Bayesian analysis of linear phased-array radar

Andrew G. Green* and David J.C. MacKay†
Cavendish Laboratory
Cambridge, CB3 0HE. United Kingdom.

ABSTRACT. A number of methods have been developed to analyze the response of the linear phased array radar. These perform remarkably well when the number of sources is known, but in cases where a determination of this number is required, problems are often encountered. These problems can be resolved by a Bayesian approach.

Here, a linear phased-array consisting of equally spaced elements is considered. Analytic expressions for the posterior probability distribution over source positions and amplitudes, and the corresponding Hessians are derived. These are integrated to give the evidence for each model order.

Tests using model data showed that performance at the second level of inference is critically determined by the accuracy of position estimation. If adequate parameter optimization is available, the Bayesian approach is demonstrated to work well, even in extreme circumstances. A commonly employed method of source location, noise subspace eigenanalysis of the correlation matrix, was tried and found to be inadequate. A Newton-Raphson optimization was then used starting from the positions predicted by eigenanalysis.

1 Introduction

We investigate the analysis of data from linear phased-array radar. Recent improvements in the speed of computers have made feasible the real-time use of more sophisticated methods than simple beam-sweep methods. One technique of interest is noise sub-space eigen-analysis of the correlation matrix [1]. This is one of a class of algorithms commonly known as super-resolution methods, because of their ability to resolve sources below the Rayleigh criterion [2, 3].

The merits of Bayesian inference have been demonstrated in many diverse fields of data analysis [4, 5, 6, 7]. Here, the improvements which may be made to position inference by eigen-analysis, with the application of Bayesian methods, is assessed. The locations of a finite number of sources are inferred, with error bars, by maximizing the integrated posterior probability. The number of sources is similarly inferred by evaluating the appropriate evidence.

2 Radar Configuration

The phased-array radar consists of a series of equally spaced elements with (ideally) isotropic far field responses. This arrangement is indicated in figure 1. For each source configuration, a number of data-sets, S, are collected in rapid succession. These are known as snapshots.

*andrewg@thphys.ox.ac.uk
†mackay@mrao.cam.ac.uk

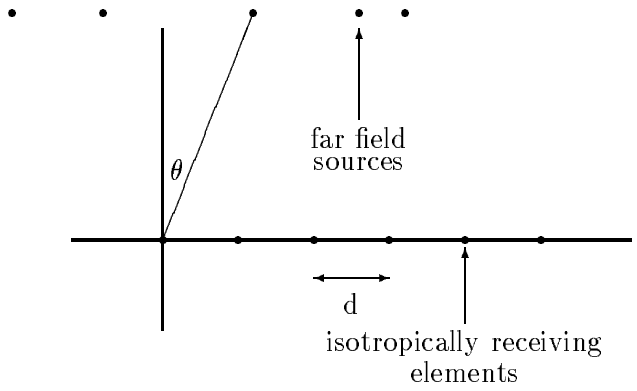

Figure 1: Antenna Configuration.

d: element separation; λ: wavelength; θ: source positions.

For a single snapshot the phase difference between the signal at adjacent elements of the antenna due to the j^{th} source is given by:

$$\phi_j = (2\pi d/\lambda)\sin\theta_j \ . \tag{1}$$

The response of the k^{th} element may be written as

$$x_k = \sum_j F_j \exp i(k\phi_j) + n_k \ . \tag{2}$$

This is more elegantly expressed in terms of matrices,

$$\underline{x} = \mathbf{V}\underline{F} + \underline{n} \ , \tag{3}$$

where $\underline{x}$ is the antenna response, $\underline{n}$, the noise vector, $\underline{F}$, the source amplitude vector and $\underline{\phi}$, the source position vector. The steering matrix $\mathbf{V}$ is given by

$$V_{kj} = \exp(ik\phi_j) \ . \tag{4}$$

The source locations are assumed not to change between snapshots, although their complex amplitudes may. We have, therefore, a single position vector $\underline{\phi}$ for all snapshots and a set of amplitude vectors $\underline{F}^s$, one for each snapshot.

