}{2} \left(\frac{d}{v} \right)^2 (\mu - 2h^T z_\theta + h^T M_\theta h), \tag{15}$$

where

$$\mu := \frac{\Delta}{\alpha} \sum_{j=1}^n y_j^2, \tag{16}$$

the superscript T denotes the transpose, $z_\theta = \text{col}(z(\theta_1), z(\theta_2), \dots, z(\theta_q))$, and M_θ is a $q \times q$ matrix whose (i, k) th entry is $r(\theta_i, \theta_k) + \delta_{ik}/d^2$. The index θ has been appended to z and M to recall that both depend on θ .

As (15) is quadratic in h , minimization with

respect to h is simple and results in the minimizing value

$$\hat{h}_\theta = M_\theta^{-1} z_\theta, \tag{17}$$

with the corresponding minimum value of (15)

$$R_\theta := \frac{1}{2} \left(\frac{d}{v} \right)^2 (\mu - z_\theta^T M_\theta^{-1} z_\theta). \tag{18}$$

The difficulty now is to minimize R_θ with respect to θ . Let us assume for the moment that the q pulses do not overlap, i.e., the interarrival time $\theta_i - \theta_k$ between any two different pulses is so large that $g(\theta_i - \theta_k)$ is zero or very close to zero. This means that $r(\theta_i, \theta_k)$ is zero or very close to zero and as a consequence M_θ is diagonal with its i th diagonal element equal to $r(\theta_i, \theta_i) + 1/d^2 \simeq 1 + 1/d^2$, if we assume $\Delta \ll \theta_i \ll I$. It follows that in the case of non-overlapping pulses we have

$$R_\theta = \frac{1}{2} \left(\frac{d}{v} \right)^2 \left[\mu - \frac{d^2}{1 + d^2} \sum_{i=1}^q z^2(\theta_i) \right]. \tag{19}$$

The estimation of the arrival times now is simple. Minimization of R_θ with respect to $\theta_1, \theta_2, \dots, \theta_q$ amounts to maximization of $\sum_i z^2(\theta_i)$. This is easily done by locating the q largest maxima of the time series $z^2(\tau)$, $\tau \in [0, I]$, which may be done by a straightforward search procedure if we restrict the maximization to a time grid with resolution Δ . Equivalently, we may locate the q largest maxima of $|z(\tau)|$, $\tau \in [0, I]$.

If the pulses overlap, which may easily be the case, this procedure fails. Let us consider the function z in more detail. We can consider z as a realization of the random process Z , given by

$$Z(\tau) = \frac{\Delta}{\alpha} \sum_{j=1}^n Y_j g(t_j - \tau), \quad \tau \in [0, I]. \tag{20}$$

Substituting for Y_j , this can be rewritten as

$$\begin{aligned} Z(\tau) &= \frac{\Delta}{\alpha} \sum_{j=1}^n \left[\sum_{i=1}^q H_i g(t_j - T_i) + W_j \right] g(t_j - \tau) \\ &= \sum_{i=1}^q H_i \frac{\Delta}{\alpha} \sum_{j=1}^n g(t_j - \tau) g(t_j - T_i) \\ &\quad + \sum_{j=1}^n W_j g(t_j - \tau) \\ &= \sum_{i=1}^q H_i r(\tau, T_i) + U(\tau) \\ &\simeq \sum_{i=1}^q H_i \rho(\tau - T_i) + U(\tau). \end{aligned} \tag{21}$$

Here

$$U(\tau) := \frac{\Delta}{\alpha} \sum_{j=1}^n W_j g(t_j - \tau), \quad \tau \in [0, I]. \tag{22}$$

If we temporarily overlook the second term on the right-hand side of (21), which represents noise, we see that the time series $Z(\tau)$, $\tau \in [0, I]$, consists of a superposition of wave forms with shape ρ , shifted to the arrival times T_i of the pulses and with amplitudes equal to the pulse heights H_i . If in the original time series the pulses overlap, also in the time series z the different wave shapes will interfere, which makes it impossible to estimate the arrival times by simply finding the q largest absolute maxima of $|z|$.

To overcome this difficulty, the following procedure may be used to eliminate the interfering pulses successively, starting with the largest. Suppose that m estimates $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$ of arrival times have been obtained. Let M_m denote the $m \times m$ matrix whose (i, k) th entry is $r(\hat{\theta}_i, \hat{\theta}_k) + \delta_{ik}/d^2$, and let z_m denote the m -dimensional column vector whose i th entry is $z(\hat{\theta}_i)$. Then an estimate $\hat{h}_m = \text{col}(\hat{h}_1^m, \hat{h}_2^m, \dots, \hat{h}_m^m)$ of the pulse height vector, corresponding to the estimated arrival times $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$, may be obtained as $\hat{h}^m = M_m^{-1} z_m$. For obtaining an estimate of the next arrival time, we first correct the time series z for the pulses already estimated by calculating

$$z^m(\tau) := z(\tau) - \sum_{i=1}^m \hat{h}_i^m \rho(\tau - \hat{\theta}_i), \quad \tau \in [0, I]. \tag{23}$$

We then obtain the next arrival time estimate $\hat{\theta}_{m+1}$ by locating the absolute maximum of the time series $|z^m(\tau)|$, $\tau \in [0, I]$.

