THE DETECTION AND ESTIMATION OF TWO SIMULTANEOUS SEISMIC EVENTS

H. Kobayashi and P. D. Welch
IBM Research Center, Yorktown Heights, N. Y.

The problem of estimating two simultaneous seismic event signals is first discussed using the signal model of Kelly and Levin and the method of maximum likelihood. In comparing the experimental results of the maximum-likelihood processor with its analytical performance, it becomes apparent that a modified signal model is required which takes into account the incoherent components of the signal. The optimum processor for rejecting an interfering event is then derived and reasonable agreement is obtained between experimental results and the analysis. With the modified model, the major change in the processor is found in the detector. The detection statistic consists of two parts: one responsive to the energy of the coherent component and one responsive to the energies of the noncoherent components.

I. INTRODUCTION

This paper discusses some analytical and experimental results in seismic array processing. It is concerned with the problem of the estimation and detection of a seismic signal in the presence of an interfering seismic signal and background noise.

The first problem that arises in an analytical study of array processing is the choice of a suitable representation for seismic signals and background seismic noise. One approach is to represent the signal and noise as stationary vector random processes with known crosscorrelation functions and apply the multidimensional Wiener filtering theory [1]. However, the finite duration of the seismogram of an event and the variations in its character as different phases arrive suggest that the representation of the signal as a stationary random process is not realistic [2].

A more appealing model for the signal is the one proposed by Kelly and Levin [2]. This model assumes that the signal waveform is completely unknown but is identical at each seismometer except for a time delay due to a finite propagation velocity. Let the k^{th} seismometer (or sensor) be located at a position $\mathbf{r}_k$ relative to some origin in the horizontal plane, $k = 1,2, \cdots, K$. Let $s(t)$ be the signal that would be observed by a seismometer at the origin in the absence of noise. Then the output of the k^{th} seismometer is

$$x_k(t) = s(t - \mathbf{u} \cdot \mathbf{r}_k + d_k) + n_k(t) \quad , \tag{1.1}$$

where $n_k(t)$ is the noise and $\mathbf{u} = [u_1, u_2]$ is the vector of delay per unit distance suffered by the signal as measured along each coordinate axis, and where d_k is

Presented at the Symposium on Computer Processing in Communications
Polytechnic Institute of Brooklyn, April 8-10, 1969.

the time anomaly associated with the k^{th} instrumentation channel. Subsequent development in this paper will formulate a model of the process which is independent of these time anomalies. The results are, therefore, applicable to data which have been preconditioned to compensate for the incurred anomalies. The $\mathbf{u}$ is often called the inverse phase velocity vector since it is related to the phase velocity vector $\mathbf{v}$ of the wave by

$$\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|^2} \quad . \tag{1.2}$$

In Section II we will discuss the problem of estimating two (or multiple, in general) simultaneous seismic event signals using the method of maximum likelihood (ML) estimation, based on the signal model (1.1). The methods of analysis used in this section are due to Kelly and Levin [2], Capon et al. [4] and Schweppe [5]. The structure of the ML-processor for two signal sources was originally obtained by Schweppe [5].

In Section III we present both analytical and experimental results of the performance of the ML-processor. The experimental results are shown to fit the general form predicted but for an interfering signal power to background noise ratio much smaller than expected. This result suggests a modified signal model in which the interfering event has both coherent and noncoherent, or noiselike, components.

In Section IV we introduce this modified signal model which takes into account the noncoherent components of the signal. By modifying the solution obtained by the method of unconditional maximum-likelihood (UML) estimation, the "optimum" processor for rejecting an interfering event is derived. Experimental results are also obtained for this processor. Reasonable agreement with the analysis is obtained, but some discrepancies are noted. These discrepancies suggest a further modification of the model in which the signal consists of coherent and partially coherent components.

In Section V the detection problem for this modified model is discussed from the maximum-likelihood point of view [7]. The structure of the detector is closely related to the ML-estimator and the UML-estimator. With our modified model, the major change in the processor is found in the detector. The detection statistic consists of two parts: one responsive to the energy of the coherent component and one responsive to the energies of the noncoherent components.

II. MAXIMUM-LIKELIHOOD ESTIMATION OF TWO SIMULTANEOUS SEISMIC EVENTS

A. Maximum-Likelihood Estimates

Let $s_i(t)$ be the signal due to a seismic event "i," and let the inverse phase velocity vector of that event be denoted by $\mathbf{u}_i$, $i = \alpha, \beta$. Then the output of the k^{th} sensor is

$$x_k(t) = \sum_{i=\alpha}^{\beta} s_i(t - \mathbf{u}_i \mathbf{r}_k) + n_k(t), \quad k = 1, 2, \cdots, K \quad . \tag{2.1}$$

Let us define the delay operator $\underline{\underline{D}}_i$ which can be represented by the following diagonal matrix function

$$\underline{\underline{D}}_i(t) = \mathrm{diag}\,[\delta(t - \mathbf{u}_i \cdot \mathbf{r}_k)], \quad i = \alpha, \beta \ . \tag{2.2}$$

Then Eq. (2.1) can be written using vector notation as

$$\underline{X}(t) = \sum_{i=\alpha}^{\beta} \underline{\underline{D}}_i(t) \circledast s_i(t)\, \underline{1} + \underline{N}(t) \ . \tag{2.3}$$

Since the detection problem is considered eventually, a hypothesis test is formulated as follows:

H_0: noise only exists

against

H_1: signals due to two events are present.

The noise components $n_k(t)$ of the sensor outputs are assumed to be Gaussian processes with mean zero and a covariance matrix

$$E[\underline{N}(t)\, \underline{N}^T(t')] = \underline{\underline{\Phi}}(t, t') \ . \tag{2.4}$$

Then the (conditional) likelihood function for H_1 against H_0 is given by

$$\Lambda(s_\alpha s_\beta, \mathbf{u}_\alpha, \mathbf{u}_\beta) = \exp\left\{ \sum_{i=\alpha}^{\beta} [\underline{X}, \underline{\underline{D}}_i \circledast s_i \underline{1}]_\Phi - \frac{1}{2} \left\| \sum_{i=\alpha}^{\beta} D_i \circledast s_i \underline{1} \right\|_\Phi^2 \right\} , \tag{2.5}$$

where

$$[\underline{X}, \underline{\underline{D}}_i \circledast s_i \underline{1}]_\Phi = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \underline{X}^T(t)\, \underline{\underline{\Phi}}^{-1}(t, t')\, \underline{\underline{D}}_i(t') \circledast s_i(t')\, \underline{1}\, dt\, dt' \ . \tag{2.6}$$

Here $\circledast$ means the convolution.

After some manipulation Eq. (2.5) can be written as

$$\Lambda(\underline{s}, \{\mathbf{u}_i\}) = \exp\left\{ < \underline{v}, \underline{s} > - \frac{1}{2} < s, \underline{\underline{\rho}}^{-1} \cdot \underline{s} > \right\} , \tag{2.7}$$

where $\underline{v}$ is a two dimensional function with components $v_i(t)$, $i = \alpha, \beta$

$$v_i(t) = \underline{1}^T\, \underline{\underline{D}}_i(-t) \circledast \int_{-\infty}^{\infty} \underline{\underline{\Phi}}^{-1}(t, t')\, \underline{X}(t')\, dt' = \sum_{k=1}^{K} \sum_{k'=1}^{K} \int_{-\infty}^{\infty} \Phi_{kk'}^{-1}(t + \mathbf{u}_i \cdot \mathbf{r}_k, t') \times x_{k'}(t')\, dt' \ , \quad i = \alpha, \beta \ , \tag{2.8}$$

and $\underline{\underline{\rho}}^{-1}$ is a 2×2 matrix function with entries $\rho_{ij}^{-1}(t, t')$ which is a function dependent on the noise covariance function and the inverse phase velocities

$$\rho_{ij}^{-1}(t,t') = \underline{1}^T \underline{\underline{D}}_i(-t) \circledast \underline{\underline{\Phi}}^{-1}(t,t') \circledast \underline{\underline{D}}_j(-t')\, \underline{1} = \sum_{k} \sum_{k'} \Phi_{kk'}^{-1}(t + \mathbf{u}_i \mathbf{r}_k, t' + \mathbf{u}_j \mathbf{r}_{k'}) \ , \quad i, j = \alpha, \beta \ . \tag{2.9}$$

