Technical Note

1965-57

J. Capon
R. J. Greenfield

Asymptotically Optimum Multidimensional Filtering for Sampled-Data Processing of Seismic Arrays

17 December 1965

ASYMPTOTICALLY OPTIMUM MULTIDIMENSIONAL FILTERING
FOR SAMPLED-DATA PROCESSING OF SEISMIC ARRAYS

J. CAPON

R. J. GREENFIELD

Group 64

TECHNICAL NOTE 1965-57

17 DECEMBER 1965

ASYMPTOTICALLY OPTIMUM MULTIDIMENSIONAL FILTERING FOR SAMPLED-DATA PROCESSING OF SEISMIC ARRAYS

J. Capon
R. J. Greenfield

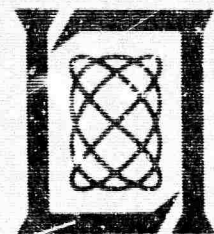Lincoln Laboratory, Massachusetts Institute of Technology
Lexington, Massachusetts

ABSTRACT

A number of asymptotically optimum multidimensional filtering methods are investigated with the purpose of determining filtering techniques which require relatively little computing time to implement with a digital computer. In particular, the asymptotic properties of the maximum-likelihood and minimum-variance unbiased multidimensional filters are investigated in the sampled-data case. These two multidimensional filters are shown to be identical since they are both based on a conditional expectation. In addition, the martingale property of conditional expectation assures that the asymptotic properties of these multidimensional filters are well defined.

An asymptotically optimum frequency domain synthesis procedure is given for two-sided multidimensional filters. This procedure is well suited to machine computation and has the advantage with respect to the exact recursive synthesis method of requiring much less computation time. A synthesis procedure for physically realizable multidimensional filters is presented which is based on a factorization of rational spectral matrices. This method is not, however, well suited to machine computation. An interpretation of optimum multidimensional filtering in terms of frequency wave-number space is also given.

BLANK PAGE

TABLE OF CONTENTS

I.   INTRODUCTION

It has been recognized that a considerable reduction of seismic noise is possible by employing multidimensional filtering for seismic arrays.[1] Some of the more important approaches to seismic array processing are the maximum-likelihood method, the minimum-variance unbiased estimator technique and multichannel Wiener filtering. In general, these multidimensional filtering methods form a single output waveform which serves as an estimator of the unknown signal which comes from a fixed direction.

The basic assumption in the analysis of multidimensional filters is that the output, $X_k(t)$, of the kth seismometer may be written as

$$X_k(t) = S(t) + N_k(t) \tag{1}$$

where $S(t)$ is the signal waveform which is assumed to be the same in each seismometer and $N_k(t)$ is the noise present in seismometer k, $k = 1, \ldots, K$. In writing Eq. (1) it is assumed that the azimuth and horizontal velocity of the event, or signal, have already been determined with sufficient accuracy to allow the signal waveforms from each seismometer to be shifted to bring them into time coincidence. In most applications, the outputs of the seismometers are given in sampled form in which case Eq. (1) becomes

$$X_{kj} = S_j + N_{kj} \qquad \begin{array}{l} k = 1, \ldots, K \\ j = 0, \pm 1, \pm 2, \ldots \end{array} \tag{2}$$

Only the sampled-data multidimensional filtering problem for seismic arrays will be considered.

It can be shown[2] that the maximum-likelihood estimator is the same as the minimum variance unbiased estimate of the signal when the noise has a multidimensional Gaussian distribution. In addition, the multichannel Wiener filter is related very simply to the minimum variance unbiased estimator. The synthesis of these optimum multidimensional filters is achieved by means of a recursive matrix inversion algorithm which is well-suited to computer application. The purpose of the present work is to point out that, asymptotically, as the memory of the filter gets large, there are synthesis procedures which can be used that in certain cases require far fewer computations than the recursive algorithm. If a large array of say 525 sensors is to be processed, then these approximate synthesis techniques offer a considerable saving in computer time required to process an event relative to the exact recursive procedure. This saving will be shown to apply only when the filter is two-sided, i.e., non-physically realizable. This restriction poses no problem in the application of the results since all waveforms are recorded on magnetic tape and non-physically realizable filters are readily implemented.

If the filters are specified to be physically realizable, i.e., one-sided with no output before any input is applied, then the synthesis problem becomes more complicated than in the two-sided case discussed previously. It is shown that a spectral matrix factorization procedure is required to synthesize the filter in the physically realizable case and a method for achieving this for rational spectral matrices is discussed. Another method, due to Wiener and Masani, is also described which is

2

valid for general spectral matrices. Unfortunately, these spectral matrix factorization techniques are not well adapted to machine computation.

An interpretation of optimum multidimensional filtering in terms of frequency wave-number space is given. The structure of the filter in frequency wave-number space is also presented.

II. DERIVATION OF THE MAXIMUM-LIKELIHOOD AND MINIMUM-VARIANCE UNBIASED ESTIMATORS

The derivation of the maximum-likelihood estimator of the signal requires the assumption that the noise components have a multidimensional Gaussian distribution. We assume, for simplicity, that the noise components have zero mean, and that the covariance matrix is

$$\rho_{kk_1}(j,j_1) = \rho_{k_1 k}(j_1,j) = E\{N_{kj}N_{k_1 j_1}\} \ , \qquad \begin{array}{l} 1 \le k, k_1 \le K \\ -\nu \le j, j_1 \le \nu \ , \end{array} \tag{3}$$

where E denotes expectation, and it is assumed that the estimator is to use $2\nu + 1$ samples extending in time from $-\nu$ to ν. Thus, the likelihood function can be written

$$L = (2\pi)^{-\frac{k}{2}(2\nu+1)} |\rho|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \sum_{k,k_1=1}^{K} \sum_{j,j_1=-\nu}^{\nu} \right.$$

$$\left. \rho_{kk_1}^{-1}(j,j_1)(X_{kj}-S_j)(X_{k_1 j_1}-S_{j_1}) \right\} \ , \tag{4}$$

3

where $|\rho|$ denotes the determinant of the matrix ρ which is a matrix of $K \times K$ submatrices, the jj_1^{th} submatrix has the elements $\rho_{kk_1}(j,j_1)$, $k, k_1 = 1,\ldots,K$, $j,j_1 = -\nu,\ldots,\nu$, with a corresponding notation for the inverse matrix ρ^{-1} whose elements are $\rho_{kk_1}^{-1}(j,j_1)$. We assume throughout that the matrix ρ is positive definite. Differentiating the logarithm of the likelihood function with respect to S_{j_1} and equating the result to zero, we obtain

$$\sum_{k,k_1=1}^{K} \sum_{j=-\nu}^{\nu} \rho_{kk_1}^{-1}(j,j_1)\, (\hat{S}_{\nu j} - X_{kj}) = 0 \; , \qquad\qquad j_1 = -\nu,\ldots,\nu \; , \qquad (5)$$

where $\hat{S}_{\nu j}$ is the maximum-likelihood estimator for S_j.

