THE INTERACTION BETWEEN ADAPTIVE AND MULTIPLICATIVE PROCESSING

Ian Roebuck

(Admiralty Underwater Weapons Establishment, Portland, Dorset)

Paper presented at the Institute of Acoustics Symposium on "Adaptive Signal Processing" held at the Technological University of Loughborough, 14 June, 1977.

INTRODUCTION.

The fundamental task in signal processing is to take inputs composed of both wanted ("signal) and unwanted data (noise, interference), and to produce an output consisting only of an accurate replica of the signal, uncorrupted by artefacts of the input noise and interference. To achieve this implies knowledge of some characteristic which may be used to discriminate the signal from unwanted data; the nature of this characteristic will determine the most effective means of rejecting unwanted data without suppressing the signal - the optimal processor for that specific situation.

In many applications, however, the information about the environment in which the processor is to work is incomplete. The objective then becomes to produce a processor which will give an adequate performance over a range of possible environments, while not being too sub-optimum in the most likely ones. (An example of such a processor is the Chebyshev shading scheme in sum-square beamforming, which provides guaranteed rejection of signals from an off-beam direction to a pre-set degree, at the expense of the gain against uncorrelated noise). Where the number of degrees of freedom (independent data inputs) is large and the range of permitted environments is small, we can provide preselected near-optimal processors; the concept of adaptive processing arises when the system is small and the range of environments large, so the optimal processor for one permitted environment is grossly sub-optimum for another, and vice versa. The adaptive procedure consists essentially of using an initially chosen processor to discover further information about the environment, and from this determine a more nearly optimal processor with which to discover further information, and hence iteratively converge to the optimal processor for the specific environment being operated in.

The current paper is prompted by the observation that adaptive schemes originated in the field of geophysics, and have exclusively followed the direction dictated by that origin, in that the initial processor to which adaptive variation is applied has been of the sum-square form. In the passive sonar case, at least, other starting points could be chosen, in particular the cross-correlation schemes which have been implemented as "multiplicative" processing. To justify this comment, the paper begins by a diversion into Detection Theory, and shows that although the sum-square processor may be optimal for a hypothesis which is appropriate (as an idealisation) to active sonar as well as geophysical time series, an alternative hypothesis is more valid for passive sonar, and this leads to the result that a multiplicative processor is more nearly optimal in the passive sonar case.

A brief "physical" discussion of multiplisative processing follows (derived largely from Horton (ref. 1) and Risness (ref. 2)), in which the relationship between multiplicative and sum-square beamforming is explored, in particular, the concept of the "intra-class correlator" as an optimal generalised multiplicative processor is introduced, and it is shown that this becomes a simple split-array multiplier in the limit of a small array, and approaches the sum-square beamformer for a large array. We then consider the procedure by which the sum-square beamformer is generalised into an adaptive scheme, and show that attempts to follow the same prescription for a split-array multiplier fail.

However, an analogous adaptive procedure can be defined for the intra-class correlator. The form of this is established, and it is shown that it is closely related to the sum-square adaptive scheme, but with different constraints. The effect of these changed constraints is discussed, and it is demonstrated that they have the property of automatically and "uniformly" providing robustness in the response for cases where sensor error or misalignment leads to signal suppression in the "classical" adaptive scheme - and thus, at least for large arrays, include in a single constraint the protective mechanisms which Hudson (ref.3) and Owsley (ref.4) provide on a more ad hoc basis by subsidiary constraints.

LIKELIHOOD RATIO AND PASSIVE SONAR.

The textbook definition of an optimal processor rests on the following piece of basic probability theory. Suppose we have a hypothesis H, and a series of observations x_i (i=1, 2 .. N), which provide evidence on the truth or falseness of the hypothesis; a priori, the probability that the hypothesis is true is p(O;T). Then, from Bayes' Theorem

$$p(1;T) = p(O;T) . \frac{p(x_i,(i=1,2 \ldots N); T)}{p(x_i,(i=1,2 \ldots N))}$$

$$\text{and similarly } 1 - p(1;T) = \{1 - p(O,T)\} \frac{p(x_i; F)}{p(x_i)} \qquad *1$$

which, as 1 - p(n;T) is, by definition, p(n;F), gives

$$\frac{p(1;T)}{p(1;F)} = \frac{p(O;T)}{p(O;F)} . \quad L(x_i) \qquad *2$$

where $L(x_i)$, the LIKELIHOOD RATIO, is defined as

$$L(x_i) = \frac{p(x_i;T)}{p(x_i;F)} \qquad *3$$

The likelihood ratio is therefore a measure of the increase in confidence in the hypothesis which may be engendered by the set of observations made; a processor which extracts the likelihood ratio for this set of observations and hypothesis is therefore optimal, and any processor which provides a function which can be transformed into the likelihood ratio by a set of monotonic (1,1) operations is equally optimal.

To clarify the concept, we will take the standard example (c.f. Horton (ref 1) and many others) of the detection of signal in noise, given two observations, as this is the basis of the choice of the sum-square scheme as the optimal processor for energy detection. The hypothesis to be tested is: HYPOTHESIS 1.

given that the two observations v_1, v_2 are each subject to random noise which is Gaussianin form and of mean power $\bar{V}_N$ and that this noise is totally uncorrelated between the two observations;

that a signal (which is also Gaussian and of mean power $\bar{V}_s^2$) is either present and completely correlated between the two observations, or is totally absent;

then that the signal is present when the observations v_1, v_2 are made. Considering the denominator of $L(x_i)$ from equation *3 we see that as, if the hypothesis is false, v_1 and v_2 are completely independent and in each case taken from the ensemble of Gaussian samples with power ∇_N^2, then

$$p(v_1, v_2;F) = p(v_1;F)\ p(v_2;F) = (2\pi\nabla_N^2)^{-1} \exp\left\{-\frac{(v_1^2 + v_2^2)}{2\nabla_N^2}\right\} \qquad *4$$

To evaluate the numerator, we let the instantaneous signal contribution be u; then there must be Gaussian noise $(v_1 - u)$, $(v_2 - u)$ in channels 1,2 respectively.

