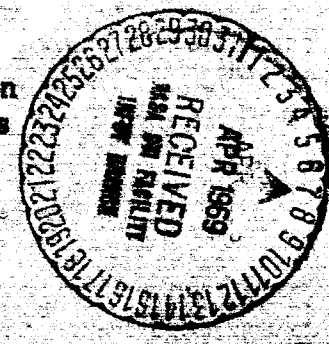
DIGITAL MATCHED FILTERS FOR DETECTING GAUSSIAN SIGNALS IN GAUSSIAN NOISE

by

T. L. Henderson
D. G. Lainiotis



Technical Memorandum No. 3
September 12, 1968

INFORMATION SYSTEMS RESEARCH LABORATORY

ELECTRONICS RESEARCH CENTER
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712

DIGITAL MATCHED FILTERS FOR DETECTING GAUSSIAN SIGNALS IN GAUSSIAN NOISE*

By

T. L. Henderson**
D. G. Lainiotis***

Technical Memorandum No. 3
September 12, 1968

INFORMATION SYSTEMS RESEARCH LABORATORY

ELECTRONICS RESEARCH CENTER
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712

* Research sponsored in part by the Joint Services Electronics Program under the Research Grant AF-AFOSR-766-67, Mod. 1 AFOSR-67-0766A and by NASA Grant NGR 44-012-066

** Graduate Student in Electrical Engineering

*** Associate Professor of Electrical Engineering, The University of Texas at Austin

ABSTRACT

The use of a set of digital matched filters is presented as an alternative to direct computation of the likelihood-ratio, for the problem of detecting a random signal in random noise. It is assumed that a random process of Gaussian background noise and (with probability P) a zero-mean Gaussian signal is sampled at N instants, the samples being corrupted by additive Gaussian measurement noise. The samples are processed by K << N digital correlation filters which are structured so that the signal can be detected with minimum Bayes risk. The optimum filters are shown to be matched to the most relevant components of the simultaneously orthogonal expansion of the set of sampled data. State variable techniques are used to find a very practical method for determining the optimum filter structures.

I. INTRODUCTION

Suppose that we are required to detect a random signal in the presence of noise; more specifically suppose

$$r(t) = \begin{cases} n(t) + s(t) & \text{when the signal is present} \\ n(t) & \text{otherwise} \end{cases} \tag{1}$$

where the signal s(t) and the background noise n(t) are random processes, and we must decide whether the signal is present by observing a sequence of noisy measurements

$$z_i = r(t_i) + m_i, \quad i = 1, \ldots, N, \tag{2}$$

taken at N sampling instants $t_1 < t_2 < \ldots < t_N$. (The random variables m_i represent measurement noise.)

The following additional assumptions will be made:

A. s(t) and n(t) are Gaussian random processes and $\{m_i\}_{i=1}^{N}$ is a sequence of Gaussian random variables. Furthermore, s(t), n(t), $m_1, m_2, \ldots, m_N$ are statistically independent.

B. s(t), n(t), and $\{m_i\}_{i=1}^{N}$ are zero-mean.

C. The covariances

$$R_s(i,j) \triangleq E[s(t_i)s(t_j)]; \; R_n(i,j) \triangleq E[n(t_i)n(t_j)]; \text{ and } M(i) \triangleq E[m_i^2]$$

are known. (Note that $E[m_i m_j] = 0$ when $i \neq j$ because of A and B.)

D. $M(i) > 0$ for $i = 1, \ldots, N$.

The assumption that n(t) and $\{m_i\}_{i=1}^{N}$ are zero-mean is made with no loss in generality, since any non-zero means would always be present and could be subtracted from the observations with no loss in our ability to detect the signal.

We shall denote by H_1 (H_0) the hypothesis that the signal is (is not) present, and define a data vector

$$\underline{z} = (z_1, z_2, \ldots, z_N)^T. \tag{3}$$

Then under either hypothesis $\underline{z}$ is a zero-mean Gaussian random vector, and its hypothesis-conditional covariance matrices are

$$\underline{R}_0 \triangleq E[\underline{z}\underline{z}^T/H_0] = \underline{R}_n + \underline{M} \tag{4}$$

$$\underline{R}_1 \triangleq E[\underline{z}\underline{z}^T/H_1] = \underline{R}_s + \underline{R}_n + \underline{M} = \underline{R}_0 + \underline{R}_s \tag{5}$$

where the elements of the matrices $\underline{R}_s$, $\underline{R}_n$ and $\underline{M}$ are defined[+]

$$(\underline{R}_s)_{ij} \triangleq R_s(i,j); \quad (\underline{R}_n)_{ij} \triangleq R_n(i,j); \quad \text{and} \quad (\underline{M})_{ij} \triangleq M(i)\delta_{ij}. \tag{6}$$

The detector can be represented as a decision function $\mathcal{D}(\underline{z})$ which, given a data sample $\underline{z}$, chooses either H_0 or H_1; i.e. $\mathcal{D}: R^N \to \{H_0, H_1\}$ where R^N is the N-dimensional observation space. The performance criterion most often used to appraise a given detector $\mathcal{D}(\underline{z})$ is the risk,

$$J \triangleq P \cdot \mathcal{C}_m \cdot \text{Prob}\{\mathcal{D}(\underline{z}) = H_0 \text{ when } H_1 \text{ is true}\}$$
$$+ (1-P) \cdot \mathcal{C}_f \cdot \text{Prob}\{\mathcal{D}(\underline{z}) = H_1 \text{ when } H_0 \text{ is true}\}. \tag{7}$$

where P is the a priori probability that the signal is present (i.e. that H_1 is true), and $\mathcal{C}_m$ and $\mathcal{C}_f$ are the assigned costs (both > 0) of a "miss" and a "false alarm" respectively. (When $\mathcal{C}_m = \mathcal{C}_f = 1$, J is simply the probability of an error).

The "optimum" detector, i.e. the one which minimizes J, is one based on the well-known Bayes test,

$$\mathcal{D}^*(\underline{z}) = \begin{cases} H_1 \text{ whenever } \ell(\underline{z}) > \eta \\ H_0 \text{ otherwise} \end{cases} \tag{8}$$

where $\ell(\underline{z})$ is the likelihood ratio and η is defined as

$$\eta \triangleq \frac{(1-P)\mathcal{C}_f}{P\mathcal{C}_m} \quad \cdot \text{ (See [1] p. 26)} \tag{9}$$

The resulting risk is called the "Bayes Risk", and we shall denote it by $J^*[\underline{z}]$. It should be noted that $J^*[\underline{z}]$ is a function of the conditional statistics of $\underline{z}$ and not of $\underline{z}$ itself.

[+]Parentheses with subscripts will be used to denote elements of vectors and matrices.