In order to find the maximum-likelihood estimate (MLE) of $\underline{s} = \mathrm{col}[s_i(t)]$ and the $\mathbf{u}_i$'s, we first fix the $\mathbf{u}_i$'s and maximize $\Lambda(\underline{s}, \{\mathbf{u}_i\})$ over $\underline{s}$. Equation (2.7) is maximized by the choice

$$\tilde{\underline{s}} = \underline{\underline{\rho}} \cdot \underline{v} \ , \tag{2.10}$$

or

$$\tilde{s}_i(t; \mathbf{u}_\alpha, \mathbf{u}_\beta) = \sum_{j=\alpha}^{\beta} \int_{-\infty}^{\infty} \rho_{ij}(t, t')\, v_j(t')\, dt' \ , \quad i = \alpha, \beta \ . \tag{2.11}$$

We substitute (2.10) into (2.7) and call the result $\hat{\Lambda}(\{\mathbf{u}_i\})$

$$\Lambda(\{\mathbf{u}_i\}) = \exp\left\{\frac{1}{2} < \underline{v}, \underline{\underline{\rho}}\, \underline{v} >\right\} = \exp\left\{\frac{1}{2} < \underline{v}, \tilde{\underline{s}} >\right\} \ . \tag{2.12}$$

The MLE of $\mathbf{u}_i$'s are those numbers which maximize (2.12) and the MLE of $s_i(t)$'s are obtained by substituting those results into (2.11)

$$\hat{s}_i(t) = \tilde{s}_i(t; \mathbf{u}_\alpha, \mathbf{u}_\beta) \ , \quad i = \alpha, \beta \ . \tag{2.13}$$

An important property of the MLE $\hat{s}_i(t)$ is that it is an unbiased estimate when the true values of parameters $\mathbf{u}_i$'s are known, i.e.,

$$E[\hat{s}_i(t)] = s_i(t) \ , \quad i = \alpha, \beta \ . \tag{2.14}$$

This can be shown by taking the expectation of $\underline{v}$ and by substituting the result into Equation (2.10). Equation (2.14) indicates that in estimating $s_\beta(t)$, the other event $s_\alpha(t)$ has no effect, that is, a complete null-steering is possible insofar as no estimation error of the $\mathbf{u}_i$'s exists.

It is not difficult to show that the error matrix function of the MLE $\hat{\underline{s}}$ is given by $\underline{\underline{\rho}}$. Moreover, it can be shown that the covariance function of any linear unbiased estimate of $\underline{s}$ cannot be smaller than $\underline{\underline{\rho}}$: let $\underline{s}^*$ be any unbiased estimate of $\underline{s}$, then

$$E[(\underline{s}^* - \underline{s})(\underline{s}^* - \underline{s})^T] \geq \underline{\underline{\rho}} \ . \tag{2.15}$$

A proof can be provided by applying the Cramer-Rao inequality to the conditional likelihood function (2.7) (see also Section IV). Equations (2.14) and (2.15) are clearly a version of the result obtained by Capon et al. [3] for a single event case. When the inverse phase velocities are unknown parameters, the above properties do not hold since $\underline{s}$ is not linearly dependent on those parameters.

III. IMPLEMENTATION OF THE MAXIMUM-LIKELIHOOD PROCESSOR AND ITS PERFORMANCE

A. Structure of the Maximum-Likelihood Processor

In order to see clearly the structure of the maximum-likelihood estimator let us assume that the background noise is stationary in time and is uncorrelated between sensors with common covariance function $\varphi(t - t')$; i.e.,

$$\underline{\underline{\Phi}}(t, t') = \varphi(t - t')\,\underline{\underline{I}} \ . \tag{3.1}$$

Substitution of (3.1) into (2.8) and (2.9) yields

$$v_i(t) = \varphi^{-1}(t) \circledast \sum_k x_k(t + \mathbf{u}_i \cdot \mathbf{r}_k) \ , \quad i = \alpha, \beta \ , \tag{3.2}$$

and

$$\rho_{ij}^{-1}(t, t') = \sum_k \varphi^{-1}(t - t' + (\mathbf{u}_i - \mathbf{u}_j) \cdot \mathbf{r}_k) \ , \quad i, j = \alpha, \beta \ . \tag{3.3}$$

Equation (3.3) can be written as

$$\rho_{ij}^{-1}(t) = K \cdot \varphi^{-1}(t) \circledast c_{ij}(t) \ , \quad i, j = \alpha, \beta \ , \tag{3.4}$$

where $c_{ij}(t)$ is a function depending only on inverse phase velocities defined by

$$c_{ij}(t) = \frac{1}{K} \sum_{k=1}^{K} \delta(t + (\mathbf{u}_i - \mathbf{u}_j)\,\mathbf{r}_k) \ , \quad i, j = \alpha, \beta \ , \tag{3.5}$$

and indicates closeness (coupling) between two events.

Inversion of the 2×2 matrix function (3.4) is given by

$$\underline{\underline{\rho}}(t) = \frac{1}{K} \cdot \varphi(t) \circledast \begin{bmatrix} \delta(t) & -c_{\alpha\beta}(t) \\ -c_{\beta\alpha}(t) & \delta(t) \end{bmatrix} \circledast g_{\alpha\beta}(t) \ , \tag{3.6}$$

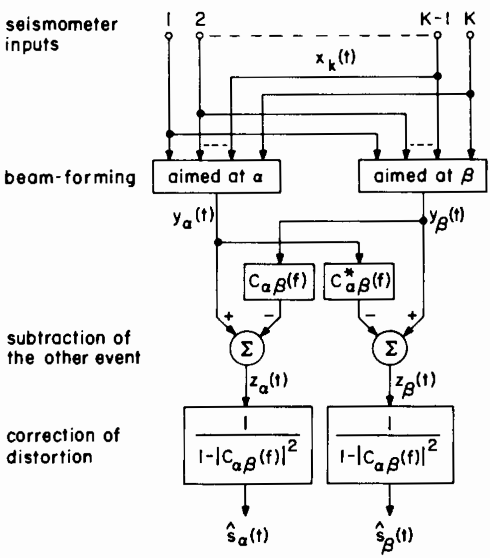


Fig. 1. The maximum-likelihood processor.

where $g_{\alpha\beta}(t)$ is the inverse of $\delta(t) - c_{\alpha\beta}(t) \circledast c_{\alpha\beta}(-t)$ and is an even function

$$g_{\alpha\beta}(t) \circledast [\delta(t) - c_{\alpha\beta}(t) \circledast c_{\alpha\beta}(-t)] = \delta(t) \quad . \tag{3.7}$$

Then substituting (3.2) and (3.3) into (2.10), one obtains

$$\tilde{s}_i(t) = \{y_i(t) - c_{ij}(t) \circledast y_j(t)\} \circledast g_{\alpha\beta}(t) \quad , \quad i = \alpha, \beta \quad , \tag{3.8}$$

where

$$y_i(t) = \frac{1}{K} \sum_{k=1}^{K} x_k(t + \mathbf{u}_i \mathbf{r}_k) \quad , \quad i = \alpha, \beta \quad . \tag{3.9}$$

Note that the noise covariance function $\varphi(t)$ does not come into the structure of the estimator at all. This is a result of the assumption that the noise is uncorrelated between sensors with common spectrum. The structure of the ML-processor is diagrammatically shown in Fig. 1, where the role of the filter $g_{\alpha\beta}(t)$ is to correct distortion caused on the signal waveshape by subtraction of the other event.

