UNLIMITED

AD-A168 180

2

RSRE
MEMORANDUM No. 3896

ROYAL SIGNALS & RADAR ESTABLISHMENT

SPECTRAL ANALYSIS FOR RADAR USING LINEAR PREDICTIVE FILTERING

Author: J F Cross

PROCUREMENT EXECUTIVE,
MINISTRY OF DEFENCE,
RSRE MALVERN,
WORCS.

RSRE MEMORANDUM No. 3896

UNLIMITED

ROYAL SIGNALS AND RADAR ESTABLISHMENT

Memorandum 3896

Title: SPECTRAL ANALYSIS FOR RADAR USING LINEAR PREDICTIVE FILTERING

Author: J F Cross

Date: November 1985

SUMMARY

The main ideas underlying the use of linear predictive filtering as a spectrum analysis tool are developed in this paper. Following this, detection and resolution performance of the method compared with that of the DFT are examined by the presentation of the results of a Monte Carlo experimentation.

RSRE MEMORANDUM 3896

SPECTRAL ANALYSIS FOR RADAR USING LINEAR PREDICTIVE FILTERING

J F Cross

CONTENTS

I INTRODUCTION

Modern airborne radars rely on spectral analysis for a variety of functions including detection in clutter and velocity measurement. Also fine detail in target spectral returns may give clues to target identity. Hence there is great interest in spectral analysis techniques with improved detection or resolution performance. An obvious candidate for the spectral analysis is the well known and well proved Discrete Fourier Transform (DFT). The fast implementation of this transform (FFT) is computationally efficient in terms of the number of non-trivial arithmetic operations it must perform. Moreover, in its unwindowed form the DFT provides a matched filter for sinusoids in a white noise background and is therefore required for optimum detection. However in radar systems the DFT seldom, if ever, can be used in its unwindowed form. The existence of very large clutter signals would limit the detection of small targets with a different Doppler frequency unless steps are taken to avoid the high spectral sidelobe levels from the clutter. A suitable window function is used to control these leakage levels to an acceptable level. This however must be regarded as a compromise because the windowed DFT is not a matched filter and losses in detectability and resolution occur.

Over the last decade or so alternative techniques for spectrum estimation have been researched and published. In fact, there seem to be a large variety of techniques based on the idea of Linear Predictive Filtering. The techniques are reported to overcome the main difficulties associated with the DFT of a time limited series, namely essentially zero or repetitive extension of the known data. Removal of these unrealistic constraints is claimed to improve resolution significantly. It is the promise of improved resolution which has prompted this short study. Although vast amounts of literature on these new techniques now exist, it was found difficult to find sufficient quantitive information which addressed the sort of questions which the radar engineer requires to be answered before abandoning the faithful FFT. This paper makes a start at answering some of the questions for one particular technique - the Covariance Method of Linear Predictive Filtering (CMLPF). Although at first sight this may seem restrictive, initial experience with this method and with Burg's entropy method (MEM) and the maximum likelihood method (MLM) indicated that it had potentially higher resolution than MLM and did not suffer some of the peculiarities of MEM. Also, because the techniques are similarly based, their performance will be similar in a large variety of situations. It was felt therefore that CMLPF was representative of the new techniques and would adequately indicate the trends.

II BASIS OF THE LINEAR PREDICTIVE FILTERING METHOD OF SPECTRAL ESTIMATION

Figure 1 is an example of a tapped delay line filter with an impulse given by the weights, a,

$$a_0 - a_4$$

and delays, T,:

$$h(z) = a_0 + a_1 z^{-1T} + a_2 z^{-2T} + a_3 z^{-3T} + a_4 z^{-4T}$$

where

$$z^{-nT}$$

is the z-transform representing a unit impulse delayed by nt.

The frequency response of the filter is the Fourier transform of the impulse response which can most conveniently be found by replacing z by

$$e^{jw}$$

Suppose the waveform input to the filter is given. Then if we can find the weights, a, which produce "white noise" at the output of the filter we know that the frequency response of the filter is related to the spectrum of the input. More specifically the power density spectrum of the output waveform is the product of the power density spectrum of the input waveform and the filter spectral response. If the output is "white" then the estimate of the power spectrum of the input is

$$p(w) = \frac{E}{|\text{Fourier transf'm of filter impulse resp'se}|^2} \tag{2.1}$$

$$p(w) = \frac{E}{\left|a_0 + a_1 e^{-jwT} + a_2 e^{-2jwT} + a_3 e^{-3jwT} + a_4 e^{-4jwT}\right|^2} \qquad (2.2)$$

where E is the energy of the output sequence.

This gives us a method of estimating the spectrum of a signal which is alternative to the DFT.

The link between the tapped delay line filter of figure 1 and linear predictive filtering is made when the weights are selected to minimize the output energy. In order to constrain the minimisation, we choose to fix $a_0 = 1$. (Note an unconstrained minimisation would simply result in all $a = 0$.) We now redraw figure 1 as figure 2. Here we can see that a current input sample is added to a filtered version of the previous four input samples to form the output. Clearly, if $y = -x$ then $e = 0$ and we can think of the filter $a_1 - a_4$ as predicting the next sample. Perfect prediciton is not possible in the presence of noise and we settle for minimum mean square difference between predicted and actual samples over the data set available.

Obviously it is only possible to perform a predictive process if there is some correlation between successive data samples. This occurs if the signal input is the sum of a (finite) number of sinusoidal waveforms plus noise. The correlation between data points is described by a covariance which can be statistically estimated from the data samples available.

The limitations of this method when the data sequence is short arise because a short data set limits the accuracy to which the weights can be estimated in the presence of noise. Also the number of weights is limited by the data length (usually to somewhat less than half the number of data samples) and this may be less than the number required to obtain the desired detail in the spectral estimate.