3 Eigen-analysis of correlation matrix

There are a number of inter-related techniques in spectrum analysis based upon the eigen-analysis of the data correlation matrix [2, 3]. The method described in brief below, due to Reilly *et al.* [1], is one of several super-resolution algorithms known as "maximum likelihood". These can be shown to depend upon the orthogonality relationship given at the end of this section [2, 3].

The correlation matrix is formed by taking the outer product of the data vector with itself. Using equation (3) we find

$$\underline{x}\underline{x}^\dagger = \mathbf{V}\underline{F}(\mathbf{V}\underline{F})^\dagger + \underline{n}\underline{n}^\dagger + \mathbf{V}\underline{F}\underline{n}^\dagger + \underline{n}(\mathbf{V}\underline{F})^\dagger \; . \tag{5}$$

Averaging over snapshots we note that the final two terms will ideally go to zero, since the sources and noise are uncorrelated. The matrix $\mathbf{V}$ may be removed from the averaging since the source positions are the same for all snapshots. Equation (5) may then be written

$$\mathbf{R} = \mathbf{V}\,\mathbf{Rs}\,\mathbf{V}^\dagger + 2\sigma^2\mathbf{I} \; , \tag{6}$$

where $\mathbf{R} = \langle \underline{x}\underline{x}^\dagger \rangle$ is the correlation matrix, $\mathbf{Rs} = \langle \underline{F}\underline{F}^\dagger \rangle$ is the source correlation matrix, $\mathbf{I}$ is the identity matrix, σ^2 is the variance of real and imaginary parts of the noise, and $\langle\,\rangle$ indicates averaging over snapshots. Note that the noise power is assumed identical for each element of the antenna.

$\mathbf{R}$ is Hermitian and positive definite by construction. For an N element antenna in the presence of k uncorrelated sources, $\mathbf{R}$ will have a series of real positive eigenvalues of decreasing magnitude, that is,

$$\lambda_1 + 2\sigma^2 \geq \lambda_2 + 2\sigma^2 \geq \ldots \geq \lambda_k + 2\sigma^2 > 2\sigma^2 = \ldots = 2\sigma^2 \; . \tag{7}$$

Multiplying equation (6) by one of the $(N-k)$ noise subspace eigenvectors (*i.e.* an eigenvector $\underline{e}_{noise}$ of $\mathbf{R}$ corresponding to an eigenvalue $2\sigma^2$), one obtains the orthogonality relation

$$\mathbf{V}^\dagger \underline{e}_{noise} = 0 \; . \tag{8}$$

If an average correlation matrix of order $(k+1)$ is formed by averaging $(k+1)\times(k+1)$ minors along the leading diagonal of $\mathbf{R}$, there will be only one noise subspace eigenvector, corresponding to the lowest eigenvalue. In this case, taking the Z-transform of the noise eigenvector, one obtains a polynomial whose solutions are $z_k = \exp(i\phi_k)$ for each of the k sources present.

There exist more sophisticated ways of averaging $\mathbf{R}$ to form an order $(k+1)$ matrix [3]. These have the property of resolving correlated sources (*i.e.* where F_i^s is correlated with F_i^{s+1} or F_j^s) as well as uncorrelated sources. These are disregarded here for simplicity. An alternative method due to Burg et al. [8] involves finding the maximum likelihood Toeplitz structure matrix from the data.

4 Bayesian analysis

Single snapshot

The inferred parameters divide naturally into: the source amplitudes $\{\underline{F}^s\}$ (which are different for each snapshot), the source positions $\underline{\phi}$, and the number of sources k, (specified by the hypothesis H). Initially a single snapshot will be considered. This is extended to several snapshots in the following subsection.

We write Bayes' theorem for a series of levels of inference, as follows;

$$P(\underline{F}|H,D,\underline{\phi}) = \frac{P(D|H,\underline{F},\underline{\phi})P(\underline{F}|H,\underline{\phi})}{P(D|H,\underline{\phi})} \; , \tag{9}$$

$$P(\underline{\phi}|H,D) = \frac{P(D|H,\underline{\phi})P(\underline{\phi}|H)}{P(D|H)} \ , \tag{10}$$

$$P(H|D) \propto P(D|H)P(H) \ . \tag{11}$$

We note that $P(\underline{F}|H,\underline{\phi}) = P(\underline{F}|H)$, since $\underline{F}$ and $\underline{\phi}$ are independent.