B. Performance of the ML-Processor

In this section we will present some numerical results on the performance of the ML-processor compared with the simple delay-and-sum (DS) processor [3]. Event β is regarded as the event of interest and the event α is considered to be a unidirectional noise added to the background noise $n_k(t)$. Let us assume that the noise is stationary Gaussian, independent between sensors with common power spectrum $P(f)$, i.e., $\mathcal{F}(\varphi(t)) = P(f)$. We also assume that event α is a second order stationary process with spectrum $P_\alpha(f)$. The variance of the output of a DS-processor aimed at the desired event is obtained as

$$P_{DS} = \text{var}\,\{y_\beta(t)\} = \frac{1}{K} \int_{-\infty}^{\infty} \{P(f) + K\,|\,C_{\alpha\beta}(f)\,|^2\, P_\alpha(f)\}\, df \quad , \tag{3.10}$$

where $C_{\alpha\beta}(f)$ is the Fourier transform of $c_{\alpha\beta}(t)$ and $y_\beta(t)$ is defined by (3.9). The variance of the MLE $\hat{s}_\beta(t)$ is obtained from (2.15) and (3.6) as

$$\begin{aligned} P_{ML} &= \text{var}\,\{\hat{s}_\beta(t)\} = \rho_{\beta\beta}(0) \quad , \\ &= \frac{1}{K} \int_{-\infty}^{\infty} \frac{P(f)}{1 - |\,C_{\alpha\beta}(f)\,|^2}\, df \quad . \end{aligned} \tag{3.11}$$

We define a processing gain of the MLE over the DS-processor by

$$G_{ML} = 10 \log_{10} \frac{P_{DS}}{P_{ML}} \qquad \text{(dB)} \quad . \tag{3.12}$$

We now give the results of some calculations to determine the ability of the ML processing. In this calculation we assume for simplicity $P_\alpha(f)$ and $P(f)$ have the same shape, i.e.,

$$P_\alpha(f) = M_\alpha\, P(f) \quad . \tag{3.13}$$

We took as the shape of $P_\alpha(f)$ and $P(f)$ a smooth approximation to the spectrum of the output of an individual seismometer for Event α, defined below. This spectrum is plotted in Figure 2.

For a seismic array, we considered the 21 subarrays of the Montana LASA (Large Aperture Seismic Array). The locations of 21 sensors are given in Table I. Two seismic events used in this study are the following: Event α: date, 4/8/66; time observed at LASA center subarray, 1 : 46 : 45.4; Event β: date, 10/29/69; time observed at LASA center subarray, 21 : 08 : 31.65. We assumed that the array was pointed in the direction of Event β.

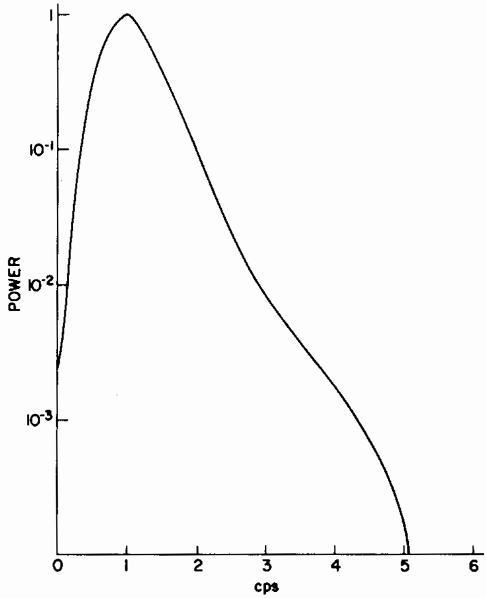


Fig. 2. The power spectrum shape used for the calculation.

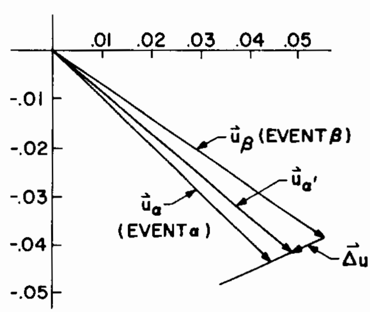


Fig. 3. The inverse phase velocity vectors of the seismic events used in experiments.

k	$r_{k,x}$ [Km]	$r_{k,y}$ [Km]
1	9.975	7.079
2	-66.607	-79.205
3	-54.447	80.571
4	0.0	0.0
5	-7.250	-3.268
6	-11.592	5.199
7	-1.592	8.831
8	7.226	16.772
9	16.049	-2.109
10	4.552	-5.958
11	-2.108	-12.722
12	-19.791	-15.406
13	-12.267	28.175
14	25.386	16.852
15	16.226	-20.605
16	-8.648	-59.925
17	-53.052	7.715
18	12.596	52.726
19	78.155	76.357
20	65.808	-19.189
21	57.157	-86.340

Table I. The locations of 21 sensors of the Montana LASA.

This location corresponds to the inverse phase velocity

$$\mathbf{u}_\beta = [0.057185, -0.038984] \text{ sec/Km} . \tag{3.14}$$

The interfering event had inverse phase velocity

$$\mathbf{u}_\alpha = [0.044631, -0.042763] \text{ sec/Km} . \tag{3.15}$$

By adjusting the time delays we moved this interfering event along the line, $\Delta\mathbf{u}$, connecting $\mathbf{u}_\alpha$ and $\mathbf{u}_\beta$ (see Fig. 3)

$$\Delta\mathbf{u} = \frac{\mathbf{u}_\alpha - \mathbf{u}_\beta}{|\mathbf{u}_\alpha - \mathbf{u}_\beta|} \cdot |\Delta\mathbf{u}| . \tag{3.16}$$

We calculated G_{ML} for $|\Delta\mathbf{u}| = 0. \sim 3.0$ sec/Km choosing M_α as a parameter. These results are valid for this range of values of $|\Delta\mathbf{u}|$ and this argument of $\Delta\mathbf{u}$ no matter where the array is pointed. This value of the argument, of $\Delta\mathbf{u}$ was chosen so that a specific comparison could be made to the performance on the two actual events. This is done in the next section. The results of this calculation of G_{ML} are given in Figure 4.

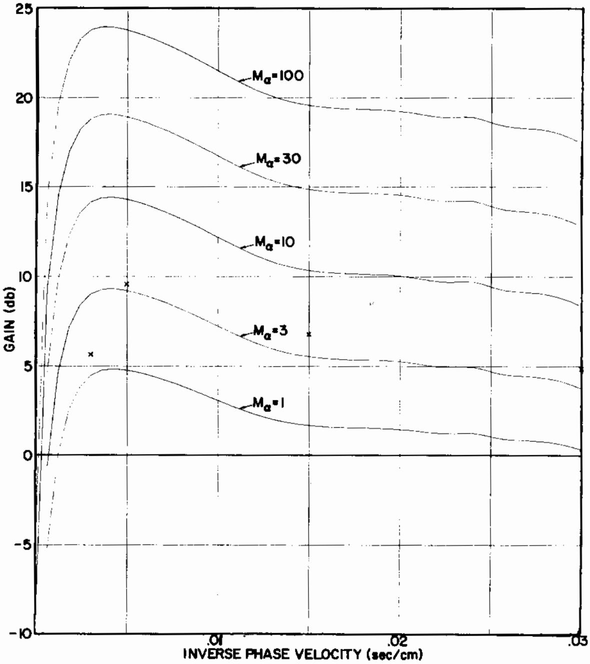


Fig. 4. Gain of the maximum-likelihood processor over the delay-and-sum processor.

