Proceedings of the Institutue of Acoustics

CHARACTERISATION OF THE ITERATIVE MULTI-PARAMETER (IMP) ALGORITHM

J L Mather

Royal Signals and Radar Establishment, St Andrews Road, Malvern, Worcs, WR14 3PS

1. SUMMARY

The incremental multi-parameter (IMP) algorithm will be described. This algorithm is used to analyse data from arrays of sensors in order to estimate signal parameters. It makes use of the conventional beamformer as a key component of an iterative analysis of the data, yet is able to outperform alternative "high resolution" estimators such as MUSIC by 10-15dB. The algorithm is suitable for use with general array geometries, and is also capable of handling single snapshot data and coherent signals. This paper outlines the algorithm and compares the performance of IMP, MUSIC and RootMUSIC, as assessed by Monte Carlo trials.

2. INTRODUCTION

Conventional processing (beam scanning, or beamforming) of received data from sensor arrays may be looked upon as a form of "spectral" analysis. If the array is linear, with $\lambda/2$ interelement spacing, then the conventional estimate for the far field distribution of received power is a padded Fourier transform of the data. Major peaks in the spectrum are assigned to localised signals, the spectral frequency corresponding to direction of arrival. The spectral peaks (or lobes) are broad, and a large antenna aperture is required to "resolve" signals which are close to each other.

This view of array processing as an analogue of spectral estimation has been largely responsible for the current emphasis on high resolution algorithms such as MLM [1,2] and more recently MUSIC [3,2]. These techniques result in "spectra" with much narrower, sharper peaks. Thus signals may be resolved with much smaller antenna apertures than for conventional analysis. In addition, "sidelobe" levels may be suppressed, apparently without the need for effectively arbitrary aperture weighting functions, such as Hamming or Chebychev [4]. As in the conventional case, spectral peak locations are used as indications of signal angles of arrival.

However, the resolution enhancements realised by many algorithms such as MUSIC, are brought about by throwing away information in order to "improve" the resulting spectrum. In fact, the principal practical distinction between algorithms of this class lies in the effectively arbitrary weighting functions used to emphasise or discard components of the data [5]. The conventional beamformer, on the other hand, does not discard information.

It should be remembered that the aim here is to extract details of signal positions and powers, and not to estimate a spatial spectrum. The spectrum is simply an aid to interpretation of the data, and has no physical significance, unlike a transmit beam. Recognition of this fact is leading increasingly to the study of approximate maximum likelihood parameter extraction techniques, which are designed specifically to extract the desired information. Examples include alternating projection [6], nonlinear regression [7], EM

[8], and the recently developed iterative multi-parameter (IMP) algorithm [9]. IMP uses the information preserving conventional beamformer as a basic processing component.

This paper describes the IMP algorithm (section 3) and presents results which illustrate its performance (section 4). These results have shown that, under ideal conditions, its resolution performance can exceed that of the popular MUSIC algorithm by the equivalent of a 10 to 15 dB increase in signal to noise ratio. Although similar performance sometimes may be achieved by Root MUSIC [10], this algorithm is only applicable to data from regular linear arrays. IMP may be applied to data from general non-linear arrays, and is more appropriate to continuous update tracking problems. In addition IMP is inherently capable of resolving coherent signals [11].

3. THE INCREMENTAL MULTI-PARAMETER (IMP) ALGORITHM

3.1. THE DATA MODEL

The usual data model employed [2] is as follows

$$\underline{d}(t) = \mathbf{M}\ \underline{f}(t) + \underline{w}(t) \tag{1}$$

where $\underline{f}(t)$ is a vector representing the input which is to be reconstructed, **M** is a linear transformation matrix, $\underline{w}(t)$ is a vector sample of zero mean Gaussian white noise, and $\underline{d}(t)$ is the resulting data vector, or "snapshot", at time t. We will assume for simplicity that the matrix **M** (often referred to as the array manifold [3]) is known to within a negligible calibration error [2]. For example, in the case of an array of n sensors expected to receive signals from independent point sources, **M** will be an $(n \times N)$ matrix, whose N columns (denoted $\underline{m}(\theta_i)$, $i = 1$ to N) represent the independent spatial transformations of *calibration* signals from N possible discrete angles, θ_i. Thus, **M** contains a representative subset of the continuum of *possible* received waveforms: it provides calibration information about the array manifold rather than details of specific signal sources. If $\underline{f}(t)$ represents the complex amplitudes of the signals associated with m independent point sources, as measured at a given instant, $\underline{d}(t)$ will be given by the linear combination of m corresponding columns of **M**, scaled by the signal amplitudes and perturbed by additive noise. From a reconstruction of $\underline{f}(t)$, we hope to locate the m sources and estimate their powers.

