Royal Norwegian Council
for Scientific and Industrial Research
(NTNF)

NORSAR

NORSAR Scientific Report No. 1-89/90

Semiannual Technical Summary

1 April — 30 September 1989

Kjeller, December 1989

7.7 Statistically optimal event detection using small array data

A generalization of Capon's maximum likelihood technique for detection and estimation of seismic signals is introduced. By using a multidimensional autoregressive approximation of seismic array noise, we have developed a technique to use Capon's group filter for on-line processing. Such autoregressive adaptation to the current noise matrix power spectrum yields good suppression of mutually correlated array noise processes. An example is shown of this technique as applied to detection of the a small Semipalatinsk underground explosion recorded at the ARCESS array.

Nuclear explosion monitoring using seismic data is faced with the problem that signals of small explosions are masked by noise, and thus have to be extracted using features of both the noise and the signal. Small arrays appear to be especially suited for that purpose. This is due to the strong correlation of noise between different closely located receivers which gives us an opportunity to obtain significant noise suppression. To realize this opportunity, special software is needed.

We have to solve two main problems: (1) to detect event signals, and (2) to classify detected signals as originating from either an explosion or an earthquake. The second task is very complex. The proper identification needs first of all estimation of signal waveform and then estimation of signal parameters such as onset times of different phases, power and spectral features, and so on. While the detection of a signal can be done in a relatively narrow frequency band (for example using a high frequency band only) the classification must in principle be based on wide band methods. This is so because bandpass filtering distorts not only the noise but also the signal, thus possibly eliminating useful classification features.

The problems mentioned can be formulated in terms of mathematical statistics, and optimal decision rules for these mathematical problems have to be found. The first task is to detect the signal. The time series received by the array has the following structure: $X_t = S_t + \xi_t, t \in \ldots - 1, 0, 1, \ldots$, where S_t is signal and ξ_t is noise, X_t is a vector of different receiver outputs. These outputs are observed through a moving window. Using data in the window we must make a decision: does it contain a signal or not. This problem can be solved in terms of statistical hypothesis testing theory. It is necessary to test hypothesis H_0: observations in the moving window are pure noise, versus hypothesis H_1: they comprise signal plus noise. We consider hypothesis H_0 to be simple, but hypothesis H_1 to be complex. This is because the statistical characteristics of noise can be measured before the signal arrival in a stage of adaptation, but the statistical features of seismic signals are almost completely unknown.

The problem is now to choose a decision function which would provide the smallest average error probability for all possible signals. This problem can be solved using a Bayesian approach and finite dimensional parameterization of the signal (Kushnir and Lapshin, 1984) (see Appendix). The decision has the form of an algorithm as shown in Fig. 7.7.1. It consists of a group filter followed by autocorrelators, calculation of a quadratic form and a trigger which compares the quadratic form with a threshold. The transfer function of the group filter is a vector described as a product of the inverse matrix power spectrum F^{-1} of the noise ξ_t and a vector G of phase shift factors due to signal delays τ_1. The algorithm shown has a form which can be easily realized as an on-line procedure. This is mostly due to the use of multidimensional autoregressive (AR) estimation of the matrix power spectrum density F of the noise. This allows us to avoid direct inversion of spectral power matrices and is very convenient in an adaptive procedure (Haykin, 1979; Kushnir *et al*, 1980).

The second problem mentioned above is to estimate properly the signal waveform. The main purpose of this estimation is to make a correct decision of whether we have an explosion or an earthquake. We must do it as precisely as possible. The model of observation here has the following form: $X_t = G_t * u_t + \xi_t$, where $*$ is a sign of convolution, u_t is a signal waveform to be estimated — (the particle motion along the seismic ray), G_t is the transfer function on the path from the seismic source to the receiver. For plane waves, this vector is defined by time delays only. For solving the estimation problem, a conditional Wiener filter can be constructed. This filter minimizes the variance of the estimate $E\{(\hat{u}_t - u_t)^2\}$ under the condition that the mathematical expectation of the estimate coincides with the real signal: $E\{\hat{u}_t\} = u_t$.

It so happens that the Wiener filter we are looking for consists of the same group filter which is used for the detection followed by a restitution filter (which makes it possible to obtain the signal undistorted). Capon was the first to propose this filter for seismic signal extraction from array data (Capon, 1970). The complete array procedure for detection and classification is shown in Fig. 7.7.2. We adapt to the noise matrix power spectrum by estimating its AR parameters and computing vector coefficients of the group filter. Then we perform group filtration in the moving window and detect the signal. The first two operations are made periodically according to the interval of noise stationarity. The third — group filtration — can be devised as an on-line procedure. After the signal is detected it must be filtered by the restitution filter, which refines its shape. And finally classification may be done using the estimates of signal parameters.

The software designed was tested using simulated data with the aim of comparing its actual efficiency with the theoretical one. The results of these tests are shown in Fig. 7.7.3. The depicted curves are the gains in power signal-to-noise ratio of undisturbed optimal group filtration relative to a conventional beam versus the coherence coefficient of the noise. The latter is defined as the ratio between the largest and the smallest eigenvalues of the noise power spectrum matrix. We see that the mentioned gain may be very significant if the array noise is coherent enough. This happens in practice at small aperture arrays. The calculations were made for the central subarray of the NORSAR array.

Highly promising results were obtained by the use of ARCESS data for signals from one of the smallest nuclear tests known to have been conducted at the Semipalatinsk test site. In Fig. 7.7.4 we display the records for four ARCESS channels and note that the signal is obscured by the noise. For the conventional beam trace (Fig. 7.7.5) the signal is likewise not seen, but inspecting the output of the undisturbed group filter we can see the signal clearly. The power signal-to-noise ratio gain relative to the beam is approximately a factor of 70–80 and it is 140–160 when compared with a single channel. The trace shown is calculated using 6 matrix AR parameters of noise. We also used other numbers of AR parameters and the results seemed to be stable. Such high suppression of noise is achieved mainly due to the high correlation of noise records in the inner ARCESS stations (see Fig. 7.7.4, traces A0,A3).

Fig. 7.7.6 shows that the group filter used does in fact retain the shape of the original waveform. The first trace is a wide frequency band waveform used for simulating the plane wave arriving at ARCESS. These simulated data were processed by the group filter used for the previously shown signal extraction. The resulting (second) trace practically coincides with the first. So, if the real signal is plane wave, it will be undisturbed by the group filter in the frequency band from 0.5 to 5 Hz.