C. Experimental Results

The ML-processor outlined above was then applied to data from the Montana LASA. The outputs of the twenty-one center seismometers were used. In this experiment, estimation of the inverse phase velocities was omitted and processing was done only for signal waveform extraction. Figure 5 is a plot of processed outputs for the case where Event β was superposed on Event α with a delay of 25 seconds. The six curves are, from the top, $y_\alpha(t)$, $y_\beta(t)$, $y_\alpha(t) - c_{\alpha\beta}(t) \circledast y_\beta(t)$, $y_\beta(t) - c_{\alpha\beta}(-t) \circledast y_\alpha(t)$, $\hat{s}_\alpha(t)$ and $\hat{s}_\beta(t)$. As is seen from this result the ML-processor performs quite satisfactorily in suppressing interference from the other event.

Now let us compare the result with the performance curves of Figure 4. The background noise in the data of Event α is extremely weak compared with the signal level. This can be noticed from Fig. 5 in which $y_\alpha(t)$ and $y_\beta(t)$ show practically no fluctuations before the arrival of Event α. Choice of $M_\alpha = 100$ (or 20 dB) in terms of (3.13) would be conservative. Then we anticipate a gain of at least 19 dB. However, the value calculated from $y_\beta(t)$ and $\hat{s}_\beta(t)$ of Fig. 4 over the period of the first 25 seconds after the arrival of Event α was 7.8 dB. This value corresponds to $M_\alpha \cong 3$ in the performance curves of Figure 4. This apparent decrease of the interfering signal to the background noise ratio is

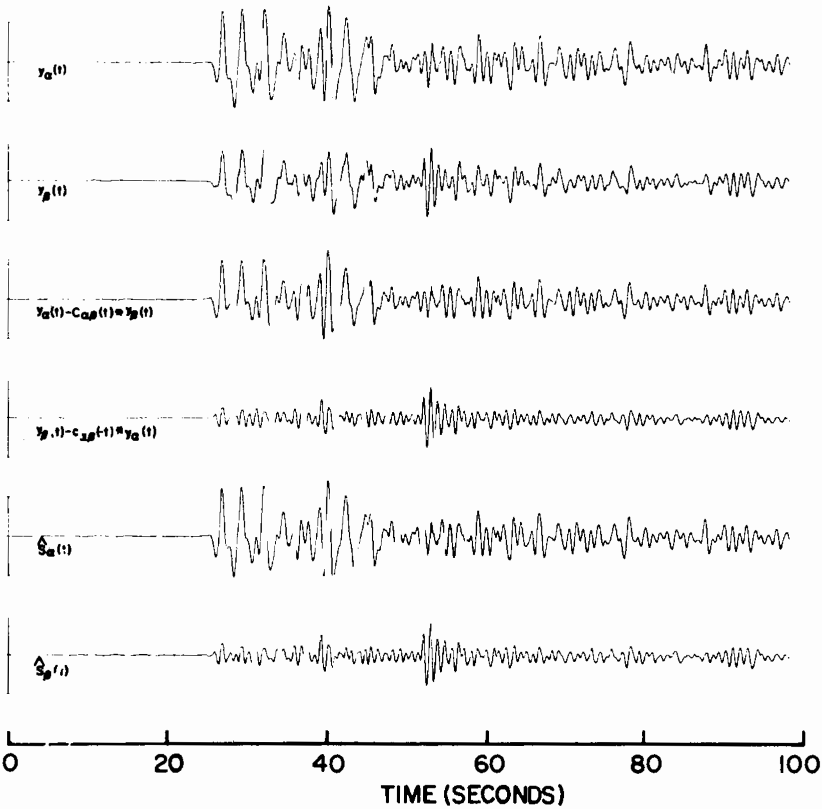


Fig. 5. Processed outputs of the maximum-likelihood processor.

mostly ascribed to a deviation of the signal characteristic from the idealized signal model on which our analysis has been based. The fluctuation observed in traces in Fig. 4 in the interval before the start of Event β is not the amplified background noise, but is due to noiselike components contained in Event α signal.

In the theoretical curves of Fig. 4 we calculated noise power as a function of $\Delta u = u_\alpha - u_\beta$. In our experiments we also moved Event α over the range $|\Delta u| = 0.0015 \sim 0.030$ sec/Km by adjusting time delays. We essentially synthesized data of an artifical event which we call α'. The event α' is the one which we might have observed if Event α had occurred at a point whose inverse phase velocity differs from that of Event β by Δu (see Figure 3). The results are plotted in Fig. 6 for $|\Delta u| = 0.03$, 0.015, 0.005, and 0.003 sec/Km. The gain of the MLE over the SD-processor outputs is as follows (averaged over 25 seconds after the start of Event α)

$\lvert\Delta u\rvert$:	0.003	0.005	0.015	0.030	(sec/Km) ,
Gain :	5.66	9.06	6.81	4.98	(dB) .

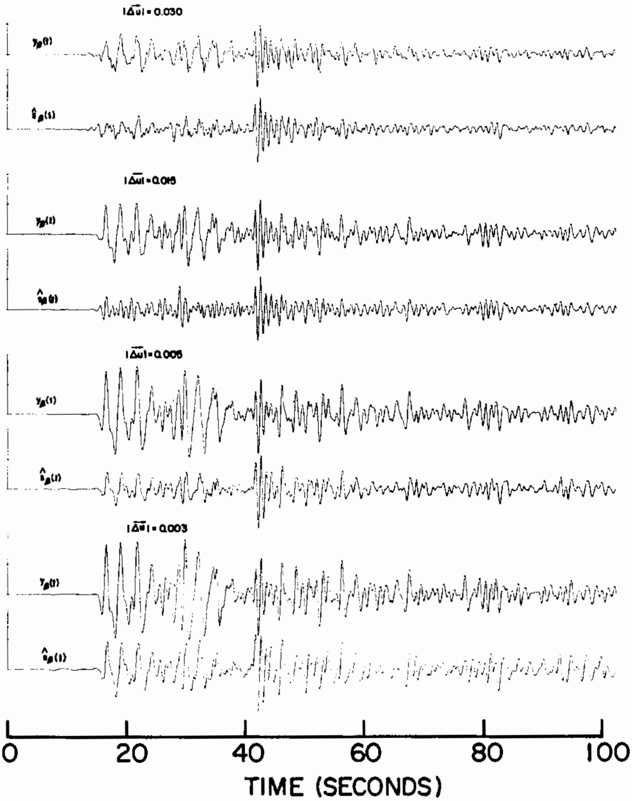


Fig. 6. Outputs of the delay-and-sum processor and the maximum-likelihood processor for various Δu.

These results are plotted in Figure 4 and are seen to correspond roughly to the case $M_{\alpha} = 2 \sim 4$.

In the next section we will develop a modified signal model in which the event has coherent and noncoherent components. We will see there that the ratio of coherent to noncoherent power for Event α is consistent with an M_{α} in the $2 \sim 4$ range.

IV. A MODIFIED SIGNAL MODEL AND ASSOCIATED OPTIMUM PROCESSOR

A. Unconditional Maximum-likelihood Estimator

In the last two sections our analysis has been based on the idealized signal model that the signal is identical at each seismometer except for a time delay. We have been neglecting dispersive effects and local inhomogeneities in the medium which would cause the signal to differ in waveform from sensor to sensor. In the present section a more realistic signal model is considered. It will enable us to more reasonably compare experimental and analytical results.

Consider the case of a single seismic event. According to our modified model, the input to each seismometer consists of three parts

$$x_k(t) = s(t - \mathbf{r}_k \cdot \mathbf{u}) + z_k(t) + n_k(t) \quad , \tag{4.1}$$

where $s(t)$ is a component common to all sensor outputs and will be called the *coherent* component of a signal. The second term of (4.1), $z_k(t)$, is an additional component due to inhomogeneities in the medium and will be called the *noncoherent* component of a signal. From the view point of estimating the coherent component, $s(t)$, the $z_k(t)$'s may be regarded as additional noises. But from the detection viewpoint they should be regarded as stochastic signals and be taken into account to serve for detection of an event. Detection problems will be discussed in a later section.