As this is exactly the situation considered above, we can say that this particular distribution of Gaussian noise occurs with probability $p(v_1 - u, v_2 - u; F)$; the

probability of the particular distribution of Gaussian noise occurs with probability $p(v_1 - u, v_2 - u;F)$; the probability of the particular signal value u is itself normally distributed with standard deviation ∇_s, so that

$$p(v_1, v_2;\ T) = \int_{\infty}^{\infty} p(v_1 - u, v_2 - u;\ F).\ p(u)\ du$$

$$= (2\pi\nabla_N)^{-1}\ (\nabla_N^2 + 2\nabla_s^2)^{-\frac{1}{2}}\ \exp\left\{\frac{-(v_1^2 + v_2^2)}{2\nabla_N^2} + \frac{\nabla_s^2}{2\nabla_N^2}\frac{(v_1 + v_2)^2}{(\nabla_N^2 + 2\nabla_s^2)}\right\} \qquad *5$$

so that $$L(v_1, v_2) = \frac{\nabla_N}{(\nabla_N^2 + 2\nabla_s^2}\ \tfrac{1}{2}\ \exp\ \frac{\nabla_s^2}{2\nabla_N^2}\ \frac{(v_1 + v_2^2)}{(\nabla_N^2 + 2\nabla_s^2)} \qquad *6$$

It is clear that this can be generalised to n observations, and the form is obvious by inpsection. In fact

$$L(v_i,\ i = 1,2\ldots n) = \frac{\nabla_N}{(\nabla_N^2 + n\ \nabla_s^2)^{\frac{1}{2}}}\exp\left\{\frac{\nabla_s^2}{2\nabla_N^2}\ \frac{(\sum_{i=1}^{n} v_i)^2}{(\nabla_N^2 + n\nabla_s^2)}\right\}$$

$$\text{Log } L(v_i) = A + B\ (\sum_{i=1}^{n} v_i)^2 \qquad *7$$

where A, B are functions of the input signal/noise ratio of the observations only.

The value of equation *7 is that it shows explicityly that $(v_i)^2$ may be obtained from $L(v_i)$ by the set of (1,1) monotonic transforms (takes logs; subtract A; divide by B), and that therefore:

The sum-square processor is optimal for testing hypothesis 1, even when generalised to n observations, and independently of the input signal-to-noise ratio of the raw observations.

Before extrapolating from this statement to the conclusion that the sum-square processor is also a valid optimal processor for all, or even most, detection situations, however, it is necessary to examine in more detail the relevance of the model in the hypothesis.

The point to which attention should be drawn is that Hypothesis 1 envisages a situation in which a noise field is already established, and then the wanted signal is switched on, without disturbing the noise field. This is a valid model for many applications (it is clearly correct if the noise against which the signal has to be detected originates in the actual processor, or for the case when weak transients are to be detected - both characteristic of the analysis of seismic records; it is also valid for modelling the detection of echoes in clutter in both radar and active sonar) but its relevance to passive sonar is questionable - provided that one is not limited by "system noise".

In passive sonars, the sensors receive an input of known power, which consists either of random noise of that power, or of random noise of some lesser power, together with a correlated signal contribution. However, even if a signal is present, its correlation will only be apparent if the array is steered so that the signal lies in the beam of the receiver. Accordingly, it is suggested that an alternative hypothesis, formulated below, is more appropriate to passive sonar detection than the signal on/signal off criterion of Hypothesis 1.

HYPOTHESIS 2 ("signal on-beam/signal off-beam")

given that the two observations v_1, v_2 are each subject to random noise which is Gaussian in form and of mean power $\overline{V_N^2}$ and that this noise is totally uncorrelated between the two observations;

that a signal(which is also Gaussain and of mean power $\overline{V_s^2}$,)is always present and is either totally correlated between the two observations (signal on-beam) or completely uncorrelated between the two observations (signal off-beam);

then that the signal is on-beam when the two observations are made.

For the case when the hypothesis is true, equation *5 again holds, as the situation is then identical with that under hypothesis 1. To calculate $p(v_1,v_2;F)$ for hypothesis 2, we note that in these circumstances the two observations are entirely independent, and that each is composed of the sum of two independent Gaussian processes - i.e. is itself a Gaussian process with variance the sum of the two component variances.

Hence $p(v_1,v_2;F) = p(v_1;F).p(v_2;F) = \left[2\pi(\overline{V}_N^2 + \overline{V}_s^2)\right]^{-1} \exp\left\{\frac{-(v_1^2 + v_2^2)}{2(\overline{V}_N^2 + \overline{V}_s^2)}\right\}$ *8

and so $$L(v_i,v_2) = \frac{\overline{V}_N^2 + \overline{V}_s^2}{\overline{V}_N(\overline{V}_N^2 + 2r_s^2)^{\frac{1}{2}}} \exp\left\{\frac{\overline{V}_s^2}{2\overline{V}_N^2(\overline{V}_N^2 + \overline{V}_s^2)}\left(2v_1v_2 - \frac{\overline{V}_s^2}{\overline{V}_N^2 + 2\overline{V}_s^2}(v_1^2 + v_2^2)\right)\right\} \quad *9$$

and thence to

$$\text{Log } L(v_1) = A' + B'\left(\sum_{i \neq j}\sum v_i v_j - \frac{e}{1+ne}\sum v_i^2\right) \quad *10$$

where e is the signal-to-noise ratio $\frac{\sigma_s^2}{\overline{V}_N}$.