For our problem, the optimum detector can be expressed in a more explicit form, (See [1] p. 107).

$$\mathcal{D}^*(\underline{z}) = \begin{cases} H_1 \text{ whenever } \underline{z}^T(\underline{R}_0^{-1} - \underline{R}_1^{-1})\underline{z} > 2\log\eta + \log\dfrac{\det\underline{R}_1}{\det\underline{R}_0} \\ H_0 \text{ otherwise} \end{cases} \tag{10}$$

Unfortunately, this decision scheme becomes prohibitively complex when N, the number of sampling instants, is large. When the signal and noise processes admit state variable models, Schweppe [2] has shown that the likelihood ratio can be obtained by recursive computation. However, for many applications, even this simplified scheme proves impractical.

A technique which can be used to avoid difficulty is to use a decision function which depends only upon the values of K linear combinations of the measurements, [3]

$$y_k = \sum_{i=1}^{N} b_{ki} z_i, \text{ for } k = 1, \ldots, K, \tag{11}$$

where K is considerably smaller than N. Each y_k may be regarded as the output of a digital filter which correlates the measurement sequence $\{z_i\}_{i=1}^N$ with a predetermined weighting sequence $\{b_{ki}\}_{i=1}^N$. If we define vectors

$$\underline{b}_k \triangleq (b_{k1}, b_{k2}, \ldots, b_{kN})^T \text{ for } k = 1, \ldots, K, \tag{12}$$

then

$$y_k = \underline{b}_k^T \underline{z} \qquad \text{for } k = 1, \ldots, K. \tag{13}$$

Another interpretation is provided by defining a vector

$$\underline{y} \triangleq (y_1, y_2, \ldots, y_K)^T \tag{14}$$

and a KxN matrix

$$\underline{B} \triangleq (\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_K)^T \tag{15}$$

so that

$$\underline{y} = \underline{B}\underline{z}. \tag{16}$$

Clearly the data vector $\underline{z}$ has simply been "reduced" by a linear transformation. Thus $\underline{y}$ is, under either hypothesis, a zero-mean Gaussian random vector, and it may be regarded as a new data vector. We can associate with $\underline{y}$ its Bayes Risk, $J^*[\underline{y}] = J^*[\underline{B}\underline{z}]$, which is the least risk that can be attained when $\underline{y}$ is used for signal detection.

Our problem will be to find the optimum $\underline{B}$ for a given value of K, i.e. to find a $\underline{B}$ which minimizes $J^*[\underline{Bz}]$. We will show that our optimum choice, designated $\underline{B}^*$, has the property that its row vectors are "matched" to the most relevant components in the simultaneously orthogonal expansion of the data vector $\underline{z}$. Furthermore, $\underline{B}^*$ depends neither upon the a priori probability P nor upon the costs $\mathcal{C}_m$, $\mathcal{C}_f$, and is invariant with respect to scalings of s(t) by a constant factor. We will present a very practical method for finding $\underline{B}^*$ when N is large, by employing state variable models for the signal and noise processes; this is perhaps our most important result. Before proceeding, we note that the independent variable "t" need not represent time; it could just as well correspond to position along a line in space.

II. SIMULTANEOUSLY ORTHOGONAL EXPANSION OF z

It is a well known fact that there exists a linear, invertible transformation which sends $\underline{z}$ into a random vector whose elements are statistically independent under either hypothesis. We shall develop this transformation in a form slightly different from that found in most texts, and interpret it in terms of a simultaneously orthogonal expansion of the data vector $\underline{z}$. (Our procedure will be similar to that of Kadota and Shepp [4].)

From assumption D, the matrix $\underline{M}$ must be positive definite, so $\underline{R}_0$ is positive definite and $\underline{R}_0^{-\frac{1}{2}}$ exists, is positive definite, and symmetric. Thus $\underline{R}_0^{-\frac{1}{2}}\,\underline{R}_1\underline{R}_0^{-\frac{1}{2}}$ is positive definite, symmetric, and has N positive eigenvalues

$$D_1 \geq D_2 \geq \ldots, \geq D_N \tag{17}$$

and associated eigenvectors

$$\underline{R}_0^{-\frac{1}{2}}\underline{R}_1\underline{R}_0^{-\frac{1}{2}}\underline{\varphi}_i = D_i\underline{\varphi}_i \text{ for } i = 1,\ldots,N, \tag{18}$$

which may be taken orthonormal,

$$\underline{\varphi}_i^T\underline{\varphi}_j = \delta_{ij} \text{ for } i, j = 1,\ldots,N. \tag{19}$$

If we define

$$\underline{d}_i \triangleq \underline{R}_0^{-\frac{1}{2}}\underline{\varphi}_i \text{ for } i = 1,\ldots,N,$$

Eqs. (18) and (19) can be expressed in the equivalent forms

$$\underline{R}_0^{-1}\underline{R}_1\underline{d}_i = D_i\underline{d}_i \text{ for } i = 1,\ldots,N \tag{20}$$

$$\underline{d}_i^T\underline{R}_0\underline{d}_j = \delta_{ij} \text{ for } i, j = 1,\ldots,N. \tag{21}$$

Alternatively, given D_i's and $\underline{d}_i$'s which satisfy Eqs.(20) and (21), we can define $\underline{\varphi}_i \triangleq \underline{R}_0^{-\frac{1}{2}}\underline{d}_i$ which will satisfy Eqs. (18) and (19).

For a given set $(D_1,\underline{d}_1)$, $(D_2,\underline{d}_2),\ldots,(D_N,\underline{d}_N)$ which satisfy Eqs. (17), (20), and (21), define

$$\alpha_i \triangleq \underline{d}_i^T\underline{z} \quad \text{for } i = 1,\ldots,N. \tag{22}$$

Then the α_i's are, under either hypothesis, zero-mean Gaussian random variables, and have hypothesis-conditional covariances, for $i, j = 1,\ldots,N$,

$$E[\alpha_i\alpha_j/H_0] = \underline{d}_i^T\underline{R}_0\underline{d}_j = \delta_{ij}, \tag{23}$$

$$E[\alpha_i\alpha_j/H_1] = \underline{d}_i^T\underline{R}_1\underline{d}_j = \underline{d}_i^T\underline{R}_0\underline{R}_0^{-1}\underline{R}_1\underline{d}_j = D_j\underline{d}_i^T\underline{R}_0\underline{d}_j = D_j\delta_{ij}. \tag{24}$$

Define an Nxl vector

$$\underline{\alpha} \triangleq (\alpha_1,\alpha_2,\ldots,\alpha_N)^T, \tag{25}$$

and an NxN matrix

$$\underline{T} \triangleq (\underline{d}_1,\underline{d}_2,\ldots,\underline{d}_N)^T, \tag{26}$$

so that

$$\underline{\alpha} = \underline{T}\underline{z}. \tag{27}$$

Eq. (21) implies that the $\underline{d}_i$'s are linearly independent, so $\underline{T}$ is invertible. Thus

$$\underline{z} = \underline{T}^{-1}\underline{\alpha} = \sum_{i=1}^{N}\alpha_i\underline{\psi}_i, \tag{28}$$

where $\underline{\psi}_1,\underline{\psi}_2,\ldots,\underline{\psi}_N$ are the column vectors of $\underline{T}^{-1}$.