The remainder of this paper sets out to show how the weights, a, can be estimated from a short sample of data and how the limitations of the method affect radar performance when compared with the DFT.

III ESTIMATING THE WEIGHTS

This section shows how the weights of a of the linear predictive filter can be found from a covariance matrix which is itself estimated from the data. The theoretical outline given below is simple, aiming only to give a general understanding of the process.

Referring to figure 2, at any time instant, j, the output e is a linear combination of the input samples x.

$$e_j = x_j + \sum_{i=1}^{4} x_{j-i}a_i \qquad (3.3)$$

Thus if we have a set of samples:

$$x_0 \text{ to } x_T$$

we can write the filtering operation in matrix form as:

$$\begin{bmatrix} x_4 & x_3 & x_2 & x_1 & x_0 \\ x_5 & x_4 & x_3 & x_2 & x_1 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ x_T & \cdots\cdots & & & x_{T-4} \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} e_4 \\ e_5 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ e_T \end{bmatrix}$$

or

$$X.\underline{a} = \underline{e} \qquad (3.4)$$

The optimum prediction filter is then defined as that filter which minimises the output energy E. Now E is simply given by:

$$E = \sum_{i=4}^{T} e_i{}^{*}\, e_i = \underline{e}^H.\underline{e}$$

where * means complex conjugate and H is the Hermitian operator.

Differentiating with respect to a and equatiang to zero (for minimum) results in the so called Normal Equations which give the solutions we require:

$$\begin{bmatrix} x_4 & x_5 & \cdots & x_T \\ x_3 & & \cdots & \\ x_2 & & \cdots & \\ x_1 & & \cdots & \\ x_0 & & \cdots & \end{bmatrix} \begin{bmatrix} x_4 & x_3 & x_2 & x_1 & x_0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ x_T & \cdot & \cdot & \cdot & x_{T-4} \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} E_{min} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

or more compactly

$$R \cdot \underline{a} = \begin{bmatrix} E(\min) \\ \cdots\cdots \\ 0 \end{bmatrix} \tag{3.5}$$

where the matrix R, which has the form of a covariance matrix, is defined as:

$$X^H.X$$

And E(min) is the energy of the residue. *Using equation (3.5) the* prediction filter coefficients:

$$a_1 - a_4$$

can be found by direct inversion of the matrix R. (Usually direct inversion is avoided and more suitable methods used).

Strictly speaking, the covariance matrix is a matrix of expected values and the matrix R, as defined above, is just one possible estimate given the limited data set {x}. The difference in the procedure for estimating the covariance matrix account for some of the variety of techniques. However we do not wish to confuse the issue and this method for developing R will be the only one used in this paper.

A few points are worthy of note:

a. Because E(min) > = 0 in equation (3.5) the matrix R is non-negative definite. Indeed in most cases of interest where the signals are corrupted by noise R will be guaranteed positive definite. (In physical terms this means that the estimated power spectrum is positive or zero at all frequencies.)

b. In general R is Hermitian but non-Toeplitz. Toeplitz covariance matrices result from stationary statistics and although many of the processes found in radar systems may be regarded as stationary, with a limited data set we have found that R is significantly non-Toeplitz. It is convenient to regard a short sample of data as converging to stationarity. This implies that an estimate of the covariance matrix formed from such data will converge to the Toeplitz form. It is interesting to note that some methods, for example Burg's MEM, force R to be Toeplitz and this may account for some of the peculiarities of the method.

c. If our signal consists of q complex exponentials (Phase and quadrature sinusoids) in white noise then a filter of order p > = q is required to leave a white noise residue. (By definition white noise is unpredictable!)

d. If our signal is coloured noise then the LPF still minimises the output leaving a white noise residue.

e. R as defined above makes no assumptions about data which is not available.

f. The LPF method does not window the data, that is there is no weighting of the available samples.

IV SPECTRUM ANALYSIS

The last section has shown that given a data set

$$\{x_i : i = 0, T\}$$

then a p'th order filter, a, can be found which minimises the output energy from the filter. In particular, if {x} consists of q complex exponentials in white noise, then if $p \geq q$ the output residue will be white. Because of this property the inverse of the power response of the filter is an estimate of the signal power density spectrum. ie

$$p(f) = e_{min} / | \text{ Fourier transform}(a) \, |^2 \qquad (4.1)$$

Notice that the resulting spectral estimate is a continuous function of frequency. In practice only a sampled version of the continuous transform is produced by forming a DFT on the vector a; the number of spectral samples determining how finely the spectrum is sampled. The required size of the DFT is an interesting question which will be discussed later.

What order of filter is needed for a given data set? The lack of a priori information makes this also a matter of compromise. There are methods for estimating the best filter order from the data ([3], [4]). However we have found them entirely unsatisfactory especially at high signal to noise ratio. Instead we are interested in finding what (if any) compromise is incurred when we fix the filter order at a relatively high value (to give capability on as many sinusoids as possible), for example at one quarter of the number of data samples. This is a much higher relative order than is found in much of the literature.

To summarise, a spectral estimate is created by the following steps

a. Form the p x p covariance matrix estimate (R) from the data.

b. Invert R to solve for E(min) and filter weights a.
(Matrix inversion is avoided at this step).

c. Form the spectral estimate p(w) using equation (4.1) or (2.2).

It is clear from above that the required processing is considerably increased over a straight DFT particularly when it is realised that parts of the processing should be done at high accuracy and dynamic range (eg 36 bit floating point arithmetic.) From a systems engineering point of view then the process must show considerable improvement over the DFT if is is to be a practical alternative. The following sections describe some work which attempts to find where and when the process might be seriously considered.