Thus again in this case we can perform the series of (1,1) monotonic transformations (take logs; subtract A'; divide by B') to obtain an optimal processor from $L(v_i)$. In this case its precise form then depends on the signal-to-noise ratio, but at least conceptually, we can go one stage further and remove the v_i^2 terms by subtraction (as these are again positive definite and hence this transformation is also monotonic), to derive the result

> A processor which forms all the cross-product terms between pairs of observations and sums the result, is an optimal processor for testing hypothesis 2, generalised to n observations.

Such a processor, which we call an "intra-class correlator", is not as generally well-known as is the sum-square processor, so we continue by a discussion in "physical" terms of some of its properties, and of those of a related sub-optimal but more easily implementable system, the split-array multiplicative processor, which has been used in some sonar applications.

MULTIPLICATIVE PROCESSING

In Figure 1a we illustrate, for a four element array, the operations needed for a straightforward implementation of the intra-class correlator defined above; Figure 1b demonstrates the relative simplicity of the sum-square scheme for the same array. This relative complexity of the intra-class correlator, which increases as the array size rises, leads to the query whether or not a simpler related system exists, which retains near-optimal properties under Hypothesis 2. If, from equation *10, one takes the view that many of the optimal properties are associated with the absence of terms of the form v_i^2, then one is led to the split-array multiplier, which is illustrated in Figure 1c; here, the array is split into two halves, for each of which the sensor outputs are summed, and these two sums are multiplied together. Only a selection of the element cross-products is thereby formed, and so this is not a true optimal processor; however, its simplicity of implementation may compensate for this in several practical situations.

Symbolically, the operations illustrated in Figure 1, for an array of 2N sensors, may be written as:

sum-square may be written as:

Sum-square beamformer $$\left(\sum_{i=1}^{2N} v_i\right)^2$$

Split-array multiplier $$\left(\sum_{i=1}^{N} v_i\right)\left(\sum_{j=N+1}^{2N} v_j\right) \qquad *11$$

Intra-class correlator $$\sum_{i=1}^{2N} \sum_{j=1}^{2N} v_i v_j \quad (i \neq j)$$

Figure 2 shows the response of each of these to a sinusoidal plane wave input, as a function of angle, in the special case of a four-element line array with sensors ½-wavelength apart; Figure 3 repeats the same information but with the sensitivities normalised to be equal in the on-beam direction for the three cases.

Certain features are immediately apparent from these figures, among them:

(a) whereas the response of the sum-square processor is positive for all directions of the input, the other two systems give negative outputs at certain angles, in particular those flanking the main beam direction.

(b) the multiplicative processor has the narrowest, and the sum-square processor the broadest, main beam;

(c) the positive side-lobes of the multiplicative and intra-class correlator processors are smaller, in both absolute and relative terms, than those of the sum-square system;

(d) the first negative-going side lobes of the multiplicative and the intra-class correlation processors are larger in magnitude than the first (positive) side lobe of the sum-square system;

(e) all the negative-going side-lobes of the intra-class correlator have equal peak amplitude.

We note that the first three of these make the simple split-array multiplier a very attractive scheme for sonar implementation, if the processor is followed by a rectifier before any display stage is reached.

All the above features can be readily explained in "physical" terms if we return to equations *11 and expand the processors into elemental sub-systems, from which they are built up by linear superposition. Two types of fundamental processor emerge - the auto-multiplier v_1^2 and the cross-multi-plier $v_i v_j$. Of these the first has the property that it has no

directional discrimination - responding equally sensitively to on-beam signals, off-beam interferences, and "noise"(as it involves only one sensor, it cannot take advantage of the lack of correlation between sensors by these latter inputs to reject them). Thus these self-product terms, which appear only in the sum-square processor, have the effect of a constant positive bias on the angular output.

The directional properties of these processors therefore derive exclusively from the cross-product terms $v_i v_j$. For an array with the spacing between adjacent elements set at ½-wavelength, the response is given by

$$v_i \, v_j = \cos\,(m\pi \sin\Theta),\ \text{where } m = j-i \qquad *12$$

Figure 4 explicityly demonstrates the narrowing of the main beam and the intrusion of repeat lobes into real angles as element separation increases by plotting the angles at which equation *12 has particular values, as a function of m.

Armed with this data, we are in a position to explain the observed features quoted in earlier paragraphs, as follows:

(a) from equation *11, the sum-square system must give a positive output as it produces the square of a real number; this implies that at some angles (where its output is zero), the net effect of the cross-product terms is strictly negative, and it is balanced by the non-directional positive bias provided by the self-product terms. The other systems, lacking this bias, will accordingly give negative outputs in directions where the sum-square processor has a zero or small positive output, in particular at the flanks of the main beam.

(b) equation *11 for the intra-class correlator can be rewritten in the form

$$\sum_{i=1}^{2N} \sum_{j=1}^{2N} v_i v_j \; (i \neq j) \; \equiv \left(\sum_{i=1}^{2N} v_i \right)^2 - \sum_{i=1}^{2N} v_i^{\,2} \qquad *13$$

i.e. the intra-class correlator consists precisely of the sum-square beam-former with all the self-product terms removed. Thus in going from sum-square to intra-class correlator, we remove the omnidirectional component; this must make directional changes more rapid in relative terms, and in particular narrow the main beam. Similarly, to achieve the split-array multiplier, we then remove the further terms corresponding to "within sub-array" cross-products, 2(N-m) of them for each separation m, up to m=N; these are all small separation, so will further narrow the main beam.