In case of two (or multiple, in general) simultaneous events, the outputs of the sensors are now given, instead of (2.3) by

$$\underline{X}(t) = \sum_{i=\alpha}^{\beta} \{\underline{\underline{D}}_i(t) \circledast s_i(t)\,\underline{1} + \underline{Z}_i(t)\} + \underline{N}(t) \quad . \tag{4.2}$$

We assume that the noncoherent components $\underline{Z}_i(t)$'s are Gaussian vector processes with zero mean and with covariance functions

$$\underline{\underline{R}}_i(t, t') = E[\underline{Z}_i(t)\, \underline{Z}_i^T(t')] \quad , \quad i = \alpha, \beta \quad . \tag{4.3}$$

Although the assumption that the background noise and the noncoherent components of the signal are Gaussian will be fairly realistic, we know little about the shape of a coherent component. As a matter of fact a novelty of the signal model set by Levin and Kelly [2] lies in the assignment of the signal waveform as a completely unknown time function, thus no *a priori* assumption regarding the shape of the signals are made. However, as was discussed in the last section, a seismic signal contains noise-like noncoherent components with a magnitude which is not negligible compared with the coherent component. In order to take this fact into consideration we need to tie down some of the parameters of the model, at least the average power of the coherent component relative to the noncoherent components.

To make our problem tractable and have a closed form solution, we assume that the coherent signal component $s_i(t)$ is also Gaussian (possibly nonstationary) with covariance function $K_i(t, t')$, $i = \alpha, \beta$. It is to be noticed that the MLE discussed in Sec. II is equivalent to the result which will be obtained by assuming $K_i(t, t') = \infty$ in the following discussion. Although the Gaussian assumption about the signal is set for ease of mathematical treatment, the optimum solution obtained in the following argument holds for the non-Gaussian case also under the minimum mean squared error criterion. This will be discussed in the course of the argument.

The probability density function of the signal $s_i(t)$ is then given by

$$\sigma_i(s_i) = c_i \exp \left\{ -\frac{1}{2} \iint s_i(t)\, K_i^{-1}(t,t')\, s_i(t')\, dt\, dt' \right\} , \quad i = \alpha, \beta . \tag{4.4}$$

The optimum procedure, with those statistics given, will be the unconditional maximum-likelihood estimate (UMLE) [8] or the maximum a posterior probability computer [10]. The unconditional likelihood ratio is given by

$$L(\underline{s}, \{\mathbf{u}_i\}) = \left\{ \prod_{i=\alpha}^{\beta} \sigma_i(s_i) \right\} \Lambda(\underline{s}, \{\mathbf{u}_i\}) , \tag{4.5}$$

where Λ is the (conditional) likelihood ratio given by

$$\Lambda(\underline{s}, \{\mathbf{u}_i\}) = c \exp \left\{ \frac{1}{2} \|\underline{X}\|^2_{\Phi} - \frac{1}{2} \|\underline{X}\|^2_{\Phi^*} + \langle \underline{v}^*, \underline{s} \rangle - \frac{1}{2} \langle \underline{s}, \underline{\underline{\rho}}^{*-1} \cdot \underline{s} \rangle \right\} , \tag{4.6}$$

where $\underline{v}^*$ is a 2×1 matrix function with entries $v_i^*(t)$, $i = \alpha, \beta$ and $\underline{\underline{\rho}}^{*-1}$ is a 2×2 matrix function with entries $\rho_{ij}^{*-1}(t,t')$. These functions take the same form as $v_i(t)$ and $\rho_{ij}^{-1}(t,t')$ of Eqs. (2.8) and (2.9) except that the noise covariance function $\underline{\underline{\Phi}}$ is now replaced by $\underline{\underline{\Phi}}^*$ which is defined by

$$\underline{\underline{\Phi}}^*(t,t') = \underline{\underline{\Phi}}(t,t') + \sum_{i=\alpha}^{\beta} \underline{\underline{R}}_i(t,t') . \tag{4.7}$$

Defining a 2×2 matrix function $\underline{\underline{K}}(t,t')$ by

$$\underline{\underline{K}}(t,t') = \begin{bmatrix} K_\alpha(t,t') & 0 \\ 0 & K_\beta(t,t') \end{bmatrix} , \tag{4.8}$$

the logarithm of the unconditional likelihood ratio takes the following form

$$\ell(\underline{s}, \mathbf{u}_\alpha, \mathbf{u}_\beta) = 2 \log L = \|\underline{X}\|^2_{\Phi} - \|\underline{X}\|^2_{\Phi^*} + 2 \langle \underline{v}^*, \underline{s} \rangle - \langle \underline{s}, (\underline{\underline{\rho}}^{*-1} + \underline{\underline{K}}^{-1})\, \underline{s} \rangle . \tag{4.9}$$

The unconditional maximum-likelihood estimates (UMLE) of s, $\mathbf{u}_\alpha$ and $\mathbf{u}_\beta$ are found by first fixing the $\mathbf{u}_i$'s and maximizing $\ell(\underline{s}, \mathbf{u}_\alpha, \mathbf{u}_\beta)$ over $\underline{s}$. Such a value $\underline{\tilde{s}}$ is readily obtained as

$$\underline{\tilde{s}} = \underline{\underline{\xi}}\, \underline{v}^* , \tag{4.10}$$

where a 2×2 matrix function $\underline{\underline{\xi}}$ is defined by

$$\underline{\underline{\xi}}^{-1} = \underline{\underline{\rho}}^{*-1} + \underline{\underline{K}}^{-1} . \tag{4.11}$$

The UMLE of $\mathbf{u}_i$'s are those numbers $\check{\mathbf{u}}_i$'s which maximize the quantity $\langle \underline{v}^*, \underline{\tilde{s}} \rangle$ or $\langle \underline{v}^*, \underline{\underline{\xi}}\, \underline{v}^* \rangle$. When these estimates are substituted in (4.10), the result is the UMLE of $\underline{s}$ which we denote by $\underline{\check{s}}$.

Let us discuss some important properties of the UMLE assuming the true values of parameters u_i's are known. Taking the conditional expectation of $\check{\underline{s}}$, we have

$$\underset{\underline{X}/\underline{s}}{E}[\check{\underline{s}}] = \underline{\underline{\xi}} \cdot \underset{\underline{X}/\underline{s}}{E}[\underline{v}^*] = \underline{\underline{\xi}} \cdot \underline{\underline{\rho}}^{*-1} \cdot \underline{s} = (\underline{\underline{I}} - \underline{\underline{\xi}} \cdot \underline{\underline{K}}^{-1})\underline{s} \;. \tag{4.12}$$

Thus $\check{\underline{s}}$ is a *biased* estimate with the bias

$$\underline{b}(\underline{s}) = -\underline{\underline{\xi}} \cdot \underline{\underline{K}}^{-1} \cdot \underline{s} \;. \tag{4.13}$$