The conventional method used for the detection of weak signals is the filtration of the array beam in a band of optimum signal-to-noise ratio (Kværna, 1989). The two traces at the bottom of Fig. 7.7.5 show the signal filtered in the frequency band 2.5–4 Hz after beamforming and after undisturbed group filtration. The gain here is not as large as in broad band, but still exceeds a factor of 5 in power SNR. Fig. 7.7.7 (at the top) shows the same signals, but in another time scale. The chosen frequency band seems to be the best for filtering the signal after beamforming. For comparison we have plotted two traces at the bottom of Fig. 7.7.7 presenting the same signals filtered in the frequency band 3–5 Hz.

For the detection of signals in our experiments, 4 different variants of the optimal detector previously described were used. All of these detectors are sensitive not only to the increase in trace power due to signal arrival, but also to changes of the trace spectrum (Kushnir *et al*, 1983). The first is optimal in a statistical sense, the second is a modification of STA/LTA using prewhitening of the noise, and the last two are components of the first. Fig. 7.7.8 shows how the detectors work when applied to the beam and group filter outputs. The gain due to optimal group filtration is evident.

The final picture (Fig. 7.7.9) shows the results of the signal onset time estimation. Estimation is performed by an algorithm based on the maximum likelihood method applied to the problem of estimating the moment in time when parameters of the AR process are abruptly changed (Pisarenko *et al*, 1987). One can see that the likelihood function maximum coincides exactly with the beginning of the signal.

Conclusions

1. Application of an adaptive optimal group filtration technique to small aperture arrays can provide large gains in SNR in comparison with conventional beamforming due to high mutual correlation of array noise.
2. By using AR estimation of the power noise spectrum for group filter adaptation, we greatly reduce time and memory needed for the adaptation procedure while providing high quality of noise suppression.
3. Optimal group filtration does not distort the signal. Thus, when seen in connection with 1. and 2. above, it is clear that optimum filtration has great advantages as a preprocessor to be applied prior to subsequent broad band operations such as source classification.

In future work, it is recommended that the following studies be undertaken:

1. To perform a large-scale experiment concerning detection and classification capability of optimal group filtration applied to seismic signals recorded at NORESS and ARCESS. Comparison of false alarm rates of the optimum detector versus conventional beam detectors should form part of this investigation.
2. To develop algorithms for the compensation of signal frequency dependent waveform distortions due to propagation in real media under the array. This will give us an opportunity to equalize signal waveforms in different array receivers and improve signal extraction at high frequencies.
3. To implement these algorithms in an operational environment at the NORSAR data center.

4. To develop algorithms for designing optimal array geometry on the basis of the optimal group filtration features.

5. To use optimal group filtration for holographic investigations of the earth's interior using NORSAR data.

We plan to continue further work along these lines in future co-operative projects between NORSAR and the Institute of Physics of the Earth, Moscow.

A.F. Kushnir, Inst. of Physics of the Earth, Moscow
V.I. Pinsky, Inst. of Physics of the Earth, Moscow
J. Fyen, NORSAR

References

Capon, J. (1970): Application of signal detection and estimation theory to the large-array seismology, *Proc. IEEE*, 58.

Haykin, S. (ed.) (1979): *Nonlinear methods of spectral analysis.* Topics in Applied Physics, 34, Springer-Verlag, 234 pp.

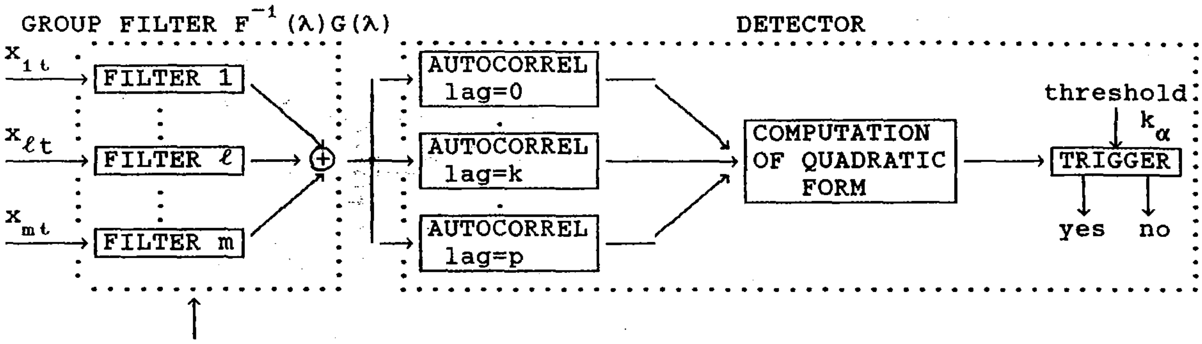
Kushnir, A.F., and V.M. Lapshin (1984): Optimal processing of the signals received by a group of spatially distributed sensors. *Computational Seismology*, 17, Allerton Press, Inc., 163–174.

Kushnir, A.F., V.F. Pisarenko and T.A. Rukavishnikova (1980): Noise compensation in multidimensional geophysical observations. I. Theory and method of data processing. *Computational Seismology*, 13, Allerton Press, Inc., 146–151.

Kushnir, A.F., I.V. Nikiforov and I.V. Savin (1983): Statistical adaptive algorithms for seismic signals automatic detection. *Computational Seismology*, 15, Allerton Press, Inc., 145–162.

Kværna, T. (1989): On exploitation of small-aperture NORESS type arrays for enhanced P-wave detectability. *Bull. Seism. Soc. Am.*, 79, 888–900.

Pisarenko, V.F., A.F. Kushnir and I.V. Savin (1987): Statistical adaptive algorithms for estimations of onset moments of seismic phases. *Phys. Earth Planet. Inter.*, 47, 4–10.



$$F^{-1}(\lambda)=\sum_{k=-q}^{q} L_k e^{-ik\lambda} = \left(\sum_{l=0}^{q} a_l e^{-il\lambda}\right) D^{-1} \left(\sum_{m=0}^{q} a_m e^{-im\lambda}\right)^*$$

$$G(\lambda)= \left(\exp(-i\tau_l\lambda),\ l=\overline{1,m}\right)$$

$$h_\theta(\lambda)=\sum_{k=0}^{p} \theta_k \cos(k\lambda)$$

Fig. 7.7.1 Optimal group detector flowchart.