3.2. CONVENTIONAL BEAMFORMING

The conventional (beamforming or beam scanning) method of solution is to use the calibration matrix, **M**, to form a set of filters which are "matched" to each of the potential signal directions, θ_i, and to evaluate

$$P(\theta_i,t) = |f'(\theta_i,t)|^2 = \underline{m}^H(\theta_i)\ \underline{d}(t)\ \underline{d}^H(t)\ \underline{m}(\theta_i)\ /\ \{\underline{m}^H(\theta_i)\ \underline{m}(\theta_i)\}\ , \quad i = 1 \text{ to } N. \tag{2}$$

The superscript H denotes the complex conjugate (Hermitian) transpose, $|\underline{x}|^2$ denotes the squared magnitude of the individual elements of the vector, $\underline{x}$, and $\underline{m}^H(\theta_i)$ is a row of the matrix $\mathbf{M}^H$. This is a simple estimate of the spatial power distribution of the input, $\underline{f}(t)$. Such processing may be considered as "scanning" the data with the beamforming weight vector, $\underline{m}^H(\theta_i)$. $P(\theta,t)$ has the familiar broad multiple lobed pattern of classical analysis, with consequent poor discrimination of multiple signals, resulting from the wide beamwidth and high sidelobes.

If P data snapshots have been taken, and are represented by the matrix **D**, then the above equation may be re-written and evaluated for the block of data (averaging over the P

snapshots),

$$P(\theta_i) = \underline{m}^H(\theta_i)\ D\ D^H\ \underline{m}(\theta_i)\ /\ \{\underline{m}^H(\theta_i)\ \underline{m}(\theta_i)\}\ ,\quad i = 1 \text{ to } N. \quad (3)$$

The location of the principal peak of this "spectrum" gives the optimal (maximum likelihood) estimate of the angular location (spatial frequency) of a single point source of signal in an additive isotropic noise background [12].

3.3. IMP

The conventional beam scan, described in section 3.2, essentially constitutes the first stage of the IMP algorithm. Assuming that only a single signal is present, the beamformer will give the best estimate of its position, $\underline{\hat{m}}_1$ (dropping the θ notation for clarity). Equation (3) may be modified for the case of a known non-isotropic noise background [5,9].

Subsequent stages of the algorithm make use of the modified beam scan defined by

$$P(\theta_i) = \underline{m}^H(\theta_i)\ Q\ D\ D^H\ Q\ \underline{m}(\theta_i)\ /\ \{\underline{m}^H(\theta_i)\ Q\ \underline{m}(\theta_i)\}\ ,\quad i = 1 \text{ to } N. \quad (4)$$

Q is a projection matrix, which projects the data into the subspace orthogonal to that spanned by the steering vectors which correspond to the latest estimates of signal angles of arrival. Thus, in order to test for the presence of a second signal we use

$$Q = I - \underline{\hat{m}}_1\ \underline{\hat{m}}_1^H\ /\ \{\underline{\hat{m}}_1^H\ \underline{\hat{m}}_1\}\ , \quad (5)$$

where I is an identity matrix. The normalisation term in the denominator of equation (4) is central to the algorithm. However, application of equation (5) creates nulls in both the numerator and the denominator of equation (4). The resulting instances of 0/0 must be trapped, and in practice it seems sufficient to limit the denominator to 10^{-6}. If the scan of equation (4) produces no significant peaks as a function of θ, then the algorithm terminates. If sufficient residual power exists in the modified beamformed output, the principal peak with respect to θ is taken as the initial estimate of the location of the second signal, $\underline{\hat{m}}_2$.

