JITTER REDUCTION IN ADAPTIVE SONAR SIGNAL PROCESSING

DT Hughes(1) and IJ Clarke(1).

(1) DRA Malvern, Malvern Worcs. WR14 3PS. UK.

1. INTRODUCTION

The sonar environment is well known to be a difficult one in which to use adaptive processing techniques. Many difficulties are present such as multipath, channel effects and reverberation. Due to the complexity of the water environment these problems are often considerably worse than the analogous conditions in radar.

In this paper we are concerned with a different set of problems which are encountered when adaptive processing is used in the sonar environment. In sonar, due to the necessary sampling conditions it is often only possible to have a comparatively small number of samples or snapshots with which the adapted weights must be calculated. This condition, as explained below, leads to weight jitter and a consequent degradation in performance.

We shall concentrate on a sonobuoy system in which preprocessing has already been carried out making a narrowband adaptive approach appropriate. Stress will be given to such a situation with a jammer present. The source of this interference may be due to a deliberate counter measure in the water or due to a nearby noise source such as a ship.

2. WEIGHT JITTER

We are interested in a sample matrix inversion approach to adaptive beamforming. Figure 1. shows a schematic of a narrowband beamformer. The scalars $\{x_1(t), x_2(t)......x_N(t)\}$ formed from the observations taken at the N sensors at a time 't' are processed by the weights $\{w_1, w_2......w_N\}$. The output from the summer is given by $w^H x(t)$ where H denotes the hermitian transpose and the vectors w and x are formed from the scalar weights and the samples respectively. The beamformer output is then $w^H R w$ where R is the $N \times N$ covariance matrix given by the expression $E\{xx^H\}$ where $E\{\}$ denotes the expectation operator. We shall assume that $x(t)$ is widesense stationary and in the estimation process the expectation operation can be approximated by time averages.

It is normal to require that the beamformer output be minimised subject to constraints. We begin by assuming a single look direction gain constraint:

$$c^H w = 1 \qquad \textit{equation 1}$$

where c is termed the constraint vector. It is well known [1] that the weight vector which maximises SNR whilst satisfying equation 1 is given by

$$w_{opt} = \frac{R^{-1}c}{c^H R^{-1} c}. \qquad \textit{equation 2}$$

We begin by briefly discussing the result of using this equation.

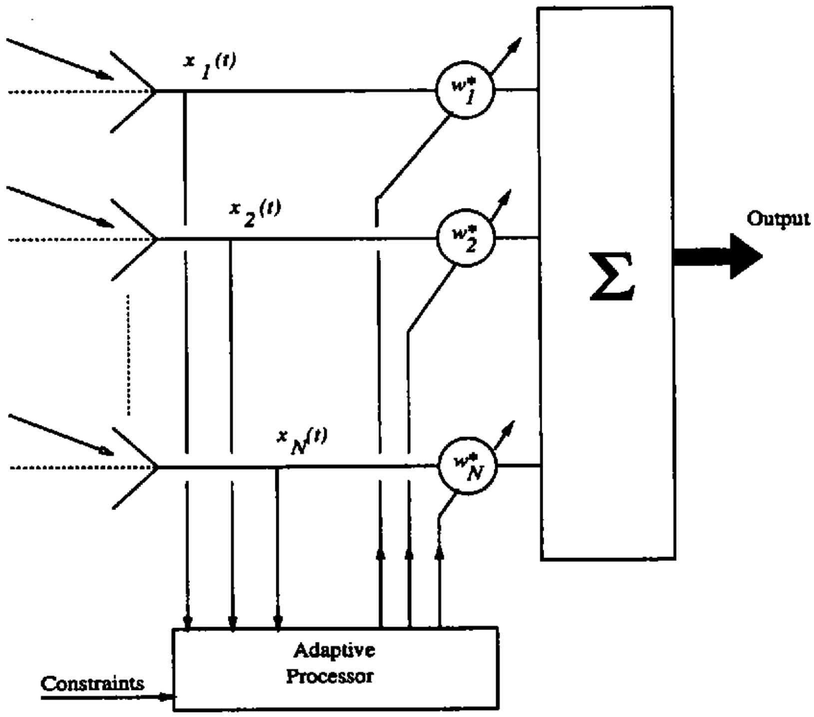

Figure 1. Schematic of narrowband beamformer

2.1 Weight jitter with estimated covariance matrix

The array considered is shown in plan view in figure 2 and is chosen to be no more than 'typical' of sonobuoy configurations in use. No effort has been made to optimize hydrophone positions since such an analysis would take us outside the scope of this paper. Neither do we investigate the role of taper weights in the beampatterns. However a method for optimizing taper weights for an arbitrary array is discussed elsewhere[2]. The array consists of three concentric circles with a third of a wavelength spacing between sensors on the same spar. All taper weights are taken to be unity.