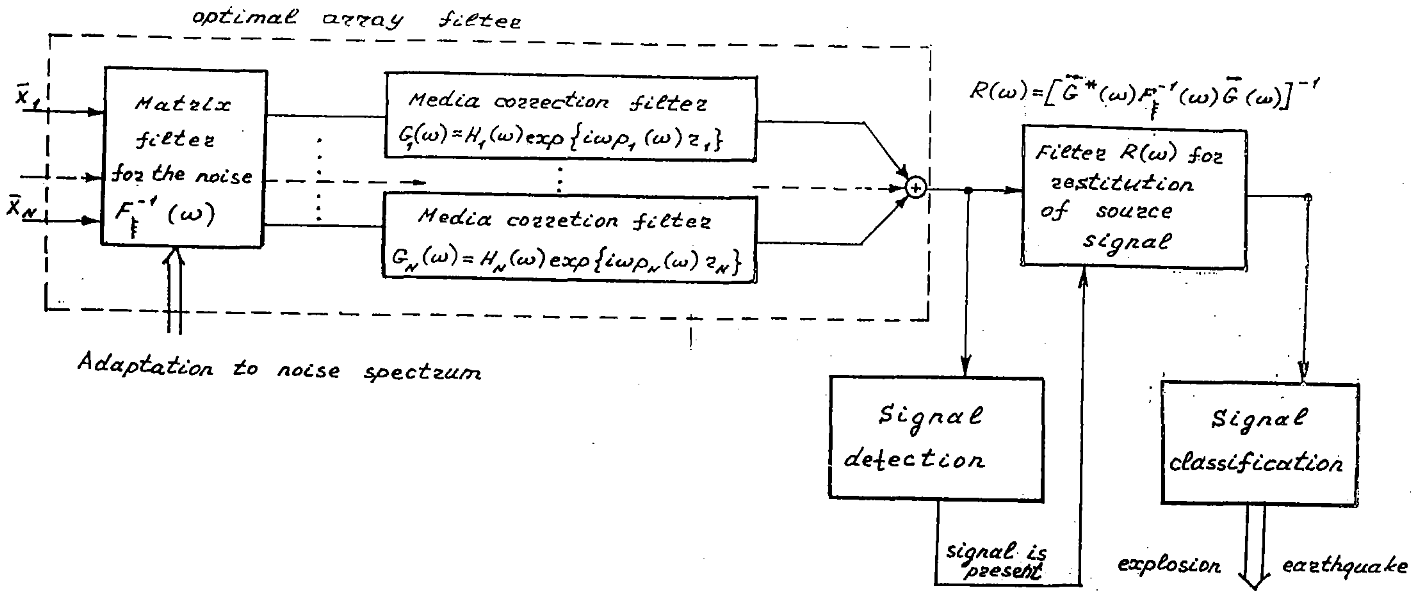
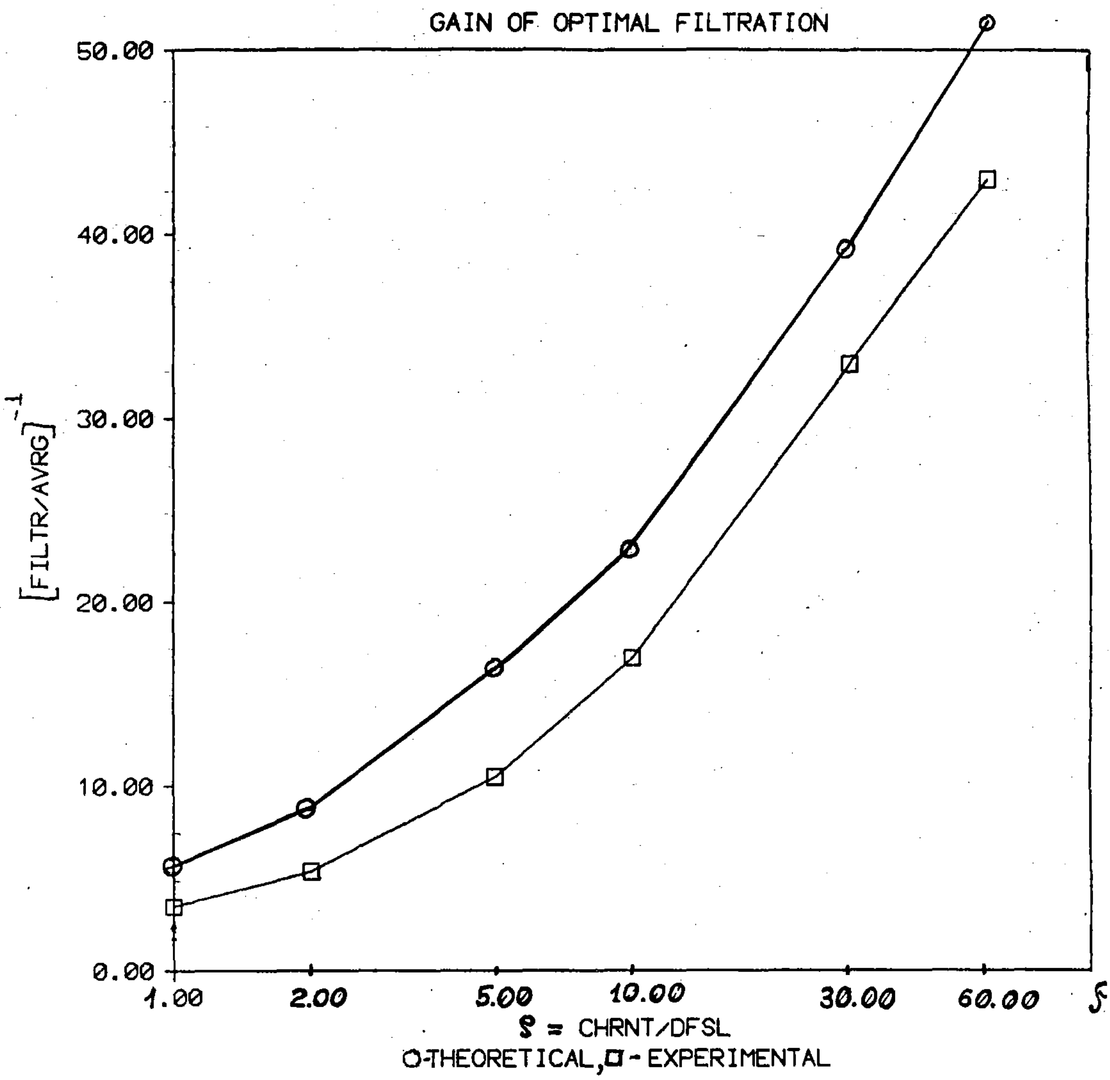

optimal array filter
$\bar{x}_1$
$\bar{x}_N$
Matrix filter for the noise $F_\xi^{-1}(\omega)$
Media correction filter $G_1(\omega)=H_1(\omega)\exp\{i\omega p_1(\omega)z_1\}$
Media correction filter $G_N(\omega)=H_N(\omega)\exp\{i\omega p_N(\omega)z_N\}$
Adaptation to noise spectrum
$R(\omega)=[\vec{G}^*(\omega)F_\xi^{-1}(\omega)\vec{G}(\omega)]^{-1}$
Filter $R(\omega)$ for restitution of source signal
Signal detection
signal is present
Signal classification
explosion
earthquake