V STATISTICS OF SPECTRAL NOISE

Of considerable interest to the radar engineer is the behaviour of the spectral estimator when stimulated by "white" receiver noise. It is this situation which determines the system detection threshold (by consideration of false alarms) and therefore ultimately the detection performance of the radar.

In order to find the properties of the spectral estimator with noise input we presented white Gaussian noise to a program running the spectrum analysis. In all cases single precision (36 bit) floating point complex arithmetic was used and in the main 64 complex data samples were used. To some extent predicting the outcome, the statistics were performed on the log power spectrum. Using this, the pdf of the noise spectrum was generated for a wide variety of filter orders.

Figure 3 shows some examples of the resultant pdf's for $p = 4, 8, 16$. It can be seen that they look generally "normal" with a mode which decreases and variance which increases with increasing p. Perhaps of more interest is the cumulative distribution function. Figure 4 plots some examples of the pfa (probability of false alarm). For small values of P these curves exhibit Gaussian behaviour. However for $P > 8$ and small values of pfa the behaviour is non Gaussian. The curves indicate that although the main body of the distribution appears Gaussian it has considerably more area in the tails.

It was mentioned earlier that solutions to the normal equations can be found which avoid the necessity to invert the matrix R. On such method is an iterative method developed by Morf [1]. A byproduct of this method is the production of a backward solution as well as a forward solution. The backward soluton is simply the optimum filter with the data set running backward in time. It is possible therefore to generate a pair of spectra from the data set: one corresponding to the forward predictor and the other to the backward. Observation of the behaviour of these spectra led us to believe that there might be some benefit from averaging the two log power spectra. This was done and the resulting pfa curves shown in figure 5. It can be seen that for low order filters and for high pfa the curves are very similar. However at high order P and low pfa the tails of the distribution seem much better controlled. The agreement with normality (represented by the circled points) now seems very good over the range of parameters measured. These plots were used in subsequent work to estimate suitable thresholds for false alarm control.

VI SINGLE SIGNAL DETECTABILITY

The primary advantage of the new methods is generally one of enhanced resolution. However in the current radar context detectability cannot be ignored. There are many reports which suggest that improved resolution can only be achieved at high signal to noise ratios without being specific as to how high. It is the purpose of this section to try to supply some quantitative answers to questions about the effect of signal to noise ratios.

Unlike the case of noise alone, detectability of a single sinusoid in white noise is affected by the transform size used to sample the continuous spectrum. This is because extremely narrow peaks can be produced in the spectrum (particularly at high signal to noise ratios). Clearly if the sampling interval of the spectrum is too large then peaks can be missed. The approach taken here to decide on how large a transform size is needed was a follows. At each (random) trial the frequency and phase of the sinusoid were chosen randomly and single look detection probabilities were derived for a variety of transform sizes. Throughout the trials the amplitude of the signal was held at unity and the input noise power varied to span a range of different signal to noise ratios.

Before examing the detection results it is worth looking at figure 6. This shows some log-power spectra for a particular case. In this example a single high signal to noise ratio sinusoid was subjected to spectrum analysis by unwindowed FFT, windowed FFT and the LPF techniques under consideration here. There were 64 complex data samples and the same data was used for each technique. The sinusoid had half a cycle in the data set (of size 64). It is not necessary to examine the results in great detail to recognise why the modern methods are appealing. The linear prediction method has a very sharp peak at the correct frequency and shows no sign of leakage. (Only the first 37 points of the 256 point spectrum are shown.) It is also clear why under-sampling of the spectrum could result in complete loss of a signal.

To date the simulations of detectabililty have been limited in number. One particular example is shown in figure 7. The data set consists of 64 complex samples. A fixed prediction filter length of 16 and a fixed thresholds was used (derived from the pfa curves given earlier). The figure corresponds to a transform size of 256 and each figure shows the single look detectability as a function of the ideal signal to noise ratio. (The ideal signal to noise ratio, SNR, here is defined as 64 times the input SNR which would be achievable by perfect coherent integration). The leftmost curve is derived from Meyer and Meyer [2], the middle curve is the result of the averaged forward and backward spectra and the rightmost is that of a single direction predictor. Points to note are:

a. The effective loss in SNR for the averaged process over the optimum linear detector is about 2dB. This is the same order of loss as might be expected from a windowed DFT.

b. The single direction predictor seems to suffer a further loss of about 3dB.

Reducing the transform size to 128 increased the loss of the "averaging" case by 0.3dB and the other by 1.5dB. Increasing the transform size to 512 made no difference to the "averaging case" and .4dB improvement in the other. It can be seen therefore that averaging reduces the sensitivity to transform size, indicating that averaging broadens the spectral peaks.

These results are very encouraging for the averaged version technique. A natural conclusion to draw is that, because the peaks are less sharp in the averaged spectrum, the resolution might be inferior. However it is our experience that the sharpness of the peaks gives little indication of the achievable resolution which is much less that might be expected. Indeed the sharpness of the peaks seems an embarrassment and in a practical system the probability of every obtaining the actual peak is quite small.

The conclusion to be drawn from this section is that the averaging process is beneficial not only because detectability is improved but also because the continuous spectrum need not be sampled as often.

VII *MULTIPLE SIGNAL RESOLUTION*

Finding a metric which adequately describes resolving power on multiple signal detectability is always difficult as it can depend on many parameters. In this paper, we avoid to some extent this complicated problem. The approach taken is once again based on Monte Carlo simulation where the M signals (each of arbitrary phase but constant amplitude) are scattered throughout the spectrum in a statistically uniformly distributed manner. We then compute the probabilities of detecting zero to M targets. When the individual target signal to noise ratios are high enough to give unity detection probability then losses is detectablity can only be attributed to targets merging through poor resolving power.