We assume that these distributions are strongly peaked and may be approximated by Gaussians about their peaks. In fact given the choice of likelihood and priors to be made later, equation (9) is exactly Gaussian. By making a Gaussian expansion of equation (9) and integrating, we obtain the expression for the evidence,

$$P(D|H,\underline{\phi}) = P(D|H,\underline{F}_M,\underline{\phi})P(\underline{F}_M|H)(2\pi)^k \det{}^{-1}\mathbf{A}(\underline{\phi}) \ , \tag{12}$$

where $\underline{F}_M = \underline{F}_M(\underline{\phi})$ is the $\underline{F}$ that maximizes (9) and the Hessian $\mathbf{A}$ is given by

$$\mathbf{A}(\underline{\phi}) = -\nabla_{\underline{F}}\nabla_{\underline{F}} \ln P(\underline{F}|H,D,\underline{\phi}) \ . \tag{13}$$

Substituting for $P(D|H,\underline{\phi})$ from equation (12) in equation (10), with the Hessian,

$$\mathbf{B} = -\nabla_{\underline{\phi}}\nabla_{\underline{\phi}} \ln P(\underline{\phi}|H,D) \ , \tag{14}$$

gives upon integration,

$$P(H|D) \propto P(D|H,\underline{F}_M,\underline{\phi}_M)P(\underline{F}_M|H)P(\underline{\phi}_M|H)(2\pi)^{\frac{3k}{2}} \det{}^{-1}\mathbf{A} \det{}^{-\frac{1}{2}}\mathbf{B} \ , \tag{15}$$

where

$$\mathbf{A}(\underline{\phi}) = -\nabla_{\underline{F}}\nabla_{\underline{F}} \Big[\ln P(D|H,\underline{F},\underline{\phi}) + \ln P(\underline{F}|H)\Big] \tag{16}$$

and

$$\mathbf{B} = -\nabla_{\underline{\phi}}\nabla_{\underline{\phi}} \Big[\ln P(D|H,\underline{F}_M,\underline{\phi}) + \ln P(\underline{F}_M|H) + \ln P(\underline{\phi}|H) - \ln \det{}^{-1}\mathbf{A}(\underline{\phi})\Big] \ . \tag{17}$$

Equations (16) and (17) have been obtained from (13) and (14) by substitution from (9) and (10), noting that the normalizing factors, $P(D|H,\underline{\phi})$ and $P(D|H)$ are constant with respect to the differentiating variables, $\underline{F}$ and $\underline{\phi}$ respectively. The determinants of $\mathbf{A}$ and $\mathbf{B}$ appear to different powers in equation (15) because $\underline{F}$ is a complex vector and $\underline{\phi}$ is a real vector.

Extension to several snapshots

In the case of several data sets or snapshots the above theory must be modified. The positions of the sources, given by $\underline{\phi}$, are the same for all snapshots. Their inference is based upon the data from all the snapshots taken together. The complex source amplitudes, however, may be different for each snapshot, giving rise to S source amplitude vectors $\{\underline{F}^s\}$, where S is the number of snapshots.

It follows from the product rule of probability that one must take the product of the likelihood and priors over snapshots. The amplitude vectors, $\{\underline{F}^s\}$, are assumed, as a rather crude first approximation, to have independent priors between snapshots. Then the likelihood $P(D|H,\underline{F},\underline{\phi})$ is replaced by $P^\pi(D|H,\underline{F}^s,\underline{\phi}) = \prod_{s=1}^{S} P(D|H,\underline{F}^s,\underline{\phi})$, and the prior

$P(\underline{F}|H)$ is replaced by $P^{\pi}(\underline{F}^s|H) = \prod_{s=1}^{S} P(\underline{F}^s|H)$. The prior on the source positions $\underline{\phi}$ is unchanged.