If the equivalent true steering vectors, $\underline{m}_1$ and $\underline{m}_2$, are correlated, then the initial estimates will be biased and need to be refined. This is done by repeatedly recomputing equation (4) in the region of interest, forming Q from each of the latest estimates, $\underline{\hat{m}}_1$ or $\underline{\hat{m}}_2$, in turn. Thus, the first iteration re-estimates $\underline{\hat{m}}_1$, using a projection based on $\underline{\hat{m}}_2$, and so on until the estimates become stable. We now have the best estimate of angles of arrival assuming the presence of two signals.

To test for higher model orders (larger numbers of signals), the same procedure is followed. At the i^{th} stage, Q takes the form

$$Q = I - M_s\ (M_s^H\ M_s)^{-1}\ M_s^H \quad , \quad (6)$$

where M_s is a matrix whose i-1 columns are estimated steering vectors. The entire algorithm is summarised in Fig. 1. As further illustration, a typical sequence of beamscans is shown in Fig. 2. IMP is shown resolving three signals located at 0, 0.3 and 0.7 beamwidths from broadside of a 16 element linear array. Fig. 3 shows the convergence history for this particular experiment.

The decision to increment model order at each stage will depend on the definition of "significance" in the test for peaks in the residual spectrum. Clearly, the spectrum will only be completely flat when all degrees of freedom are used, and this is likely to be much greater than the number of signals present. In practice, some knowledge of the noise statistics, as perceived through a limited number of samples, is required. This may be used to specify a criterion for "flatness" in the residual spectrum. This test enables the algorithm to determine the number of signals [9] without the need for tests such as the "minimum description length" measure [15].

THE BASIC IMP ALGORITHM

```
N = 0              ! N is the estimated number of targets, initialised to 0.
Beamform           ! Evaluate equation (4), with Q = I.
IF no significant peak THEN
    Exit           ! If P(θ) has no significant peaks, assume no signals present and quit.
ELSE
    Assign peak to m̂_1
    N = 1
END IF
FOR i = 1 to Maxorder       ! Maxorder is the maximum number of signals, usually
                            ! 1 less than the number of sensor array elements.
    FOR j = 1 to N_it       ! N_it is the number of iterations at each stage.
        FOR k = 1 to N
            Calculate M_s       ! M_s is the matrix [m̂_1 ... m̂_N].
            Calculate Q         ! Using equation (6),
            Beamform            ! try to find (N+1)^th signal.
            IF no significant peak THEN      ! If none, then assume no
                Estimate signal powers       ! more signals and exit.
                Exit
            ELSE
                Assign peak to m̂_N+1
            END IF
            IF j < N_it THEN
                FOR p = 1 to N
                    m̂_p = m̂_p+1
                NEXT p
            END IF
        NEXT k
    NEXT j
    N = N + 1
NEXT i
Estimate signal powers          ! See [2] or [13].
END
```

Fig. 1. A summary of the IMP algorithm. The detailed implementation is sensitive to the method employed for estimation of peak positions, the number of iterations at each stage, and the criterion used for termination of the iteration.