We can rewrite Eq. (5) as follows

$$\sum_{j=-\nu}^{\nu} \hat{S}_{\nu j} \sum_{k,k_1=1}^{K} \rho_{kk_1}^{-1}(j,j_1) \;\; = \;\; \sum_{k,k_1=1}^{K} \sum_{j=-\nu}^{\nu} X_{kj}\, \rho_{kk_1}(j,j_1) \; , \qquad (6)$$

$$j_1 = -\nu,\ldots,\nu \; .$$

Let us define the $(2\nu+1) \times (2\nu+1)$ matrix $G(j,j_1)$ in terms of its inverse $G^{-1}(j,j_1)$ whose elements are defined as

$$\alpha^{-1}(j,j_1) = \sum_{k,k_1=1}^{K} \rho_{kk_1}^{-1}(j,j_1) \; , \qquad\qquad j,j_1 = -\nu,\ldots,\nu \; . \qquad (7)$$

We can solve for $\hat{S}_{\nu j'}$ as

$$\hat{S}_{\nu j'} = \sum_{j=-\nu}^{\nu} \sum_{k=1}^{K} \theta_k(j|j')\, X_{kj} \; , \qquad (8)$$

where

4

$$\theta_k(j|j') = \sum_{k_1=1}^{K} \sum_{j_1=-\nu}^{\nu} \rho_{kk_1}^{-1}(j,j_1)\, \alpha(j_1,j') \qquad\qquad \begin{array}{l} k = 1,\ldots,K \\ j = -\nu,\ldots,\nu\,. \end{array} \qquad (9)$$

It should be noted that a different set of weights is obtained from Eq. (9) for different values of ν and also for different values of j', which can vary from $-\nu$ to ν. The fact that $\theta_k(j|j')$ depends on ν has not been brought out by the notation. It is easily seen from the quadratic nature of the logarithm of the likelihood function that the solution given in Eq. (9) is unique.

We now consider the minimum-variance unbiased estimator of $S_{j'}$, denoted by $\hat{S}'_{\nu j'}$, expressible as

$$\hat{S}'_{\nu j'} = \sum_{j=-\nu}^{\nu} \sum_{k=1}^{K} \theta'_k(j|j')\, X_{kj} \qquad (10)$$

with the constraints

$$\sum_{k=1}^{K} \theta'_k(j|j') = \delta_{jj'}\,, \qquad\qquad j = -\nu,\ldots,\nu\,, \qquad (11)$$

where

$$\begin{aligned} \delta_{jj'} &= 1\,, \quad j = j' \\ &= 0\,, \quad \text{otherwise}\,. \end{aligned} \qquad (12)$$

The variance of $\hat{S}'_{\nu j'}$ is

$$E(\hat{S}'_{\nu j'})^2 = \sum_{j,j_1=-\nu}^{\nu} \sum_{k,k_1=1}^{K} \theta'_{k_1}(j_1|j')\, \theta_k(j|j')\, \rho_{kk_1}(j,j_1)\,. \qquad (13)$$

Using the calculus of variations, we obtain that the minimum-variance unbiased estimator has weights which satisfy the system of equations

5

$$\sum_{k_1=1}^{K} \sum_{j_1=-\nu}^{\nu} \theta'_{k_1}(j_1|j') \, \rho_{kk_1}(j,j_1) + \lambda_{\nu j} = 0, \qquad \begin{array}{l} k = 1, \ldots, K \\ j = -\nu, \ldots, \nu, \end{array} \tag{14}$$

where the $\lambda_{\nu j}$ are $2\nu+1$ Lagrangian multipliers chosen to satisfy the constraints given in Eq. (11).

If we define

$$\theta'_{K+1}(j|j') = \lambda_{\nu j} \tag{15}$$

$$\rho_{k,K+1}(j,j_1) = \rho_{K+1,k}(j,j_1) = \delta_{jj_1}(1 - \delta_{k,K+1}), \tag{16}$$

$$\begin{array}{l} k = 1, \ldots, K+1 \\ j,j_1 = -\nu, \ldots, \nu, \end{array}$$

then Eqs. (11) and (14) may be written as a single set of equations as follows

$$\sum_{k_1=1}^{K+1} \sum_{j_1=-\nu}^{\nu} \theta'_{k_1}(j_1|j') \, \rho_{kk_1}(j,j_1) = \delta_{k,K+1} \, \delta_{jj'} \tag{17}$$

$$\begin{array}{l} k = 1, \ldots, K+1 \\ j = -\nu, \ldots, \nu. \end{array}$$

This set of equations is equivalent to the set given previously, cf. Reference 3, pp. 21-24. The reason for writing the equations in the above form is that we now have a Toeplitz matrix of submatrices and a recursive procedure[4] can be used in the solution of filter weighting coefficients.

We can write the system of Eqs. (14) as

6

$$\theta'_k(j|j') = -\sum_{k_1=1}^{K} \sum_{j_1=-\nu}^{\nu} \rho_{kk_1}^{-1}(j,j_1)\lambda_{\nu j_1}, \qquad \begin{array}{l} k = 1,\ldots,K \\ j = -\nu,\ldots,\nu. \end{array} \tag{18}$$

Using Eq. (11) in (18) we get

$$-\sum_{k} \sum_{k_1=1}^{K} \sum_{j_1=-\nu}^{\nu} \rho_{kk_1}^{-1}(j,j_1)\lambda_{\nu j} = \delta_{jj''}, \qquad j = -\nu,\ldots,\nu, \tag{19}$$

and using Eq. (7)

$$\lambda_{\nu j} = -\sum_{j_1=-\nu}^{\nu} \alpha(j,j_1)\,\delta_{j,j'}$$

$$= -\alpha(j,j'), \qquad\qquad j = -\nu,\ldots,\nu. \tag{20}$$

Therefore, we see from Eqs. (18) and (20) that

$$\theta'_k(j|j') = \sum_{k_1=1}^{K} \sum_{j_1=-\nu}^{\nu} \rho_{kk_1}^{-1}(j,j_1)\,\alpha(j_1,j'), \qquad \begin{array}{l} k = 1,\ldots,K \\ j = -\nu,\ldots,\nu, \end{array} \tag{21}$$

which is the same as Eq. (9). Hence the maximum-likelihood and minimum-variance unbiased estimators are identical. It is easily seen from the quadratic nature of the expression in Eq. (13) that the solution given in Eq. (21) is unique.

Using Eqs. (11), (13), and (14) we obtain

$$E\{\hat{S}^2_{\nu j'}\} = \sum_{k=1}^{K} \sum_{j=-\nu}^{\nu} \theta_k(j|j')\,(-\lambda_{\nu j})$$

$$= -\sum_{j=-\nu}^{\nu} \lambda_{\nu j}\,\delta_{jj'} = -\lambda_{\nu j'} = \alpha(j',j'), \qquad j' = -\nu,\ldots,\nu. \tag{22}$$

7

In addition, we have

$$E\{\hat{S}_{vj'}, \hat{S}_{vj''}\} = \sum_{k=1}^{K} \sum_{j=-v}^{v} \theta_k(j|j') [\alpha(j,j''),$$

$$= \alpha(j',j''), \qquad\qquad j',j'' = -v, \ldots, v . \qquad (23)$$

It should be noted that

$$E\{\hat{S}^2_{vj}\} \leq E\{\hat{S}^2_{v'j}\} , \qquad\qquad v' \leq v . \qquad (24)$$

This follows from the fact that the minimum value of a quadratic form, subject to certain constraints, in a v'-dimensional space cannot be increased if the dimension of the space is increased to $v \geq v'$.