Figure 3 shows the quiescent beampattern for the array of figure 2 appropriate to a look direction at 0° azimuth. All sidelobes are below -11dB and the beampattern exhibits a 3dB beamwidth of about 28°. Figure 4 shows the adapted beampatterns obtained using equation 2. The covariance matrix has been estimated by using 36 snapshots of data (corresponding to twice the number of

JITTER REDUCTION IN ADAPTIVE SONAR SIGNAL PROCESSING

hydrophones present). 5 separate sets of noise have been used. For the purpose of simulation the noise is taken to be Gaussian, white and uncorrelated from sensor to sensor. The noise power is taken to be 0dB at a sensor. Also a jammer of power 30dB (at a sensor) has been introduced at 60°. It is clear from figure 4 that the array maintains a 0dB response in the constraint direction and introduces a null in the direction of the jammer. However, it is also clear that, in general, the sidelobes fluctuate widely from one simulation to the next.

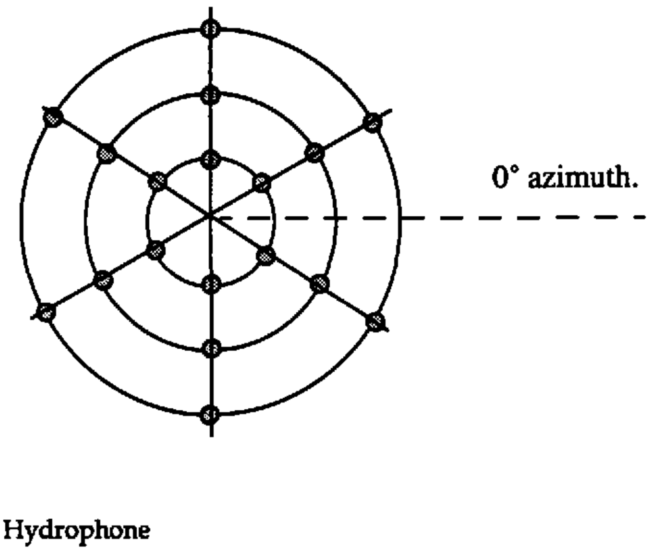

Figure 2. Sonobuoy configuration considered. 0° azimuth is taken to lie along the dotted line.

The presence of jitter can be shown to be extremely deleterious in signal processing resulting in appreciable reductions in the signal to noise ratio as well as altering the output noise statistics.[3] The reason for this sidelobe jitter can be traced back to fluctuations in the calculated weights of equation 1. These fluctuations can be thought of as being related to spurious correlations in the covariance matrix. Consequently one way to reduce these sidelobe fluctuations is to increase the number of snapshots used in the formation of $\boldsymbol{R}$. This, of course is not always a viable approach and other methods of reducing the sidelobe jitter must be found.

3. METHODS OF REDUCING JITTER EFFECTS

In this paper we shall discuss two methods of reducing sidelobe jitter: the penalty function method and an eigen-decomposition approach.

3.1 Penalty function approach

The use of this method has been shown to be effective, straightforward to implement and conceptually simple to understand [4]. The main thrust of the technique is to maximise the output SNR of an array subject to a set of secondary conditions which need be satisfied only approximately. The degree to which these *soft* constraints are satisfied can be thought of as being determined by a set of user defined parameters.