Using these distributions with Bayes' theorem we obtain the multi-snapshot analogue of equation (9),

$$P^{\pi}(\underline{F}^s|H, D, \underline{\phi}) = \frac{P^{\pi}(D|H, \underline{F}^s, \underline{\phi}) P^{\pi}(\underline{F}^s|H)}{P(D|H, \underline{\phi})} \,. \tag{18}$$

Equations (10) and (11) are used without alteration. Expanding the distributions as Gaussians about their maxima and integrating, as in section 4, we derive the final result,

$$P(D|H) \propto P^{\pi}(D|H, \underline{F}^s_M, \underline{\phi}) P^{\pi}(\underline{F}^s_M|H) P(\underline{\phi}_M|H) (2\pi)^{k(S+1/2)} \det{}^{-S}\mathbf{A} \, \det{}^{-1/2}\mathbf{B} \,, \tag{19}$$

where

$$\mathbf{A}(\underline{\phi}) = -\nabla_{\underline{F}} \nabla_{\underline{F}} \left[\ln P(D|H, \underline{F}^s, \underline{\phi}) + \ln P(\underline{F}^s|H) \right] \,, \tag{20}$$

and

$$\mathbf{B} = -\nabla_{\underline{\phi}} \nabla_{\underline{\phi}} \sum_{s=1}^{S} [\ln P(D|H, \underline{F}^s_M, \underline{\phi}) + \ln P(\underline{F}^s_M|H) + \ln P(\underline{\phi}|H) - \ln \det \mathbf{A}(\underline{\phi})] \,. \tag{21}$$

Equation (20) is identical to equation (16), and equation (21) is simply the sum over snapshots of equation (17).

Application to Phased-Array radar

For a single snapshot, assuming Gaussian noise, the likelihood function is

$$P(D|H, \underline{F}, \underline{\phi}) = \left(\frac{1}{2\pi\sigma^2} \right)^{N} \exp \left[-\frac{|\underline{n}|^2}{2\sigma^2} \right] \,. \tag{22}$$

The noise vector $\underline{n}$ is defined for each snapshot as the difference between the data and mock data, *i.e.*,

$$\underline{n} = \underline{x} - \mathbf{V}\underline{F} \,. \tag{23}$$

The noise variance σ^2 is assumed to be known for a particular antenna rather than included as a hyperparameter. Noise is generated in the antenna and can be measured. Taking the product over snapshots, the likelihood is given by

$$P^{\pi}(D|H, \underline{F}^s, \underline{\phi}) = \left(\frac{1}{2\pi\sigma^2} \right)^{SN} \exp \left[-\sum_{s=1}^{S} \frac{|\underline{n}|^2}{2\sigma^2} \right] \,. \tag{24}$$

The prior on positions is simply a uniform distribution between $\pm\frac{2\pi d}{\lambda}$, $P(\underline{\phi}|H) = \left(\frac{\lambda}{4\pi d}\right)^k$. The prior on source amplitudes is taken to be Gaussian with variance δ^2 on real and imaginary parts, where δ^2 is entered as a user defined parameter, *i.e.*,

$$P^{\pi}(\underline{F}^s|H) = \left(\frac{1}{2\pi\delta^2} \right)^{Sk} \exp \left[-\frac{1}{2\delta^2} \sum_{s=1}^{S} |\underline{F}^s|^2 \right] \,. \tag{25}$$

Using the above priors, the Hessian matrices $\mathbf{A}$ and $\mathbf{B}$ defined in equations (20) and (21) are found to be:

$$\mathbf{A}(\underline{\phi}) = \frac{\mathbf{V}^\dagger \mathbf{V}}{\sigma^2} + \frac{\mathbf{I}}{\delta^2}, \tag{26}$$

and

$$\begin{aligned} B_{ij} &= \frac{1}{2\sigma^2} \sum_{s=1}^{S} \underline{x}^{s\dagger} \Big[\mathbf{B1}_{(ij)} + \mathbf{B1}_{(ij)}^\dagger - \mathbf{B2}_{(ij)} - \mathbf{B2}_{(ij)}^\dagger - \mathbf{B5}_{(ij)} \\ &\quad - \mathbf{B2}_{(ji)} - \mathbf{B2}_{(ji)}^\dagger + \mathbf{B3}_{(ij)} + \mathbf{B3}_{(ij)}^\dagger + \mathbf{B4}_{(ij)} + \mathbf{B4}_{(ij)}^\dagger \Big] \underline{x}^s \\ &\quad + \mathrm{Tr} \left[-\mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \phi_j} \mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \phi_i} + \mathbf{A}^{-1} \frac{\partial^2 \mathbf{A}}{\partial \phi_i \partial \phi_j} \right] \end{aligned} \tag{27}$$