Basically we wish to compare the performance of windowed DFT with two LPF techniques. Although we have not done exhaustive tests the work done so far has indicated that there is no significant different in the resolution of the two LPF techniques. We therefore focus our attention on the results of the averaging technique in comparison to the (windowed) DFT.

Equal Amplitude Signal Pairs

In this particular experiment we again took 64 data samples and fixed the filter order at 16. For both the processes a 256 point DFT was used to form the spectrum. The thresholds were set to produce a pfa of 7 in 1000 per data sample. Single signal detectability was then checked at an ideal output SNR of 10dB. With these particular parameters the LPF method proved to be a marginally better detector. Taking this difference into account the effective resolutions are roughly equivalent up to 10dB SNR. As the SNR is increased the resolution of the LPF method improves over the DFT. At 30dB SNR the LPF is about two and a half times better and at 50dB SNR about three and a half times better. At this latter SNR the LPF method produced a correct signal pair on 98% of the trials and on the windowed DFT on 93%.

Sixteen Equal Amplitude Signals

The analysis of signal pairs does suggest that the LPF method provides superior resolution. In order to check how this improvement might reflect on the multiple signal situation the same approach was used for sixteen randomly positioned and phased signals. The results shown in figures 8, 9 and 10 indicate that the 10dB S/N ratio the DFT and LPF behave similarly whilst at 20dBs and above the LPF is significantly better than the DFT, the latter not improving with S/N ratio.

Unequal Signal Pairs

One of the main criticisms of the DFT method is the requirement to window the data to reduce spectral leakage to an acceptable level. The LPF method uses no such window and it is therefore worthwhile to observe the effects when a large and small signal coexist. Leaving all other parameters the same two targets, one of 10dB and the other of 50dB SNR were subject to the previously described Monte-Carlo tests with a uniform random distribution of frequencies over the band. If the large target had represented clutter then we are interested in the loss of detectability of the small target through its merging into the clutter on the random occasions that they were close together and not resolved. The results showed that detectability of the small target was reduced to 90% (of the value without the large target) using the windowed DFT and 93% using the LPF method. The DFT method suffered a small increase in false alarms whilst the LPF method enjoyed a small decrease in false alarms. These results are convincing evidence that windowing is not required for the LPF method and also that spurious signal generation is not evident.

VIII CONCLUSIONS

This work, though not exhaustive, has given some useful perspective on the performance of spectral analysis (for pure sinusoids in white noise) using the Covariance Method of linear predictive filtering. In particular it has been found significantly advantageous to take the geometric mean of the power spectra produced from the forward and backward filters. Given this approach, the LPF method can match the detection performance of a windowed DFT for a single sinusoid in white noise. Moreover, in all the work done so far there is no evidence of spurious peak production. That is to say, the false alarm probability does not increase above the noise only case when signals are introduced. Above about 10dB SNR improvements in resolution over a windowed DFT are observable and are small factor improvements rather than order of magnitude.

The main conclusion must be that at high SNR significant improvements in multiple signal detectability can be expected from the LPF method. However this conclusion must be tempered with a full realisation of the rather large computational increase over the DFT methods. It will not be until the full cost of reducing word lengths are understood that the method is likely to be used for real time spectrum analysis.

ACKNOWLEDGMENT

The author wishes to acknowledge two colleagues, namely M M Hayward and Dr I D Longstaff who, through lengthy discussions, made a significant contribution in this work.

REFERENCES

1. M Morf, B Dickinson, T Kailathe, A Vieira. "Efficient Solution of Covariance Equations for Linear Prediction". IEEE Trans Acoust, Speech, Signal Processing, vol ASSP-225, pp 429-433, 1977.

2. Meyer and Meyer. "Radar Target Detection". Academic Press 1973, ISBN 0-12-492850-1.

3. H Akaike. "Fitting Autoregressive Models for Prediction". Ann Inst Statist Math, Vol 21, pp 243-247, 1969.

4. H Akaike. "Statistical Predictor Identification". Ann Inst Statist Math, Vol 22, pp 203-217, 1970.

5. Modern Spectral Analysis Edited by Donals G Childers IEEE Press 1978.

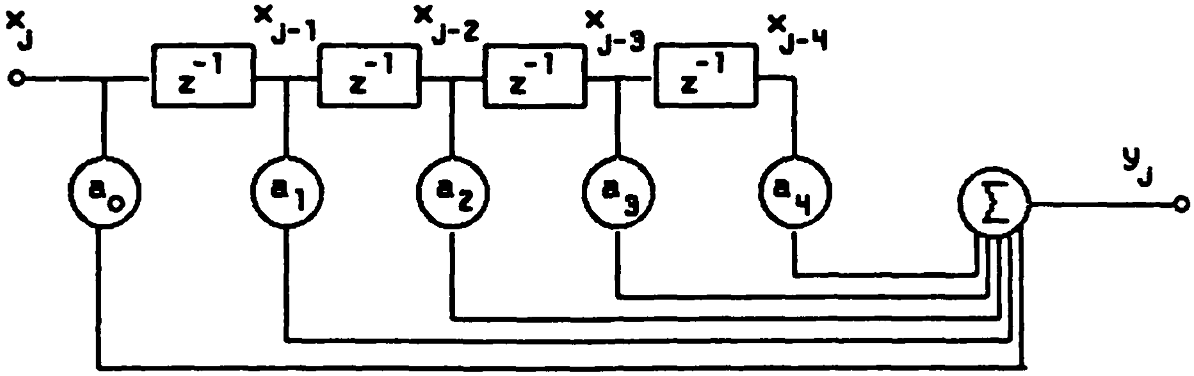

x_j
x_{j-1}
x_{j-2}
x_{j-3}
x_{j-4}
z^{-1}
z^{-1}
z^{-1}
z^{-1}
a_0
a_1
a_2
a_3
a_4
Σ
y_j