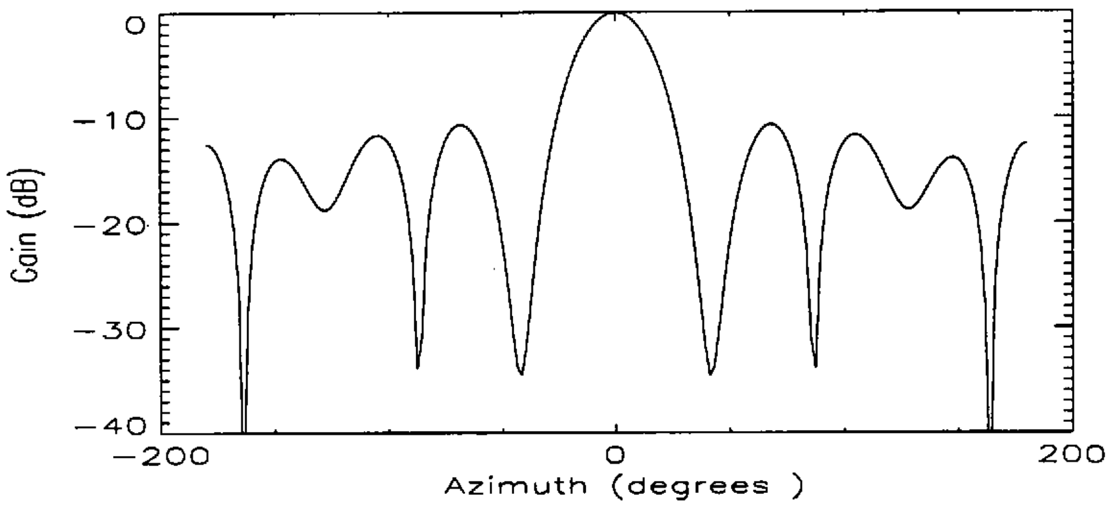

Figure 3. Quiescent beampattern for array of figure 2. All taper weights are taken to be unity.

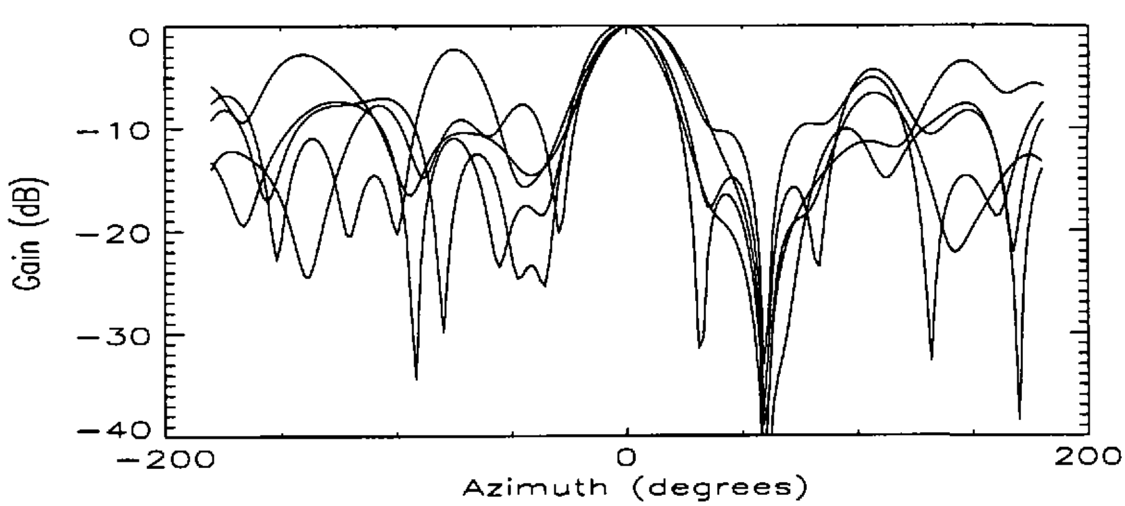

Figure 4. Adapted beampatterns for array of figure 2. A jammer of 30dB is present at 60°. 5 independent white Gaussian noise samples have been used. The sidelobe variation can be seen to be extreme from set to set.