where

$$\begin{aligned} \mathbf{B1}_{(ij)} &= \frac{1}{\sigma^2} \left(\frac{\partial^2 \mathbf{V}}{\partial \phi_i \partial \phi_j} \mathbf{A}^{-1} \mathbf{V}^\dagger \right) \\ \mathbf{B2}_{(ij)} &= \frac{1}{\sigma^2} \left(\frac{\partial \mathbf{V}}{\partial \phi_i} \mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \phi_j} \mathbf{A}^{-1} \mathbf{V}^\dagger \right) \\ \mathbf{B3}_{(ij)} &= \frac{1}{\sigma^2} \left(\frac{\partial \mathbf{V}}{\partial \phi_i} \mathbf{A}^{-1} \frac{\partial \mathbf{V}^\dagger}{\partial \phi_j} \right) \\ \mathbf{B4}_{(ij)} &= \frac{1}{\sigma^2} \left(\mathbf{V} \mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \phi_j} \mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \phi_i} \mathbf{A}^{-1} \mathbf{V}^\dagger \right) \\ \mathbf{B5}_{(ij)} &= \frac{1}{\sigma^2} \left(\mathbf{V} \mathbf{A}^{-1} \frac{\partial^2 \mathbf{A}}{\partial \phi_i \partial \phi_j} \mathbf{A}^{-1} \mathbf{V}^\dagger \right) \end{aligned} \tag{28}$$

5 Results

Code was written to simulate the antenna response $\{\underline{x}^s\}$ for up to thirty-two snapshots of data. The simulated source environment consisted of up to five uncorrelated sources, with arbitrary position and amplitude, and Gaussian noise of arbitary amplitude. Using these data, the eigenanalysis procedure was tested, and the ability of equation (19) to evaluate the evidence for different model orders was determined.

Eigenanalysis

Without noise, and given the correct model order (*i.e.*, number of sources), the eigenanalysis predicted the position and amplitude of sources to within the computer accuracy as was expected. If the procedure was used assuming a model order greater than the actual, the extra sources were predicted to have zero amplitude to within the computer accuracy.

With Gaussian noise added to the data, eigenanalysis predicted positions and powers well, as long as the sources were well separated. Estimates of resolution for a range of noise powers were made by moving two unit amplitude sources together until the eigenanalysis predicted a single source of twice their amplitude at their average position. Resolution reduced with increasing noise power.

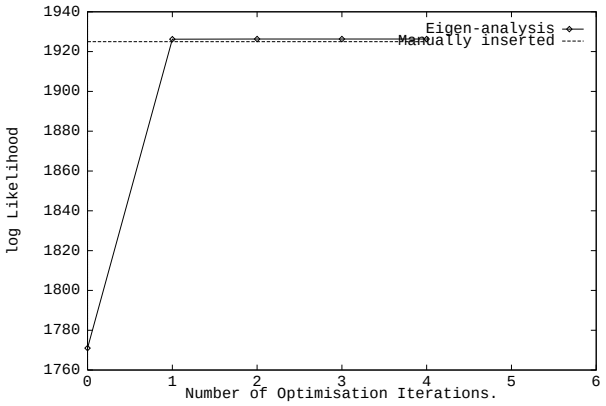

Figure 2: Variation of log Likelihood with Optimization iterations.

Eigenanalysis as described here depends upon the noise-subspace eigenvector, $\underline{e}_{\text{noise}}$, corresponding to the lowest eigenvalue of $\mathbf{R}$. For an S snapshot data-set with Gaussian noise, this eigenvalue is $\sim 2\sigma^2(1 \pm 1/\sqrt{S})$. For a source configuration containing two narrowly separated sources, the lowest source subspace eigenvalue, λ_k, reduces with separation and amplitude of the sources. It is evident from equation (7), that the resolution limit will occur when $\lambda_k \simeq \frac{2\sigma^2}{\sqrt{S}}$.

Does Eigenanalysis Maximize the Likelihood?

The parameters predicted by eigenanalysis do not maximize the likelihood. This conclusion was drawn for two reasons:

1. Manual insertion of the source parameters used to generate the data gave higher likelihoods than the parameters generated by eigenanalysis.