Eq. (28) is the simultaneously orthogonal expansion of the data vector $\underline{z}$. The basis vectors $\underline{\psi}_1,\underline{\psi}_2,\ldots,\underline{\psi}_N$ are, in general, not orthogonal. The words "simultaneously orthogonal" refer to the fact that the random coefficients α_i are statistically orthogonal, i.e. uncorrelated, under both hypotheses. Note that since $\underline{T}\underline{T}^{-1} = I$,

$$\underline{d}_i^T\underline{\psi}_j = \delta_{ij} \quad \text{for } i, j = 1,\ldots,N. \tag{29}$$

III. OPTIMUM CHOICE OF B

Given a zero-mean Gaussian random vector with two alternative positive definite covariance matrices $\underline{R}_0$ and $\underline{R}_1$, the problem of finding the

the best $\underline{y} = \underline{B}\underline{z}$ is closely related to the simultaneously orthogonal expansion. For the case $K = 1$ (so that $\underline{B}$ is simply a row vector), it is shown in Kullback (see [5] p. 198) that the $\underline{B}$ for which the divergence[+] is maximum is given by

$$\underline{B} = \underline{d}_{j^*}^T$$

where j^* is the j for which $f_j \triangleq D_j + 1/D_j$ has its largest value.

Kadota and Shepp [3] generalized this result to the case $K > 1$, showing that the $\underline{B}$ which maximizes the divergence is

$$\underline{B} = (\underline{d}_{j_1^*}, \underline{d}_{j_2^*}, \ldots, \underline{d}_{j_K^*})^T$$

where j_1^* is the j for which f_j has its largest value.
j_2^* is the j for which f_j has its second largest value, etc.
They also showed that the same choice maximizes the Bhattacharya distance. (Actually, they showed that it minimizes the Hellinger integral, which is equivalent.)

Since the statistically independent α_i's have unit variances under hypothesis H_0, and variances D_i under hypothesis H_1, it follows that the expansion components most relevant for hypothesis determination are those for which the D_i's are farthest away from the value one. (This is in agreement with the use of $f_j = D_j + 1/D_j$ in the cases just described). However, for our problem $\underline{R}_1 = \underline{R}_0 + \underline{R}_s$. Using this fact and premultiplying Eq. (20) by $\underline{d}_i^T\underline{R}_0$, we achieve the following results:

$$D_i = 1 + \frac{\underline{d}_i^T\underline{R}_s\underline{d}_i}{\underline{d}_i^T\underline{R}_0\underline{d}_i} \qquad \text{for } i = 1, \ldots, N.$$

Thus $D_i \geq 1$ for $i = 1, \ldots, N$. Hence the most relevant components of the expansion are those with the largest values of D_i. The following theorem is, therefore, not surprising.

THEOREM: Given a set $(D_1, \underline{d}_1), (D_2, \underline{d}_2), \ldots, (D_N, \underline{d}_N)$, which satisfy Eqs. (17), (20), and (21), define

$$B^* \triangleq (\underline{d}_1, \underline{d}_2, \ldots, \underline{d}_K)^T. \tag{30}$$

Then for any KxN matrix $\underline{B}$,

$$J^*[\underline{B}^*\underline{z}] \leq J^*[\underline{B}\underline{z}] \tag{31}$$

[+]See Kullback [5] p. 6 for the definition of divergence.

i.e. $\underline{B}^*$ minimizes the Bayes risk. The proof of this theorem is given in Appendix A.

If we define

$$\underline{y}^* \triangleq \underline{B}^* \underline{z}, \tag{32}$$

then the optimum decision rule is simply

$$\mathcal{D}^*(\underline{y}^*) = \begin{cases} H_1 \text{ whenever } \sum_{i=1}^{K} (1-1/D_i)(\underline{y}^*)_i^2 > 2\log\eta + \sum_{i=1}^{K} \log D_i \\ H_0 \text{ otherwise} \end{cases} \tag{33}$$

This result is obtained by replacing $\underline{z}$ by $\underline{y}^*$ and substituting the appropriate covariance matrices into Eq. (10).

Certainly, $\underline{B}^*$ does not depend $P, \mathcal{C}_m$, or $\mathcal{C}_f$. Furthermore, Eq. (20) is equivalent to

$$\underline{R}_0^{-1}\underline{R}_s\underline{d}_i = (D_i - 1)\underline{d}_i \quad \text{for } i = 1, \ldots, N. \tag{34}$$

Thus if $s(t)$ becomes $\beta \cdot s(t)$, where β is a constant scaling factor, then $\underline{R}_s$ becomes $\beta^2\underline{R}_s$, the eigenvectors remain the same, and the ordering of the eigenvalues is unchanged; i.e. $\underline{B}^*$ is unaffected when the signal is scaled by a constant factor.

Clearly, in order to find $\underline{B}^*$ we must find the eigenvectors of $\underline{R}_0^{-1}\underline{R}_1$ corresponding to the largest eigenvalues; so the fundamental problem is that of finding solutions $(D, \underline{d})$ to the equation

$$\underline{R}_0^{-1}\underline{R}_1\underline{d} = D\underline{d}. \tag{35}$$

However, it should be noted that any solution $(D, \underline{d})$ with $D = 1$ would yield a component of $\underline{y}^*$ which would be useless for signal detection (since it would have the same variance under either hypothesis). From Eq. (34) it is clear that a useless solution (i.e. $D = 1$) exists if (and only if) $\underline{R}_s$ is singular.

The "orthogonalization" of the solution vectors required by Eq. (21) is not difficult to achieve. If $(D_a, \underline{d}_a)$ and $(D_b, \underline{d}_b)$ are two solutions to Eq. (35) and $D_a \neq D_b$ then $\underline{d}_a^T\underline{R}_0\underline{d}_b = 0$; thus we need only insure that Eq. (21) is satisfied when $D_i = D_j$.