FIG 1

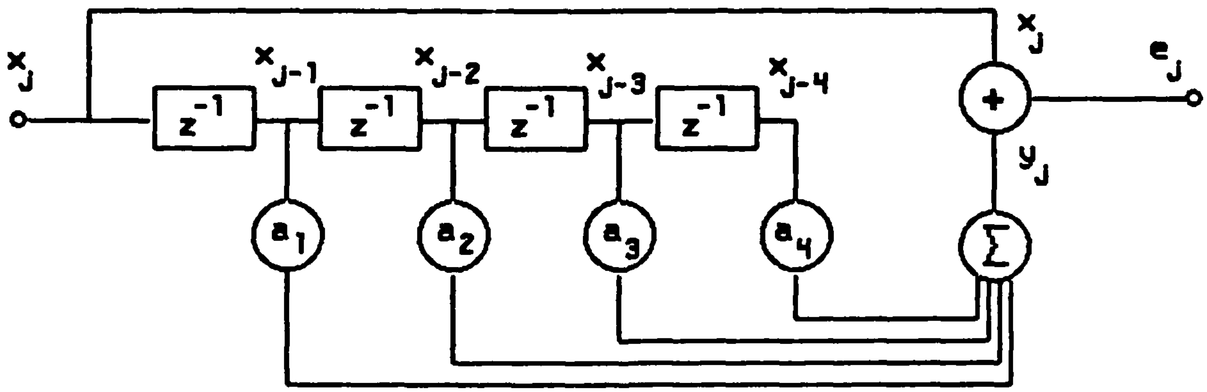

x_j
x_{j-1}
x_{j-2}
x_{j-3}
x_{j-4}
z^{-1}
z^{-1}
z^{-1}
z^{-1}
x_j
e_j
+
y_j
a_1
a_2
a_3
a_4
Σ