JITTER REDUCTION IN ADAPTIVE SONAR SIGNAL PROCESSING

The constraint imposed in this soft fashion is one which requires the beampattern obtained after adaptive beamforming to approximate to the quiescent beampattern. This desired quiescent beampattern is defined by a weight vector w_q and the adaptive beampattern obtained by a weightvector w. The difference between the two beampatterns at a particular direction (θ, ϕ) can be defined as (where for completeness we consider azimuth and elevation):

$$e(\theta, \phi) = \left|s^H(\theta, \phi)w - s^H(\theta, \phi)w_q\right|^2 \qquad \textit{equation 3}$$

where $s(\theta, \phi)$ is the steering vector appropriate to the direction (θ, ϕ). The actual important quantity is a weighted integral of this expression over all directions of interest. If we denote this total error by E and the weighting function by $k^2h(\theta, \phi)$ where k is a scalar weighting and $h(\theta, \phi)$ is a scalar, well behaved, function defined over the region of interest. $h(\theta, \phi)$ gives emphasis to parts of the beampattern around (θ, ϕ) where $h(\theta, \phi)$ is large. The total weighted error is given by

$$E = k^2(w - w_q)^H\left[\int_S h(\theta, \phi)s(\theta, \phi)s^H(\theta, \phi)(d\theta)d\phi\right](w - w_q) \qquad \textit{equation 4}$$

The term in square brackets defines an $N \times N$ matrix which we denote by Z. If we also impose a look direction gain constraint as defined in equation 1 it can be shown that [4] the optimum weightvector is given by

$$w_{opt} = \frac{A^{-1}c}{c^HA^{-1}c} + k^2A^{-1}Zw_q - \frac{k^2A^{-1}cc^HA^{-1}Zw_q}{c^HA^{-1}c} \qquad \textit{equation 5}$$

where for simplicity of notation we have defined the matrix $A = R + k^2Z$.

It can be seen that

(i) as $k \to 0$ we re-obtain the result of equation 2. Equivalently this can be seen as making the constraints on beampattern negligibly small so that we return to what may be thought of as the classical textbook case and hence exhibiting jitter as in figure 4.

(ii) as $k \to \infty$ the optimum weight vector $w_{opt} \to w_q$. This is intuitively to be expected: as the constraint on the beampattern is increased the obtained beampattern approaches the desired or quiescent pattern until in the limit the two patterns become identical.

It is worth commenting at this point that case (ii) is not necessarily desirable since in such a situation although the sidelobe jitter is driven towards zero the array becomes unable to adapt and the system becomes equivalent to conventional beamforming.

It has been found that often the full complexity of equation 5 need not be used and in fact the matrix Z can be taken to be equal to the identity matrix. The details of this approximation will not be discussed here but the consequences for a linear array are discussed in reference 4. This simplified form for Z will be used in the results section presented below.

3.2 Eigen-decomposition approach

The basic principle of this approach involves adapting separately to the two sets of interference components *i.e.* the noise and the jammers. Partitioning uses an eigen-decomposition $R \Rightarrow U\Lambda U^H$ where again R is the signal free covariance matrix, U is a unitary matrix and Λ is a diagonal matrix containing the eigenvalues of the decomposition. Two possible approaches are discussed in this section to utilise this decomposition

i) the noise partition is simply replaced by a normalised subspace on a batch to batch basis.

ii) an average of the noise partition is performed over a number of batches to improve the estimate and thus reduce weight jitter.

We observe that the first method, though simple in concept, implicitly assumes an *a priori* expectation of noise (without interferers) defined by $\sigma^2 I$ whereas the second method involves estimating the prevailing noise subspace statistics.

3.2.1 Rank-reduced adaptive cancellation

Discriminating between the larger and smaller eigenvalues of a symmetric eigenvalue decomposition $R \Rightarrow U\Lambda U^H$ provides a convenient means of partitioning a spatial sample matrix R, into two orthogonal subspaces:

$$U_J \Lambda_J U_J{}^H + U_n \Lambda_n U_n{}^H \qquad \textit{equation 6}$$

where the subscript J refers to the set of larger eigenvalues and subscript n refers to the smaller values. The first, jammer, subspace, defined by the column vectors U_J, should span the same subspace as that defined by the spatial direction vectors of strong interferers. These can then be processed separately from the subspace defined by the smaller noise, eigenvalues Λ_n and their corresponding eigenvectors U_n. Clearly, strong interference must be heavily suppressed by the adaptive cancellation or prewhitening transform. Any wanted signals in this subspace are also suppressed because the jamming is, in this case, essentially 'mainbeam'. Signals can only be detected easily in the complementary 'noise' subspace spanned by the set of vectors that form the columns of the matrix U_n.