Eq. (35) may be expressed in another form,

$$\underline{\mathcal{K}}(D)\underline{d} = 0 \tag{36}$$

where $$\underline{\mathcal{K}}(D) \triangleq (1-D)(\underline{R}_n + \underline{M}) + \underline{R}_s. \tag{37}$$

Solutions therefore occur at values of D for which $\underline{\mathcal{K}}(D)$ is singular; i.e. at the zeros of $\det \underline{\mathcal{K}}(D)$. Unfortunately the dimension of the matrix $\underline{\mathcal{K}}(D)$ is equal to the number of sampling instants, and if this number is large (say in excess of fifty) it becomes practically impossible to find the zeros of $\det \underline{\mathcal{K}}(D)$. In the next two sections of this paper we will assume that the signal and noise processes admit state variable models, and will develop a technique by which the problem is reduced to finding solutions $(D,\underline{g})$ of

$$\underline{\mathcal{L}}(D)\underline{g} = 0, \tag{38}$$

where $\underline{\mathcal{L}}(D)$ is a square matrix which depends upon D and whose dimension is not more than twice the sum of the dimensions of the signal and noise state vectors. The values of D for which Eq. (38) has a solution are the same as those for Eq. (36), and the solution vectors $\underline{d}$ for Eq. (36) are related to the solution vectors of Eq. (38) by a simple transformation.

Before proceeding, we note that the rows $\underline{d}_i^T$ of $\underline{B}^*$ are matched to the components of the simultaneously orthogonal expansion of $\underline{z}$ in the sense that $\underline{d}_i^T \underline{\psi}_j = \delta_{ij}$ (from Eq. (29)). Also, $(\underline{y}^*)_i$ is just z_i, the $i^{\underline{th}}$ coefficient in the expansion.

IV. STATE VARIABLE MODELS FOR THE SIGNAL AND NOISE PROCESSES

We will assume that the noise and signal processes can be represented in terms of an $L_n \times 1$ noise state vector $\underline{x}_n(t)$ and an $L_s \times 1$ signal state vector $\underline{x}_s(t)$ as follows,

$$n(t) = \underline{c}_n^T(t)\underline{x}_n(t); \quad s(t) = \underline{c}_s^T(t)\underline{x}_s(t) \tag{39}$$

where the state vectors obey the following dynamic equations,

$$\frac{d\underline{x}_n(t)}{dt} = \underline{F}_n(t)\underline{x}_n(t) + \underline{u}_n(t); \quad \frac{d\underline{x}_s(t)}{dt} = \underline{F}_s(t)\underline{x}_s(t) + \underline{u}_s(t) . \tag{40}$$

$\underline{u}_n(t)$ and $\underline{u}_s(t)$ are zero-mean, white Gaussian vector processes, independent of each other, with covariances

$$E[\underline{u}_n(t)\underline{u}_n^T(\tau)] = \underline{U}_n(t)\delta(t-\tau); \quad E[\underline{u}_s(t)\underline{u}_s^T(\tau)] = \underline{U}_s(t)\delta(t-\tau) . \tag{41}$$

Furthermore we will assume that the initial statistics for the state vectors are $E[\underline{x}_n(t_1)] = E[\underline{x}_s(t_1)] = 0$, and

$$E[\underline{x}_n(t_1)\underline{x}_n^T(t_1)] = \underline{K}_n(1,1); \quad E[\underline{x}_s(t_1)\underline{x}_s^T(t_1)] = \underline{K}_s(1,1) . \tag{42}$$

We assume that the vectors $\underline{c}_n(t)$, $\underline{c}_s(t)$ and the matrices $\underline{F}_n(t)$, $\underline{F}_s(t)$, $\underline{U}_n(t)$, $\underline{U}_s(t)$, $\underline{K}_n(1,1)$, and $\underline{K}_s(1,1)$ are known.

We now convert from continuous state variable models to discrete models. Since the models for n(t) and s(t) have the same form, we will refrain from writing equations for both models, and use only the Greek letter θ which may be taken as either n or s.

Define the (square) transition matrix $\underline{\Phi}_\theta(t,\tau)$ through the relations

$$\frac{\partial \underline{\Phi}_\theta(t,\tau)}{\partial t} = \underline{F}_\theta(t)\,\underline{\Phi}_\theta(t,\tau); \quad \underline{\Phi}_\theta(\tau,\tau) = \underline{I} \tag{43}$$

where $\underline{I}$ is the identity matrix.

Using the transition matrix,

$$\underline{x}_\theta(t_{i+1}) = \underline{\Phi}_\theta(t_{i+1},t_i)\underline{x}_\theta(t_i) + \int_{t_i}^{t_{i+1}} \underline{\Phi}_\theta(t,\tau)\underline{u}_\theta(\tau)d\tau \quad \text{for } i = 1,\ldots,N-1.$$

If we define

$$\underline{A}_\theta(i) \triangleq \underline{\Phi}_\theta(t_{i+1},t_i); \quad \underline{q}_\theta(i) \triangleq \int_{t_i}^{t_{i+1}} \underline{\Phi}_\theta(t,\tau)\underline{u}_\theta(\tau)d\tau \tag{44}$$

then the equation can be written

$$\underline{x}_\theta(t_{i+1}) = \underline{A}_\theta(i)\underline{x}_\theta(t_i) + \underline{q}_\theta(i) \quad \text{for } i = 1,\ldots,N-1 \tag{45}$$

which is the discrete version of (40). Note that the $\underline{q}_\theta(i)$'s are zero-mean, Gaussian random vectors. From (41) and (44) it can easily be shown that

$$E[\underline{q}_\theta(i)\underline{q}_\theta^T(j)] = \underline{Q}_\theta(i)\,\delta_{ij} \tag{46}$$

where $\underline{Q}_\theta(i)$ is defined

$$\underline{Q}_\theta(i) \triangleq \int_{t_i}^{t_{i+1}} \underline{\Phi}_\theta(t,\tau)\underline{U}_\theta(\tau)\underline{\Phi}_\theta^T(t,\tau)d\tau \tag{47}$$

Since $\underline{u}_n(t)$ and $\underline{u}_s(t)$ are mutually independent, $\{\underline{q}_n(i)\}_{i=1}^N$ and $\{\underline{q}_s(i)\}_{i=1}^N$ will be mutually independent random sequences.

As a final comment before proceeding to the next section, we note that $\underline{A}_\theta(i)$ always has an inverse since it is a transition matrix.

V. METHOD OF SOLUTION

Since we are only interested in solutions of Eq. (35) such that $D \neq 1$, we may define $D' \triangleq \frac{1}{D-1}$ and write Eq. (35) in the equivalent form

$$\underline{d} = \underline{M}^{-1}[D'\underline{R}_s\underline{d} - \underline{R}_n\underline{d}].$$

By defining $d(i) \triangleq (\underline{d})_i$, this equation can be written as

$$d(i) = M^{-1}(i)[D' \sum_{j=1}^{N} R_s(i,j)d(j) - \sum_{j=1}^{N} R_n(i,j)d(j)] \quad \text{for } i = 1,\ldots,N \tag{48}$$

But, for $i, j = 1,\ldots,N$,

$$R_\theta(i,j) \triangleq E[\theta(t_i)\epsilon(t_j)] = \underline{c}_\theta^T(t_i)\underline{K}_\theta(i,j)\underline{c}_\theta(t_j) \tag{49}$$

where

$$\underline{K}_\theta(i,j) \triangleq E[\underline{x}_\theta(t_i)\underline{x}_\theta^T(t_j)] \tag{50}$$

(Note that this definition of $\underline{K}_\theta(i,j)$ is compatible with the previous definition of $\underline{K}_\theta(1,1)$).