For those who prefer a symbolic proof, we offer the inequality, true for all positive a, x and y<x $\frac{a+y}{a+x} > \frac{y}{x}$, and identify a as the terms to be removed at each stage, x the on-axis contribution from the terms which are retained, and y the response from these same terms at the angle under consideration.

(c) and (d) then follow immediately from the same inequality, where in this case a is specifically the constant self-product term reponse, x is the on-axis response of the directional terms, and y the response of these same terms at the angle under consideration. In the case of (d), note also that the inequality is valid for y negative.

(e) follows from the form of the intra-class correlator shown in equation *13 - a sum-square beamformer with the (non-directional) self-product terms removed. Hence the peaks of the negative lobes in the intra-class correlator coincide with the minima of the sum-square beamformer output, which are zeroes as the latter is a squaring system. Accordingly the maximum amplitude of negative response in the intra-class correlator is equal in magnitude (though opposeite in sign) to the summed self-product terms of all the array elements.

Two further points arise from this identification of the relationship between the sum-square beamformer and the intra-class correlator. First, as the advantage of the latter must derive solely from the absence of self-product terms, the suggestion that the split-array multiplier might be close to optimal because of its exclusion of self-product terms acquires some validity. Second, in the limit of a large array, the number of self-product terms becomes negligible in comparison with the cross-product terms, so that the sum-square system approaches the optimum. Noting that a two-element intra-class correlator is identical with a two element split-array multiplier, we see therefore that the optimal processor for Hypothesis 2 transforms smoothly from being a split-array multiplier for small arrays to becoming a sum-square beamformer for large arrays.

ADAPTIVE SUPPRESSION OF CORRELATED INTERFERENCE

So far, the discussion in this paper has been exclusively concerned with OPTIMAL processing, for two carefully stated hypothesis. In each, deliberately, we have totally excluded the factor which limits performance in many real situations - the presence of interferences which provide unwanted correlated inputs to the information channels. We now address the question of how to suppress these interferences without removing the wanted signal, and for the moment will concentrate on sum-square processing schemes.

For arrays with a large number of degrees of freedom, the solution is a preselected weighting scheme (i.e. maintaining the sum-square form, but giving different emphasis to individual relements before forming the sum) which provides a guaranteed rejection of correlated interferences while retaining the sensitivity to wanted signals, at the cost of slightly sub-optimum performance against noise and a reduction of the ability to resolve signals within the main beam. Representative of such schemes is the Dolph-Chebyshev shading; a useful rule of thumb for this is that the performance against noise starts to be significantly degraded if side lobe levels more than 20 log (n-1) dB down are to be guaranteed for an n element array.

Thus, for a small array, this trade-off is less than satisfactory and we are led to the requirement for a more sophisticated, adaptive system. However, the philosophy remains the same. We start with the equal weight scheme which is optimal against noise, and vary the weights while maintaining the form of the processor. The motivation for the variation is to reduce the influence of the correlated interference on the output, while we impose the constraint that the response of the varied system to the wanted signal shall be the same as that of the inital optimal processor.

This is particularly simple for a processor of sum-square form, as the response to each independent input, be it noise, interference, or signal, is positive. Therefore, we achieve the desired result by minimising the beamformer output by varying element weights, subject to the restriction that the response to a unit signal from the look direction remains constant. If we take, for simplicity, the special case where the beam axis is normal to the array, this process may be represented symbolically as:

$$\text{minimise} \quad \left(\sum_i w_i V_i \right)^2_{(\text{all } w_i)} \quad \text{subject to} \left(\sum_i w_i \right)^2 = K \ (\text{i.e. } \sum_i w_i = K' \ ) \qquad *14$$

and this minimised output represents the optimum estimate of the look-direction input which can be obtained using a sum-square system, in this particular noise and interference environment.

As we have indicated that a sum-square system is not necessarily the optimum starting point in the passive sonar case, we now ask whether an analogous procedure can be defined, to give enhancement for either the split-array multiplier, or the intra-class correlator, by adaptive choice of sensor weightings. An immediate problem arises in deciding, irrespective of constraints, what is the analogue of minimising the sum-square output, for a correlating processor. As distinct from the situation modelled by *14, the correlator output is not a positive definite quantity, so that strict minimisation means maximising the output amplitude, though with a negative sign; far from suppressing the effect of coherent off-beam interferences, this has the effect of enhancing them. Minimising the modulus of the correlator output is equally not a valid solution - it drives the output to zero irrespective of the existence or not of an on-beam signal, and this zero value can always be attained as it does not represent an extreme value for the system.

Thus we conclude, slipping back into the language of detection theory, that if the concept of adaptive enhancement of multiplicative processors along the lines of *14 is to be retained, then the minimisation must be performed on some positive definite function which is related to the final system output by a (1,1) monotonic transform. For a split-array system, this requirement has no obvious means of fulfilment; the only positive definite functions immediately available are the sum-square outputs of the half arrays and/or the full, undivided array, and these do not lead to the multiplier output by a straightforward monotonic transform. In physical terms, this can be seen if we examine briefly the initially plausible ideas which could be put forward, viz:

(a) minimise full array sum-square output, then form split-array cross products; this is a completely fallacious approach, as it is possible to set up situations in which the half-arrays have large but opposed responses to correlated interference, whereupon the sum-square array will have the influence of the interference removed, but the multiplier derived from it by this scheme will have the influence enhanced.

(b) minimise independently the half-array sum-square outputs, then form the multiplicative output of these two arrays; although, with appropriate constraints, this will give an extreme value which is in some sense an optimum, this procedure cannot give as good a result as the original full array sum-square, as the minimisation of each half-array output against the same background effectively halves the number of degrees of freedom which may be used.