The procedure is repeated until estimates $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_q$ have been obtained. The procedure is initialized for $m=1$, choosing $\hat{\theta}_1$ as the location of the absolute maximum of $|z(\tau)|$, $\tau \in [0, I]$. Note that the consecutive arrival time estimates $\hat{\theta}_1, \hat{\theta}_2, \dots$ are numbered in the order in which they are found, not in temporal order.

By this procedure, the largest pulses are eliminated first, so that the smaller pulses can be estimated without interference from the larger ones. The whole procedure is an adaptation of the familiar matched filtering technique.

The computing load for this procedure is modest. The computations are prepared by first calculating the time series z and the function r , using the approximation (13) for the latter, both on a time grid with resolution Δ . Then the recursive procedure described above is started. Numerical results for a simulated example will be described in Section 5.

As the matrix M_m is Toeplitz, special methods could be employed for its inversion, especially as

its dimensions are increased by one at each step. It was not felt that it was necessary to go to this length for the example considered in Section 5, where m was at most 20.

3. ESTIMATION OF THE NUMBER OF PULSES

In the preceding section it has been described how the pulse heights and arrival times may be estimated for a given number of pulses. Estimating this number is a problem in itself. The difficulty is that the minimum of the neg-log-likelihood function

$$\min_{h, \theta} N(h, \theta, q) \tag{24}$$

is a monotonically decreasing function of the number of pulses q . This phenomenon is caused by overfitting: by increasing the number of free parameters (for a fixed number of observations) the corresponding increase in the number of degrees of freedom is used to artificially improve the fit. This situation is only partly remedied by including in the neg-log-likelihood function an extra term $-\log[P(Q=q)]$ for the *a priori* likelihood of the number of pulses. Simulation shows that the minimum of the extended neg-log-likelihood function thus obtained still is at a much too high value of q .

This problem is quite similar to that encountered in estimating the system order in system identification. For this difficulty several solutions have been proposed. We shall discuss those of Akaike (1974), Rissanen (1978) and Fine and Hwang (1979).

Akaike's approach consists of estimating the amount of overfitting, and compensating for this. The formula he obtains is particularly simple and therefore appealing. Application of the formula to the present situation shows that the number of pulses is still estimated far too high. Moreover, the formula is not really applicable since it is derived under the assumption that the estimated parameter values are close to the true values. This assumption is not fulfilled in the pulse estimation problem since the procedure described in the preceding section leads to the estimation of false pulses, i.e., pulses that do not exist.

A more promising method, which has been given serious attention, is Rissanen's shortest data description method. We shall give a condensed exposition of its application to the present problem. Rissanen concludes from elementary information theory that the symbol length (expressed in nats) minimally needed to encode the data $q, \theta_1, \theta_2, \dots, \theta_q, h_1, h_2, \dots, h_q, w_1, w_2, \dots, w_n$ is

given by

$$S = - \sum_{j=1}^n \log [p_w(w_j)\varepsilon] - \sum_{i=1}^q \log [p_H(h_i)\eta_i] - \log [p_{T|Q}(\theta|q)\delta_1 \delta_2 \dots \delta_q] - \log [P(Q=q)]. \tag{25}$$

Here ε is the resolution with which the measurement noises w_j are digitized, η_i the resolution for digitizing the pulse height h_i , and δ_i the resolution for digitizing the arrival time θ_i . The choice of the resolutions is to be discussed later. Rissanen's approach to estimation consists in minimizing S with respect to the unknown quantities $q, \theta_1, \theta_2, \dots, \theta_q, h_1, h_2, \dots, h_q, w_1, w_2, \dots, w_n$, subject to the constraint imposed by the observations, in this case,

$$y_j = \sum_{i=1}^q h_i g(t_j - \theta_i) + w_j, \quad j=1, 2, \dots, n. \tag{26}$$

These constraints can be taken care of by substituting $w_j = y_j - \sum_{i=1}^q h_i g(t_j - \theta_i)$ into S as given by (25), which leads to the expression

$$S = N_e(h, \theta, q) - n \log(\varepsilon) - \sum_{i=1}^q \log(\eta_i) - \sum_{i=1}^q \log(\delta_i), \tag{27}$$

where $N_e = N - \log [P(Q=q)]$, with N the neg-log-likelihood function of the preceding section. We note that Rissanen's data length S is the neg-log-likelihood function including an *a priori* term on q , augmented with a number of additional terms.