III. THEORETICAL JUSTIFICATION FOR THE ASYMPTOTIC APPROACH

We now wish to show that the use of filters based on only a part of either the past, future, or past and future, can be used to approximate the performance of filters based on, respectively, the full past, the full future, or the full past and future. In order to do this it will be convenient to introduce new random variables, as indicated by Rozanov,[5]

$$Y_{k*j} = X_{k*j}$$

$$Y_{kj} = X_{k*j} - X_{kj} \qquad\qquad k \neq k^*, \; k = 1, \ldots, K \qquad (25)$$
$$j = 0, \pm 1, \pm 2, \ldots$$

8

Let us denote the set of random variables Y_{kj}, $k \neq k^*$, $j = -\nu, \ldots, \nu$, by Y_ν. It should be noted that the random variables Y_{kj}, $k \neq k^*$, do not depend on S_j, i.e.,

$$Y_{kj} = N_{k^*j} - N_{kj}, \qquad\qquad k \neq k^*, \qquad\qquad (26)$$

and that

$$Y_{k^*j} = S_j + N_{k^*j} . \qquad\qquad (27)$$

It therefore follows that the minimum-variance unbiased estimator of S_j, may be written as

$$\hat{S}_{\nu j'} = X_{k^*j'} - \hat{N}_{\nu k^*j'} , \qquad\qquad (28)$$

where $\hat{N}_{\nu k^*j'}$ is a linear combination of the Y_{kj}'s, $k \neq k^*$, $j = -\nu, \ldots, \nu$, i.e., $\hat{N}_{\nu k^*j'}$ depends only on Y_ν. When written in this form, it follows easily that $\hat{N}_{\nu k^*j'}$ is given by

$$\hat{N}_{\nu k^*j'} = E\{N_{\nu k^*j'} | Y_\nu\}, \qquad\qquad (29)$$

i.e., $\hat{N}_{\nu k^*j'}$ is the conditional expectation of $N_{k^*j'}$ with respect to Y_ν, cf. reference 6, pp. 150-155. However, $\hat{N}_{\nu k^*j'}$ is a martingale since, cf. reference 6, pp. 91-94,

$$E\{\hat{N}_{\nu+1, k^*j'} | Y_\nu\} = E\{E[N_{k^*j'} | Y_{\nu+1}] | Y_\nu\}$$

$$= E\{N_{k^*j'} | Y_\nu\} = \hat{N}_{\nu k^*j'} . \qquad\qquad (30)$$

9

Thus, it follows from a martingale convergence theorem, cf. reference 6, p. 167, Theorem 7.4, and also pp. 560-562,

$$\hat{N}_{k*j'} = \underset{\nu \to \infty}{\ell. i. m.} \hat{N}_{\nu k*j'} , \tag{31}$$

for fixed j'. If we wish to consider sequences of physically realizable filters for which $j' = \nu$, we obtain in a similar manner

$$\hat{N}_{k*\infty} = \underset{\nu \to \infty}{\ell. i. m.} \hat{N}_{\nu k*\nu} , \tag{32}$$

and for filters using future values, $j' = -\nu$, and

$$\hat{N}_{k*, -\infty} = \underset{\nu \to \infty}{\ell. i. m.} \hat{N}_{\nu k*, -\nu} . \tag{33}$$

The minimum-variance unbiased estimators for the above three cases are, respectively,

$$\hat{S}_{j'} = \underset{\nu \to \infty}{\ell. i. m.} S_{\nu j'} = X_{k*j'} - \hat{N}_{k*j'} , \tag{34}$$

$$\hat{S}_{\infty} = \underset{\nu \to \infty}{\ell. i. m.} S_{\nu\nu} = X_{k*\infty} - \hat{N}_{k*\infty} , \tag{35}$$

$$\hat{S}_{-\infty} = \underset{\nu \to \infty}{\ell. i. m.} S_{\nu, -\nu} = X_{k*, -\infty} - \hat{N}_{k*, -\infty} . \tag{36}$$

Thus, we have shown that filters based on a large part of either the past, future, or past and future, can approximate the performance of filters based on, respectively, the full

10

past, the full future, or the full past and future. It is easily seen that all results remain valid when $k*$ is any integer between 1 and K.

We have already shown that the maximum-likelihood estimator for $S_{j'}$ is identical with the minimum-variance unbiased estimator for $S_{j'}$, cf. Eqs. (9) and (21). However, at this point it is extremely simple to show that this is true. The joint probability density of $X_{k*j'}$, Y_ν is

$$p_{S_{j'}}(X_{k*j'}, Y_\nu) = p_{S_{j'}}(X_{k*j'}|Y_\nu)\, p(Y_\nu)\,, \tag{37}$$

since Y_ν is independent of $S_{j'}$. In addition, we have for the Gaussian multivariate case

$$p_{S_{j'}}(X_{k*j'}|Y_\nu) = (2\pi)^{-1/2}\left[\sigma^2(X_{k*j'}|Y_\nu)\right]^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left[\frac{X_{k*j'} - E(X_{k*j'}|Y_\nu)}{\sigma(X_{k*j'}|Y_\nu)}\right]^2\right\} \tag{38}$$

where $\sigma(X_{k*j'}|Y_\nu)$ is independent of $S_{j'}$ and

$$E\{X_{k*j'}|Y_\nu\} = E\{N_{\nu k*j'}|Y_\nu\} + S_{\nu j'}\,. \tag{39}$$

The maximum-likelihood estimator for $S_{j'}$ is obtained by differentiating, with respect to $S_{j'}$, the probability density function in Eq. (37), or equivalently that in Eq. (38), from which we get, when using Eq. (39),

$$\hat{S}_{\nu j'} = X_{k*j'} - E\{N_{\nu k*j'}|Y_\nu\}\,, \tag{40}$$

11

which agrees with Eqs. (28) and (29). Thus, all asymptotic results derived for the minimum-variance unbiased estimator remain true, of course, for the maximum-likelihood estimator.

IV. APPROXIMATE FREQUENCY-DOMAIN SYNTHESIS PROCEDURE FOR TWO-SIDED FILTERS

An approximate frequency-domain synthesis procedure will now be presented for maximum-likelihood filters which use a large part of the past and future. Such filters which use the past as well as future values will be termed two-sided filters. We begin by assuming the noise is wide-sense stationary and by letting $\nu = \infty$, $j' = 0$, so that Eqs. (13) and (14) may be written as

$$E\{\hat{S}_0^2\} = \lim_{\nu \to \infty} E\{\hat{S}_{\nu 0}^2\}$$

$$= \int_{-\pi}^{\pi} \left[\sum_{j,k=1}^{K} f_{jk}(x) A_j^*(x) A_k(x) \right] \frac{dx}{2\pi} \ , \tag{41}$$

and

$$\int_{-\pi}^{\pi} \sum_{k=1}^{K} f_{jk}(x) A_k(x) \epsilon^{-i\ell x} \frac{dx}{2\pi} + \lambda_\ell = 0 \ , \tag{42}$$

$$\ell = 0, \pm 1, \pm 2, \ldots ,$$
$$j = 1, \ldots, K,$$

where

$$A_k(x) = \sum_{j=-\infty}^{\infty} \theta_k(j|0) \epsilon^{ijx} \ , \qquad k = 1, \ldots, K, \tag{43}$$

12

$$\rho_{jk}(\ell) = \int_{-\pi}^{\pi} f_{jk}(x) \, \epsilon^{-i\ell x} \, \frac{dx}{2\pi} \, , \qquad\qquad \begin{array}{l} j, k = 1, \ldots, K \\ \ell = 0, \pm 1, \pm 2, \ldots \end{array}$$

and

$$f_{jk}(x) = \sum_{\ell = -\infty}^{\infty} \rho_{jk}(\ell) \, \epsilon^{i\ell x} \, , \qquad\qquad j, k = 1, \ldots, K \, , \qquad (44)$$

is the sampled cross power spectral density function, $x = \omega T$, T is the sampling

interval. We have from Eq. (43) that

$$\theta_k(j|0) = \int_{-\pi}^{\pi} A_k(x) \, \epsilon^{-ijx} \, \frac{dx}{2\pi} \, , \qquad\qquad \begin{array}{l} k = 1, \ldots, K \\ j = 0, \pm 1, \pm 2, \ldots \end{array} \qquad (45)$$