FIG 2

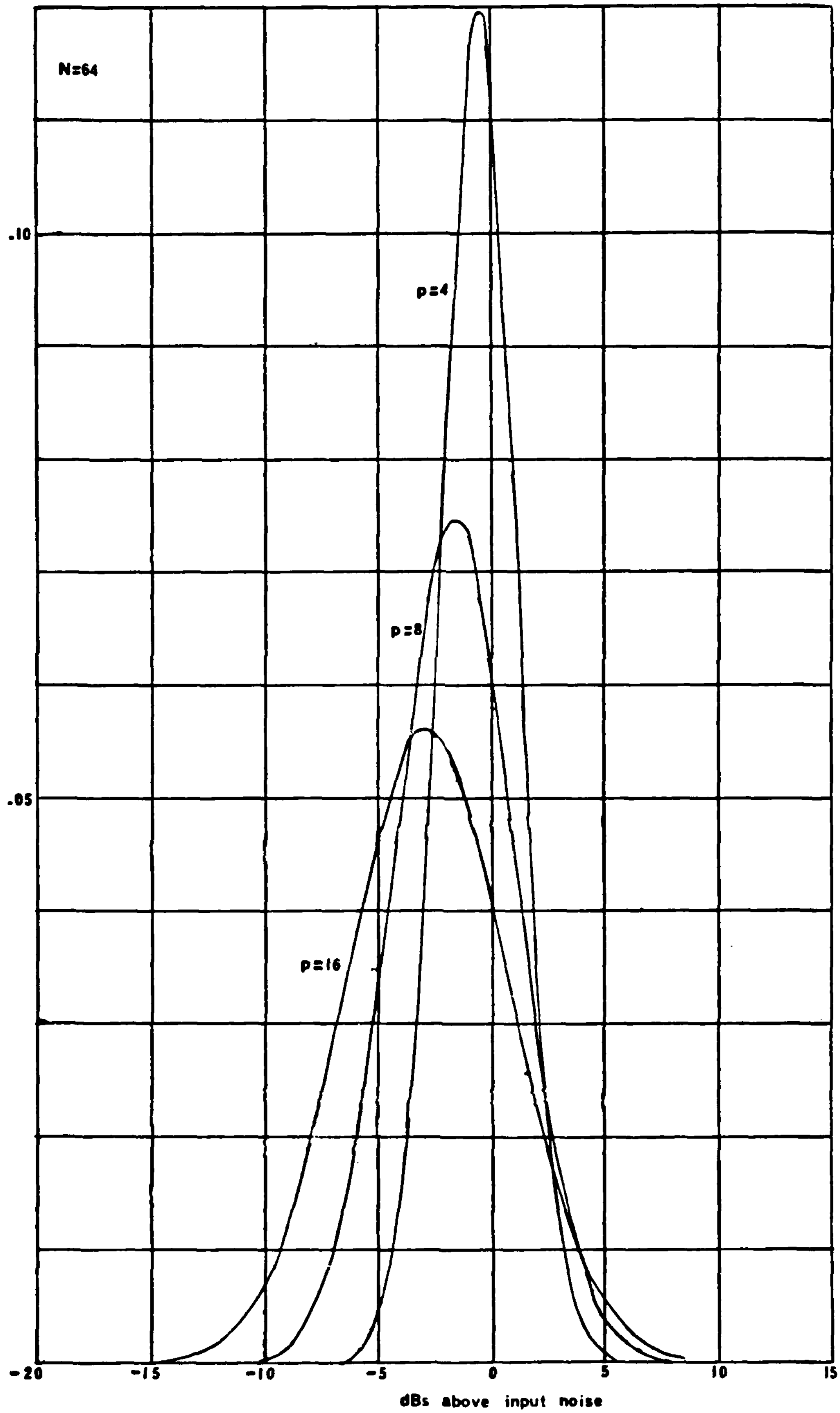
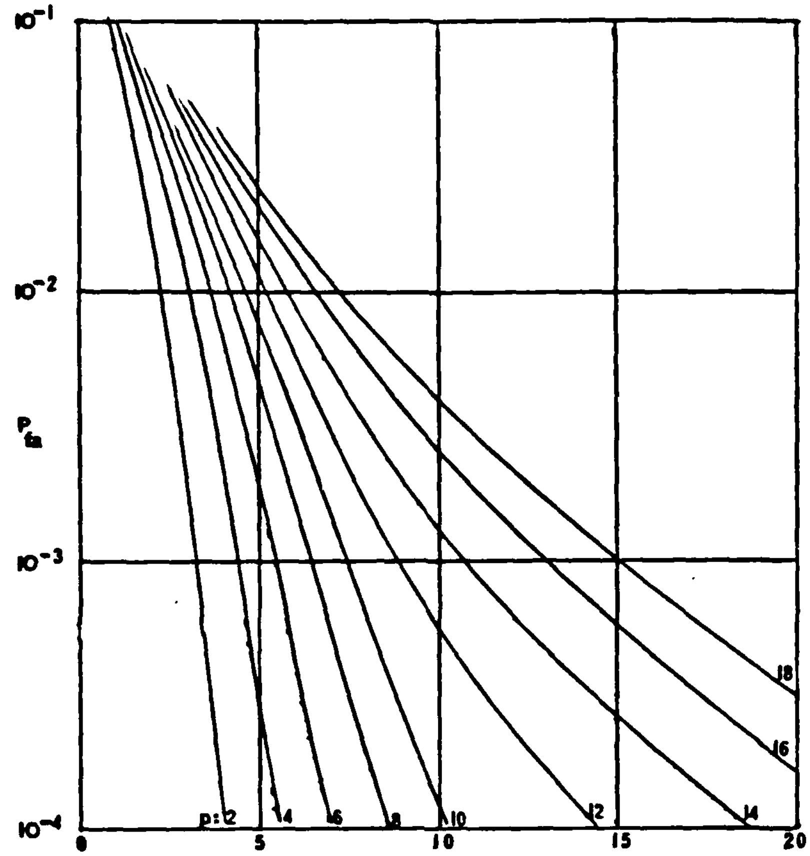
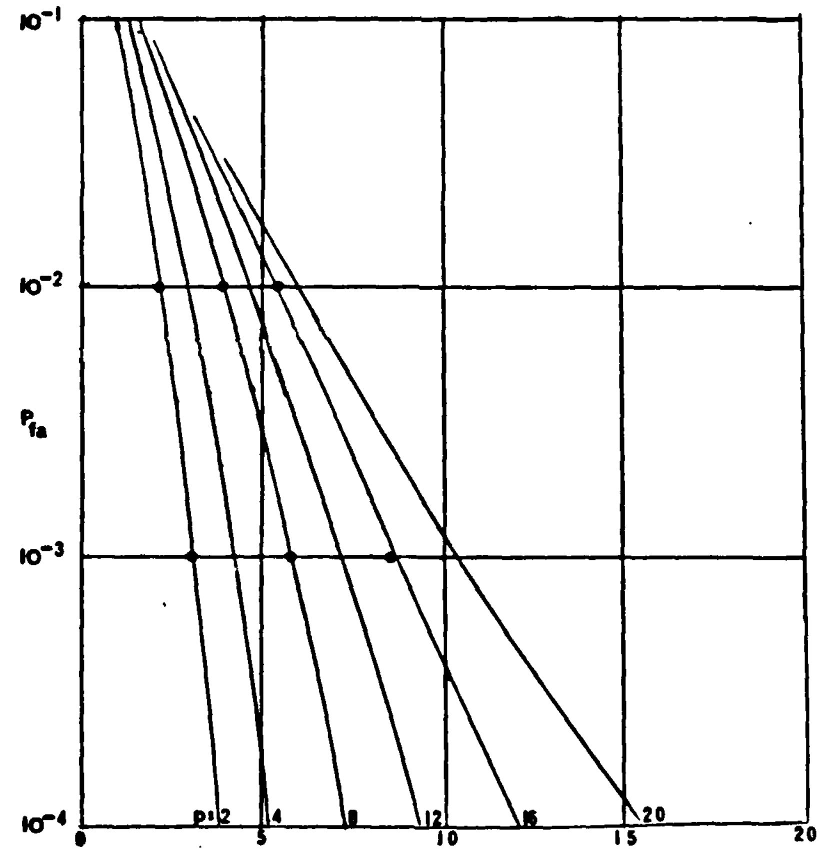