By defining, for $i = 1,\ldots,N$,

$$\underline{v}_\theta(i) \triangleq \sum_{j=1}^{N} \underline{K}_\theta(i,j)\underline{c}_\theta(t_j)d(j), \tag{51}$$

Eq. (48) becomes

$$d(i) = M^{-1}(i)[D'\underline{c}_s^T(t_i)\underline{v}_s(i) - \underline{c}_n^T(t_i)\underline{v}_n(i)] \quad \text{for } i = 1,\ldots,N. \tag{52}$$

In Appendix B it is shown that $\underline{v}_\theta(i)$ may be obtained as the solution of a system of linear difference equations,+

$$\left.\begin{aligned} \underline{v}_\theta(i+1) &= \underline{A}_\theta(i)\underline{v}_\theta(i) + \underline{Q}_\theta(i)\underline{A}_\theta^{-T}(i)\underline{w}_\theta(i) \qquad (53)\\ \underline{w}_\theta(i+1) &= \underline{A}_\theta^{-T}(i)\underline{w}_\theta(i) - \underline{c}_\theta(t_{i+1})d(i+1) \qquad (54)\end{aligned}\right\} \text{for } i = 1,\ldots,N-1$$

with boundary conditions $\underline{w}_\theta(N) = 0$ (55)

and $$\underline{v}_\theta(1) = \underline{K}_\theta(1,1)[\underline{c}_\theta(t_1)d(1) + \underline{w}_\theta(1)] \tag{56}$$

It is also shown in Appendix B that, given $d(i)$ for $i = 1,\ldots,N$, Equations (53)-(56) have a unique solution.

+The superscript "-T" denotes inverse transpose.

The procedure by which the summations of Eq. (48) have been expressed in terms of "state" vectors $\underline{v}_\theta$ is similar to that employed by Baggeroer [5] to transform integral equations into state-differential equations.

The models for $\theta = s$ or n can be combined by defining,

$$\text{for } i = 1, \ldots, N, \quad \underline{v}(i) \triangleq \begin{bmatrix} \underline{v}_s(i) \\ \hline \underline{v}_n(i) \end{bmatrix} \; ; \quad \underline{c}(i) \triangleq \begin{bmatrix} \underline{c}_s(t_i) \\ \hline \underline{c}_n(t_i) \end{bmatrix} \; ;$$

$$\underline{C}(i) \triangleq \underline{c}(i)\underline{c}^T(i); \quad K(1,1) \triangleq \left[\begin{array}{c|c} \underline{K}_s(1,1) & \underline{0} \\ \hline \underline{0} & \underline{K}_n(1,1) \end{array}\right] \; ;$$

$$\underline{D}(i) \triangleq \left[\begin{array}{c|c} D'M^{-1}(i)\underline{I}_s & \underline{0} \\ \hline \underline{0} & -M^{-1}(i)\underline{I}_n \end{array}\right] \quad \text{where } \underline{I}_\theta \text{ is an identity matrix of dimension}$$

L_θ, and by defining, for $i = 1, \ldots, N-1$,

$$\underline{A}_\#(i) \triangleq \left[\begin{array}{c|c} \underline{A}_s(i) & \underline{0} \\ \hline \underline{0} & \underline{A}_n(i) \end{array}\right] \; ; \; \underline{Q}(i) \triangleq \left[\begin{array}{c|c} \underline{Q}_s(i) & \underline{0} \\ \hline \underline{0} & \underline{Q}_n(i) \end{array}\right] \; .$$

Eq. (52) may then be written simply as

$$d(i) = \underline{c}^T(i)\underline{D}(i)\underline{v}(i) \quad \text{for } i = 1, \ldots, N, \tag{57}$$

and (53) and (54) become

$$\left.\begin{aligned} \underline{v}(i+1) &= \underline{A}_\#(i)\underline{v}(i) + \underline{Q}(i)\underline{A}_\#^{-T}(i)\underline{w}(i) \qquad (58) \\ \underline{w}(i+1) &= \underline{A}_\#^{-T}(i)\underline{w}(i) - \underline{c}(i+1)d(i+1) \qquad (59) \end{aligned}\right\} \text{for } i = 1, \ldots, N-1 \; .$$

But $d(i+1)$ can be expressed in terms of $\underline{v}(i+1)$ by using (57), and $\underline{v}(i+1)$ can be obtained from (58), so that Eq. (59) is equivalent to

$$\underline{w}(i+1) = -C(i+1)\underline{D}(i+1)\underline{A}_\#(i)\underline{v}(i) + [\underline{I} - \underline{C}(i+1)\underline{D}(i+1)\underline{Q}(i)]\underline{A}_\#^{-T}\underline{w}(i) \tag{60}$$

$$\text{for } i = 1, \ldots, N-1.$$

Eqs. (58) and (59) can be combined,

$$\begin{bmatrix} \underline{v}(i+1) \\ \hline \underline{w}(i+1) \end{bmatrix} = \underline{A}(i) \begin{bmatrix} \underline{v}(i) \\ \hline \underline{w}(i) \end{bmatrix} \qquad \text{for } i = 1, \ldots, N-1, \tag{61}$$

where, for $i = 1, \ldots, N-1$,

$$\underline{A}(i) \triangleq \left[\begin{array}{c|c} \underline{A}_\#(i) & \underline{Q}(i)\underline{A}_\#^{-T}(i) \\ \hline -\underline{C}(i+1)\underline{D}(i+1)\underline{A}_\#(i) & [\underline{I} - \underline{C}(i+1)\underline{D}(i+1)\underline{Q}(i)]\underline{A}_\#^{-T}(i) \end{array}\right] \tag{62}$$

The boundary conditions (55) and (56) can be written as

$$\underline{w}(N) = 0 \tag{63}$$

and

$$[\underline{I} - \underline{K}(1,1)\underline{C}(1)\underline{D}(1)]\underline{v}(1) - \underline{K}(1,1)\underline{w}(1) = 0, \tag{64}$$

where we have substituted for d(1) from Eq. (57).