It would therefore appear that an impasse has been reached, and the split-array multiplier is not suitable for adaptive enhancement; to some extent, this should not be regarded as too surprising a conclusion, as it is not, in the unadapted case, an optimal processor. However, the intra-class correlator is an optimal processor, and, moreover, one which is very closely related to the sum-square beamformer, as evidenced by equation *13. Starting from *13, recognising that $(w_i V_i)^2$ is positive, and keeping the requirement that the adapted system retains the same sensitivity to the wanted signal as the unadapted system from which it derives, we immediately find, as the analogue for *14 in the case of the intra-class correlator, the adaptive process:

$$\underset{(\text{all } W_i)}{\text{minimise}} \left(\sum_i W_i V_i \right)^2 \qquad \text{subject to} \sum_i \sum_j \left(W_i W_j \right) = C$$

$$\text{i.e} \left(\sum W_i \right)^2 = C + \left(\sum_i W_i \right)^2 \qquad \text{Form as output} \left(\sum_i W_i V_i \right)^2 - \sum \left(W_i V_i \right)^2 \qquad *15$$

DISCUSSION OF THE PROPOSED ADAPTIVE SYSTEM *15.

By a somewhat tortuous route, we have now achieved a proposed adaptive system which is very closely related to that commonly adopted, but which we claim to be optimum for a hypothesis which is more nearly matched to the passive sonar situation than is the classical energy detection on which the sum-square scheme is based. The question may fairly be asked, what benefits will this more complicated procedure provide? To answer this, we begin by considering the limit of a large array (for which, admittedly, the case for adaptive rather than preselected processing schemes is weakest, but where in addition, the optimal intra-class correlator and sum-square beamformer tend to the same limit in the unadapted case).

If, in the limit as N tends to infinity, the w_i remain relatively well-behaved (i.e. generally of the same order of magnitude and the same sign), then as there are N(N-1) cross-product terms and only N terms of the form w_i^2 in the constraint, these latter may be neglected and the constraint reduces to the limiting form $\left(\sum_i w_i \right)^2 = C$, identical with that of *14; equally, the correction in going from the minimised sum-square value to the final output can be neglected, and so we are left with effectively an adaptive sum-square beamformer in the precise terms of *14. This is to be expected, given the close relationship between the two unadapted schemes; the perhaps unexpected new information is the identification of the situations in which the two schemes may not converge to the same limit - the cases when the individual w_i are large in magnitude and of alternating sign, so that in *15

$$\sum_i w_i^2 \gg C, \text{ ie. } \sum_i w_i^2 \text{ is comparable with } \left(\sum_i w_i\right)^2,$$

or equivalently in *14 $w_i > Ko$ for some individual w_i.

But these situations have already been identified as cases in which the "classical" adaptive beamformer defined by *14 is in difficulties, and is prone to signal suppression (see, e.g. Hudson (ref. 3)). In physical terms they correspond to, among other things, large correlated inputs from sources close to, but distinct from, the beam axis, os equivalently slight sensor misalighment or error so that the wanted signal appears to come from an off-axis direction. In any of these cases the scheme described by *14 "goes superdirective" with weights of large magnitude relative to the total array sensitivity, whose values change rapidly and non-uniformly in response to small changes in the estimated environment - thus leading to slow convergence, hunting and stability problems. To counter this, and provide a processor which is robust to such problems, Hudson supplements *14 with an inequality norm constraint on the w_i, derived on an ad hoc basis, of the form $W_i^2 \leq M^*$ for each i.

Returning to *15, we see that it may be interpreted as including a constraint of a similar nature, in that as the magnitude of the w_i's increases so does the "sensitivity" in the sense of *14, against which the sum-square output is to be minimised. In particular, whereas the constraint in *14 allows the relative sensitivity to noise to rise as the weights grow in amplitude, the constraint in *15 increases the <u>absolute</u> sensitivity, and this, taken in conjunction with the minimisation requirement, acts as a limit to the degree of superdirectivity which can be attained. Additionally, the supplementary constraints which have been suggested for *14 make an artificial distinction between interference and noise, as they preset the maximum sensitivity of noise which can be tolerated, independent of signal/noise and interference/noise ratios; the total constraint in *15 automatically takes into account these ratios in determining its optimal form, and does not require any absolute limit to noise sensitivity, as the noise component is subtracted out to produce the final processor output.

An alternative further insight into the differing effects of the two processors of *14 and *15 is obtained by considering how, physically, they react when confronted by an input comprising: the wanted signal, noise of comparable power, and a very much stronger interference arising from an adjacent direction. This interference initially dominates the situation, and the adaptation attempts to reduce its influence to zero - i.e. to make the total effect of the cross-product terms negative to balance the positive bias from self-product terms. In the case where the interference is within the main beam, this can only be done by giving some elements negative weights, hence increasing the mean modulus of the weights, and accordingly the effect of the self-product terms, which are sensitive to the noise input also. As, in addition, giving negative weights reduces the positive contribution of the cross-product terms for the on-beam signal contribution (while in compensation increasing the self-product effect), we see that the overall mechanism of minimisation is to reduce the relative influence of the cross-product (i.e. directionally discriminating) components to the sum-square output.