The constraint Eq. (11) becomes

$$\sum_{k=1}^{K} A_k(x) = 1 \, . \qquad (46)$$

If we let

$$\Lambda(x) = \sum_{k=-\infty}^{\infty} \lambda_k \, \epsilon^{ikx} \, , \qquad (47)$$

then Eq. (42) may be written as

$$\int_{-\pi}^{\pi} \epsilon^{-i\ell x} \left[\sum_{k=1}^{K} f_{jk}(x) \, A_k(x) + \Lambda(x) \right] dx = 0 \, , \qquad \begin{array}{l} \ell = 0, \pm 1, \pm 2, \ldots \\ j = 1, \ldots, K \, . \end{array} \qquad (48)$$

According to Eq. (48), the Fourier coefficients of the quantity in the brackets

in E   ) must all be zero. It follows that the quantity itself must be zero so that

13

$$\sum_{k=1}^{K} f_{jk}(x) A_k(x) + \Lambda(x) = 0 , \qquad \begin{array}{l} -\pi \leq x \leq \pi \\ j = 1,\ldots,K . \end{array} \qquad (49)$$

Thus

$$A_k(x) = \frac{\displaystyle\sum_{j=1}^{K} q_{kj}(x)}{\displaystyle\sum_{j,k=1}^{K} q_{kj}(x)} , \qquad k = 1,\ldots,K , \qquad (50)$$

$$- \Lambda(x) = \left[\sum_{j,k=1}^{K} q_{kj}(x) \right]^{-1} , \qquad (51)$$

where $\{q_{jk}(x)\}$ is the inverse of the spectral matrix $\{f_{jk}(x)\}$, $j,k = 1,\ldots,K$. We note that

$$q_{jk}(x) = q_{kj}^*(x) , \qquad (52)$$

since

$$f_{jk}(x) = f_{kj}^*(x) , \qquad (53)$$

and

$$q_{jk}(x) = q_{jk}^*(-x) , \qquad (54)$$

since

$$f_{jk}(x) = f_{jk}^*(-x) . \qquad (55)$$

Therefore

14

$$A_k(x) = A_k^*(-x), \qquad\qquad k = 1, \ldots, K. \qquad (56)$$

We obtain from Eqs. (41) and (50)

$$E\{\hat{S}_o^2\} = \int_{-\pi}^{\pi} \frac{dx}{2\pi} \left[\sum_{j,k=1}^{K} q_{kj}(x) \right]^{-1}. \qquad (57)$$

The filter weighting coefficients are obtained from Eq. (45) as

$$\theta_k(j|0) = \int_{-\pi}^{\pi} \frac{dx}{2\pi} \left[\mathrm{Re}\, A_k(x) \cos jx + \mathrm{Im}\, A_k(x) \sin jx \right], \quad \begin{matrix} k = 1, \ldots, K \\ j = 0, \pm1, \pm2, \ldots \end{matrix} \qquad (58)$$

It is easily seen that the constraint conditions are satisfied since

$$\sum_{k=1}^{K} \theta_k(j|0) = \int_{-\pi}^{\pi} \cos jx \, \frac{dx}{2\pi}$$

$$= \delta_{jo}. \qquad (59)$$

The filter weights given by Eq. (58) would have to be obtained by inverting the spectral matrix at all points on the frequency axis between $-\pi$ and π. This is clearly impractical so that an approximation would have to be used in practice. Such an approximation is most easily obtained by approximating the integral in Eq. (58) by a finite sum

15

$$\theta_k(j|0) = \frac{1}{2\nu} \sum_{\ell=-\nu+1}^{\nu} [\text{Re } A_k(\ell \frac{\pi}{\nu}) \cos j\ell \frac{\pi}{\nu} + \text{Im } A_k(\ell \frac{\pi}{\nu}) \sin j\ell \frac{\pi}{\nu}] \, ,$$

$$k = 1, \ldots, K$$
$$j = -\nu, \ldots, \nu \, . \qquad (60)$$

The symmetry properties of $A_k(x)$ expressed in Eq. (56) enable us to simplify the above equation

$$\theta_k(j|0) = \frac{1}{2\nu} [A_k(0) + (-1)^j A_k(\pi) + 2 \sum_{\ell=1}^{\nu-1} \text{Re } A_k(\ell \frac{\pi}{\nu}) \cos j\ell \frac{\pi}{\nu} + \text{Im } A_k(\ell \frac{\pi}{\nu}) \sin j\ell \frac{\pi}{\nu}]$$

$$k = 1, \ldots, K$$
$$j = -\nu, \ldots, \nu \, . \qquad (61)$$

The constraint equations are still satisfied since

$$\sum_{k=1}^{K} \theta_k(j|0) = \frac{1}{2\nu} \left[(-1)^j + 2(\frac{1}{2} + \sum_{\ell=1}^{\nu-1} \cos j\ell \frac{\pi}{\nu}) \right]$$

$$= \frac{1}{2\nu} \left[(-1)^j + \frac{\sin (\nu - \frac{1}{2}) j \frac{\pi}{\nu}}{\sin \frac{1}{2} j \frac{\pi}{\nu}} \right]$$

$$= \frac{1}{2\nu} [+1 + 2\nu - 1] = 1 \, , \qquad\qquad j = 0$$

$$= \frac{1}{2\nu} \left[(-1)^j - \frac{\sin (\frac{1}{2} j \frac{\pi}{\nu} - j\pi)}{\sin \frac{1}{2} j \frac{\pi}{\nu}} \right]$$

$$= \frac{1}{2\nu} [(-1)^j - (-1)^j] = 0 \, , \qquad\qquad j \neq 0 \, , \qquad (62)$$

where we have used the identity

16

$$\frac{1}{2} + \sum_{k=1}^{n} \cos ku = \frac{1}{2} \frac{\sin (n + \frac{1}{2}) u}{\sin \frac{1}{2} u} \quad . \tag{63}$$

It is easy to see that as $\nu \to \infty$, the estimator based on the approximate filter weights given by Eq. (61) converges in the mean to $\hat{S}_0$.

In order for the asymptotic results to hold, we must have

$$2\nu TW > 1 \tag{64}$$

where T is the sampling interval, and W is the bandwidth of the major spectral peak of the $f_{jk}(x)$ and is assumed to be roughly the same for all j, k. It has been found experimentally that for microseismic signals T should be about 0.1 seconds so that $W > 5/\nu$ cps. For 21 filter points, $\nu = 10$, so that $W > \frac{1}{2}$ cps. It appears from preliminary spectral analysis that W is approximately $\frac{1}{2}$ cps, or slightly smaller, so that it is difficult to say whether the time-bandwidth product given in Eq. (64) is sufficiently large for the asymptotic results to hold for 21 filter points. In general, the rate with which the asymptotic results are achieved would have to be determined experimentally by means of computer calculations.