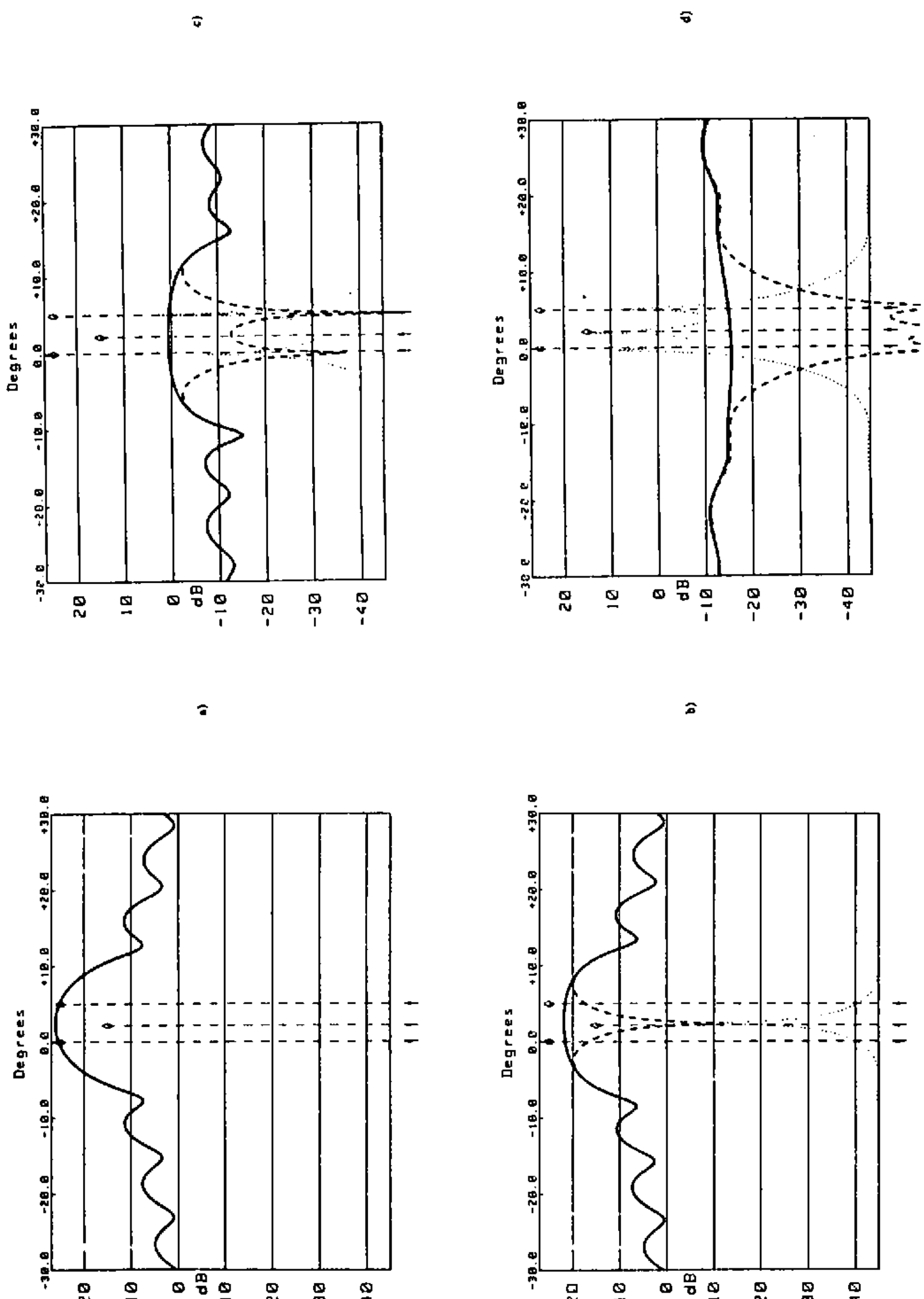

Fig.2. Estimation of angles of arrival of three signals, using IMP. Data is from a 16 element 0.5λ spaced linear array. Signals are as indicated by vertical broken lines, located at 0, 0.3 and 0.7 beamwidths away from broadside, and with S/N ratio per element indicated by height of lines. a) Initial beamscan, assuming isotropic noise background. First signal estimate is at peak position. b) Broken line shows numerator of beamscan (equation 4) with first null; dotted line shows (inverted and normalised to base of graph for clarity) denominator of equation 4; solid line shows resulting normalized beamscan at end of first stage. Peak of solid curve is initial estimate of second signal angle of arrival. c) At end of second stage of algorithm, broken line shows numerator of beamscan (equation 4) with first and second nulls; dotted line shows (inverted and normalised to base of graph for clarity) denominator of equation 4; solid line shows resulting normalized beamscan. Peak of solid curve is initial estimate of third signal angle of arrival. d) All three angles of arrival have been estimated. Since the residual beamscan (solid line) is approximately flat and lies close to the integrated noise level, the algorithm is terminated.