The minimization of S has to be performed within the appropriate digitizations of the pulse heights and arrival times. This is done by first minimizing S with respect to h and θ continuously, and then rounding off the minimizing values with resolutions η_i for the pulse heights h_i and resolutions δ_i for the arrival times θ_i . Let us denote the minimizing values of h and θ as \hat{h} and $\hat{\theta}$, and the rounded values as h^* respectively θ^* . Assuming that the rounding errors are independent and uniformly distributed, the rounding process results in the following approximate expression for the minimum data length:

$$S^* = N_e(\hat{h}, \hat{\theta}, q) + \frac{1}{24} \sum_{i=1}^q \hat{N}_{h_i, h_i} \eta_i^2 + \frac{1}{24} \sum_{i=1}^q \hat{N}_{\theta_i, \theta_i} \delta_i^2 - n \log(\varepsilon) - \sum_{i=1}^q \log(\eta_i) - \sum_{i=1}^q \log(\delta_i). \tag{28}$$

Here $\hat{N}_{h_i h_i}$ is the second partial derivative of N with respect to h_i in the point $(\hat{h}, \hat{\theta})$ and $\hat{N}_{\theta_i \theta_i}$ the second partial derivative with respect to θ_i . Now the best values of the resolutions η_i and δ_i are found by minimizing S^* with respect to η_i and δ_i . The resolution ε for the measurement noise, incidentally, should be chosen equal to the (supposedly given) accuracy with which the observations $y_j, j=1, 2, \dots, n$, are recorded; this allows reconstruction of the measured data from the estimated data with the correct accuracy.

The optimal resolutions may be found to be given by

$$\hat{\eta}_i = (12/\hat{N}_{h_i h_i})^{1/2}, \quad \hat{\delta}_i = (12/\hat{N}_{\theta_i \theta_i})^{1/2}, \quad (29)$$

both for $i=1, 2, \dots, q$. The partial derivatives are

$$\hat{N}_{h_i h_i} = \frac{1+d^2}{v^2}, \quad \hat{N}_{\theta_i \theta_i} = \left(\frac{\hat{h}_i}{v}\right)^2 \frac{\beta}{\alpha} d^2, \quad (30)$$

where in $\hat{N}_{\theta_i \theta_i}$ we have introduced a certain approximation. The quantities d and α are as defined previously, while

$$\beta := \Delta \sum_{j=-x}^x g'^2(t_j) \simeq \int_{-x}^x g'^2(s) ds. \quad (31)$$

Substituting these results we obtain

$$\hat{\eta}_i = v \sqrt{\frac{12}{1+d^2}}, \quad \hat{\delta}_i = \frac{1}{d} \sqrt{\frac{12\alpha}{\beta}} \left| \frac{v}{\hat{h}_i} \right|, \quad (32)$$

both for $i=1, 2, \dots, q$. With this we find from (28) for the minimum data length

$$S^* = N_e(\hat{h}, \hat{\theta}, q) + q - n \log(\varepsilon) - \frac{1}{2} q \log\left(\frac{12v^2}{1+d^2}\right) - \frac{1}{2} q \log\left(\frac{12\alpha}{\beta d^2}\right) + \sum_{i=1}^q \log\left(\left|\frac{\hat{h}_i}{v}\right|\right), \quad (33)$$

where

$$N_e(h, \theta, q) = \frac{1}{2\sigma^2} \sum_{j=1}^n \left[y_j - \sum_{i=1}^q h_i g(t_j - \theta_i) \right]^2 + \frac{1}{2v^2} \sum_{i=1}^q h_i^2 + n \log(\sigma \sqrt{2\pi}) + q \log(v \sqrt{2\pi}) - q \log(\lambda) + \lambda I. \quad (34)$$

Thus, the estimation procedure consists in evaluating S^* as a function of q , to which end $N(h, \theta, q)$ has to be minimized with respect to h and θ for each q , and selecting that value of q that minimizes S^* .

It is seen that Rissanen's approach augments the likelihood function. It is interesting to note

that the first of the additional terms, q , is precisely the Akaike correction term. It is also interesting to observe that the optimal pulse height and arrival time resolutions $\hat{\eta}_i$ and $\hat{\delta}_i$ are upon omission of the factor $\sqrt{12}$ equal to the Cramer-Rao lower bounds for the standard deviations of the estimation error of the pulse heights respectively arrival times. Rissanen's method seems to offer an alternative approach to the maximum likelihood philosophy, with the advantage that it can also handle order estimation problems. As in the maximum-likelihood approach, however, one can never trust an estimator before its properties have been investigated. Therefore, we shall study in Section 4 the consistency of Rissanen's estimation method as applied to the problem at hand.

Finally, another approach to order determination is offered by Fine and Hwang (1979). Their idea is very basic. Transposed to the present context, it amounts to selecting a clip level γ , and deciding that a pulse is really absent if $|\hat{h}_i| < \gamma$. The problem then is how to choose the clip level. We shall consider this in the next section. There we shall also see that Rissanen's method is essentially equivalent to the clip level approach, where the level is implicitly produced through the use of the augmented neg-log-likelihood function.