We now compare the asymptotic results obtained for the maximum-likelihood filter, with those for the Wiener filter. The Wiener filter functions are given by (see Reference 2, p. 10),

$$H_k(x) = A_k(x) \frac{G(x)}{G(x) - \Lambda(x)} \quad ,$$

17

here $G(x)$ is the assumed spectral density function for the signal, when the signal is taken to be a stationary random process. These filters are the same as the maximum-likelihood filters multiplied by the common filter function

$$\frac{G(x)}{G(x) - \Lambda(x)} \quad .$$

The filter weighting coefficients are

$$\alpha_k = \int_{-\pi}^{\pi} \frac{G_0(x) \, \epsilon^{-ikx}}{G_0(x) - \Lambda(x)} \, \frac{dx}{2\pi} \, , \qquad k = 0, \pm 1, \pm 2, \ldots \qquad (65)$$

which may be approximated as

$$\alpha_k = \frac{1}{2\nu + 1} \sum_{j=-\nu}^{\nu} \frac{G_0(j\frac{\pi}{\nu})}{G_0(j\frac{\pi}{\nu}) - \Lambda(j\frac{\pi}{\nu})} \, \cos jk \frac{\pi}{\nu} \qquad k = -\nu, \ldots, \nu \qquad (66)$$

We now compare the computing times required by the exact recursive and approximate frequency domain procedures. It is easily seen that the exact recursive procedure requires approximately, cf. Reference 4, pp. 98-105,

$$10\nu^2 \, K^3 \, (\mu + \alpha) \text{ seconds} , \qquad (67)$$

where the amount of computing time required to invert a $K \times K$ matrix is taken to be $K^3(\mu + \alpha)$, μ and α are the multiplication and addition times, respectively, in seconds. If ν and K are large, the approximate frequency domain synthesis procedure requires essentially the inversion of $(\nu + 1)$ $K \times K$ Hermitian spectral matrices plus a Fourier transform operation. The matrix inversion requires

18

$$(v + 1) \cdot \frac{1}{2} \cdot 4K^3 (\mu + \alpha) \cong 2vK^3 (\mu + \alpha) \text{ seconds} \tag{68}$$

and the Fourier transform operation requires

$$4v^2 K(\mu + \alpha) \text{ seconds} \tag{69}$$

for a total of

$$2vK^3 (\mu + \alpha) [1 + 2 \frac{v}{K^2}] \cong 2vK^3 (\mu + \alpha) \text{ seconds} , \qquad \frac{K^2}{2v} \gg 1 . \tag{70}$$

In Eq. (68) the factor $4K^3$ is required for the inversion of matrices with complex elements and the factor of $\frac{1}{2}$ enters due to the symmetry of the matrices. Thus, the approximate frequency domain synthesis procedure requires $5v$ times less computing time than the recursive procedure.

A convenient way to estimate the f_{jk}'s is to transform the estimate of the correlation coefficients, i.e.,

$$\hat{f}_{jk}(x) = \sum_{\ell=-2v}^{2v} (1 - \frac{|\ell|}{2v + 1}) \, \hat{\rho}_{jk}(\ell) \, \epsilon^{i\ell x} , \qquad j,k = 1,\ldots,K , \tag{71}$$

where $\hat{\rho}_{jk}(\ell)$ is the estimated correlation coefficient obtained in the following way,

$$\hat{\rho}_{jk}(\ell) = \frac{1}{L} \sum_{\substack{0 \le i \le L \\ 0 \le i+\ell \le L}} N_{j,i+\ell} N_{k,i} , \qquad \begin{array}{l} j,k = 1,\ldots,K \\ \ell = 0, 1,\ldots,2v , \end{array}$$

where L is the number of samples used to estimate the correlation coefficients. This method for estimating $f_{jk}(x)$ leads to estimated spectral matrices which are nonnegative-definite for all frequencies, $-\pi \le x \le \pi$. Note that $\hat{\rho}_{jk}(-\ell) = \hat{\rho}_{kj}(\ell), \; \ell = 0, 1,\ldots,2v.$

19

The estimated covariance matrix is nonnegative-definite since

$$\sum_{j,k=1}^{K} \sum_{m,n=-\nu}^{\nu} \hat{\rho}_{jk}(m,n)\, \alpha_{jm}\, \alpha_{kn}^{*}$$

$$= \frac{1}{L} \sum_{j,k=1}^{K} \sum_{m,n=-\nu}^{\nu} \sum_{\substack{0 \le i \le L \\ 0 \le i+m-n \le L}} N_{j,i+m-n}\, N_{ki}\, \alpha_{jm}\, \alpha_{kn}^{*}$$

$$= \frac{1}{L} \sum_{j,k=1}^{K} \sum_{m,n=-\nu}^{\nu} \sum_{\substack{0 \le i+m \le L \\ 0 \le i+n \le L}} N_{j,i+m}\, N_{k,i+n}\, \alpha_{jm}\, \alpha_{kn}^{*}$$

$$= \frac{1}{L} \sum_{i=0}^{L} \sum_{j,k=1}^{K} \sum_{m,n=-\nu}^{\nu} e(i+m)\, e(i+n)\, N_{j,i+m}\, N_{k,i+n}\, \alpha_{jm}\, \alpha_{kn}^{*}$$

$$= \frac{1}{L} \sum_{i=0}^{L} \left| \sum_{j=1}^{K} \sum_{m=-\nu}^{\nu} e(i+m)\, \alpha_{jm}\, N_{j,i+m} \right|^{2} \ge 0 \ ,$$

where

$$e(i+m) = 1, \quad 0 \le i+m \le L$$

$$= 0, \quad \text{otherwise} \ .$$

We now show that the estimated spectral matrix $\{\hat{f}_{jk}(x)\}$ is nonnegative-definite for all x, by noting

$$\hat{f}_{jk}(x) = \sum_{\ell=-2\nu}^{2\nu} \left(1 - \frac{|\ell|}{2\nu+1}\right) \hat{\rho}_{jk}(\ell)\, \epsilon^{i\ell x}$$

$$= \frac{1}{2\nu+1} \sum_{m,n=-\nu}^{\nu} \hat{\rho}_{jk}(m,n)\, \epsilon^{i(m-n)x}$$

so that

20

$$\sum_{j,k=1}^{K} \alpha_j \alpha_k^* \hat{f}_{jk}(x) = \sum_{j,k=1}^{K} \alpha_j \alpha_k^* \frac{1}{2\nu+1} \sum_{n,m=-\nu}^{\nu} \hat{\rho}_{jk}(m,n) \, \epsilon^{i(m-n)x}$$

$$= \frac{1}{2\nu+1} \sum_{j,k=1}^{K} \alpha_j \alpha_k^* \sum_{m,n=-\nu}^{\nu} \epsilon^{i(m-n)x} \left[\frac{1}{L} \sum_{\substack{0 \le i \le L \\ 0 \le i+m-n \le L}} N_{j,i+m-n} N_{ki} \right]$$

$$= \frac{1}{L(2\nu+1)} \sum_{i=0}^{L} \left| \sum_{j=1}^{K} \sum_{m=-\nu}^{\nu} e^{(i+m)} \alpha_j \, \epsilon^{imx} \, N_{j,i+m} \right|^2$$

$$\ge 0 \ .$$

Ordinarily, the amount of computer time required to perform the Fourier transformation in Eq. (71) would be

$$(2 \cdot 4\nu)(\nu+1)\frac{K^2}{2}(\mu+\alpha) \cong 4\nu^2 K^2 (\mu+\alpha) \text{ seconds },$$

and this time should be added to that in Eq. (70). If ν and K are of the same order of magnitude then the above computer time is comparable to that in Eq. (70). However, it is possible to compute the Fourier transform in Eq. (71) in less than, cf. Reference 7,

$$(8\nu \log_2 4\nu)\frac{K^2}{2}(\mu+\alpha) \cong 4\nu K^2 \log_2 4\nu (\mu+\alpha) \text{ seconds.}$$

Thus, if ν is large so that

$$\frac{\log_2 4\nu}{\nu} \ll 1 \ ,$$

a considerable saving in the amount of computer time required to perform the Fourier

transform in Eq. (71) is achieved. In addition, if we have

$$\frac{K}{2 \log_2 4\nu} \gg 1 \; ,$$

then the amount of computer time required for the Fourier transformation becomes

negligible compared to that in Eq. (70), and may therefore be neglected.