2. The log likelihood is expected to increase by ~ 0.5 for each additional parameter beyond the correct number. This was not the case. (In this case, an increase in k of 1 introduces $(1 + 2S)$ extra parameters giving an expected increase in log likelihood of about 32 between models.)

Although the predicted parameters do not maximize the likelihood, they do give a fair first approximation. A simple Newton-Raphson procedure was used to optimize the parameters predicted by eigenanalysis. (This was easily done, since the relevant Hessian matrix, $\mathbf{B}$, has already been evaluated using equation (27).) Initially, the integrated likelihood $P(D|H, \underline{\phi})$ was optimized by setting δ^2 to a very large value. Figure 2 shows typically how the likelihood increased with the number of iterations of the optimization routine. The likelihood increases above that for manual insertion of the parameters after only one iteration and remains fairly constant at this value through the subsequent iterations. Reassuringly, this indicates that the distributions are indeed Gaussian at their peaks.

Figure 3 shows the variation of log likelihood with model order. The expected increase is observed for each additional parameter beyond the correct model order.

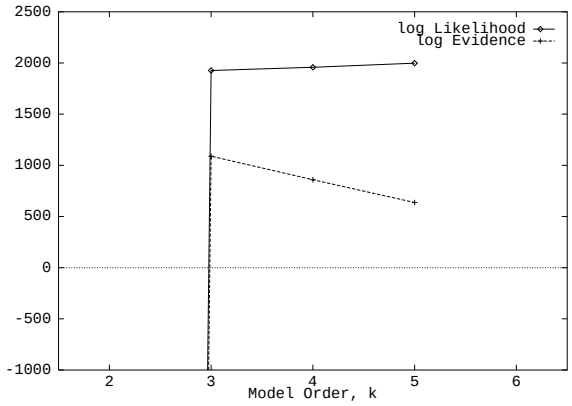

Figure 3: Variation of Likelihood and Evidence (Parameters Generated by Eigenanalysis and Newton-Raphson Optimization.)

$$\begin{aligned} \theta &= -0.2 \quad 0.1 \quad 0.3 \\ |F|^2 &= \quad 1 \quad\quad 1 \quad\quad 1/4 \end{aligned}$$
$$S = 32,\ N = 32,\ \sigma = 0.1,\ \delta = 1.0$$

Estimated Evidence.

Figure 3 shows the variation of log likelihood and log evidence with model order for a well resolved source environment with three sources. The evidence has a maximum at the correct model order. Figure 4 similarly shows the variation of log likelihood and log evidence, but for a situation where two of the sources have not been resolved by the eigenanalysis. Newton-Raphson optimization of the eigenanalysis parameters did not resolve these sources. The model order predicted by the evidence is correspondingly reduced.

Note that in cases such as that shown in figure 4, where the unresolved sources have large amplitude, the peak evidence is greatly reduced (as compared with cases where parameters are correctly evaluated). It is tempting to interpret this as an indication of error in the inferring of the parameters, however, it is not at all clear that such deductions may be drawn consistently.

Limitations of Model Comparison.

In all cases discussed up to now, Bayesian model comparison has worked well. No severe test of this level of inference has been made, due to the limitations of the techniques used to determine the source positions. In lieu of a good optimization method the following test of the model comparison was made.

Several well separated sources of unit amplitude were generated and a source at the noise level introduced at a small angle from one of these. Below the correct model order eigenanalysis was used to seed the Newton-Raphson optimization of parameters. At and above the correct model order, optimization was seeded with actual source positions cou-

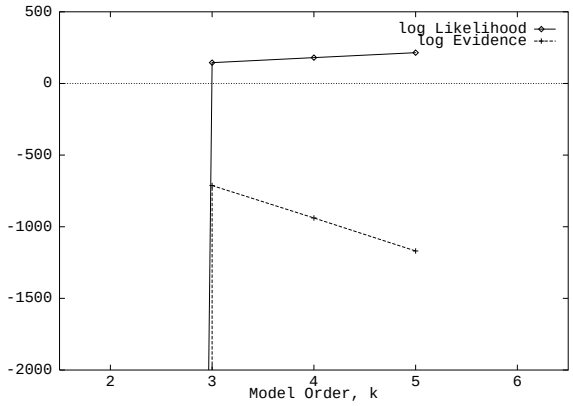

Figure 4: Variation of Likelihood and Estimated Evidence (Sources Unresolved by Eigenanalysis.)