FIG 3



dBs above input noise

FIG 4



dBs above input noise

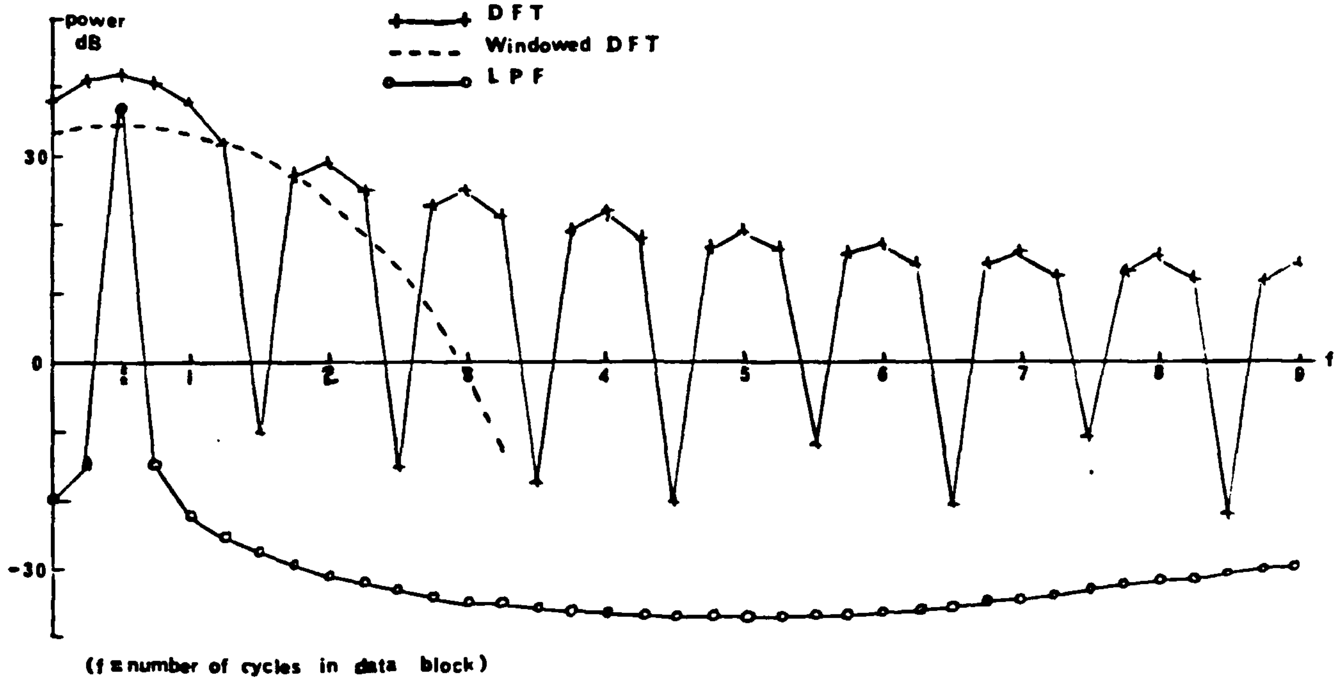

power
dB
DFT
Windowed DFT
LPF
30
0
-30
1
2
3
4
5
6
7
8
9
f
(f = number of cycles in data block)

FIG 6

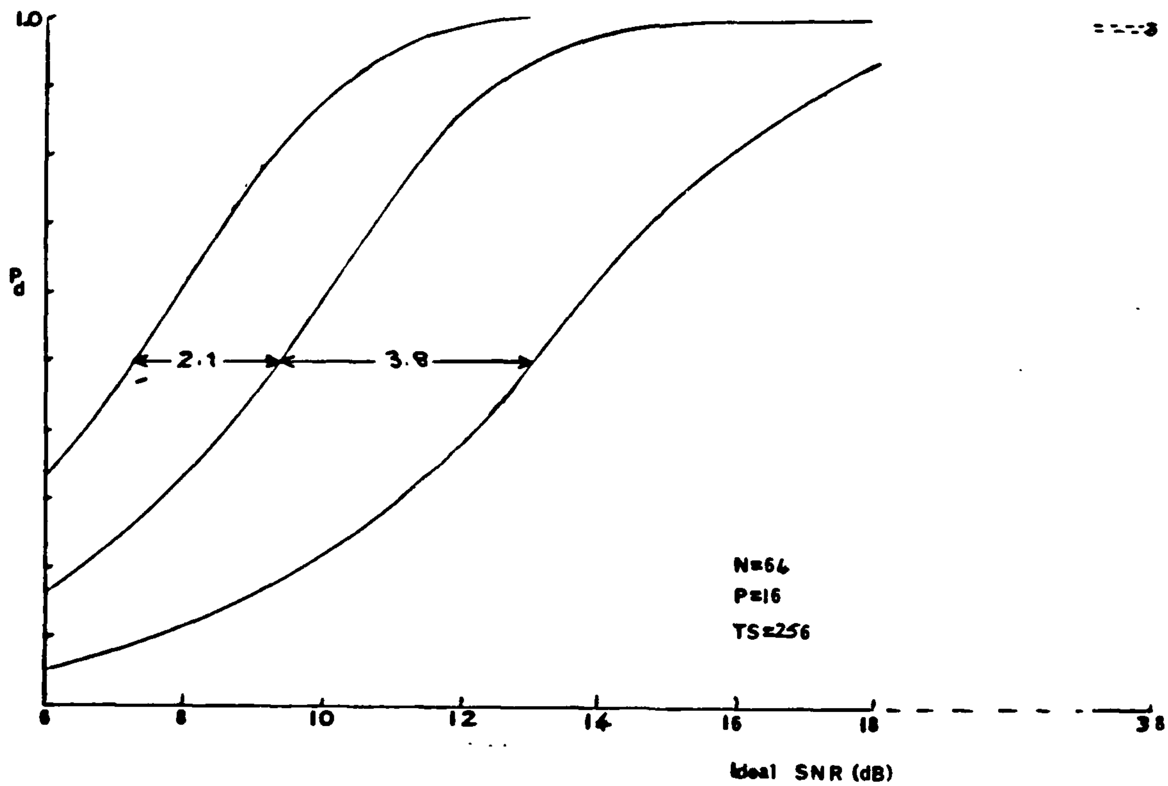
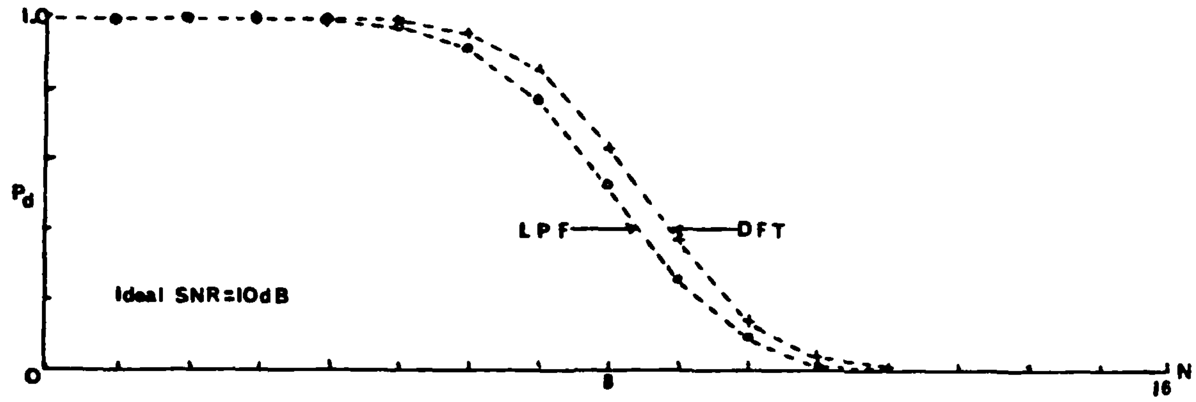
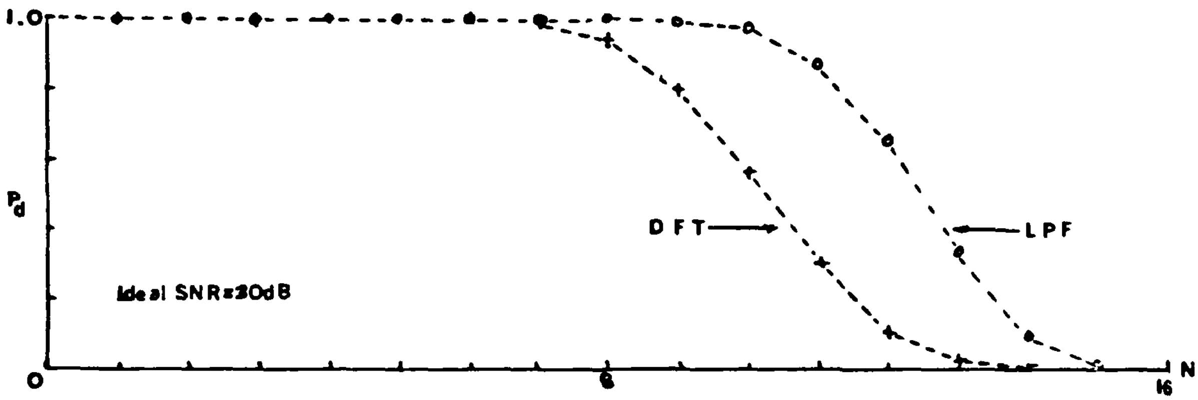