V. SYNTHESIS OF PHYSICALLY REALIZABLE FILTER BY MEANS OF A SPECTRAL MATRIX FACTORIZATION METHOD

We now consider the synthesis of the filter for the physically realizable case, $j' = v$, $v \to \infty$. In this case we have

$$A_k(x) = \sum_{j=0}^{\infty} \theta_k(j|\infty) \epsilon^{-ijx} , \tag{72}$$

and

$$\int_{-\pi}^{\pi} \epsilon^{i\ell x} \left\{ \sum_{k=1}^{K} f_{jk}(x) A_k(x) + \Lambda(x) \right\} \ dx = 0 \qquad \begin{array}{l} \ell = 0, 1, 2, \ldots \\ j = 1, \ldots, K. \end{array} \tag{73}$$

Let us define $z = \epsilon^{-ix}$ and

$$\psi_j(z) = \sum_{k=1}^{K} f_{jk}(z) A_k(z) + \Lambda(z) , \qquad j = 1, \ldots, K . \tag{74}$$

We have from Eq. (72)

$$A_k(z) = \sum_{j=0}^{\infty} \theta_k(j|\infty) z^j , \qquad k = 1, \ldots, K . \tag{75}$$

Since $A_k(z)$ is a power series in ascending powers of z, $A_k(z)$ must be analytic in the unit circle of the z-plane. In order to have $\psi_k(z)$ satisfy Eq. (73) we must have

$$\psi_k(z) = \sum_{j=1}^{\infty} b_j z^{-j} , \qquad k = 1, \ldots, K , \tag{76}$$

23

so that $\psi_k(z)$ is analytic outside and on the unit circle of the z-plane. If we subtract the first equation in (74) from all the others and substitute

$$A_1(z) = 1 - \sum_{k=2}^{K} A_k(z), \tag{77}$$

we obtain a new system of equations

$$\sum_{k=2}^{K} [f_{11}(z) + f_{jk}(z) - f_{j1}(z) - f_{1k}(z)] A_k(z) = f_{11}(z) - f_{j1}(z) + \psi_1(z) - \psi_j(z),$$
$$j = 2, \ldots, K. \tag{78}$$

The functions $\psi_1(z) - \psi_j(z)$, $j = 2, \ldots, K$ are analytic outside and on the unit circle.

Thus, we have a system of (K-1) equations in (K-1) unknowns from which we can solve for $A_2(z), \ldots, A_K(z)$. We then obtain $A_1(z)$ from Eq. (77). The system of equations in (78) may be written in matrix notation as

$$A(z) f(z) = h(z) + \psi(z). \tag{79}$$

The matrix f(x) is easily seen to be a spectral matrix. It will be assumed that f(z) has a spectral matrix factorization

$$f(z) = P(z) P'(z^{-1}), \tag{80}$$

where the matrices P(z), $[P(z)]^{-1}$ have matrix Laurent expansions on $|z| = 1$ with no negative powers of z and F'(z) denotes transpose. We have

24

$$A(z) P(z) P'(z^{-1}) = h(z) + \psi(z) , \qquad (81)$$

and

$$A(z) P(z) = h(z) [P'(z^{-1})]^{-1} + \psi(z) [P'(z^{-1})]^{-1} . \qquad (82)$$

The matrix on the left-hand side of Eq. (82) has a matrix Laurent expansion with no negative powers of z which is assumed to converge in some annulus containing the unit circle. The second term oi the right has only negative powers of z. Equating coefficients we obtain, cf. Whittle,[8] pp. 67, 100,

$$A(z) = \left\{ h(z) [P'(z^{-1})]^{-1} \right\}_{+} [P(z)]^{-1} , \qquad (83)$$

where the operation $\{ \quad \}_{+}$ indicates that only the non-negative powers of z in the Laurent expansion of the matrix within the braces are to be retained. Equation (83) represents the complete solution to the synthesis problem for multidimensional physically realizable filters.

It is now necessary to show how the factorization in Eq. (80) can be obtained. A procedure for obtaining the spectral matrix factorization for rational spectral matrices has been given by Whittle,[8] pp. 101-103, and is similar to that used in the one-dimensional case, cf. also Rozanov.[9,10]

As an example, let us consider that K = 2 and

$$f_{11}(z) - f_{21}(z) = \frac{1}{(1 - \alpha z)(1 - \alpha z^{-1})} \qquad , \ |\alpha| < 1, \qquad (84)$$

25

$$f_{11}(z) + f_{22}(z) - f_{12}(z) - f_{21}(z) = \frac{(1 - \beta z)(1 - \beta z^{-1})}{(1 - \alpha z)(1 - \alpha z^{-1})} , \quad |\beta| < 1 . \tag{85}$$

We have

$$P(z) = \frac{1 - \beta z}{1 - \alpha z} , \tag{86}$$

and

$$A_2(z) = \left[\frac{1}{(1 - \alpha z)(1 - \beta z^{-1})} \right]_+ \frac{1 - \alpha z}{1 - \beta z} . \tag{87}$$

Since $|\alpha| < 1$, the Laurent expansion of $1/(1-\alpha z)$ may be made in positive powers of z and will converge in the circle $|z| < \frac{1}{|\alpha|}$, which encloses the unit circle. The Laurent expansion of $1/(z-\beta)$ must be made in negative powers of z, i.e., $z^{-1} + \beta z^{-2} + \ldots$, and will converge in the annulus $|\beta| < |z| < \infty$, which also encloses the unit circle. Thus, in order to perform the required operation on the term in the brackets in Eq. (87) we perform a partial fraction expansion and neglect the term $1/(z-\beta)$, to obtain

$$A_2(z) = \frac{1}{1-\alpha\beta} \frac{1}{1-\beta z} \tag{88}$$

$$A_1(z) = \frac{\beta}{1-\alpha\beta} \frac{z(\alpha\beta-1) - \alpha}{1-\beta z} . \tag{89}$$

We now wish to give an example in which we will compare the performance of the physically realizable filter with that of the two-sided filter. For simplicity we consider that $K = 2$ and

$$f_{11}(z) = \frac{1}{(1 - \alpha z)(1 - \alpha z^{-1})} \quad , \qquad 0 < \alpha < 1$$

$$f_{22}(z) = -\frac{1}{(1 - \beta z)(1 - \beta z^{-1})} \quad , \qquad 0 < \beta < 1$$

$$f_{12}(z) = f_{21}(z) = 0 \quad ,$$

so that

$$f_{11}(z) - f_{21}(z) = \frac{1}{(1 - \alpha z)(1 - \alpha z^{-1})}$$

$$f_{11}(z) + f_{22}(z) - f_{12}(z) - f_{21}(z) = \frac{(1 - \alpha z)(1 - \alpha z^{-1}) + (1 - \beta z)(1 - \beta z^{-1})}{(1 - \alpha z)(1 - \alpha z^{-1})(1 - \beta z)(1 - \beta z^{-1})} .$$

In addition, let us define

$$(1 - \theta z)(1 - \theta z^{-1}) \equiv (1 - \alpha z)(1 - \alpha z^{-1}) + (1 - \beta z)(1 - \beta z^{-1})$$

so that

$$\theta = \alpha + \beta$$

27

and

$$\theta^2 = 1 + \alpha^2 + \beta^2 .$$

Thus, α and β must satisfy the equation

$$2\alpha\beta = 1$$

and

$$\theta = \alpha + \frac{1}{2\alpha} .$$

Hence α and β must lie in the open interval $(\frac{1}{2}, 1)$, θ is in the open interval $(\sqrt{2}, \frac{3}{2})$, and we have

$$f_{11}(z) + f_{22}(z) - f_{12}(z) - f_{21}(z) = \frac{\theta^2(1 - \theta^{-1}z)(1 - \theta^{-1}z^{-1})}{(1 - \alpha z)(1 - \alpha z^{-1})(1 - \beta z)(1 - \beta z^{-1})} .$$