If we can assume that the expectation of the noise components with no jammers present is spatially white, then the relevant subspace eigenvalues can be set to be σ^2. In this simple method of stabilization, R^{-1} is replaced by:

$$U_J \Lambda_J{}^{-z} U_J{}^H + U_n U_n{}^H / \sigma^2 \qquad \textit{equation 7}$$

The value of the exponent z is not critical and can be set between 1 and infinity according to the

JITTER REDUCTION IN ADAPTIVE SONAR SIGNAL PROCESSING

degree of interference cancellation required. Higher values of z lead to a greater depth of nulls. These are required if, for example, the pdf of interferers is significantly non-Gaussian (impulse-like). In this case the eigenvalues Λ_J could be considerably lower than the peak power of interferers and, if z is set to 1, the adaptive nulls would not be sufficiently deep. Other non-linear functions of Λ_J may be considered if necessary.

Since, in equation 7, the noise subspace is simply scaled by σ^2, adaptive prewhitening is restricted to the subspace of large eigenvalues and we can regard the proposed method as providing reduced-rank adaptive cancellation. The concept is similar to that of sidelobe cancellation where the number of cancellation loops is often restricted to avoid a similar 'weight jitter' problem.

It is well known that the sample matrix $\boldsymbol{R}$ is rank deficient (singular) if the number of samples used is less than the number of sensors. A unique sample matrix inverse $\boldsymbol{R}$ does not then exist but, we realise from equation 7, that the missing components in $\boldsymbol{R}$ can be replaced using default eigenvalues σ^2. Clearly this solution to the singular problem relies on the assumption, rightly or wrongly depending on the application, that the true **expectation** of the noise covariance (excluding strong interferers) is spatially white.

3.2.2 The double average method

Ideally, if the noise is not white, then to stabilise 'weight jitter' in the noise subspace, we need to devise a 'double-average' method where the weaker subspace components are averaged over more samples. It is intended to fully investigate this more general method of stabilising weight jitter in a later paper[5]. Briefly, it is essential that the second average is performed on a full rank noise-only matrix. We therefore suggest averaging $\boldsymbol{U}_J\tilde{\sigma}^2\boldsymbol{U}_J^H + \boldsymbol{U}_n\Lambda_n\boldsymbol{U}_n^H$ over several batch estimates of $\boldsymbol{R}$ where $\tilde{\sigma}^2$ is itself an average of Λ_n both within a batch and over the several batches. The resulting estimate of the noise only covariance matrix $\bar{\boldsymbol{R}}$ for the group of batches can easily be updated from batch to batch. The prewhitening transform for each of the batches in turn is evaluated using:

$$\boldsymbol{U}_J\Lambda_J^{-z}\boldsymbol{U}_J^H + \boldsymbol{U}_n(\mathrm{diag}(\boldsymbol{U}_n^H\bar{\boldsymbol{R}}\boldsymbol{U}_n))^{-1}\boldsymbol{U}_n^H \qquad \textit{equation 8}$$

where the diagonal matrix replaces $1/\sigma^2$ in equation 7. A recursive procedure with a forget factor can be adopted in place of the batch method described above.

4. RESULTS

We shall, for concreteness, concentrate on how the penalty function method behaves when applied to the weight jitter problem as experienced by the array of figure 2 as exemplified by the beampatterns of figure 4.