4. CONSISTENCY AND ESTIMATION OF THE STATISTICAL PARAMETERS

In the problem at hand, a consistency analysis of the statistical estimation procedure for the case where the number of observations n increases indefinitely does not make much sense, since from a certain point on new observations carry no new information about the arrival time and size of a given pulse. We shall therefore perform a different kind of consistency analysis, namely for the case where the detection index d goes to infinity. The detection index d , which is defined by $d^2 = (\alpha v^2)/(\Delta \sigma^2)$, is a measure for the signal-to-noise ratio.

It is not necessary to make this consistency analysis very complicated. We restrict ourselves to the case where the pulses do not overlap. Because of the successive elimination procedure developed in Section 2, we believe that the analysis also has relevance for the overlapping case. Let us first consider the Fine-Hwang approach, where a clip level is chosen to determine whether a pulse is really present or not. As we saw in Section 2, the pulses are detected by considering the maxima of the time series $|Z|$, for which we have with good approximation

$$Z(\tau) = \sum_{i=1}^x H_i \rho(\tau - T_i) + U(\tau), \quad \tau \in [0, I], \quad (35)$$

where the H_i are the pulse heights, the T_i the arrival times, and

$$U(\tau) = \frac{\Delta}{\alpha} \sum_{j=1}^n W_j g(t_j - \tau), \quad \tau \in [0, I], \quad (36)$$

is to be considered as noise. It is easily found from (36) that for each τ , the random variable $U(\tau)$ is Gaussian with zero expectation and variance v^2/d^2 . Now, if τ is a true arrival time, it is seen from (35) that $Z(\tau)$ is a Gaussian random variable with expectation zero and variance $v^2 + v^2/d^2 = (1 + 1/d^2)v^2$. Since for the non-overlapping case, if $\hat{\theta}_i$ is an estimate of an arrival time, the corresponding pulse height estimate is $\hat{h}_i = (d^2/(1 + d^2))z(\hat{\theta}_i)$, application of a clip level γ to the estimated pulse heights is equivalent to applying a clip level $((1 + d^2)/d^2)\gamma$ to the maxima of the time series $|Z|$ for deciding whether a pulse is present or not. Now there are two possibilities: either a maximum of the time series $|Z|$ corresponds to a true pulse, or it is a spurious maximum caused by the noise U . If we let the detection index d approach infinity, clearly we shall in the limit detect the true pulses with probability one, and reject all false pulses with probability one, if at the same time

$$\frac{\frac{1 + d^2}{d^2} \gamma}{\text{standard deviation of pulse plus noise}} \rightarrow 0, \quad (37)$$

$$\frac{\frac{1 + d^2}{d^2} \gamma}{\text{standard deviation of noise alone}} \rightarrow \infty.$$

The first condition ensures that in the limit all true pulses exceed the clip level with probability one, while the second condition ensures that all spurious maxima of the noise do not exceed the clip level with probability one. Substituting the appropriate expressions for the standard deviations, it may be seen that the conditions are equivalent to

$$\frac{\gamma}{v} \rightarrow 0, \quad \frac{\gamma}{v} d \rightarrow \infty. \quad (38)$$

This means that asymptotically the clip level should approach zero, but not too fast. The asymptotic behaviour $\gamma/v \sim \log(d)/d$ would be satisfactory, for example.

The trouble with a consistency analysis of this type is that it gives information about the desirable asymptotic behaviour, but, given a particular

detection index, does not help to choose a particular clip level. The type of reasoning employed above, though, may be used to find crude estimates for various probabilities of error, and guide in choosing the clip level.

Let us now study the consistency of the Rissanen method. In the non-overlapping case, as q is increased in steps of one, the successively estimated pulse heights decrease monotonically. Let us consider the effect on S^* as given by (33) of increasing q to $q + 1$; in doing this, we denote $|\hat{h}_{q+1}| = \gamma$, where for the moment γ has nothing to do with a clip level. It may be found that when increasing q to $q + 1$, the corresponding change in S^* is given by

$$\Delta S^* = -\frac{1}{2} \frac{1 + d^2}{d^2} d^2 \left(\frac{\gamma}{v}\right)^2 + \frac{1}{2} \log\left(\frac{\pi}{6} (1 + d^2)\right) + \frac{1}{2} \log\left(\frac{\beta d^2}{12\alpha\lambda^2}\right) + 1 + \log\left(\frac{\gamma}{v}\right). \quad (39)$$

Note that this expression does not depend on q but only on γ . Now S^* assumes a minimum if $\Delta S^* = 0$. We can consider the value of γ that makes $\Delta S^* = 0$ as the clip level implicitly generated by the procedure of minimizing S^* with respect to q . By letting d approach infinity, it is easily found that the solution of $\Delta S^* = 0$ that corresponds to a minimum of S^* asymptotically behaves as

$$\frac{\gamma}{v} \sim \frac{1}{d} \sqrt{2 \log(d)}. \quad (40)$$

This shows that the Rissanen estimator is consistent in the sense discussed before.