This means that the effect of the noise must be enhanced, relative to the interference-free situation; as, in *14, the constraint requires the total response to "noise from signal direction" (i.e. the non-directional terms) plus correlated signal to be constant, this implies that although the output may be a good estimte, in terms of energy level, of the signal input, its structure is dominated by the enhanced noise response, and the relative importance of the correlated signal structure is degraded. In contrast, in *15, although the effect of noise on the minimised function is equally enhanced, it is then identically removed from the final output; here the limiting factor lies in the effect of the constraint, demanding constant sensitivity to correlation on-axis, in forcing very rapidly varying directional properties if the interference is to be nulled out. This then means large w_i with consequently the increase in noise contribution to the sum-square function to be minimised preventing complete elimination of nearby correlated interference. That is, the system contains its own, built-in, procedure for preventing the effective beamwidth becoming too narrow. Further, beyond this new, adapted, beamwidth limit, consider what happens. The minimisation goes to completion with the total contribution from the interference zero; we then subtract the self-product terms, and in the absence of any look-direction signal to give a positive output, the final correlator output is negative. That is, the adapted intra-class correlator retains the characteristics of its unadapted forerunner in providing a flanking white-out strip on either side of the narrowed main beam.

All the discussion above has been in the context of a large array, for which the neglect of the self-product terms in all but pathological cases might appear reasonable. However, for a small array, which by virtue of its limited degrees of freedom more requires adaptive enhancement, the self-product terms are a significant portion. The effect of the constraint *15 is uniform, independent of the number of elements, whereas supplementary constraints to make *14 robust will have a progressively more restrictive effect on the environments for which adaptation can be carried out, as the array size decreases. We therefore conclude that the modified constraint system described by equation *15 does represent a more general, and more effective scheme of adaptation for the passive sonar scenario, although it is conceded that there is a price to be paid in greater complexity of the required minimisation algorithms.

ACKNOWLEDGEMENT

This paper appears by permission of the Ministry of Defence (PE).

REFERENCES

(1) C. W. Horton : Principles of Sonar

(2) E. J. Risness : Thoughts on Multiplicative Processing, 1967 Internal A.U.W.E. Memorandum.

(3) J. E. Hudson : An Adaptive Array with Beamwidth Control. IOA, Adaptive Processing Meeting, 1977.

(4) N. L. Owsley Extent Invariant Signal Extraction, IOA Adaptive Processing Meeting, 1977.