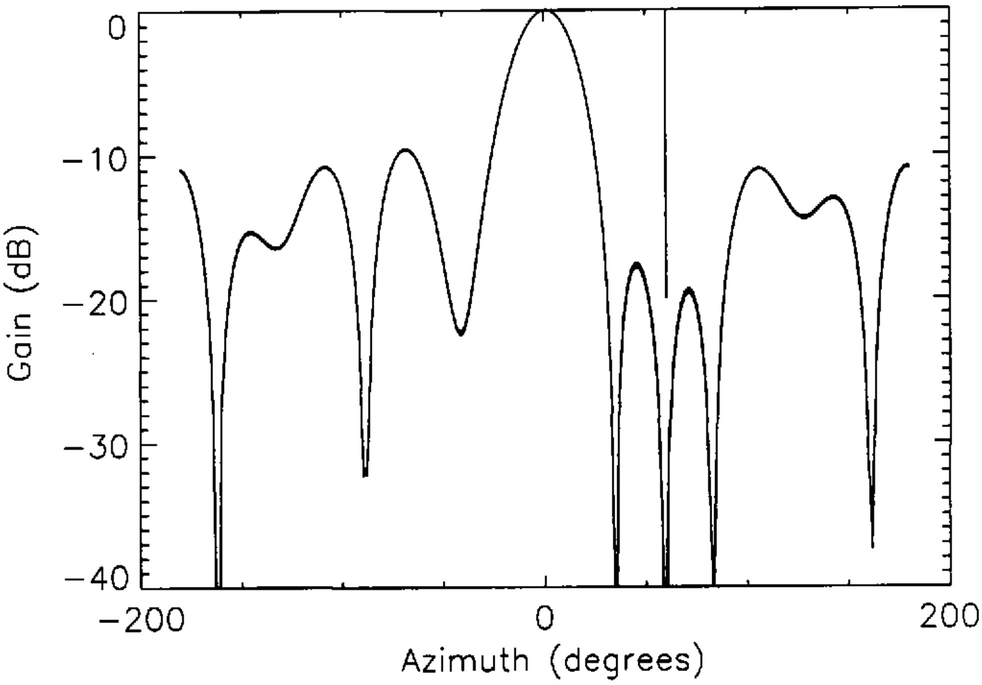

Figure 5. Adapted beampatterns for array of figure 2. A jammer of 30dB is present at 60° (indicated by he vertical line). 5 independent white Gaussian noise samples have been used. The penalty function method has been used as explained in the text.

Various parameter selections need to be considered when applying the penalty function. The most important are the selection of $h(\theta, \phi)$ and k in expressions 4 and 5. Since we are dealing with a sonobuoy with no vertical resolution we can ignore resolution in the θ direction. Consequently we are only interested in azimuthal discrimination. It has been shown elsewhere that for certain arrays the $\mathbf{Z}$ matrix of equation 5 can be chosen to be diagonal with the consequence that the adapted beampattern is most closely fitted to the quiescent beampattern in regions close to the look direction. The value of k is chosen to be 10.0, although, as will be shown and discussed below acceptable values actually fall within quite a broad range.

We choose a relatively simple signal/jammer scenario in order to allow straightforward analysis of the results. A look direction gain constraint is imposed in the direction $\phi = 0°$. A jammer of 30dB is present at 60°. The noise is again taken to be white, Gaussian and uncorrelated from hydrophone to hydrophone with a power of 0dB at each hydrophone. A 0dB signal is present during adaption and is uncorrelated with the jammer.

Figure 5 shows the results of such a simulation for five different noise samples. The visible reduction in jitter is marked. We shall not make a quantitative investigation of jitter reduction at this point but rather refer the reader to reference 4. In fact for the resolution shown the jitter has become almost unobservable. However, the beampattern has adapted in the required way:

i) A null has been formed in the direction of the jammer.
ii) The quiescent beampattern has been retained at directions away from the jammer.
iii) The 0dB response is present at 0°.