Fig. 7.7.2



GAIN OF OPTIMAL FILTRATION
50.00
40.00
30.00
20.00
10.00
0.00
[FILTR/AVRG]^-1
1.00
2.00
5.00
10.00
30.00
60.00
ξ
ξ = CHRNT/DFSL
O-THEORETICAL, □ - EXPERIMENTAL

Fig. 7.7.3

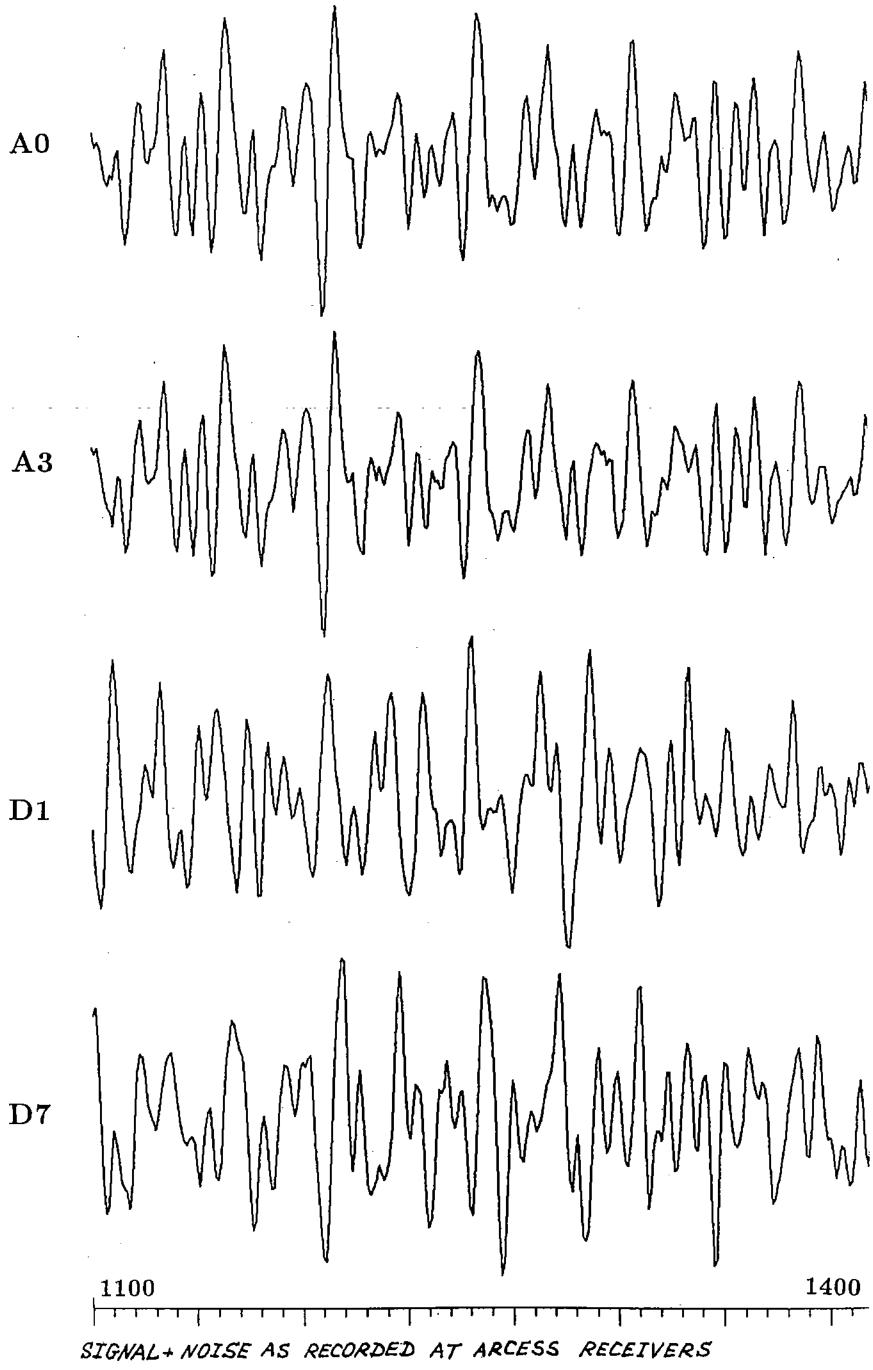

A0
A3
D1
D7
1100
1400
SIGNAL+NOISE AS RECORDED AT ARCESS RECEIVERS

Fig. 7.7.4

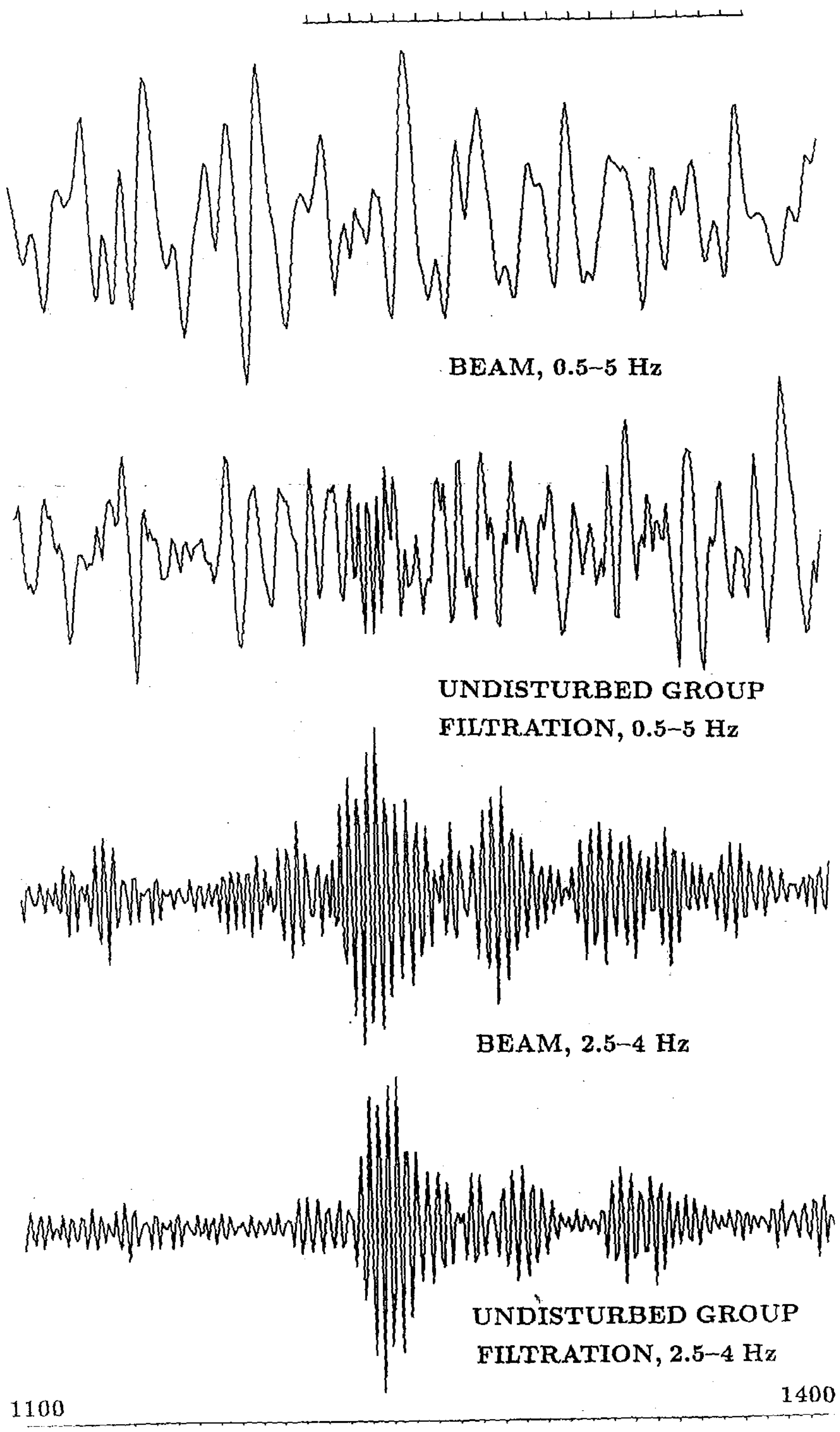

BEAM, 0.5–5 Hz
UNDISTURBED GROUP
FILTRATION, 0.5–5 Hz
BEAM, 2.5–4 Hz
UNDISTURBED GROUP
FILTRATION, 2.5–4 Hz
1100
1400

Fig. 7.7.5

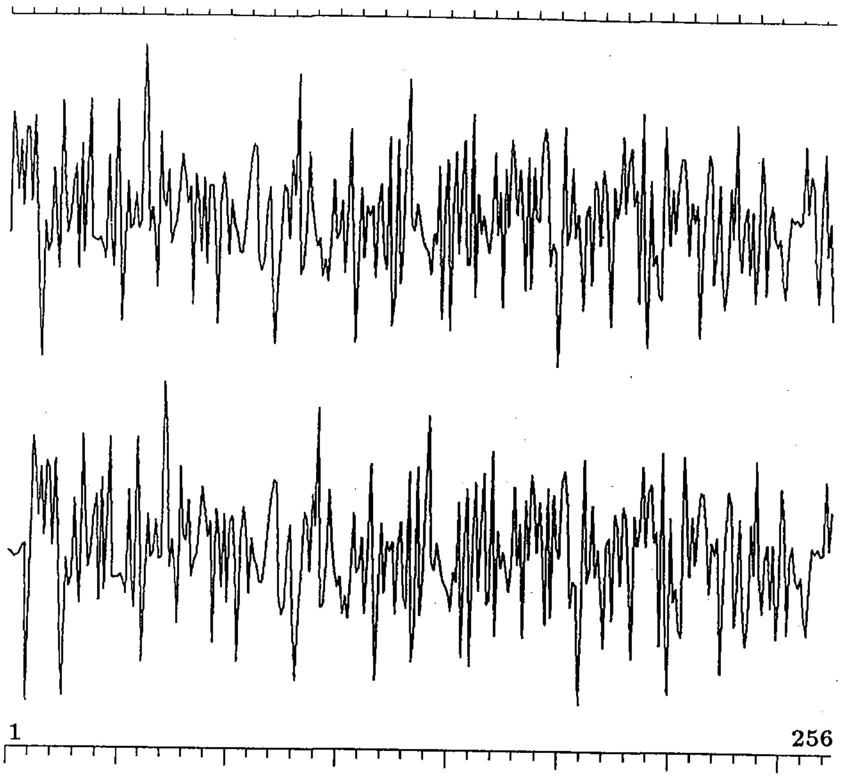


Fig. 7.7.6

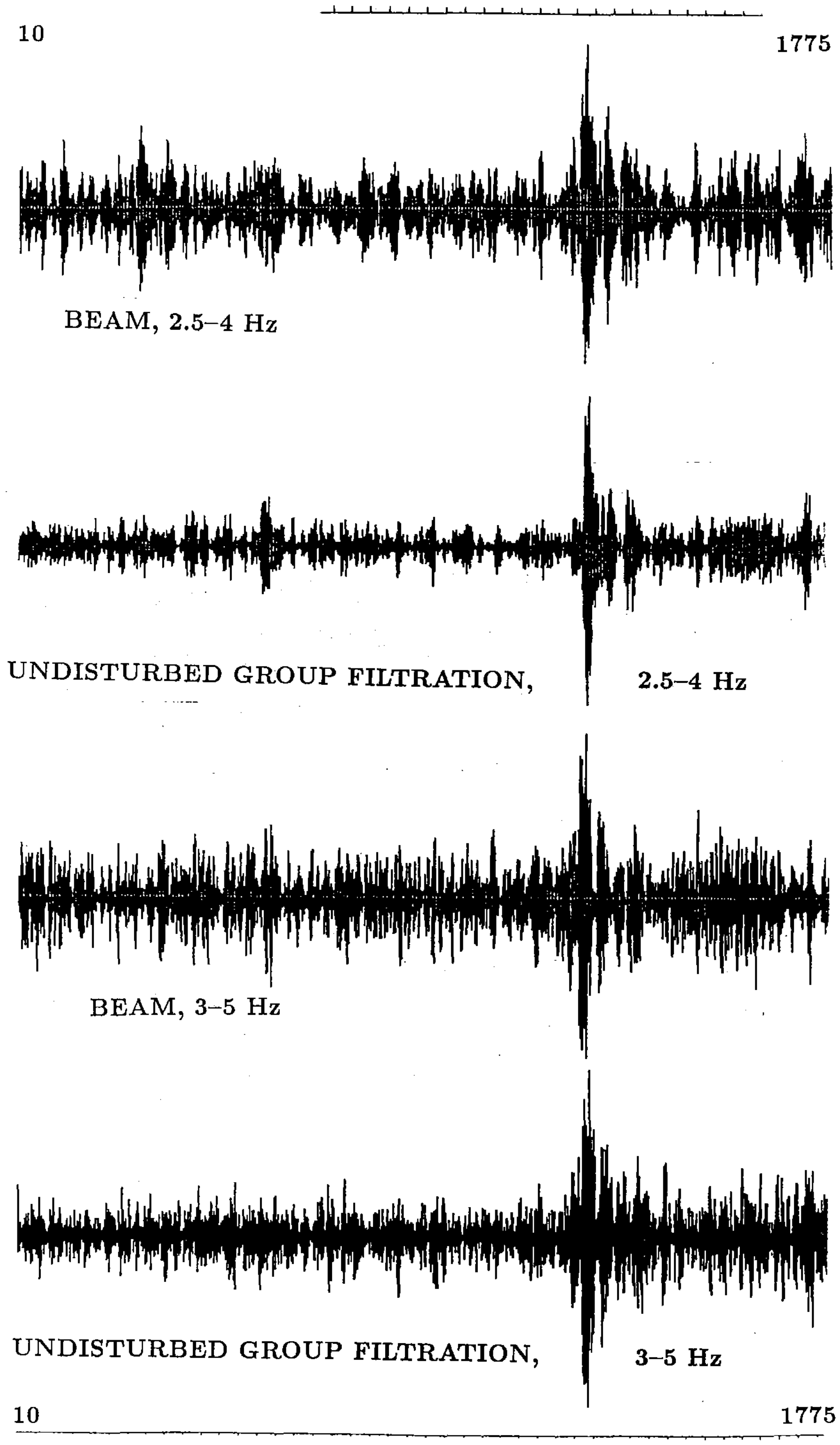

10
1775
BEAM, 2.5–4 Hz
UNDISTURBED GROUP FILTRATION,
2.5–4 Hz
BEAM, 3–5 Hz
UNDISTURBED GROUP FILTRATION,
3–5 Hz
10
1775

Fig. 7.7.7

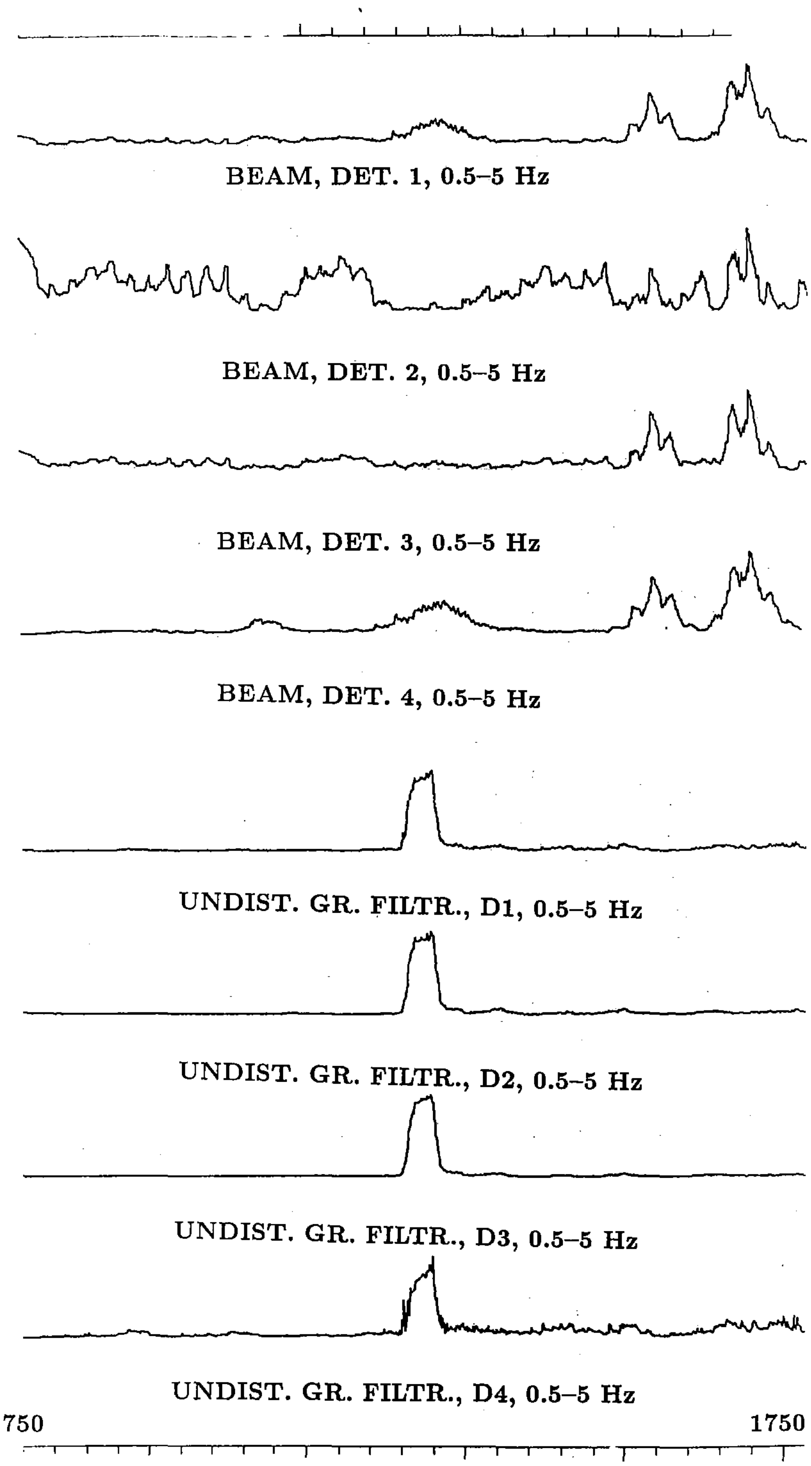