From the linear recursion relation of (61) it follows that

$$\begin{bmatrix} \underline{v}(N) \\ \hline \underline{w}(N) \end{bmatrix} = \underline{\Lambda} \begin{bmatrix} \underline{v}(1) \\ \hline \underline{w}(1) \end{bmatrix} , \tag{65}$$

where $\underline{\Lambda} \triangleq \underline{A}(N-1)\underline{A}(N-2)\cdots\underline{A}(1)$. If we partition $\underline{\Lambda}$ (66)

into four square matrices,

$$\underline{\Lambda} = \left[\begin{array}{c|c} \underline{\Lambda}_{vv} & \underline{\Lambda}_{vw} \\ \hline \underline{\Lambda}_{wv} & \underline{\Lambda}_{ww} \end{array}\right] , \tag{67}$$

then the boundary conditions (63) and (64) can be combined;

$$\left[\begin{array}{c|c} \underline{\Lambda}_{wv} & \underline{\Lambda}_{ww} \\ \hline \underline{I} - \underline{K}(1,1)\underline{C}(1)\underline{D}(1) & -\underline{K}(1,1) \end{array}\right] \begin{bmatrix} \underline{v}(1) \\ \hline \underline{w}(1) \end{bmatrix} = 0. \tag{68}$$

If we define

$$\underline{q} \triangleq \begin{bmatrix} \underline{v}(1) \\ \hline \underline{w}(1) \end{bmatrix} ; \quad \underline{\mathcal{L}}(D) \triangleq \left[\begin{array}{c|c} \underline{\Lambda}_{wv} & \underline{\Lambda}_{ww} \\ \hline \underline{I} - \underline{K}(1,1)\underline{C}(1)\underline{D}(1) & -\underline{K}(1,1) \end{array}\right] , \tag{69,70}$$

then (68) reduces to

$$\underline{\mathcal{L}}(D)\underline{q} = 0. \tag{71}$$

Each step that we have made toward arriving at Eq. (71) has involved only substitution, and is therefore reversible; so that Eq. (71) is, whenever $D \neq 1$, completely equivalent to Eqs. (48) and (36). If we can find a value of $D \neq 1$ for which $\det \underline{\mathcal{L}}(D) = 0$, then we can find a solution $\underline{q}$ of (71); and if we define d(i) from (57) using (61) and (69) to compute $\underline{v}(i)$, then d(i) will solve (48) (for the same value of D). Note that $\underline{\mathcal{L}}(D)$ is a square matrix of dimension $2L_s + 2L_n$.

Given a particular value of $D \neq 1$ for which Eq. (71) and (36) have solutions, Eqs. (53),(54),(55),(56),(57),(61), and (69) describe a linear, invertible, one-to-one transformation between solutions $\underline{q}$ of (71) and solutions $\underline{d}$ of (36). Hence if (71) has k linearly independent solutions $\underline{q}$ for some value of D, then (36) has k linearly independent solutions $\underline{d}$ for the same value of D.

If $\underline{K}(1,1)$ is nonsingular, then (64) implies

$$\underline{w}(1) = [\underline{K}^{-1}(1,1) - \underline{C}(1)\underline{D}(1)]\underline{v}(1), \tag{72}$$

so that (71) is equivalent to

$$\underline{\mathcal{L}}_{\#}(D)\underline{q}_{\#} = 0, \tag{73}$$

where

$$\underline{q}_{\#} \triangleq \underline{v}(1); \quad \underline{\mathcal{L}}_{\#}(D) \triangleq \underline{\Lambda}_{wv} + \underline{\Lambda}_{ww}[\underline{K}^{-1}(1,1) - \underline{C}(1)\underline{D}(1)] \tag{74,75}$$

The problem has been reduced to one of finding the largest zeros of $\det \underline{\mathcal{L}}(D)$ [or $\det \underline{\mathcal{L}}_{\#}(D)$]. For given values of D, $\underline{\Lambda}$ can be computed and used to determine $\det \underline{\mathcal{L}}(D)$; thus the zeros can be found by standard computer techniques. When a zero is found, initial conditions solving Eq. (68) can be obtained, $\underline{v}(i)$ and $\underline{w}(i)$ can be computed recursively by using Eq. (61), and the eigenvector $\underline{d}$ can be obtained through use of Eq. (52).

If the signal, background noise, and the measurement noise are stationary random processes, then $\underline{A}(i)$ is the same for all i, and

$$\underline{\Lambda} = \underline{A}^{N} \text{ where } \underline{A} \triangleq \underline{A}(i).$$

In particular, if $N = 2^n$, then $\underline{\Lambda}$ can be obtained by n matrix-squaring operations. (eg. $N = 128 = 2^7$).

VI. CONCLUSIONS AND APPLICATIONS

We have shown that the matched filter concept can be successfully applied to the problem of detecting zero-mean Gaussian signals in Gaussian noise. For cases in which the signal and noise processes admit state variable models, we have developed an efficient, easily-implemented technique for finding the optimum set of digital matched filters. Approximations to the continuous problem (i.e. finding the best set of continuous filters) may be obtained by making N large. Just as in the Kalman-Bucy filtering problem, we found it necessary to assume that the observed process contained additive white measurement noise. In order to discover what happens in a specific problem when there is no such white noise, we may determine the optimum filter set for successively smaller values of M(i).

In our treatment the use of a set of K matched filters was taken as an alternative to the direct (and difficult) computation of the likelihood-ratio. Although our technique may yield a somewhat larger Bayes risk (particularly if K is very small) its inherent simplicity may make it more practical, or even necessary, for many real problems. Since the matched filter basically involves a correlation process, many existing devices (optical, mechanical, etc.) can

be used for its implementation. Another advantage over the direct computation of the likelihood ratio is that the transformation of the data achieved by the use of matched filters is quite simple (and linear), and even if the statistical assumptions for which the filters were designed are incorrect or only approximately correct, the reduced-data vector $\underline{y}$ will still probably prove useful for signal detection purposes.

The assumption that the signal is zero mean is somewhat restrictive, but there is a large class of problems which fit this assumption, particularly problems of "passive listening" and non-coherent reception. The remark has been made previously that the independent variable "t" could represent position in space. If this is so the quantities $t_1, t_2, \ldots, t_N$ represent not sampling instants, but receiving-element positions, s(t) is the spatial random process impressed upon the receiving elements by the signal, n(t) is that due to background noise, and the m_i's represent receiver noise.

APPENDIX A

We must show that, for any KxN matrix B,

$$J^*[\underline{B}^*\underline{z}] \leq J^*[\underline{B}\underline{z}]$$

The following lemma is essential.

Lemma. Let $\underline{\alpha}$ be as defined in Section II, and let $\underline{\Gamma}$ be a KxN matrix such that, with $\underline{\gamma} \triangleq \underline{\Gamma}\underline{\alpha}$, the following conditions are satisfied:

(i) $E[\underline{\gamma}\underline{\gamma}^T/H_0] = \underline{I}$, the KxK identity matrix

(ii) $E[\underline{\gamma}\underline{\gamma}^T/H_1] = \underline{\Sigma}$, a KxK diagonal matrix having ordered diagonal elements $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_K$.