Therefore, we can write

$$P(z) = \frac{\theta(1 - \theta^{-1}z)}{(1 - \alpha z)(1 - \beta z)}$$

and

$$A_2(z) \left[\frac{\theta(1 - \theta^{-1}z)}{(1 - \alpha z)(1 - \beta z)} \right] = \left[\frac{(1 - \alpha z^{-1})(1 - \beta z^{-1})}{\theta(1 - \alpha z)(1 - \alpha z^{-1})(1 - \theta^{-1}z^{-1})} \right]_+$$

$$= \frac{\alpha}{1 - \alpha z} ,$$

28

$$_2(z) = \frac{1}{2} \frac{z - 2\alpha}{z - \theta} \quad ,$$

$$A_1(z) = \frac{1}{2} \frac{z - \alpha^{-1}}{z - \theta} \quad .$$

We have from Eq. (13)

$$E\{\hat{S}_\infty^2\} = \int_{-\pi}^{\pi} \sum_{j,k=1}^{K} A_j(x) A_k^*(x) f_{jk}(x) \frac{dx}{2\pi}$$

$$= \frac{1}{2\pi i} \oint_{\substack{\text{unit} \\ \text{circle}}} \sum_{j,k=1}^{K} A_j(z) A_k(z^{-1}) f_{jk}(z) \frac{dz}{z}$$

$$= \sum_{\substack{\text{inside} \\ \text{unit} \\ \text{circle}}} \text{Residues} \left\{ \sum_{j,k=1}^{K} \frac{1}{z} A_j(z) A_k(z^{-1}) f_{jk}(z) \right\}$$

$$= \operatorname*{Res}_{z=\theta^{-1}} \left\{ \frac{z - \alpha^{-1}}{4\alpha(z - \theta)(\theta z - 1)(1 - \alpha z)} + \frac{\alpha}{2} \frac{z - 2\alpha}{(z - \theta)(\theta z - 1)(1 - \beta z)} \right\}$$

$$= \frac{1}{\theta^2 - 1} \left[\alpha^2 + \frac{1}{4\alpha^2} \right] \quad .$$

For the two-sided filter, we obtain

$$q_{11}(z) = (1 - \alpha z)(1 - \alpha z^{-1})$$

$$q_{22}(z) = (1 - \beta z)(1 - \beta z^{-1})$$

29

$$q_{12}(z) = q_{21}(z) = 0$$

$$\left[\sum_{j,k=1}^{2} q_{jk}(z)\right]^{-1} = \frac{1}{(1 - \theta z)(1 - \theta z^{-1})} \quad .$$

Using Eq. (57) we have

$$E\{\hat{S}_o^2\} = \int_{-\pi}^{\pi} \left[\sum_{j,k=1}^{K} q_{jk}(x)\right]^{-1} \frac{dx}{2\pi}$$

$$= \frac{1}{2\pi i} \oint_{\substack{\text{unit} \\ \text{circle}}} \left[\sum_{j,k=1}^{K} q_{jk}(z)\right]^{-1} \frac{dz}{z}$$

$$= \sum_{\substack{\text{inside} \\ \text{unit} \\ \text{circle}}} \text{Res} \left\{\frac{1}{z} \left[\sum_{j,k=1}^{K} q_{jk}(z)\right]^{-1}\right\}$$

$$= \operatorname*{Res}_{z=\theta^{-1}} \frac{1}{(1 - \theta z)(z - \theta)}$$

$$= \frac{1}{\theta^2 - 1}$$

Thus

$$\frac{E\{\hat{S}_\infty^2\}}{E\{\hat{S}_c^2\}} = \alpha^2 + \frac{1}{4\alpha^2} \quad .$$

The minimum value of the above ratio is unity and occurs when $\alpha = 1/\sqrt{2}$. The maximum value, in the permissible range for α, is 5/4 and occurs when $\alpha \to \frac{1}{2}$ or $\alpha \to 1$. Thus, there can be a loss in noise variance reduction of between zero and approximately 1 db by using the physically realizable filter rather than the two-sided filter.

30

Another method for achieving the spectral matrix factorization which is valid for general spectral matrices is due to Wiener and Masani.[1i] Suppose

$$0 < mI \leq ||f(x)|| \leq MI < \infty \ ,$$

where the norm $||a||$ of the matrix $a = \{a_{jk}\}$ is equal to $\sum_{j,k} |a_{jk}|^2$, and

$$f(x) = I + g(x) \ ,$$

$$||g(x)|| < 1 \ .$$

Under these conditions we have

$$f(x) = b^{-1}(x) C^{1/2} [b^{-1}(x) C^{1/2}]' \ ,$$

where

$$b(x) = I - g^- + (g^- g)^- - [(g^- g)^- g]^- + \ldots$$

the matrix $g^-(x)$ is obtained from the matrix $g(x) = \sum_{k=-\infty}^{\infty} g_k \epsilon^{ikx}$ by omitting the positive powers of ϵ^{ix}, $g^-(x) = \sum_{k=-\infty}^{0} g_k \epsilon^{ikx}$, and the matrix C is defined by

$$C = b(x) f(x) b'(x) \ .$$

Thus, the factorization of $f(x)$ is achieved in terms of $b^{-1}(x) C^{1/2}$. It should be noted that C is a constant nonnegative-definite matrix which is independent of x so that its square root may be obtained in the usual way, i.e., by diagonalizing C with a unitary matrix and taking the square root of the resultant diagonal matrix. Unfortunately,

31

b(x) can be obtained only as an infinite sum of matrices, the rate of convergence of which is difficult to determine.

It is readily appreciated that the techniques for synthesizing physically realizable filters by means of a spectral matrix factorization method for rational matrices are impractical for machine computation due to the requirement of having to determine roots of polynomials. In addition, the first step would require approximating all measured spectral matrices with rational spectral matrices. This step could also be quite complicated and could entail a serious loss with respect to the amount of noise power which can be minimized.

It is difficult in general to determine how well a physically realizable filter performs relative to a two-sided filter. The two-sided filter must always be better than the physically realizable filter since

$$E\{\hat{S}_0^2\} \le E\{\hat{S}_\infty^2\} .$$

This follows from the fact that the optimum weighting functions in the two-sided case are subject to fewer constraints than those for the physically realizable case. However, we can obtain some results along these lines by using the theory of Toeplitz forms.[12]

We have already shown, cf. Eq. (57),

$$\lim_{\nu \to \infty} E\{\hat{S}_{\nu j}^2\} = \int_{-\pi}^{\pi} \left[\sum_{j,k=1}^{K} q_{jk}(x) \right]^{-1} \frac{dx}{2\pi} , \tag{90}$$

32

for a fixed j'. The quantity on the right in Eq. (90) represents a lower bound for the variance of the processed noise and can only be attained asymptotically as the memory of the filter, $(2\nu + 1)$, increases. We obtain from Eqs. (7), (22), in conjunction with some results from theory of Toeplitz forms,[12]

$$\lim_{\nu \to \infty} \frac{1}{2\nu + 1} \sum_{j'=-\nu}^{\nu} E\{\hat{S}^2_{\nu j'}\} = \int_{-\pi}^{\pi} \left[\sum_{j,k=1}^{K} f^{-1}_{jk}(x) \right]^{-1} \frac{dx}{2\pi}. \qquad (91)$$

This equation tells us what the average processed noise variance should be, where the averaging is done with respect to the time, j', at which the signal is to be estimated.