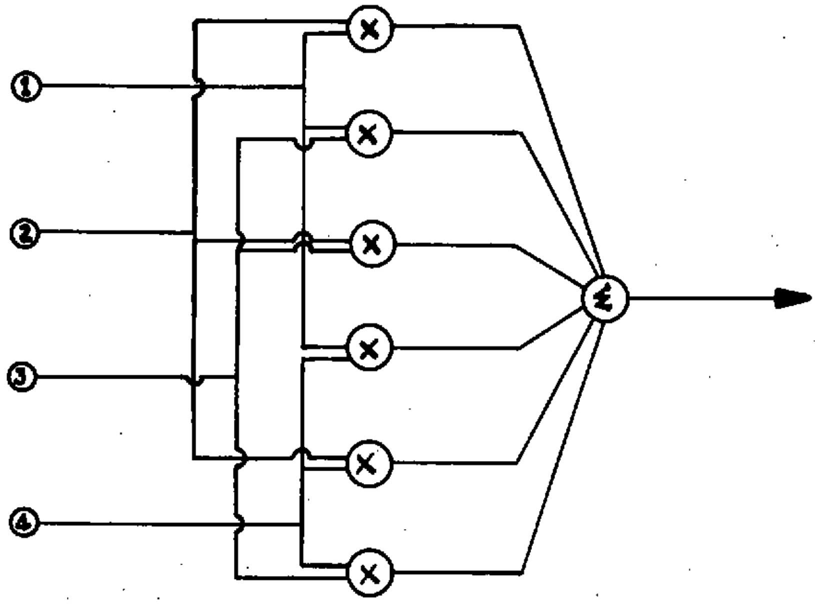

FIGURE 1a. Intra-class Correlator Schematic

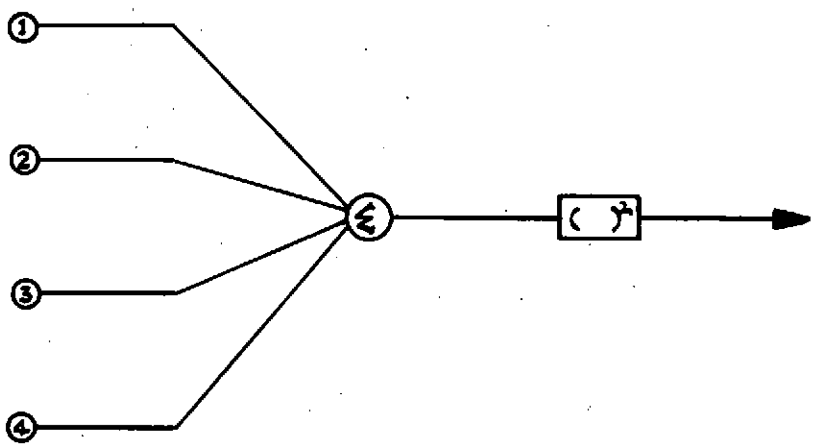

FIGURE 1b. Sum-square Beamformer Schematic

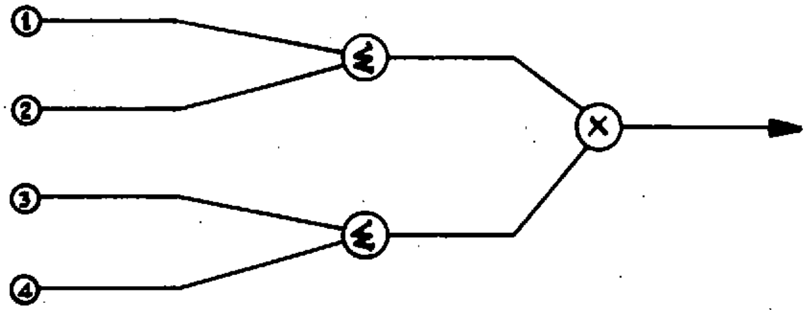

FIGURE 1c. Split-array Multiplier Schematic

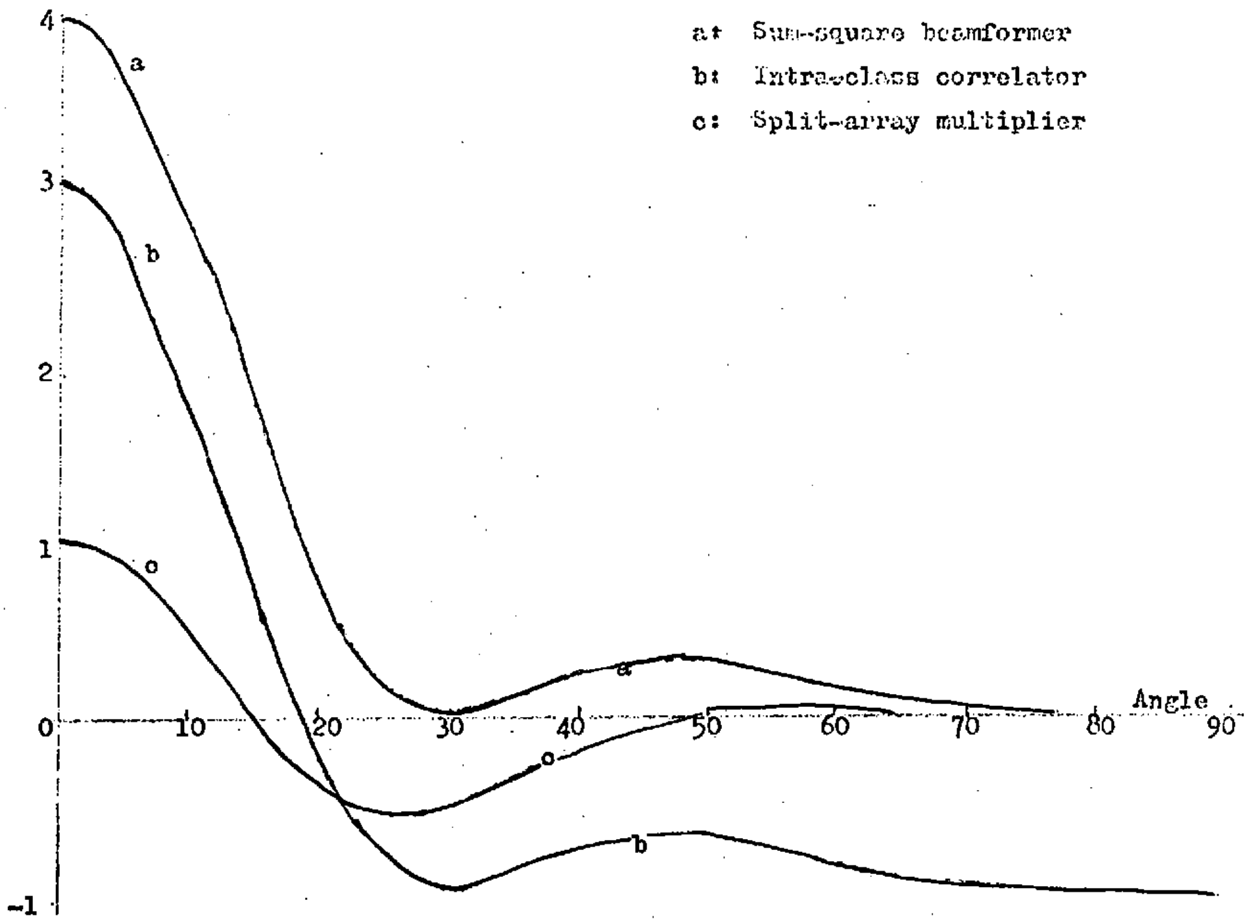

FIGURE 2. Output for a four-element array in response to a pure sinusoid, at half-wavelength element spacing, as a function of angle for the three processing schemes stated.