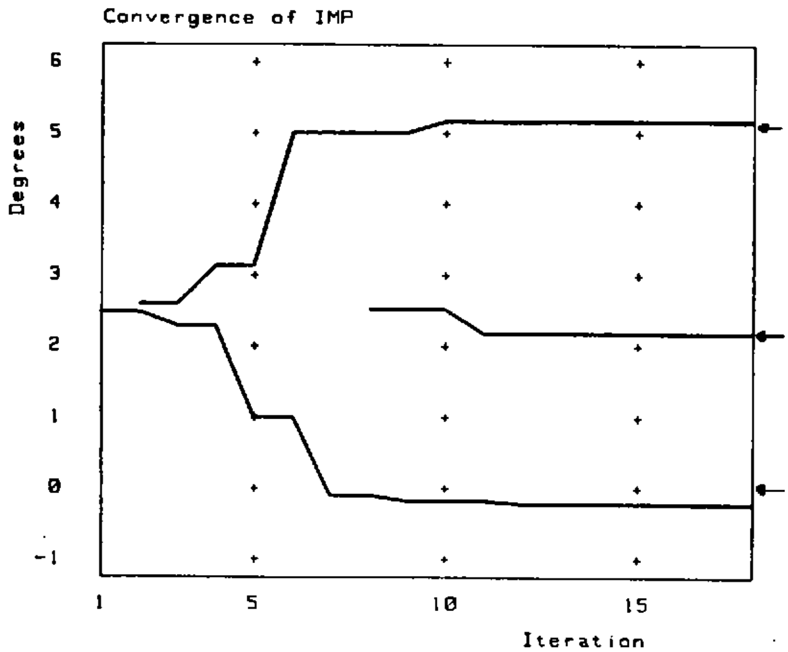

Fig. 3. Convergence of the IMP algorithm for the experiment of Fig. 2 (16 snapshots of data, with three random phase signals at 0, 2.15 and 5.02 degrees from broadside of a 16 element 0.5λ spaced linear array). The solid lines indicate the estimated signal positions for each iteration of the algorithm.

4. MONTE CARLO RESULTS

4.1. INTRODUCTION

The performance data presented here are typical of results measured in a series of Monte Carlo trials of IMP, MUSIC and Root MUSIC [11]. The data matrices, consisting of 16 snapshots, were generated from simulations of two signals in the far field with additive background noise. A perfectly calibrated 16 element 0.5 wavelength spaced linear array was assumed.

The procedure used for the collection and analysis of the Monte Carlo trial data is described in [11]. Briefly, the statistics presented here are based on 100 trials of each algorithm at each signal to noise ratio (in 3dB increments). The results presented are conditioned on resolution. That is to say, bias and variance were only calculated if both signals were detected. Resolution is said to occur if two signals are found at approximately the correct (angle, power) coordinates. Acceptability is defined by margins ΔP and $\Delta\theta$ on the known coordinates. "Array signal to noise ratio" (ASNR) in Fig. 4 is a normalised measure equal to ((instantaneous signal to noise at each element of the array) + 10 $\log_{10}(n)$).

Section 4.2 describes the results obtained. More comprehensive results for IMP, MUSIC, Root MUSIC and a number of other popular algorithms are presented in [11] and [13].

4.2. EXPERIMENTAL RESULTS.

The target scenario consists of two equal power random phase sinusoidal signals in the far field, the first located at the broadside position (perpendicular to the line of the array), and the second at 0.1 beamwidths to one side. The beamwidth referred to here is the angle from the peak of the main lobe of a matched filter (equation (2)) placed on the first target location, to the position of the first null.

Fig. 4 shows the variation of the performance statistics as a function of ASNR. The results, taken at 3dB signal to noise ratio increments, are shown for IMP (heavy solid line), MUSIC (dotted line) and Root MUSIC (broken line). The fine solid line in the plot of angular standard deviation is the Cramér Rao bound [14] for the problem, assuming uncorrelated emitters. A noise power threshold of $10\log_{10}(n)$ has been used, together with $\Delta\theta$ = ±0.5 beamwidths and ΔP = ±3dB.

As the ASNR rises, the probability of resolution increases (associated with a peak in the false alarm rate), bias of the angle and power estimates tend to zero, and the variance of the angle estimates, $\mathrm{var}(\hat{\theta}_i)$, tend towards the Cramér Rao bound. The bias and variance results are plotted for the "left-hand" signal, and thus negative angular bias indicates that the $\hat{\underline{m}}_i$ corresponding to the estimated signals have moved further apart.

As can be seen in Fig. 4, the performance of IMP clearly exceeds that of either MUSIC or Root MUSIC. A significant probability of resolution is achieved at ASNRs between approximately 10 and 15 dB lower than required by MUSIC, and up to 5dB below that required by Root MUSIC. The estimation bias, variance and false alarm rates of IMP are all either lower than or similar to those of the other algorithms.