The variance of $\check{\underline{s}}$ must satisfy the following inequality which is an extension of the Cramer-Rao inequality to unconditional estimates [9]

$$\underset{\underline{s}}{E}\;\underset{\underline{X}/\underline{s}}{E}[(\check{\underline{s}} - \underline{s} - \underline{b}(\underline{s}))(\check{\underline{s}} - \underline{s} - \underline{b}(\underline{s}))^T] \geq \underline{\underline{\Delta}} \cdot \underline{\underline{J}}^{-1} \cdot \underline{\underline{\Delta}}^T \tag{4.14}$$

where $\underline{\underline{J}}$ and $\underline{\underline{\Delta}}$ are 2×2 matrix functions defined by

$$\underline{\underline{J}} = \underset{\underline{s}}{E}\;\underset{\underline{X}/\underline{s}}{E}\left\{\frac{\partial \log L}{\partial \underline{s}}\left[\frac{\partial \log L}{\partial \underline{s}}\right]^T\right\}, \tag{4.15}$$

and

$$\underline{\underline{\Delta}} = \underline{\underline{I}} + \underset{\underline{s}}{E}\left[\frac{\partial \underline{b}(\underline{s})}{\partial \underline{s}T}\right], \tag{4.16}$$

respectively. By substituting the unconditional likelihood ratio (4.6) into the above equations and after some manipulation, we have the following result

$$\underset{\underline{s}}{E}\;\underset{\underline{X}/\underline{s}}{E}[(\check{\underline{s}} - \underline{s} - \underline{b}(\underline{s}))(\check{\underline{s}} - \underline{s} - \underline{b}(\underline{s}))^T] \geq \underline{\underline{\rho}}^* - \underline{\underline{\xi}}\,\underline{\underline{K}}^{-1}\,\underline{\underline{\rho}}^* - \underline{\underline{\xi}} \cdot \underline{\underline{K}}^{-1} \cdot \underline{\underline{\xi}}^T \;. \tag{4.17}$$

Therefore, the mean squared error of $\check{\underline{s}}$ must satisfy the inequality

$$\underset{\underline{s}}{E} \cdot \underset{\underline{X}/\underline{s}}{E}[(\check{\underline{s}} - \underline{s})(\check{\underline{s}} - \underline{s})^T] \geq \underset{\underline{s}}{E}[\underline{b}(\underline{s})\underline{b}^T(\underline{s})] + \underline{\underline{\Delta}}\,\underline{\underline{J}}^{-1}\,\underline{\underline{\Delta}}^T \;,$$

$$= \underline{\underline{\rho}}^* - \underline{\underline{\xi}}\,\underline{\underline{K}}^{-1}\,\underline{\underline{\rho}}^* = \underline{\underline{\xi}} \;. \tag{4.18}$$

It is not difficult to show that the UMLE actually attains the equality in (4.17) and (4.18).

B. The Optimum Processor for Rejecting Interfering Signal and Noise

Let us assume that the background noise $\underline{N}(t)$ and the noncoherent components $\underline{Z}_l(t)$ are stationary in time and are uncorrelated between seismometers with common covariance functions $\varphi(t\text{-}t')$ and $r_l(t\text{-}t')$, respectively.

Define $\varphi^*(t)$ as the covariance function of all noiselike components

$$\varphi^*(t) = \varphi(t) + \sum_{l=\alpha}^{\beta} r_l(t) \tag{4.19}$$

The result of Sec. III will then carry over directly to our present case (see Eqs. (3.2) to (3.9)).

Let us define a function $h_i(t)$ by

$$h_i(t) = \frac{1}{K} \varphi^*(t) \circledast K_i^{-1}(t) \quad , \quad i = \alpha, \beta \quad , \tag{4.20}$$

or

$$H_i(f) = \frac{1}{K} \frac{P^*(f)}{P_i(f)} , \quad i = \alpha, \beta \quad , \tag{4.21}$$

where $P^*(f)$ and $P_i(f)$ are power spectra of the noiselike components and of $s_i(t)$, i.e., $P^*(f) = \mathfrak{F}\{\varphi^*(t)\}$ and $P_i(f) = \mathfrak{F}\{K_i(t)\}$. The UMLE $\underline{\check{s}}$ is then obtained in a closed form

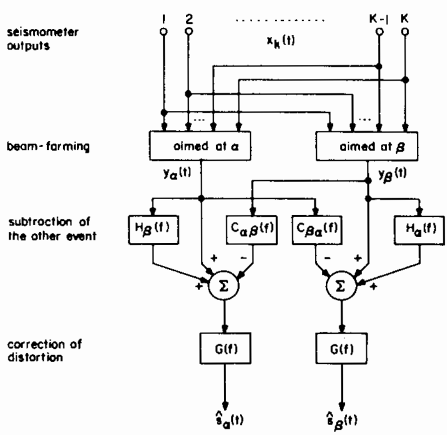
$$\check{s}_i(t) = [\{\delta(t) + h_j(t)\} \circledast y_i(t) - c_{ij}(t) \circledast y_j(t)] \circledast g^*_{\alpha\beta}(t) \quad , \quad i = \alpha, \beta \quad , \tag{4.22}$$

where $g^*_{\alpha\beta}(t)$ is a function which satisfies the equation

$$g^*_{\alpha\beta}(t) \; * \; [\{\delta(t) + h_\alpha(t)\} \circledast \{\delta(t) + h_\beta(t)\} - c_{\alpha\beta}(t) \circledast c_{\alpha\beta}(-t)] = \delta(t) \quad . \tag{4.23}$$

Figure 7 represents a diagram of the UML-processor for two simultaneous events. Notice that $\varphi^*(t)$ of (4.19) does not come into the structure. We notice also its basic similarity to the ML-processor of Figure 1. In fact the former is reduced to the latter when $h_\alpha(t)$ and $h_\beta(t)$ are set to zero.

As was mentioned earlier the UMLE is equivalent to the minimum mean



$$G(f) = \frac{1}{(1+H_\alpha(f))(1+H_\beta(f)) - |C_{\alpha\beta}(f)|^2}$$

where

$$H_i(f) = \frac{1}{K} \frac{P_n(f)}{P_i(f)} , \; i = \alpha, \beta$$

Fig. 7. The unconditional maximum-likelihood processor.

squared error linear estimate but it is not unbiased. We can now introduce another filter which compromises unbiasedness and minimization of the mean squared error. Let us assume again that $s_\beta(t)$ is a desired signal and $s_\alpha(t)$ be an interfering unidirectional noise. We now define an estimate $s_{\beta,\mathrm{opt}}(t)$ as the value of $s_\beta(t)$ of (4.22) when $h_\beta(t)$ is set equal to zero, i.e.,

$$s_{\beta,\mathrm{opt}}(t) = [\{\delta(t) + h_\alpha(t)\} \circledast y_\beta(t) - c_{\beta\alpha}(t) \circledast y_\alpha(t)] * g^*_{\alpha\beta}(t) \quad , \tag{4.24}$$

where

$$g^*_{\alpha\beta}(t) \circledast [\delta(t) + h_\alpha(t) - c_{\alpha\beta}(t) \circledast c_{\alpha\beta}(-t)] = \delta(t) \quad . \tag{4.25}$$

An alternative expression for $s_{\beta,\mathrm{opt}}(t)$ is obtained by

$$s_{\beta,\mathrm{opt}}(t) = [y_\beta(t) - d_\alpha(t) \circledast c_{\alpha\beta}(-t) \circledast y_\alpha(t)] \circledast [\delta(t) - d_\alpha(t) \circledast c_{\alpha\beta}(t) \circledast c_{\alpha\beta}(-t)]^{-1} \quad , \tag{4.26}$$

or in the frequency domain expression

$$S_{\beta,\mathrm{opt}}(f) = \frac{Y_\beta(f) - D_\alpha(f)\, C_{\beta\alpha}(f)\, Y_\alpha(f)}{1 - D_\alpha(f)\, |C_{\beta\alpha}(f)|^2} \quad , \tag{4.27}$$

where the function $d_\alpha(t)$ and its Fourier transform $D_\alpha(f)$ are defined by

$$d_\alpha(t) \circledast [\delta(t) + h_\alpha(t)] = \delta(t) \quad , \tag{4.28}$$

and by

$$D_\alpha(f) = \frac{K P_\alpha(f)}{P^*(f) + K P_\alpha(f)} \quad . \tag{4.29}$$

The mean squared error of $s_{\beta,\mathrm{opt}}(t)$ is obtained, in the frequency domain, as

$$P_{\mathrm{opt}} = \frac{1}{K} \int_{-\infty}^{\infty} \frac{P^*(f)}{1 - D_\alpha(f)\, |C_{\alpha\beta}(f)|^2}\, df \quad . \tag{4.30}$$

It should be remarked that the expression (4.27) was originally obtained by Kelly and Levin from a somewhat different approach [2]. We will call the processor which generates $s_{\beta,\mathrm{opt}}(t)$ the *optimum* processor in the sense that it is the optimum for rejecting a second seismic event when the power spectral ratio of the coherent component and the total noise-like component is known.