VI. INTERPRETATION OF OPTIMUM MULTIDIMENSIONAL FILTERING IN TERMS OF FREQUENCY WAVE-NUMBER SPACE

We now wish to show the manner in which maximum-likelihood filtering can be interpreted in the frequency wave-number space. We introduce the wave-numbers k_x, k_y and the frequency wave-number spectrum as follows

$$\bar{f}(x',k_x,k_y) = \sum_{j,k=1}^{K} \sum_{\ell=-\infty}^{\infty} \rho_{jk}(\ell) \exp\left\{i[\ell x' - k_x(x_j-x_k) - k_y(y_j-y_k)]\right\}$$

$$= \sum_{j,k=1}^{K} f_{jk}(x') \exp[i(k_x x_j + k_y y_j)] \exp[i(k_x x_k + k_y y_k)] , \qquad (92)$$

where (x_j,y_j) and (x_k,y_k) are the spatial coordinates of the jth and k^{th} seismometers, respectively, measured with respect to some reference. Since $\{f_{jk}(x')\}$ is a non-negative-definite matrix for all x', it follows from Eq. (92) that $\bar{f}$ is nonnegative for all x', k_x,k_y.

We have

$$f_{jk}(x') = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \bar{f}(x',k_x,k_y) \exp[i(k_x x_j + k_y y_j)] \exp[-i(k_x x_j + k_y y_j)] \frac{dk_x}{2\pi} \frac{dk_y}{2\pi} , \quad (93)$$

so that

$$E\{\hat{s}_j^2\} = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |B(x',k_x,k_y)|^2 \bar{f}(x',k_x,k_y) \frac{dx'}{2\pi} \frac{dk_x}{2\pi} \frac{dk_y}{2\pi} , \qquad (94)$$

where

$$B(x',k_x,k_y) = \sum_{j=1}^{K} \sum_{k=-\infty}^{\infty} \theta_j(k|j') \exp[i(kx' + k_x x_j + k_y y_j)] . \qquad (95)$$

34

We now consider only the situation where $j' = 0$, in which case the constraint Eq. (11) becomes

$$B(x',0,0) = 1 \ . \tag{96}$$

Thus, the maximum-likelihood filter, or minimum-variance unbiased estimator, has those filter weights $\theta_j(k|0)$ such that the triple integral over frequency wave-number space, in Eq. (94) is minimized subject to the constraints in Eqs. (95), (96). The solution to this problem would have to be obtained in the usual way, i.e., by using the recursive method of solution, assuming, of course, that ν is finite, the approximate frequency domain approach for two-sided filters, if ν is large, or the spectral matrix factorization technique for physically realizable filters, when ν is l rge.

However, we can visualize to a great extent what the minimizing solution for B should be in the frequency wave-number space. It is easily seen that if K is reasonably large and if the spatial coordinates (x_j, y_j) are reasonably different from (x_k, y_k), for all j, k, the minimizing solution for B will be such that $|B|$ is small whenever $\bar{f}$ is large, except at the origin of the wave-number plane, where, for all frequencies, B is equal to unity. Thus, $|B|$ would tend to be needle-shaped, with amplitude unity, at the origin of the k_x-k_y plane, for all values of the frequency variable x', and would tend to have sidelobes in those parts of the k_x-k_y plane where $\bar{f}$ would be small, for a given x'. In general, the positions and amplitudes of these lobes would be frequency dependent.

35

VII. CONCLUSIONS

It has been shown that, basically, the reason for the equivalence of the maximum-likelihood and minimum-variance unbiased estimators is that they are based on a conditional expectation. The martingale property of conditional expectation then assures that the asymptotic properties of these estimators are well defined.

An asymptotically optimum frequency domain synthesis procedure has been presented which requires much less computing time than the exact recursive procedure. However, the frequency domain synthesis technique is valid only for two-sided filters, while the recursive procedure is always valid. If two-sided filters are to be used due to their inherently better capability to suppress the noise, then the frequency domain approach is superior, assuming that ν is large enough for the asymptotic results to hold.

Synthesis procedures for physically realizable filters, in the asymptotic case, have also been presented. These procedures were based on a spectral matrix factorization technique, and are not well adapted to machine computation.

REFERENCES

1. E. J. Kelly and M. J. Levin, "Signal Parameter Estimation for Seismometer Arrays," Technical Report 339, M. I. T., Lincoln Laboratory (8 January 1964) DDC 435489.

2. E. J. Kelly, "A Comparison of Seismic Array Processing Schemes," Technical Note 1965-21, M. I. T., Lincoln Laboratory (14 June 1965).

3. M. J. Levin, "Maximum-Likelihood Array Processing," Semiannual Technical Summary Report on Seismic Discrimination, M. I. T., Lincoln Laboratory (31 December 1964).

4. E. A. Robinson, "Wavelet Composition of Time Series," contained in Econometric Model Building, H. O. A. Wold (North-Holland Publishing Co., Amsterdam; 1964).

5. Y. A. Rozanov, "One Problem of Optimal Control of a Complex of Instruments," Math. Inst. imeni V. A. Steklova, Akad. Nauk SSSR, Trudy, LXXI, 88-101, (1964). English translation available from Lincoln Laboratory Library.

6. J. L. Doob, Stochastic Processes (John Wiley and Sons, Inc., New York; 1953).

7. J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," Math. of Comp., 19 (1965).

8. P. Whittle, Prediction and Regulation by Linear Least-Square Methods (D. Van Nostrand Co., Inc., Princeton, New Jersey; 1963).

9. Y. A. Rozanov, "Spectral Theory of Multidimensional Stationary Stochastic Processes with Discrete Time," 13, 93 (1958). English translation in "Selected Translations in Mathematical Statistics and Probability," I, American Mathematical Society, Providence, R. I. (1961).

10. Y. A. Rozanov, "Spectral Properties of Multivariate Stationary Processes and Boundary Properties of Analytic Matrices," Theory Prob. and Appl., English Edition, 5 (1960).

11. N. Wiener and P. Masani, "Prediction Theory of Multivariate Stochastic Processes, Part I, The Regularity Condition," Acta Math., 98 (1957); "Part II, The Linear Predictor," Acta. Math., 99 (1958).

37

12.   U. Grenander and G. Szegö, Toeplitz Forms and Their Applications (University of California Press, Berkeley; 1958).

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Lincoln Laboratory, M.I.T.	Unclassified
	2b. GROUP None

3. REPORT TITLE

Asymptotically Optimum Multidimensional Filtering for Sampled-Data Processing of Seismic Arrays

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Note

5. AUTHOR(S) (Last name, first name, initial)

Capon, Jack   Greenfield, Roy J.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
17 December 1965	46	12

8a. CONTRACT OR GRANT NO. AF 19(628)-5167	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. ARPA Order 512	Technical Note 1965-57
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ESD-TDR-65-586
d.	

10. AVAILABILITY/LIMITATION NOTICES

None

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Advanced Research Projects Agency, Department of Defense

13. ABSTRACT

A number of asymptotically optimum multidimensional filtering methods are investigated with the purpose of determining filtering techniques which require relatively little computing time to implement with a digital computer. In particular, the asymptotic properties of the maximum-likelihood and minimum-variance unbiased multidimensional filters are investigated in the sampled-data case. These two multidimensional filters are shown to be identical since they are both based on a conditional expectation. In addition, the martingale property of conditional expectation assures that the asymptotic properties of these multidimensional filters are well defined.

An asymptotically optimum frequency domain synthesis procedure is given for two-sided multidimensional filters. This procedure is well suited to machine computation and has the advantage with respect to the exact recursive synthesis method of requiring much less computation time. A synthesis procedure for physically realizable multidimensional filters is presented which is based on a factorization of rational spectral matrices. This method is not, however, well suited to machine computation. An interpretation of optimum multidimensional filtering in terms of frequency wave-number space is also given.

14. KEY WORDS

filtering	Gaussian distribution	Lagrange
digital computers	Wiener filter	Fourier transform
spectral matrices	synthesis procedures	

39