$$\begin{array}{rccccc} \theta & = & -0.21 & -0.2 & 0.1 & 0.3 \\ |F|^2 & = & 1 & 1 & 1 & 1/4 \end{array}$$
$$S = 32,\ N = 32,\ \sigma = 0.1,\ \delta = 1.0$$

pled with spurious source positions, predicted by eigenanalysis (note that this amounts to starting the optimization in the correct place, so this is not a demonstration of the entire system; it is only a test of the model comparison part). The variations of log likelihood and log evidence obtained in this way are shown in figure 5. The Bayesian analysis correctly predicts the number of sources present.

6 Conclusions.

Limitations to the resolution of noise subspace eigenanalysis have been exposed. For real systems, where the number of snapshots is large, resolution will still be much better than the Rayleigh limit which restricts Fourier transform and beam-sweep methods.

The eigenanalysis technique as employed here does not give the maximum likelihood parameters. The predicted parameters are, however, a good approximation to the optimum. The success of Newton-Raphson optimization shows that assumptions of Gaussian probability distributions are well founded.

The application of Bayesian techniques has enabled the prediction of source positions to be given error bars. Bayesian model comparison has been shown to give consistent predictions even when positions are not well determined. In cases where parameters are well optimized, the Bayesian approach correctly infers the number of sources k.

Finally, it has been shown that the use of Bayesian techniques to make model comparisons is limited only by the standard of optimization routines employed.

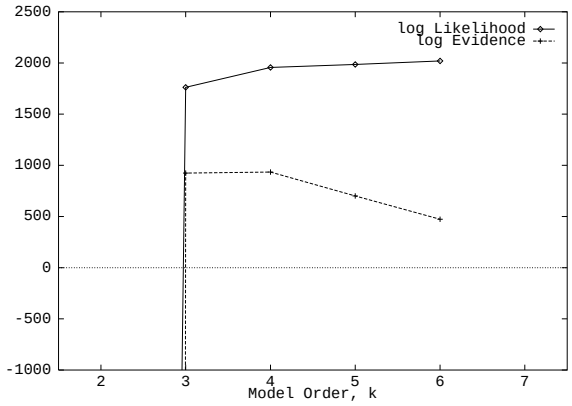

Figure 5: Variation of Likelihood and Evidence (Parameters Generated by Manually Seeded optimization)

$$\begin{array}{rcccc} \theta & = & -0.21 & -0.2 & 0.1 & 0.3 \\ |F|^2 & = & 0.01 & 1 & 1 & 1/4 \end{array}$$
$$S = 32,\ N = 32,\ \sigma = 0.1,\ \delta = 1.0$$

References

[1] J. Reilly, J. Litva, P. Bauman, "New Angle-of-Arrival Estimator : Comparative Evaluation Applied to the Low Angle Tracking Problem", *proc. IEEE* **135F**, 5, 1988.

[2] S.M. Kay, S.L. Marple, "Spectrum Analysis : A Modern Perspective", *proc. IEEE* **69**, 11, 1981.

[3] S. Alaykin, "Adaptive Filter Theory", Prentice Hall Information and System Sciences Series.

[4] G.L. Bretthorst, "Bayesian Spectrum Analysis and Parameter Estimation", Lecture notes in statistics, Springer-Verlag, 1990.

[5] S.F. Gull, "Bayesian Inductive Inference and Maximum Entropy", in *Maximum Entropy and Bayesian Methods in Science and Engineering, Vol. 1, Foundations*, ed. G.J. Erickson and C.R. Smith, pp. 53-74, Kluwer, 1988.

[6] E.T. Jaynes, "Bayesian Methods : General Background", in *Maximum Entropy and Bayesian methods in applied statistics*, ed. J.H. Justice. pp.1-25. C.U.P. 1986.

[7] D.J.C. MacKay, "Bayesian Interpolation", *Neural Computation* **4**, 415-447, 1992.

[8] J.P. Burg, D.G. Luenberger, D.L. Wenger, "Estimation of Structured Covariance Matrices", *proc. IEEE* **70**, 9, 1982.