1.0
P_d
2.1
3.8
N=64
P=16
TS=256
6
8
10
12
14
16
18
38
Ideal SNR (dB)

FIG 7



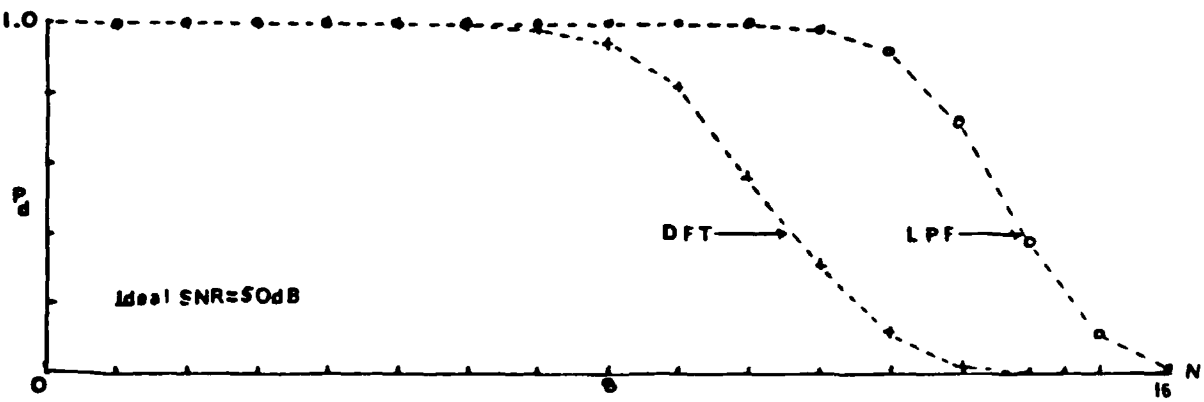
Probability of detecting at least N out of 16

FIG 8



Probability of detecting at least N out of 16

FIG 9



Probability of detecting at least N out of 16

FIG. 10

DOCUMENT CONTROL SHEET

Overall security classification of sheet ....Unclassified...

(As far as possible this sheet should contain only unclassified information. If it is necessary to enter classified information, the box concerned must be marked to indicate the classification eg (R) (C) or (S))

1. DRIC Reference (if known)	2. Originator's Reference	3. Agency Reference	4. Report Security Classification
	Memorandum 3896		U/C

5. Originator's Code (if known)	6. Originator (Corporate Author) Name and Location
	Royal Signals and Radar Establishment
5a. Sponsoring Agency's Code (if known)	**6a. Sponsoring Agency (Contract Authority) Name and Location**

7. Title

Spectral Analysis for Radar Using Linear Predictive Filtering

7a. Title in Foreign Language (in the case of translations)

7b. Presented at (for conference papers) Title, place and date of conference

8. Author 1 Surname, initials	9(a) Author 2	9(b) Authors 3,4...	10. Date	pp. ref.
Cross J F				

11. Contract Number	12. Period	13. Project	14. Other Reference

15. Distribution statement

Unlimited

Descriptors (or keywords)

continue on separate piece of paper

Abstract

The main ideas underlying the use of linear predictive filtering as a spectrum analysis tool are developed in this paper. Following this, detection and resolution performance of the method compared with that of the DFT are examined by the presentation of the results of a Monte Carlo experimentation.