Further experiments [11] have led us to believe that the greater performance of IMP stems from the fact that the algorithm is sensitive to the effective spatial "narrowband" signal to noise ratio. This corresponds to the "visibility" of the signal above the local background, and is influenced by the level of additive noise in the local region, and also by the presence of neighbouring strong signals. MUSIC, on the other hand, is rather more sensitive to the spatial "wideband" average noise level, through its requirement to separate a "noise"-only eigen-subspace.

5. CONCLUSIONS

We have reviewed the incremental multi-parameter (IMP) algorithm as an analysis tool for data collected from sensor arrays. We have shown it to have both potentially higher performance and more general applicability than the popular MUSIC algorithm. In particular, we have demonstrated

- resolution of two equal power signals, using a linear array, at signal to noise ratios 10 to 15 dB lower than required by MUSIC;
- performance comparable to that of Root MUSIC for problems involving a perfect linear array and partially uncorrelated signals;

In addition, some observations have been made regarding convergence of the algorithm.

Finally, we note that although, during the initial target acquisition stages, the computational requirement of the version of IMP described here is frequently greater than for MUSIC, its incremental and iterative structure makes it far more efficient for tracking and updating of initial estimates. It is clearly possible to build model order decrementation into the algorithm to allow for disappearing signals. It is also possible to conceive of running multiple versions of the algorithm in parallel, each "tuned" to a different model order. It would then be very simple to choose the appropriate output depending on estimated signal strengths for example. Clarke (private communication) has recently developed significantly faster implementations of the algorithm.