C. Performance of the Optimum Processor and Experimental Results

The purpose of this section is to present the numerical results of the performance of the optimum processor defined above, along with some experimental results. As we did in Sec. III B, we assume that $P^*(f)$ and $P_\alpha(f)$ are of the same shape and are given by Figure 2. Defining a constant M_α by (3.13), P_{opt} is written as

$$P_{\mathrm{opt}} = \frac{1}{K} \int_{-\infty}^{\infty} \frac{P^*(f)}{1 - D_\alpha\, |C_{\alpha\beta}(f)|^2}\, df \quad , \tag{4.31}$$

where

$$D_\alpha = \frac{K M_\alpha}{1 + K M_\alpha} \quad . \tag{4.32}$$

The processing gain of the optimum processor over the delay-and-sum processor, defined by

$$G_{opt} = 10 \log_{10} \frac{P_{DS}}{P_{opt}} \quad , \tag{4.33}$$

is a function of $\Delta \mathbf{u}$ and M_α only. G_{opt} is given in Fig. 8 for various values of M_α and for the range of $\Delta \mathbf{u}$, described in Section III B. Comparing these curves with those in Fig. 4, the difference in performances of the ML-processor and the optimum processor can be noticed especially for small $|\Delta \mathbf{u}|$ and M_α. $\Delta G = G_{opt} - G_{ML}$ is a gain due to the knowledge of the interfering signal to the noise component ratio.

This processor was also run on the data described in Section III C. The results are given in the table below and are plotted in Figure 8.

$\lvert\Delta \mathbf{u}\rvert$:	0.003	0.005	0.015	0.030	(sec/Km) ,
Gain :	6.01	9.15	6.81	4.99	(dB) .

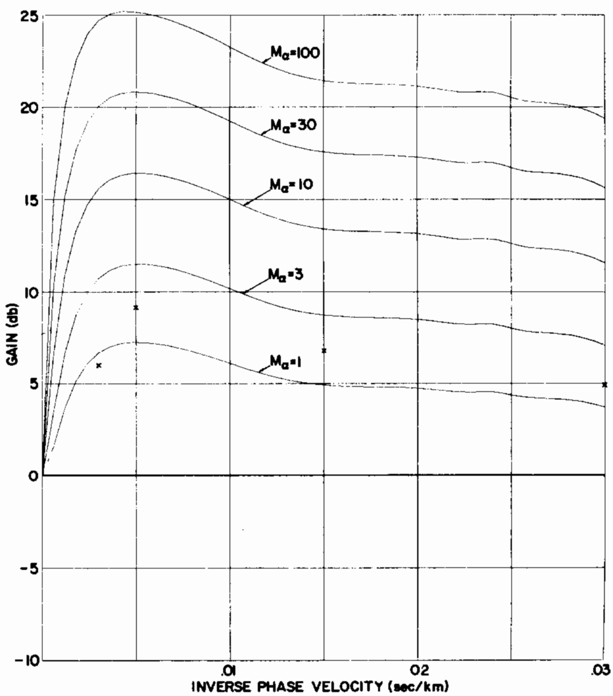


Fig. 8. Gain of the optimum processor over the delay-and-sum processor.

M_α was also estimated on this data by averaging over the first 25 seconds of Event α. The estimate obtained was 2.8. Notice that the data generally fit the model for this value of M_α. Notice also that this value of M_α gives general agreement to the output of the maximum-likelihood processor (Figure 4).

There are two discrepancies in the experimental results. First, the "optimum" processors do not perform as well as for small $\Delta \mathbf{u}$ as is predicted. Second, there is not the predicted difference between the ML processor of Section III and the optimum processor of this section. Our conjecture is that these effects are a consequence of a partial coherency or directional quality of the non-coherent component of the signal.

V. MAXIMUM-LIKELIHOOD DETECTION

In this section we consider the detection problem. In the detection of a signal of known form but with parameters which are only vaguely known (e.g., arrival time and carrier frequency in case of a radar signal), the method of the maximum likelihood detection is often used [7]. In this detection strategy one assumes that a signal of the expected form is present in the observed input and the signal parameters are estimated by the method of maximum-likelihood. These estimates are tested to determine whether they might have been obtained from an input containing noise alone. The detector structure takes the form of the maximum-likelihood estimator followed by a matched filter (or equivalently a correlator) that treats the MLE as the true value of the signal. The problem of detecting the presence or absence of a Gaussian signal in additive Gaussian noise can be handled in a similar fashion. The optimum detector in this case consists of the UML estimator (equivalent to the optimum filter under the minimum mean squared error criterion) and a correlator that treats the UMLE as the true signal waveshape. This interpretation of an optimum receiver has been found by Price [10] and Kailath [11]. In the present section a similar approach is applied to the case of two simultaneous seismic event signals.

We start from the original signal model which does not take into account non-coherent components. If we substitute the MLE $\hat{s}$ of the signal waveform s and parameters $\mathbf{u}_i$'s into (2.7), the likelihood ratio becomes

$$\Lambda(\underline{\hat{s}}, \{\hat{\mathbf{u}}_i\}) = \exp\left\{\frac{1}{2} < \underline{v}, \underline{\hat{s}} >\right\} \quad , \tag{5.1}$$

which is the maximum-likelihood detection statistic. When the background noise is uncorrelated among sensors (see (2.1)), the detection statistic is written in terms of $\varphi(t)$ and the beam output $y_i(t)$

$$T_1 = < \underline{v}, \underline{\hat{s}} > = \int \underline{\hat{s}}^T(t) \{\varphi^{-1}(t) \circledast \underline{y}(t)\} \, dt \quad . \tag{5.2}$$

Note that the ML-detector depends on the noise covariance function $\varphi(t)$ although the ML-estimator does not.

Equation (5.2) is a test statistic to detect whether two events are present or noise only is present. A more realistic situation will be the case where one knows that an event has occurred but is uncertain about the existence of the second event. This detection problem is formulated as the following hypothesis test

$$H_0 : \underline{X}(t) = \underline{\underline{D}}_\alpha(t) \circledast s_\alpha(t)\, \underline{1} + \underline{N}(t) \ , \tag{5.3}$$

$$H_1 : \underline{X}(t) = \sum_{i=\alpha}^{\beta} \{\underline{\underline{D}}_i(t) \circledast s_i(t)\, \underline{1}\} + \underline{N}(t) \ . \tag{5.4}$$

Then the conditional likelihood ratio function for H_1 against H_0 is calculated as

$$\Lambda(\underline{s}) = \exp\left\{ <v_\beta, s_\beta> - <s_\alpha, \rho_{\alpha\beta}^{-1}\ s_\beta> - \frac{1}{2}<s_\beta, \rho_{\beta\beta}^{-1}\ s_\beta> \right\} . \tag{5.5}$$

If one substitutes the MLE $\hat{\underline{s}}$ of (2.13) into (5.5), one obtains after some manipulation

$$\Lambda(\hat{\underline{s}}) = \exp \frac{1}{2} <\hat{s}_\beta, \rho_{\beta\beta}^{-1}\ \hat{s}_\beta> \ . \tag{5.6}$$