It may be inferred from the consistency analysis given above that Rissanen's modification of the neg-log-likelihood function is not the only way to achieve consistency. Kormylo (1979), for instance, bases the estimation of the number of pulses on the minimization of a neg-log-likelihood function that differs from ours in this respect that the pulse heights are considered as nuisance parameters. Thus, he chooses the neg-log-likelihood function

$$-\log[p_{Y_1, Y_2, \dots, Y_n | T_1, T_2, \dots, T_q}(y | \theta)] - \log[p_{T_1, T_2, \dots, T_q | Q}(\theta | q)] - \log[P(Q = q)]. \quad (41)$$

Here, $p_{Y_1, Y_2, \dots, Y_n | T_1, T_2, \dots, T_q}$ is obtained by integrating

$$p_{Y_1, Y_2, \dots, Y_n, H_1, H_2, \dots, H_q | T_1, T_2, \dots, T_q}(y, h | \theta) = p_{Y_1, Y_2, \dots, Y_n | H_1, H_2, \dots, H_q, T_1, T_2, \dots, T_q}(y | h, \theta) \times p_{H_1, \dots, H_q}(h) \quad (42)$$

with respect to h over R^q . It is not difficult to find that the modified neg-log-likelihood function is given by

$$N_e(\hat{h}, \theta, q) = \frac{q}{2} \log \left(\frac{2\pi v^2}{d^2} \right) + \frac{1}{2} \log [\det (M_\theta)], \quad (43)$$

with N_e are defined in (34) and the matrix M_θ as defined in Section 2. Compared with (33), this is a different modification of the neg-log-likelihood function. A consistency analysis shows that for the neg-log-likelihood function (43) γ/v shows an asymptotic behaviour that is identical to that given by (40) for the Rissanen modification. Thus, also the modification (43) yields a consistent estimation procedure.

Up to this point we have assumed that the statistical parameters are known. The statistical parameters are understood to be the observation noise variance σ^2 , the pulse height variance v^2 , and the arrival time rate λ . If these parameters are unknown, maximum likelihood estimates may be obtained by minimizing N_e as given by (34) with respect to σ , v and λ . This yields respectively

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{n=1}^n \left[y_j - \sum_{i=1}^q \hat{h}_i g(t_j - \hat{\theta}_i) \right]^2, \quad (44)$$

$$\hat{v}^2 = \frac{1}{q} \sum_{i=1}^q \hat{h}_i^2, \quad (45)$$

$$\hat{\lambda} = \frac{q}{I}. \quad (46)$$

The difficulty here is that for calculating $\hat{\sigma}$, \hat{v} and $\hat{\lambda}$ we need estimates for the pulse heights and arrival times, while for obtaining the latter the statistical parameters are required. Inspection of the procedure proposed in Section 2 shows, however, that the only statistical parameter that is really needed is d . Moreover, if d is reasonably large, its effect is minor, since it occurs in the configuration $1 + 1/d^2$. A practical method therefore is to eliminate the statistical parameters from the procedure of Section 2 by setting $d = \infty$, accordingly obtain estimates of the pulse heights and arrival times, and then estimate the statistical parameters from (44) to (46). If necessary, a corrected value for the detection index may be determined from the estimated statistical parameters, and the procedure of Section 2 repeated for this corrected value of d .

The procedure of the preceding paragraph may be used to obtain simultaneous estimates for the pulse heights and arrival times and the statistical parameters for any given value of q . The number of pulses may then be estimated by minimizing

the minimum data length as given by (33), replacing the statistical parameters with their estimates.

5. SIMULATION STUDY

To test the procedure described in the preceding sections, a numerical simulation study has been performed. For comparison purposes the example presented by Mendel (1977) was chosen, with as much as possible the same data. Simulation runs were made for the third pulse shape reported in Mendel's paper, which is given by

$$g(t) = \begin{cases} 1360te^{-500t} - 0.5e^{-15.3t} \sin\left(\frac{2\pi t}{0.06}\right), & t \geq 0, \\ 0, & t < 0 \end{cases} \quad (47)$$

The pulse shape is plotted in Fig. 1. Following

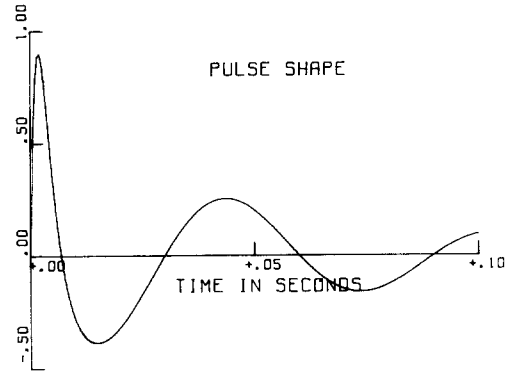


FIG. 1. The pulse shape.