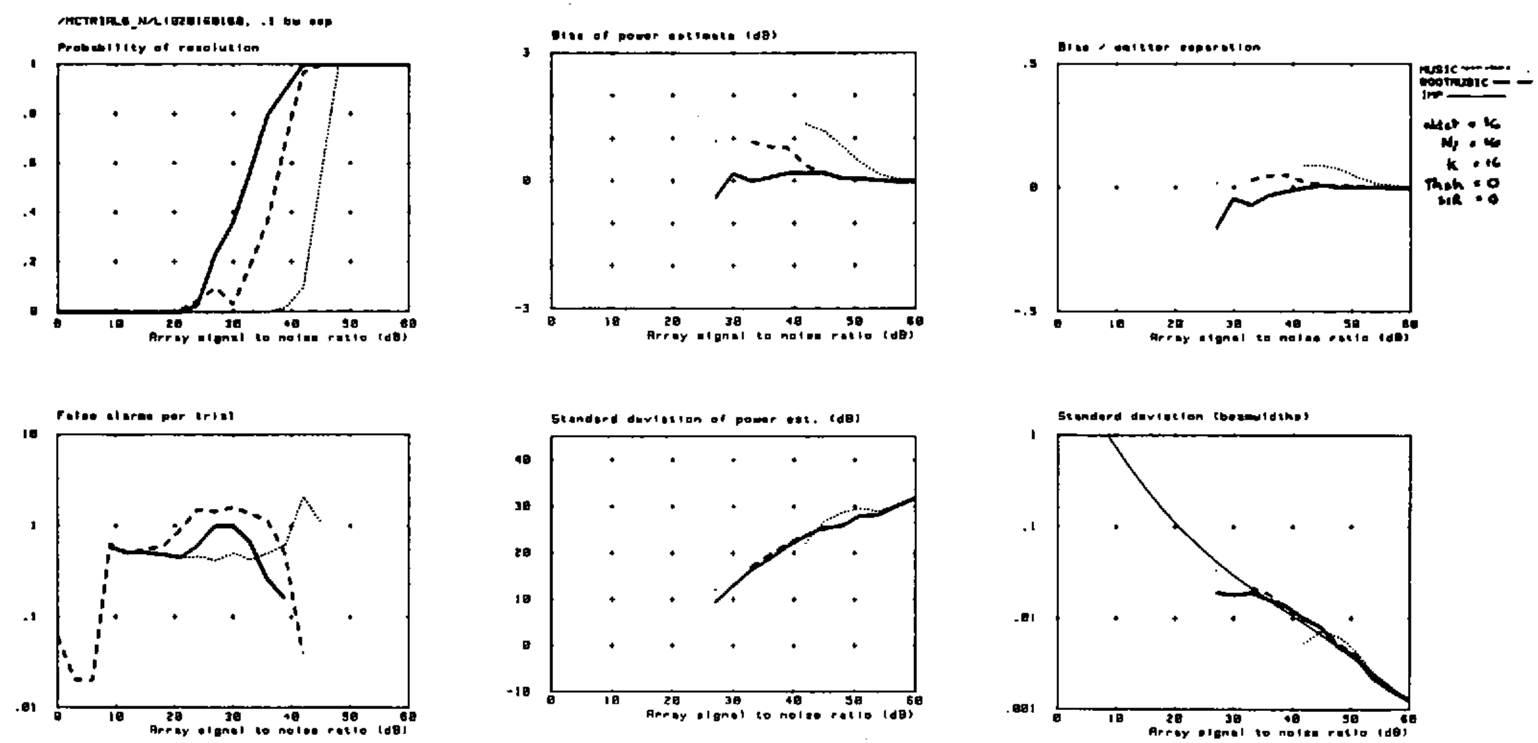

Fig. 4. Performance statistics as a function of array signal to noise ratio (ASNR) for two signals at 0 and 0.1 beamwidths from broadside. Results are based on 100 trials at each ASNR, taking 16 snapshots of data from the 16 element 0.5 wavelength spaced linear array. Dotted lines show the results of processing using the MUSIC algorithm [3]; broken lines show the results obtained using RootMUSIC [10]; the solid lines show the results obtained using IMP (equation 4). The fine continuous line in the plot of angular standard deviation is the Cramér Rao bound.

6. REFERENCES

[1] Capon, J., Greenfield, R.J., Kolker, R.J. "Multidimensional maximum likelihood processing of a large aperture seismic array." Proc. IEEE 55(2), 192-211, (1967).

[2] Mather, J.L. "Least squares solutions in signal processing using the singular value decomposition." Royal Signals and Radar Establishment Memorandum no. 3864, (1986).

[3] Schmidt, R.O. "Multiple emitter location and signal parameter estimation." RADC Spectrum Estimation Workshop, 243-258, (1979).

[4] Harris, F.J. "On the use of windows for harmonic analysis with the discrete Fourier transform." Proc IEEE 66, 51-83, (1978).

[5] Clarke, I.J. "High discrimination target detection algorithms and estimation of parameters." Underwater acoustic data processing, ed Y T Chan, Kluwer Academic, (1989).

[6] Ziskind, I., Wax, M. "Maximum likelihood localization of multiple sources by alternating projection." IEEE Trans ASSP-36(10), 1553-1560, (1988).

[7] Böhme, J.F. "Source parameter estimation by approximate maximum likelihood and nonlinear regression." IEEE J Oceanic Eng, OE-10, July 1985.

[8] Feder, M., Weinstein, E. "Optimal multiple source location estimation via the EM algorithm." IEEE Proc ICASSP, 1762-1765, (1985).

[9] Clarke, I.J. "High discrimination detection bound and model order control." Proc SPIE 975, 344-351, (1988).

[10] Barabell, A.J. "Improving the resolution performance of eigenstructure-based direction-finding algorithms." IEEE Proc. Int. Conf. Acoustics Speech Signal Processing 1, 336-339, (1983).

[11] Mather, J.L. "Performance of the IMP array processing algorithm: first results." Royal Signals and Radar Establishment Memorandum 4291, (1989).

[12] Rife, D.C., Boorstyn, R.R. "Single tone parameter estimation from discrete-time observations." IEEE Trans IT-20(5), 591-598, (1974).

[13] Mather, J.L. "A Monte Carlo performance analysis of accelerated svd-based high discrimination algorithms." Royal Signals and Radar Establishment Memorandum 4083, (1988).

[14] van Trees, H.L. "Detection, estimation, and modulation theory, I." Wiley, N.Y., (1968).

[15] Wax, M., Kailath, T. "Determining the number of signals by information theoretic criteria.", Proc IEEE Spectrum Estimation workshop II, 192-196, (1983).