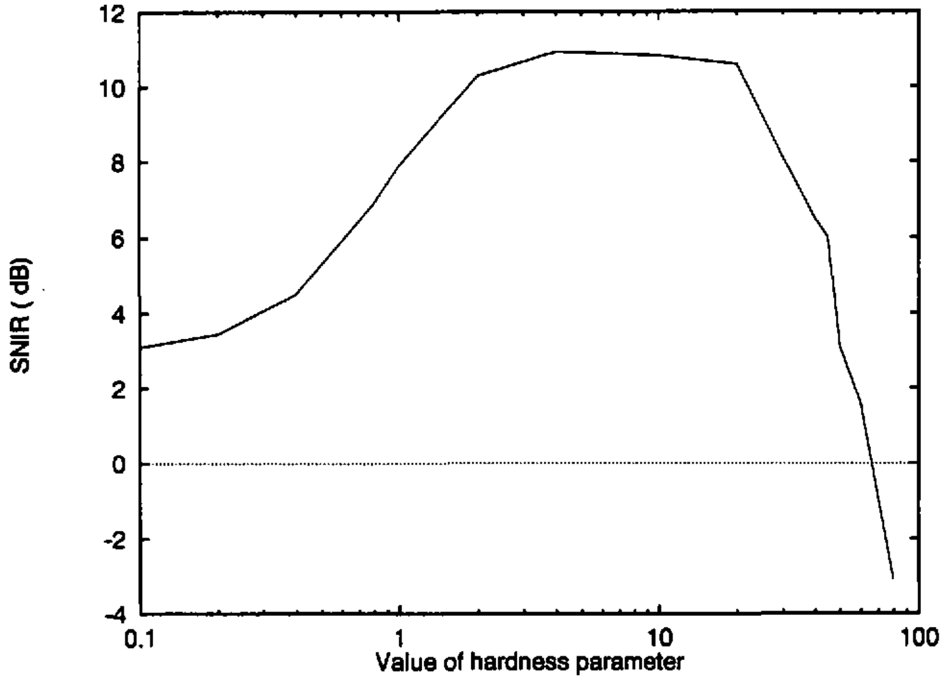

Figure 6. SNIR as a function of the parameter k (the hardness parameter).

It is interesting to consider how the penalty function method performs for varying values of the 'hardness parameter' k. Consequently, figure 6 shows the variation of SNIR for varying k for the signal/jammer configuration discussed above.

For values of k between approximately 3 and 20 the SNIR is seen to exhibit a fairly flat plateau. However, as k is decreased below 3 it can be seen that the SNIR falls, levelling off at a value of approximately 3dB as k approaches 0 and we return to the extreme weight jitter situation of figure 3.

For values of k greater than 20 there is a much more rapid decline in SNIR. This is due to a reduction in the depth of the null at 60°. As was pointed out above as k becomes larger the adapted pattern is forced closer to the quiescent pattern until in the limit as $k \rightarrow \infty$ they become identical. In this extreme case the sidelobe would be at approximately -15dB hence providing an insufficient amount of nulling for the 30dB jammer.

5. CONCLUSION

It has been shown how a poor estimate of the covariance matrix can cause weight jitter and how this may cause a deleterious effect for adaptive processing when using a sonobuoy (although this is generally true for adaptive sonar).

We have presented, in brief, two methods by which weight jitter effects can be significantly reduced.
The penalty function method which is simple to apply and understand and, at least for the examples presented, computationally simple to implement. It requires only a choice of the function $h(\theta, \phi)$ which determines the parts of the beampattern which are most closely controlled and a parameter k which determines the overall fit of the beampattern. (The best value of this parameter has been shown to be directly dependent on the noise power [4] which must be evaluated by separate means). The full power of this method by which $h(\theta, \phi)$ is chosen to use a priori information has not been investigated here.

A second, eigenvalue based, approach in which the jammer and noise subspaces are separated and the weight jitter removed by, in effect, forcing the noise eigenvalues to be equal. This simple method, however, can be usefully extended to the 'double average method' which uses an estimate of the noise gained over long periods of time combined with jammer statistics averaged over comparatively short periods of time (appropriate to stationary noise and intermittent jamming).

It was shown how the penalty function method can be used to reduce weight jitter substantially whilst still retaining nulling. Further more testing evidence of the efficacy of the method can be found in the literature. We do no more than note in passing that similar results can be achieved with the eigenvalue approach[5].

It is clear, then, that the two methods discussed can be used to decrease the effects of jitter but as distinct methods can also offer extra advantages depending on *a priori* information.

6. REFERENCES

1. See for instance BD Van Veen and KM Buckley. "Beamforming: a versatile approach to spatial filtering." *IEEE ASSP Magazine*, pp4-24 April 1988.
2. AJ Fenwick and IJ Clarke. "Low sidelobes for arbitrary arrays". See proceedings of this conference.
3. Monzingo and Miller. Introduction to adaptive arrays. J Wiley (1980).
4. DT Hughes and JG McWhirter. "Penalty function method for sidelobe control in least squares adaptive beamforming". *Advanced Signal Processing Algorithms*. pp 170-181 SPIE 1995. Editor FT. Luk.
5. DT Hughes and IJ Clarke. In preparation.