Mendel, the pulse arrival rate was taken as $\lambda = 16.67 \text{ s}^{-1}$ and the pulse height standard deviation as $v=0.15$. In Mendel's work the sampling period was chosen as 0.003 s. For reasons to be explained below it was felt that this sampling period was too large, and subsequently it was reduced to $\Delta=0.0005 \text{ s}$. For the given pulse shape and sampling period the constants α and β as given by (9) and (31) equal $\alpha=0.005778 \text{ s}$ and $\beta = 1435 \text{ s}^{-1}$.

In Table 1 the simulated pulse arrival times and corresponding pulse heights are listed. These data were taken to correspond as closely as possible to the data used by Mendel. Sufficient decimals were randomly added to the arrival times to avoid the possibility that the arrival times are multiples of the sampling period. The total observation time, which was 0.9 s in Mendel's case, was extended to 1 s to avoid truncation effects near the end of the observation interval.

Mendel parameterized his simulation study according to the signal-to-noise ratio, which he

TABLE 1. PULSE ARRIVAL TIMES AND PULSE HEIGHTS

Pulse number	Arrival time	Pulse height
1	0.03760	0.029
2	0.13743	-0.060
3	0.14094	0.080
4	0.18937	0.038
5	0.23879	-0.070
6	0.26521	-0.017
7	0.34622	0.170
8	0.40318	-0.010
9	0.41530	0.108
10	0.49108	0.059
11	0.49733	0.116
12	0.50879	-0.057
13	0.51962	0.028
14	0.66807	-0.113
15	0.67797	-0.220
16	0.73691	0.078
17	0.79491	0.155
18	0.86944	0.040
19	0.89424	-0.230

defines as the ratio of the average signal power to the observation noise variance. The average signal power equals $\lambda\alpha v^2$; if the signal-to-noise ratio in Mendel's sense is to be 10 (one of the cases considered by Mendel), this would imply that the observation noise variance σ^2 equals $1/10$ of $\lambda\alpha v^2$. Mendel's definition of signal-to-noise ratio is not appropriate, however, if the sampling time is reduced, such as we do. If the sampling time is reduced and the signal-to-noise ratio according to Mendel's definition is kept constant, the estimation situation actually becomes more favourable, because for each pulse we obtain more independent observations with the same observation error standard deviation. This is the reason that the detection index is a more suitable measure for the signal-to-noise ratio. In Mendel's work, with his signal-to-noise ratio equal to 10, the detection index is equal to $d=14.141$. We have kept the detection index the same as in Mendel's simulation by choosing the observation noise variance equal to $\sigma^2=0.001300$, or $\sigma=0.03606$.

We now state the reason for choosing the sampling interval Δ smaller than Mendel did. The resolution for the arrival times δ_i according to Rissanen's method is given in formula by (32). For the data given above, we find numerically

$$\delta_i = 0.0004916 \cdot \left| \frac{v}{h_i} \right|. \quad (48)$$

As we pointed out, this number is of the same order of magnitude as the Cramer-Rao lower bound on the error standard deviation for the arrival time estimates. For pulses with heights equal to the pulse height standard deviation v the resolution is about equal to 0.0005 s, while for smaller pulses it is larger and for larger pulses

smaller. In view of this a sampling time of 0.0005 s seems better than one of 0.003 s.

With the data adopted, the resolution for the pulse height according to Rissanen is $\hat{\eta}=0.03665$. This means that pulses whose heights are less than about 0.037 are likely to be overlooked or eliminated by the estimation procedure.

The given pulse shape is a Bohl function, i.e., there exist a row vector c , a square matrix A and a column vector b such that $g(t)=ce^{At}b$ for $t \geq 0$, and $g(t)=0$ for $t < 0$ (see Mendel, 1977). This means that for the computation of the time series z and of the function ρ , a state representation can be used to advantage. In the computations, we used the approximation $r(\theta_1, \theta_2)=\rho(\theta_1 - \theta_2)$ throughout.

For the simulation study, a seismic track was simulated according to the data listed. Then the time series z and the function ρ were computed, and the successive elimination procedure described in Section 3 was followed. In Fig. 2 we illustrate graphically the functions z^m for $m=0, 1, 2, 3, 4$. It can clearly be seen how the largest pulses are successively removed. The plots are for the case where the detection index d is assumed to be known, but they are indistinguishable from the case where d is not known and set equal to infinity.