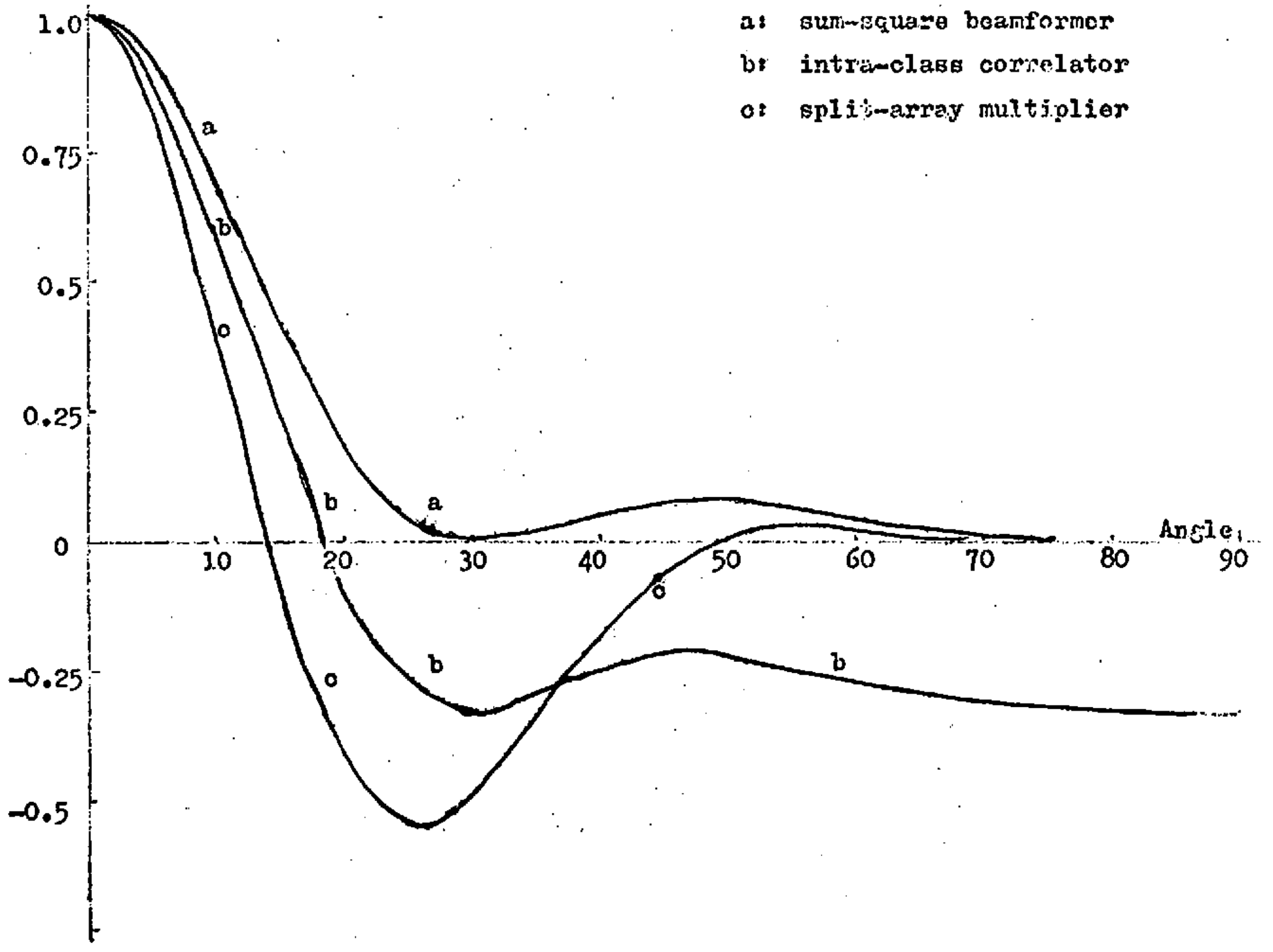

FIGURE 3. Normalised output for the half-wavelength spacing four-element array, as a function of angle for a pure sinusoid input, for the three processing schemes stated.

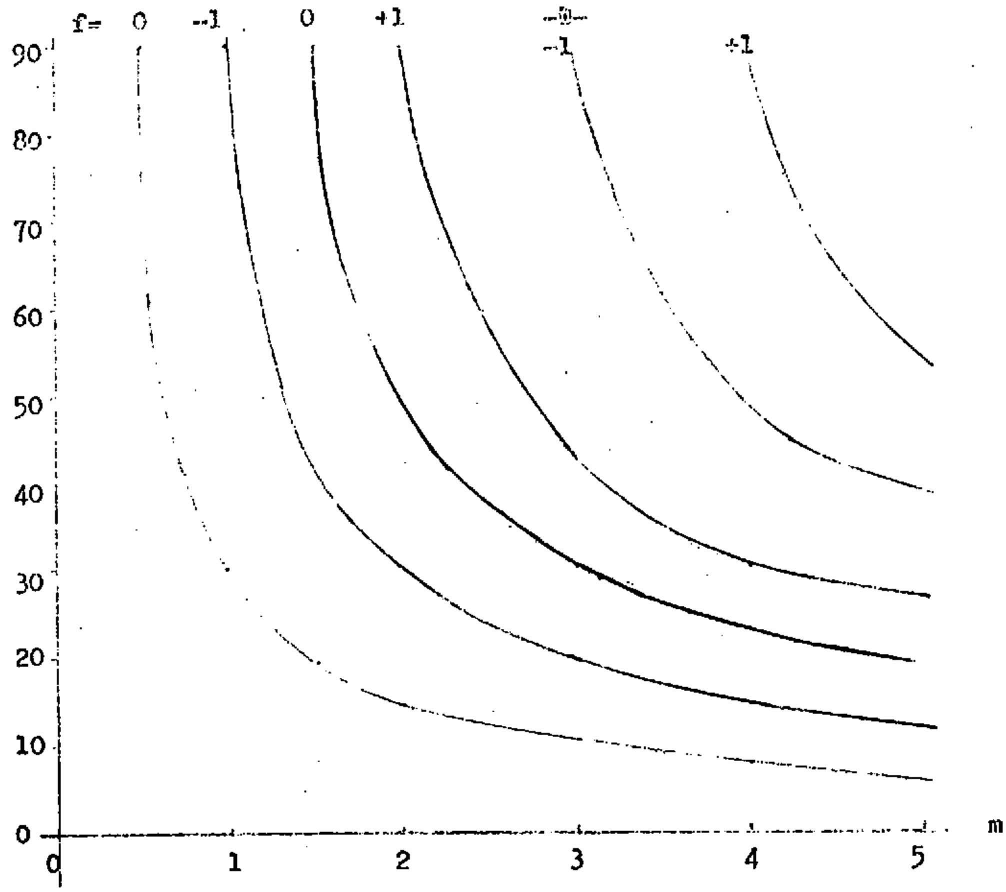

FIGURE 4. Positions in angle of nulls, repeat lobes, and negative repeat lobes as a function of element separation in wavelengths for a fundamental two-element cross-correlation processor.