Thus the detection of the event β is equivalent to comparing the following quantity with some threshold

$$T_2 = \iint \hat{s}_\beta(t)\, \rho_{\beta\beta}^{-1}(t, t')\, \hat{s}_\beta(t')\, dt\, dt' \ . \tag{5.7}$$

Now let us consider the modified signal model discussed in Section IV. If *a priori* knowledge of the signal waveform is given as in Eq. (4.4), then the optimum detector will be derived from the unconditional likelihood ratio of (4.5). By substituting the UML estimate $\underline{s}$ into (4.9), a test statistic is obtained as

$$T_3 = ||\underline{X}||_\Phi^2 - ||\underline{X}||_{\Phi^*}^2 + <\underline{v}^*, \underline{s}> \ . \tag{5.8}$$

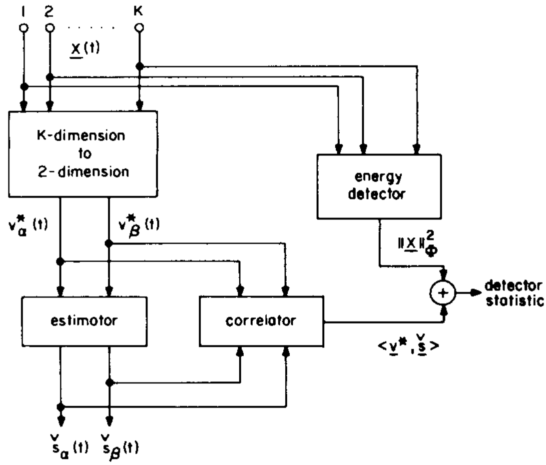
If the background noise is much weaker than the noncoherent components of the event, then $\underline{\underline{\Phi}}^{-1} >> \underline{\underline{\Phi}}^{*-1}$ and thus T_3 can be approximated as follows

$$T_3 \cong ||\underline{X}||_\Phi^2 + <\underline{v}^*, \underline{s}> \ . \tag{5.9}$$

Note that the last term of (5.8) or (5.9) is essentially the same as T_1 of (5.2) and is the estimated signal energy to noise energy ratio. The other terms of (5.8) and (5.9) are due to the noncoherent components. If the background noise is uncorrelated and white, the first term of (5.8) is simply the sum of energys of each seismometer's output. Figure 9 illustrates the structure of the maximum-likelihood detector.

Now we will discuss again the problem of detecting the second event. It can be treated in a similar way as we did in (5.3) through (5.7) by setting the following hypothesis test

$$H_0 = \underline{\underline{D}}_\alpha(t) \circledast s_\alpha(t)\, \underline{1} + \underline{Z}_\alpha(t) + \underline{N}(t) \ , \tag{5.10}$$



$$T = \|\underline{x}\|^2_{\Phi} - \|\underline{x}\|^2_{\Phi^*} + <\underline{v}^*, \check{\underline{s}}>$$

$$\cong \|\underline{x}\|^2_{\Phi} + <\underline{v}^*, \check{\underline{s}}>$$

$$\underline{\underline{\Phi}}^* = \underline{\underline{\Phi}} + \sum_{i=\alpha}^{\beta} \underline{\underline{R}}_i$$

$$v_i^*(t) = \underline{1}^T \underline{\underline{D}}(-t) * \int \underline{\underline{\Phi}}^{*-1}(t,t')\,\underline{x}(t')\,dt'$$

Fig. 9. The maximum-likelihood detector.

$$H_1 = \sum_{i=\alpha}^{\beta} \{\underline{\underline{D}}_i(t) \circledast s_i(t)\, \underline{1} + \underline{Z}_i(t)\} + \underline{N}(t) \quad . \tag{5.11}$$

The unconditional likelihood ratio function for this case can be defined in the similar way as in (4.5) and after the UMLE of (4.10) being substituted, the logarithm of the unconditional likelihood function can be written as

$$\ell(\check{\underline{s}}) = \|\underline{X}\|^2_{\tilde{\Phi}} - \|\underline{X}\|^2_{\Phi^*} - 2 < \tilde{v}_\alpha, \check{s}_\alpha > + < \check{s}_\alpha, \tilde{\xi}^{-1}_{\alpha\alpha}\, \check{s}_\alpha > + < \check{\underline{s}}, \underline{\underline{\xi}}^{*-1}\, \check{\underline{s}} > \quad , \tag{5.12}$$

where $\tilde{v}_\alpha(t)$ and $\tilde{\underline{\underline{\xi}}}^{-1}(t, t')$ are defined in the similar way as $v_\alpha(t)$ of (2.8) and $\underline{\underline{\xi}}^{-1}(t, t')$ of (4.11), respectively, except that the noise covariance function involved is replaced by $\tilde{\Phi}$ which is defined by $\tilde{\underline{\underline{\Phi}}} = \underline{\underline{\Phi}} + \underline{\underline{R}}_\alpha$. This complicated structure is somewhat simplified if we assume uncorrelatedness of noncoherent components and the noise among seismometers. In this case (5.12) can be written as

$$T_4 = \|\underline{X}\|^2_{\tilde{\Phi}} - \|\underline{X}\|^2_{\Phi^*} - <\check{\underline{s}}, (\underline{\underline{\rho}}^{-1} - \underline{\underline{\rho}}^{*-1})\, \check{\underline{s}}> + <\check{s}_\beta, (\rho^{-1}_{\beta\beta} + K^{-1}_\beta)\, \check{s}_\beta> \quad . \tag{5.13}$$

If we are not given the covariance functions $K_i(t)$, $i = \alpha, \beta$, then the optimum detector will be obtained by setting $K_i(t) = \infty$ and replacing the UMLE $\check{\underline{s}}$ by the MLE $\hat{\underline{s}}$.

ACKNOWLEDGMENT

The work reported was supported in part under contract AF 19628-67-C-0198 sponsored by the Advanced Research Projects Agency, Department of Defense.

REFERENCES

[1] J. B. Burg, "Three-Dimensional Filtering With an Array of Seismometers," *Geophysics*, 29, No. 5, 693–713, (Oct., 1964).

[2] E. J. Kelly, M. J. Levin, "Signal Parameter Estimation for Seismometer Arrays," M.I.T. Lincoln Lab., Lexington, Mass., Tech. Rept. 339, (Jan., 1960).

[3] P. E. Green, E. J. Kelly, M. J. Levin, "A Comparison of Seismic Array Processing Methods," *Geophysics J. R. Astr. Soc.*, 11, 67–84, (1966).

[4] J. Capon, R. Greenfield, R. Kolker, "Multidimensional Maximum-Likelihood Processing of a Large Aperture Seismic Array," *Proc. IEEE*, 55, 192–211, (Feb., 1967).

[5] F. C. Schweppe, "Sensor-Array Data Processing for Multiple-Signal Sources," *IEEE Trans. on Information Theory*, IT-14, No. 2, 294–304, March, 1968.

[6] C. R. Rao, "Minimum Variance and the Estimation of Several Parameters," *Proc. Camb. Phil. Soc.*, 43, 280–283.

[7] C. Helstrom, *Statistical Theory of Signal Detection*, (London England: Permagon, 1960).

[8] D. Middleton, *An Introduction to Statistical Communication Theory*, (New York, New York: McGraw-Hill Book Co., 1960).

[9] H. Kobayashi, J. B. Thomas, "The Generalized Cramer-Rao Inequality and Its Application to Parameter Estimation," *First Asilomar Conference on Circuit and Systems*, (1967), also IBM Research Report 2100, (May, 1968).

[10] R. Price, "Optimum Detection of Random Signals in Noise, with Applications to Scatter-Multipath Communication, Part I," *IRE Trans. Inf. Theory*, IT-2, 125–135, (April, 1956).

[11] T. Kailath, "Adaptive Matched Filters," R. Bellman, Ed. *Mathematical Optimization Techniques*, (Berkeley, Calif.: University of California Press, 1963) ch. 6.