To estimate the number of pulses, the Rissanen approach may be followed. Computation of the minimum data length S^* according to (33) has been performed both for the case where the statistical parameters were assumed to be known and where they were assumed unknown. The plot of S^* for the known case is given in Fig. 3; it assumes a minimum for $q=14$. In computing S^* , the value of ε , although not relevant for the eventual results, was taken equal to 0.01 . The plot of S^* for the case of unknown parameters is quite similar; it assumes its minimum for $q=13$, however. The successively estimated arrival times and pulse heights for the case of unknown statistical parameters are listed in Table 2. Also these estimates are not much different from the case of known statistical parameters. The estimates of the statistical parameters that were found for $q=13$ are $\hat{\sigma}=0.03766$, $\hat{v}=0.1281$ and $\hat{\lambda}=13.00 \text{ s}^{-1}$.

The 13 pulses that are estimated are all true pulses. It turns out that if one more pulse is estimated, this is also a true pulse, but that all pulses additionally estimated are 'false' pulses, i.e., pulses that correspond to spurious maxima of the time series $|Z|$. The remaining five true pulses, all with small amplitudes, are not recovered, and appear irretrievably lost. Clearly, at least in this case, the Rissanen method is conservative in the sense that false pulses are rejected at the expense of losing a number of true pulses.

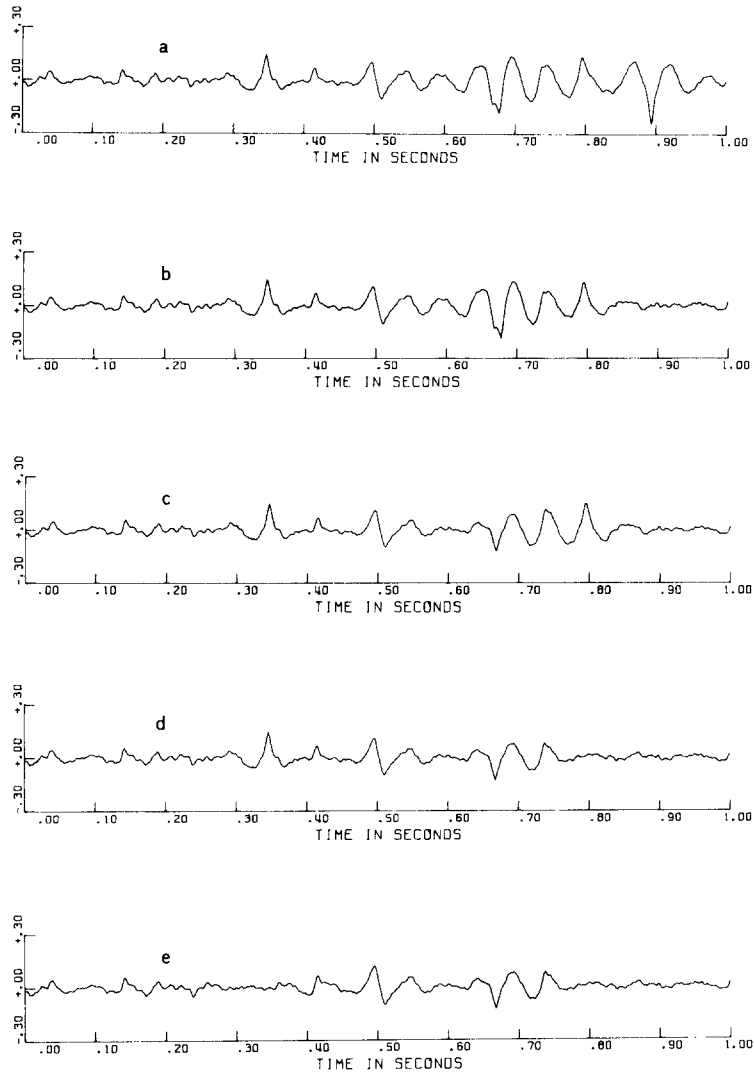


FIG. 2. The successive time series x^m , for (a) $m=0$, (b) $m=1$, (c) $m=2$, (d) $m=3$, (e) $m=4$.

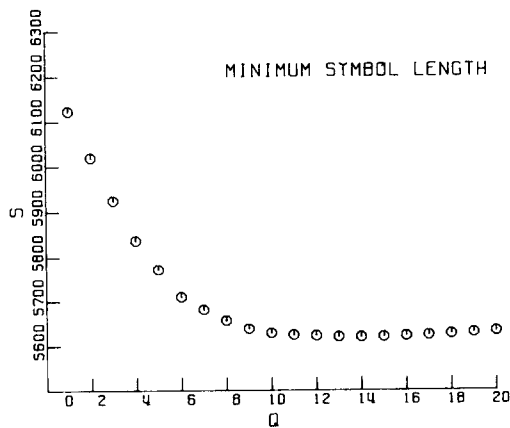


FIG. 3. The minimum data length as a function of q for the case of known statistical parameters.