Then $J^*[\underline{B}^*\underline{z}] \leq J^*[\underline{\Gamma}\underline{\alpha}]$.

Proof of Lemma: Suppose $\underline{\Gamma}$ satisfies the conditions of the lemma. Then

$$\underline{I} = E[\underline{\gamma}\underline{\gamma}^T/H_0] = \underline{\Gamma}\, E[\underline{\alpha}\underline{\alpha}^T/H_0]\,\underline{\Gamma}^T = \underline{\Gamma}\underline{\Gamma}^T.$$

Also, note that σ_j is the $j^{\underline{th}}$ largest eigenvalue of $\underline{\Sigma}$, and D_j is the $j^{\underline{th}}$ largest eigenvalue of the diagonal matrix

$$E[\underline{\alpha}\underline{\alpha}^T/H_1].$$

But $\underline{\Sigma} = E[\underline{\gamma}\underline{\gamma}^T/H_1] = \underline{\Gamma}\, E[\underline{\alpha}\underline{\alpha}^T/H_1]\,\underline{\Gamma}^T$, so since $\underline{\Gamma}\underline{\Gamma}^T = \underline{I}$ we can use the Poincare' Separation Theorem (see [7], p. 52) to obtain

$$D_{N-j+1} \le \sigma_j \le D_j \quad \text{for } j = 1, \ldots, K.$$

However, all the D_i's are ≥ 1, so

$$1 \le \sigma_j \le D_j \quad \text{for } j = 1, \ldots, K.$$

Now the optimum decision function $\mathcal{D}^*_{\underline{Y}}$ for $\underline{Y}$ is simply

$$\mathcal{D}^*_{\underline{Y}}(\underline{y}) = \begin{cases} H_1 \text{ whenever } \sum_{i=1}^{K} (1-1/\sigma_i) y_i^2 > \lambda \\ H_0 \text{ otherwise} \end{cases}$$

where $\lambda \triangleq 2 \log \eta + \sum_{i=1}^{K} \log \sigma_i$. Thus

$$J^*[\Gamma\alpha] = J^*[\underline{Y}] = P \cdot C_m \cdot J_1^*[\underline{Y}] + (1-P) \cdot C_f \cdot J_2^*[\underline{Y}]$$

where $J_1^*[\underline{Y}] \triangleq \text{Prob}\{ \mathcal{D}^*_{\underline{Y}}(\underline{y}) = H_0 \text{ when } H_1 \text{ is true}\}$

and $J_2^*[\underline{Y}] \triangleq \text{Prob}\{ \mathcal{D}^*_{\underline{Y}}(\underline{y}) = H_1 \text{ when } H_0 \text{ is true}\}$.

Now suppose we use the same decision function on $\underline{y}^* = \underline{B}^* \underline{z}$. The resulting risk, denoted $J^+[\underline{Y}^*]$, will be

$$J^+[\underline{Y}^*] = P \cdot C_m \cdot J_1^+[\underline{Y}^*] + (1-P) \cdot C_f \cdot J_2^+[\underline{Y}^*]$$

where $J_1^+[\underline{Y}^*] = \text{Prob}\{\mathcal{D}^*_{\underline{Y}}(\underline{y}^*) = H_0 \text{ when } H_1 \text{ is true}\}$

and $J_2^+[\underline{Y}^*] = \text{Prob}\{\mathcal{D}^*_{\underline{Y}}(\underline{y}^*) = H_1 \text{ when } H_0 \text{ is true}\}$.

But $\underline{y}^*$ and $\underline{y}$ have the same statistics under hypothesis H_0, (since $E[\underline{y}^* \underline{y}^{*T}/H_0] = E[\underline{y}\underline{y}^T/H_0] = I$), so $J_1^+[\underline{Y}^*] = J_1^*[\underline{Y}]$.

Also, under hypothesis H_1, $\underline{y}^*$ has the same statistics as

$$\underline{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} \triangleq \begin{bmatrix} (D_1/\sigma_1) & & 0 \\ & (D_2/\sigma_2) & \\ & & \ddots \\ 0 & & (D_K/\sigma_K) \end{bmatrix} \underline{y},$$

so

$$J_{\ell}^{+}[\underline{y}^*] = \text{Prob}\{\sum_{i=1}^{K}(1-1/\sigma_i)\,a_i^2 < \lambda \text{ when } H_1 \text{ is true}\}$$

$$= \text{Prob}\{\sum_{i=1}^{K}(1-1/\sigma_i)(D_i/\sigma_i)^2(\underline{y})_i^2 < \lambda \text{ when } H_1 \text{ is true}\}$$

$$\leq \text{Prob}\{\sum_{i=1}^{K}(1-1/\sigma_i)(\underline{y})_i^2 < \lambda \text{ when } H_1 \text{ is true}\}$$

since $D_i/\sigma_i \geq 1$.

But the last quantity is simply $J_{\ell}^{*}[\underline{y}]$ so

$$J_{\ell}^{+}[\underline{y}^*] \leq J_{\ell}^{*}[\underline{y}], \text{ so}$$

$$J^{+}[\underline{y}^*] \leq J^*[\underline{y}].$$

However, $J^*[\underline{y}^*]$ is the lowest possible risk that can be attained using $\underline{y}^*$, so

$$J^*[\underline{y}^*] \leq J^{+}[\underline{y}^*] \leq J^*[\underline{y}]$$

i.e. $J^*[\underline{B}^*\underline{z}] \leq J^*[\underline{\Gamma}\underline{\alpha}]$ Q.E.D.

In view of this lemma, the theorem will be proved if we can show that, given any KxN $\underline{B}$, there exists a $\underline{\Gamma}$ satisfying the conditions of the lemma, such that

$$J^*[\underline{\Gamma}\underline{\alpha}] \leq J^*[\underline{B}\,\underline{z}].$$

If one of the rows of $\underline{B}$ can be expressed as a linear combination of the others, then the Bayes risk $J^*[\underline{y}] = J^*[\underline{B}\,\underline{z}]$ will certainly not be increased if this row is replaced by a row vector lying outside the space spanned by the other row vectors, since this adds new information. Thus there always exists a KxN matrix $\underline{B}_{\ell}$ having linearly independent rows, such that

$$J^*[\underline{B}_{\ell}\,\underline{z}] \leq J^*[\underline{B}\,\underline{z}].$$

Now since the rows of $\underline{B}_{\ell}$ are linearly independent, $\underline{y}_{\ell} \triangleq \underline{B}_{\ell}\underline{z}$ has positive definite covariance matrices (this fact is easily proven). Then from the results of Section II we know that there is an invertible matrix transformation

$$\underline{y} = \underline{T}_{\ell}\underline{y}_{\ell}$$

such that $E[\underline{y}\underline{y}^T/H_0] = I$, and $E[\underline{y}\underline{y}^T/H_1]$ is diagonal with ordered elements.