BEAM, DET. 1, 0.5–5 Hz
BEAM, DET. 2, 0.5–5 Hz
BEAM, DET. 3, 0.5–5 Hz
BEAM, DET. 4, 0.5–5 Hz
UNDIST. GR. FILTR., D1, 0.5–5 Hz
UNDIST. GR. FILTR., D2, 0.5–5 Hz
UNDIST. GR. FILTR., D3, 0.5–5 Hz
UNDIST. GR. FILTR., D4, 0.5–5 Hz
750
1750

Fig. 7.7.8

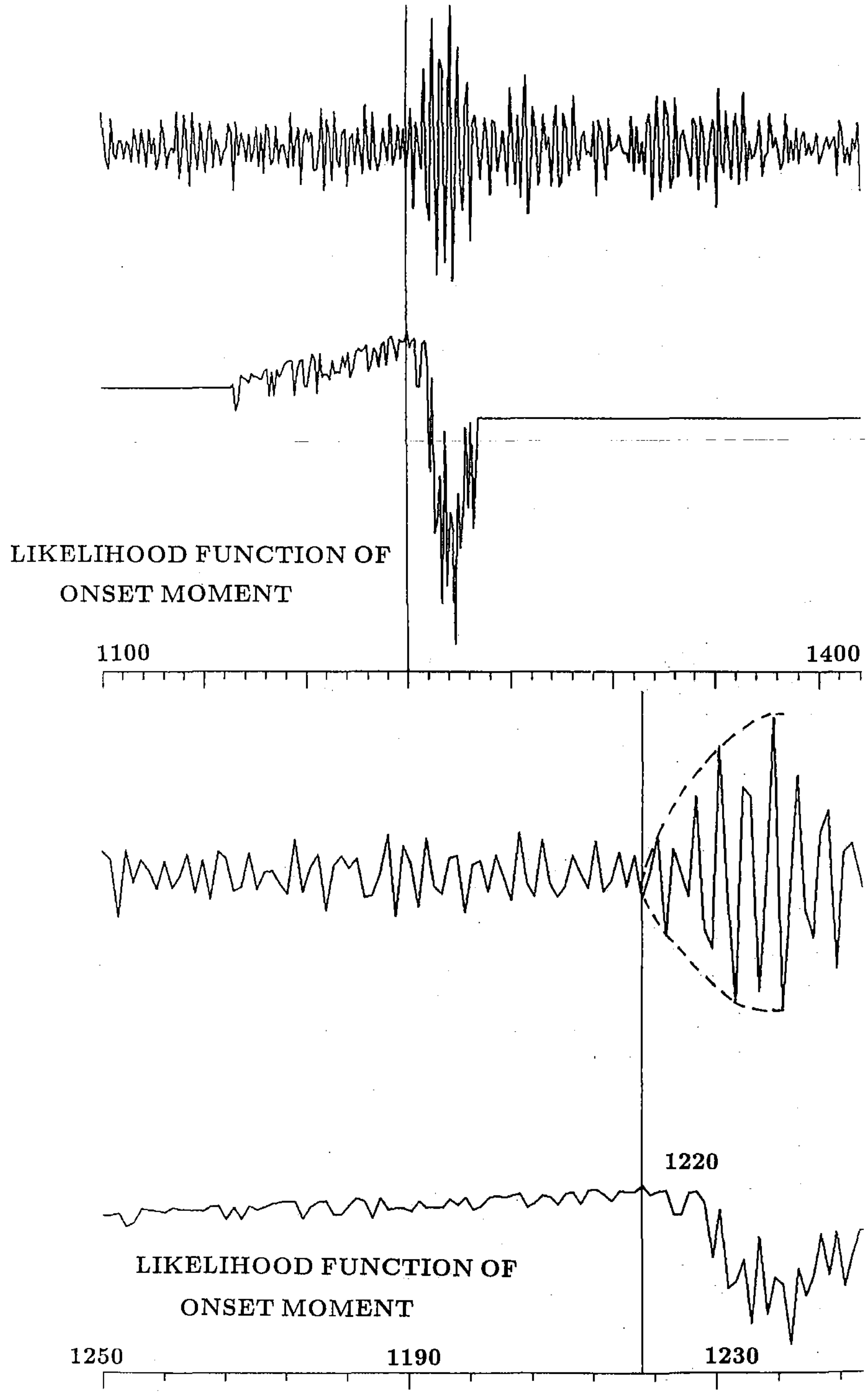

LIKELIHOOD FUNCTION OF
ONSET MOMENT
1100
1400
1220
LIKELIHOOD FUNCTION OF
ONSET MOMENT
1250
1190
1230

Fig. 7.7.9

APPENDIX
Statistically optimal event detection using array data

Let us assume that the noise $\xi_t = (\xi_{1t}, \ldots, \xi_{mt})^T$ at the m receivers of the array is a multidimensional Gaussian time series with zero mean and matrix power spectral density $F(\lambda), \lambda \in [0, 2\pi]$ and that the signal is represented as a Gaussian scalar process $\mu u_t(\Theta)$ at the seismic event source. The power spectral density of $u_t(\Theta)$ is denoted $g_\Theta(\lambda), \lambda \in [0, 2\pi]$, where Θ is an unknown vector parameter. The medium transfer functions are assumed to be linear. Then the observations $x_t = (x_{1t}, \ldots, x_{mt})^T$ become multidimensional Gaussian time series and have the form $G_t * \mu u_t + \xi_t$ and it is easy to write down the likelihood function $w(X_N \mid \mu, \Theta)$ for the moving window observations $X_N = (x_1^T, \ldots, x_N^T)^T$. Here $G_t = (G_{1t}, \ldots, G_{mt})^T$ is the vector impulse response function of the media along the paths from the seismic source to the receivers; $*$ is the sign of convolution. Then the hypothesis H_0 is that $\mu = 0$ and hypothesis H_1 is that $\mu \neq 0$. We consider the function $w(X_N \mid 0)$ to be known through adapative esitmation. If the unknown signal parameters have an *a priori* distribution $P(\Theta)$, then the best (Bayesian) test for testing these hypotheses (providing the least average signal miss probability for a given false alarm probability) has the form:

$$q(X_N) = \begin{cases} 1 \text{ (signal is present)} & \text{if } \rho(X_N) > k_\alpha \\ 0 \text{ (signal is absent)} & \text{if } \rho(X_N) < k_\alpha \end{cases} \tag{1}$$

where

$$\rho(X_N) = \int_\Theta \frac{w(X_N \mid \mu^2, \Theta)}{w(X_N \mid 0)} \, dP(\Theta) \tag{2}$$

Here, k_α is the detection threshold, determined on the basis of the given false alarm probability α.