TABLE 2. ESTIMATES OF THE PULSE ARRIVAL TIMES AND HEIGHTS FOR THE CASE OF UNKNOWN STATISTICAL PARAMETERS WITH $q=13$

Sequential number	Actual arrival time(s)	Estimated arrival time(s)	Actual pulse height	Estimated pulse height
1	0.89424	0.8940	-0.230	-0.237
2	0.67797	0.6780	-0.220	-0.214
3	0.79491	0.7950	0.155	0.137
4	0.34622	0.3460	0.170	0.160
5	0.66807	0.6680	-0.113	-0.135
6	0.49733	0.4965	0.116	0.127
7	0.41530	0.4150	0.108	0.087
8	0.73691	0.7375	0.078	0.090
9	0.50879	0.5100	-0.057	-0.078
10	0.14094	0.1420	0.080	0.048
11	0.23879	0.2380	-0.070	-0.062
12	0.03760	0.0400	0.029	0.047
13	0.18937	0.1890	0.038	0.049

6. CONCLUSIONS

In this paper, the problem has been considered of estimating the arrival times and sizes of pulses observed with additive noise. It has turned out that a crucial problem is that of estimating the number of pulses. Rissanen's approach yields an augmented likelihood function that allows estimating the number of pulses. From the simulation study and the corresponding analysis Rissanen's method appears to work well. We have noted that it is a bit conservative in the sense that it keeps on the safe side, rejecting false pulses with a high probability at the expense of losing a few (small) true pulses. The mathematical analysis has shown Rissanen's method to be consistent.

It was also found that Rissanen's method can be seen as a version of the clip level method, which implicitly generates the value of the clip level. Explicit application of the clip level technique would permit direct control over the probabilities of errors of the first kind (rejection of true pulses) and the second kind (acceptance of false pulses). The calculation of these errors requires a detailed statistical analysis, however, involving the statistical properties of the maxima and minima of a stationary time series (see e.g. Cramer and Leadbetter, 1967).

We conclude this section with a few brief and necessarily superficial comments on the robustness of the proposed estimation method. It is felt that the approach is reasonably insensitive with respect to deviations of the observation noise from the white noise model, as long as the noise remains more or less stationary, Gaussian and wide-band. Also deviations from the Poisson arrival model are quite tolerable. A crucial role, however, is played by the assumption that the pulse shape is known and fixed from one pulse to the next. If this assumption is not satisfied, the

problem has to be studied from a new point of view.

Acknowledgements—I am pleased to acknowledge helpful comments from Prof. E. W. Gröneveld of Twente and an informative discussion with Prof. J. M. Mendel of the University of Southern California.

REFERENCES

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Aut. Control* **AC-19**, 716–722.
- Au, S. and A. H. Haddad (1978). Suboptimal sequential estimation-detection scheme for Poisson driven linear systems. *Information Sci.* **16**, 95–113.
- Cramer, H. and M. R. Leadbetter (1967). *Stationary and Related Stochastic Processes*. Wiley, New York.
- Fine, T. L. and H. W. Hwang (1979). Consistent estimation of system order. *IEEE Trans. Aut. Control* **AC-24**, 387–402.
- Kormylo, J. J. (1979). Maximum-likelihood seismic deconvolution. Ph.D. Dissertation, University of Southern California.
- Kwakernaak, H. (1975). Filtering for systems excited by Poisson white noise. In: *Control Theory, Numerical Methods and Computer Systems*, (Edited by A. Bensoussan and J. L. Lions). Springer Lecture Notes in Economics and Mathematical Systems vol. 107. Springer, Berlin, pp. 468–492.
- Kwakernaak, H. and M. A. Nauta (1977). Recursive estimation of impulse-driven processes. Paper delivered at the Seventh Conference on Stochastic Processes and Their Applications, Enschede, The Netherlands, August 15 to 19.
- Liechtenstein, M. G. and T. Y. Young (1968). The resolution of closely spaced signals. *IEEE Trans. Inform. Theory* **IT-14**, 2, 288–293.
- Mars, H. J. I. (1974). Signal processing for ultrasound medical diagnosis equipment. Diploma project report, Department of Electrical Engineering, Twente University of Technology, Enschede, The Netherlands.
- Mendel, J. M. (1977). White-noise estimators for seismic data processing in oil exploration. *IEEE Trans. Aut. Control* **22**, 694–706.
- Nauta, M. A. (1977). Recursive estimation of impulse-driven processes. Diploma project report, Department of Applied Mathematics, Twente University of Technology, Enschede, The Netherlands.
- Rissanen, J. (1978). Modeling by shortest data description. *Automatica* **14**, 465–471.
- Van Trees, H. L. (1971). *Detection, Estimation and Modulation Theory. Part III. Radar-Sonar Signal Processing and Gaussian Signals*. Wiley, New York.
- Young, T. Y. (1969). A recursive method for signal resolution. *IEEE Trans. Aerospace Electron. Systems* **AES-5**, 46–51.