But $\quad J^*[\underline{y}_\ell] = J^*[\underline{T}_\ell \underline{y}_\ell]$

since the Bayes risk is not affected by invertible transformations of the "test statistic" $\underline{y}_\ell$. Define $\underline{\Gamma} = \underline{T}_\ell \underline{B}_\ell$. Clearly $\underline{\Gamma}$ satisfies the conditions of the lemma since $\underline{y} = \underline{\Gamma}\underline{z}$ has the required properties. Also

$$J^*[\underline{\Gamma}\underline{z}] = J^*[\underline{T}_\ell \underline{y}_\ell] = J^*[\underline{B}_\ell \underline{z}] \leq J^*[\underline{B}\underline{z}].$$

Q.E.D.

APPENDIX B

Proof of Eqs. (53), (54), (55), and (56).

Multiple application of Eqs. (45) and (46) yields the result that, for $i = 1, \ldots, N-1$ and $j = 1, \ldots, N$,

$$\underline{K}_\theta(i+1,j) = \begin{cases} \underline{A}_\theta(i)\underline{K}_\theta(i,j) & \text{for } j \leq i \\ \underline{A}_\theta(i)\underline{K}_\theta(i,j) + \underline{Q}_\theta(i)\underline{A}_\theta^T(i+1)\underline{A}_\theta^T(i+2)\ldots\underline{A}_\theta^T(j-1) & \text{for } j > i. \end{cases}$$

Since $\underline{v}_\theta(i) \triangleq \sum_{i=1}^{N} \underline{K}_\theta(i,j)\underline{c}_\theta(t_j)d(j)$ for $i = 1, \ldots, N$, it follows that

$$\underline{v}_\theta(i+1) = \sum_{j=1}^{N} \underline{K}_\theta(i+1,j)\underline{c}_\theta(t_j)d(j) \quad \text{for } i = 1, \ldots, N-1$$

$$= \sum_{j=1}^{N} \underline{A}_\theta(i)\underline{K}_\theta(i,j)\underline{c}_\theta(t_j)d(j) + \sum_{j+i+i}^{N} \underline{Q}_\theta(i)\underline{A}_\theta^T(i+1)\underline{A}_\theta^T(i+2)\ldots\underline{A}_\theta^T(j-1)$$

or

$$\underline{v}_\theta(i+1) = \underline{A}_\theta(i)\underline{v}_\theta(i) + \underline{Q}_\theta(i)\underline{A}_\theta^{-T}(i)\underline{w}_\theta(i) \quad \text{for } i = 1, \ldots, N-1. \tag{53}$$

where

$$\underline{w}_\theta(i) \triangleq \sum_{j=i+1}^{N} \underline{A}_\theta^T(i)\underline{A}_\theta^T(i+1)\ldots\underline{A}_\theta^T(j-1)\underline{c}_\theta(t_j)d(j) \quad \text{for } i = 1, \ldots, N-1.$$

If we define $\underline{w}_\theta(N) \triangleq 0$, then for $i = 1, \ldots, n-1$, (55)

$$\underline{w}_\theta(i+1) = \underline{A}_\theta^{-T}(i)\underline{w}_\theta(i) - \underline{c}_\theta(t_{i+1})d(i+1). \tag{54}$$

The $\underline{w}_\theta(i)$'s are uniquely determined by (54) and the boundary condition $\underline{w}_\theta(N) = 0$, provided the d(i)'s are given.

Repeated use of (45) and (46) gives

$$\underline{K}_\theta(i,j) = \underline{K}_\theta(1,1)\underline{A}_\theta^T(2)\ldots\underline{A}_\theta^T(j-1) \quad \text{for } 1 < j \leq N,$$

so that

$$\underline{v}_\theta(1) = \underline{K}_\theta(1,1)[\underline{c}_\theta(t_1)d(1) + \underline{w}_\theta(1)], \tag{56}$$

thus $\underline{v}_\theta(1)$ is determined and (53) has a unique solution.

REFERENCES

[1] H. L. Van Trees, Detection, Estimation, and Modulation Theory, New York: John Wiley and Sons, 1968.

[2] F. C. Schweppe, "Evaluation of Likelihood Functions for Gaussian Signals," IEEE Trans. on Information Theory, Vol. IT-11, No. 1, January, 1965.

[3] D. G. Lainiotis, "Optimal Feature Extraction in Pattern Recognition," Inter. Information Theory Symposium Abstracts, San Remo, Italy, September 11-15, 1967.

[4] T. T. Kadota and L. A. Shepp, "On the Best Set of Linear Observables for Discriminating Two Gaussian Signals," IEEE Trans. on Information Theory, Vol. IT-13, No. 2, April, 1967.

[5] S. Kullback, Information Theory and Statistics, New York: John Wiley and Sons, 1959.

[6] A. B. Baggeroer, "A State Variable Approach to the Solution of Fredholm Integral Equations," Technical Report No. 459, M.I.T. Res. Lab. of Electronics, Cambridge, Mass., November, 1967.

[7] C. R. Rao, Linear Statistical Inference and Its Applications, New York: John Wiley and Sons, 1965.

UNCLASSIFIED
Security Classification

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) The University of Texas at Austin Electronics Research Center Austin, Texas 78712	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

DIGITAL MATCHED FILTERS FOR DETECTING GAUSSIAN SIGNALS IN GAUSSIAN NOISE

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

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

Terry L. Henderson Demetrios G. Lainiotis

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
12 September 1968	23	7

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
AF-AFOSR-67-766A b. PROJECT NO. 4751 c. 6144501F d. 681305	JSEP, Technical Memorandum No. 3 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR 68-1919

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH, OTHER	JSEP through AF Office of Scientific Research (SREE) 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

The use of a set of digital matched filters is presented as an alternative to direct computation of the likelihood-ratio, for the problem of detecting a random signal in random noise. It is assumed that a random process composed of Gaussian background noise and (with probability P) a zero-mean Gaussian signal is sampled at N instants, the samples being corrupted by additive Gaussian measurement noise. The samples are processed by $K << N$ digital correlation filters which are structured so that the signal can be detected with minimum Bayes risk. The optimum filters are shown to be matched to the most relevant components of the simultaneously orthogonal expansion of the set of sampled data. State variable techniques are used to find a very practical method for determining the optimum filter structures.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
DETECTION						
FEATURE EXTRACTION						
FEATURE SELECTION						
FILTERS						
MATCHED FILTERS						