It is practically impossible to devise an on-line algorithm on the basis of the statistic (1) when the distribution $P(\Theta)$ is arbitrary. But in the important particular case of weak signal detection, where the SNR, μ^2/σ^2, is sufficiently small and the moving window size N is sufficiently large, the statistic $\rho(X_N)$ can be simplified. As it is shown in Kushnir and Lapshin (1984), if the likelihood ratio $w(X_N \mid \gamma\sqrt{N}, \Theta)/w(X_N, 0)$ in (2) is replaced by its exponential approximation, i.e., the statistic $\rho(X_N)$ in (2) is replaced by the more computationally convenient statistic

$$r(X_N) = \int_\Theta exp\{\gamma \Delta(X_N, \Theta) - \frac{\gamma^2}{2} \Gamma_N(\Theta)\} dP(\Theta) \tag{3}$$

the asymptotic error probability limits for the test (3), $\mu^2 = \gamma/N, N \to \infty$, stay constant.

The functions $\Delta(X_N, \Theta)$ and $\Gamma_N(\Theta)$ in (3) have the following forms

$$\Delta(X_N, \Theta) = \frac{1}{2\sqrt{N}} \sum_{j=1}^{N} g_j(\Theta) \mid W_j^* x_j \mid^2 - g_j(\Theta) V_j, \quad \Gamma_N(\Theta) = \frac{1}{4N} \sum_{j=1}^{N} g_j(\Theta) V_j,$$

where $\Delta(X_N, \Theta)$ is an asymptotically sufficient statistic of the observations X_N, and

$$W_j^* = G^*(\lambda_j) F^{-1}(\lambda_j), \qquad V_j = W_j^* G(\lambda_j), \qquad g_j(\Theta) = g_\Theta(\lambda_j)$$

$\Gamma_N(\Theta)$ is the Fisher information quantity, divided by N,

$G(\lambda)$ is the Fourier transform of G_t,

$x_j = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} x_t \, exp(i\lambda_j t)$ is the discrete Fourier transform of the observations X_N,

$\lambda_j = \frac{2}{N} \pi j$,

$F(\lambda)$ is the matrix power spectral density of the noise ξ_t.

For calculation of the integral (3) in analytic form we will assume that:

$$g_\Theta(\lambda) = 1 + \sum_{k=1}^{P} c_k cosk\lambda, \quad \Theta = (c_1, \ldots, c_P)^T \tag{4}$$

where $c_k = 2E\{u_t u_{t+k}\}/E\{u_t^2\}$ is the autocorrelation of the source signal.

We further assume that the *a priori* distribution $P(\Theta)$ in (3) is Gaussian:

$$dP(\Theta) = \{(2\pi)^P \, detB\}^{-1/2} \; exp\{-\frac{1}{2}(\Theta - b)^T B^{-1} (\Theta - b)\} \tag{5}$$

Then we have

$$log\{r(x_N)\} = K + \gamma(\Delta_0 - \Delta^T A \delta) + \frac{1}{2}\Delta^T A \Delta + \frac{1}{\gamma}\Delta^T A B^{-1} b = K + r_\gamma(x_N) \tag{6}$$

where K is a constant independent of X_N, $\Delta = (\Delta_1, \ldots, \Delta_P)^T$,

$$\Delta_0 = \frac{1}{2\sqrt{N}} \sum_{j=1}^{N} \mid W_j^* \, x_j \mid^2 - V_j$$

$$\Delta_k = \frac{1}{2\sqrt{N}} \sum_{j=1}^{N} (\mid W_j^* V_j \mid^2 - V_j) cos(k\lambda_j)$$

$$\delta = (\delta_1, \ldots, \delta_P)^T$$

$$\delta_k = \frac{1}{4N}\sum_{j=1}^{N} V_j^2\, cos(k\lambda_j)$$

$$A = (\Gamma + \gamma^{-2} B^{-1}) = (B\Gamma + \gamma^{-2} I)^{-1} B$$

$$\Gamma = [\frac{1}{4N}\sum_{j=1}^{N} V_j^2\ cos(m\lambda_j)cos(n\lambda_j); \qquad m, n \in \overline{1, P}]$$

Thereby, the asymptotically optimal Bayes test for array detection of seismic signals with unknown spectrum is:

$$q_\gamma(x_N) = \begin{cases} 1 \text{ (signal is present)} & \text{if } r_\gamma(x_N) > k_\alpha \\ 0 \text{ (signal is absent)} & \text{if } r_\gamma(x_N) < k_\alpha \end{cases} \tag{7}$$

where the threshold k_α is determined on the basis of the given false alarm error probability, and the fitting parameter γ is determined so as to provide the highest asymptotical efficiency for the test (Kushnir *et al*, 1983).

Let us consider in particular two important (diametrically opposite) cases, where:

1. The power spectrum parameters of the source signal are almost known:

$$\| B \| << \| \Gamma^{-1} \|$$

2. The *a priori* information about these parameters is negligible:

$$\| B \| >> \| \Gamma^{-1} \|$$

In these cases, the statistic $r_\gamma(x_N)$ of the test (7) is simplified and looks like, respectively:

$$\begin{aligned} 1)\ r_1(x_N) &= \Delta_0 + b^T \Delta \\ 2)\ r_{2\gamma}(x_N) &= \gamma(\Delta_0 - \Delta^T \alpha) + \frac{1}{2}\, \Delta^T \Gamma^{-1} \Delta \end{aligned} \tag{8}$$

Calculation of the test statistic (8) can be realized in the time domain as shown in Fig. 7.7.1.