BALANCED DATA FOCUSING:
DIRECTION OF ARRIVAL ESTIMATION
WITH MAXIMUM LIKELIHOOD PERFORMANCE (U)

by

William J.L. Read

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
REPORT N0. 1233

Canada

19941205 102

National   Défense
Defence   nationale

BALANCED DATA FOCUSING:
DIRECTION OF ARRIVAL ESTIMATION
WITH MAXIMUM LIKELIHOOD PERFORMANCE (U)

by

William J.L. Read
Electronic Support Measures Section
Electronic Warfare Division

Accesion For

NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	

By

Distribution /

Availability Codes

| Dist | Avail and / or Special |
| A-1 | |

DTIC QUALITY INSPECTED 3

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
REPORT NO. 1233

PCN
041LX

October 1994
Ottawa

ABSTRACT

The performance of a direction of arrival estimation procedure at low signal-to-noise ratios and limited data samples is an important characteristic. The approach based on maximum likelihood (ML) estimation is considered to be among the best for this problem as long as the underlying signal model is properly chosen. Unfortunately, in most cases, there is no closed-form solution so fast search procedures are employed. Given no a priori knowledge, selecting the initial parameter values for these search procedures can be a difficult problem, especially under low signal-to-noise conditions. In this paper, a new method for uniform linear sensor arrays which overcomes the initial value problem is introduced. This method is called Balanced Data Focusing (BDF). Simulation results are included comparing the performance of this new method to that of the ML approach using the Alternating Projection Maximization search procedure and another popular estimation approach, the root-MUSIC method.

RÉSUMÉ

La performance des techniques permettant l'estimation de l'angle d'arrivé d'un signal en présence d'un faible rapport signal sur bruit et d'un nombre limité d'échantillons est une charactéristique importante. L'approche se fondant sur l'estimation à maximum de vraisemblance (ML) est considérée comme l'une des meilleurs repondant à cette exigence en autant que le modèle soit adéquatement choisi. Malheureusement, dans la plupart des cas, une solution abrégée n'existe pas de sorte que des procédures rapides de recherche sont employées. Sans information à priori, choisir les valeurs initiales des paramètres pour ces procédures de recherche peuvent s'avérer un problème difficile, spécialement dans le cas où on a à faire à des conditions de rapports signal sur bruit particulièrement bas. Le présent rapport introduit une nouvelle méthode qui s'applique à un réseau linéaire d'antennes uniformément réparties. Cette méthode, appelée BDF, surmonte le problème du choix des valeurs initiales. Les résultats des simulations comprises dans ce rapport comparent les performances de cette nouvelle méthode à celles de deux autres méthodes populaires: la méthode ML employant la procédure de recherche fondée sur la maximization des projections alternées ainsi qu'à la méthode root-MUSIC.

EXECUTIVE SUMMARY

The accuracy of a direction of arrival estimation procedure in the presence of noise is an important characteristic, especially at low signal to noise ratios. Although numerous approaches have been proposed [1] - [5], those based on maximum likelihood (ML) estimation are considered to achieve the best performance [6] assuming the underlying signal model is properly chosen.

The idea behind the ML approach is relatively simple. Using an appropriate model of the signal environment, synthetic data can be generated and compared to the observed data. The maximum likelihood parameter estimates are taken as the values of the model parameters for which the synthetic data best fits (in the statistical sense) the observed data. Despite the simplicity of this idea, the actual derivation of the ML approach for a particular problem can often be very complex depending on the nature of the model and the measured data available.

For the bearing estimation problem, the deterministic maximum likelihood (DML) method has been shown to achieve good performance. The main difficulty with this method is that it requires an M dimensional search where M is the number of signal bearings to be estimated. A number of iterative algorithms have been developed to speed up the search process compared to a brute force search. In general, these algorithms either suffer from a requirement for good initial bearing estimates (without which they may converge to the wrong solution), or their convergence time is still too slow to be practical.

To overcome these problems a new algorithm has been developed called Balanced Data Focusing (BDF). Rather than adjust a model to fit the data, this method adjusts the data until it represents a valid form of model data. The bearings can then be easily calculated from the data using a linear interpolation method (similar to linear prediction methods). The advantage is that initial bearing estimates are not required, and this is accomplished without sacrificing convergence speed compared to other approaches. The main restriction is that the sensor array is required to be linear with uniform spacing between sensors.

In simulations the BDF algorithm was compared to the deterministic ML method and the root-MUSIC algorithm (an alternate DF estimation approach which has received considerable attention in the open literature). It was found that the BDF method performed as well as the deterministic ML approach and outperformed root-MUSIC in terms of both accuracy and threshold performance. The BDF method is, however, slower than

root-MUSIC, especially as the number of signals increases. Further research into better search algorithms (compared to the gradient descent technique used) would likely yield useful improvements in processing speed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

The accuracy of a direction of arrival estimation procedure in the presence of noise is an important characteristic, especially at low signal to noise ratios. Although numerous approaches have been proposed [1] - [5], those based on maximum likelihood (ML) estimation are considered to achieve the best performance [6] assuming the underlying signal model is properly chosen.

The idea behind the ML approach is relatively simple. Using an appropriate model of the signal environment, synthetic data can be generated and compared to the observed data. The maximum likelihood parameter estimates are taken as the values of the model parameters for which the synthetic data best fits (in the statistical sense) the observed data. Despite the simplicity of the idea, the actual derivation of the ML approach for a particular problem can often be very complex depending on the nature of the model and the measured data available.

In this report a linear array of N uniformly spaced narrowband sensors is considered. The complex baseband outputs from this array at time t can be modelled using

$$y_n(t) = \sum_{m=1}^{M} c_m(t) e^{-j\omega_m nd}, \tag{1}$$

where M represents the number of signals impinging on the array, $y_n(t)$ is the model output for sensor n ($0 \leq n < N - 1$), $c_m(t)$ represents the complex baseband amplitude of the mth signal, ω_m is the corresponding spatial frequency, and d is the spacing between sensor elements. The spatial frequency is related to the signal direction of arrival ϕ_m by

$$\omega_m = \frac{2\pi}{\lambda} \cos(\phi_m), \tag{2}$$

where λ is the signal wavelength. The direction of arrival is defined as the angle between the sensor baseline (in the direction from sensor 0 to sensor $N - 1$) and the direction of the signal (which is perpendicular to the wavefront). The model is adjusted by changing the values $c_m(t)$ and ϕ_m.

The corresponding measured data at time t for sensor n is represented by $x_n(t)$. This data will differ from the model data due to the effects of noise, mutual coupling, multipath propagation, etc.. In this report, only noise is considered; in particular, white Gaussian noise, since this is the most common type of noise in practice. The problem

then, is to estimate the bearings of the incoming signals given the noisy data.

Using these definitions of the model and measured data, various DF estimation methods based on maximum likelihood theory can be developed. In Section 2, one such method, deterministic maximum likelihood (DML), is discussed along with the difficulties associated with it. This is followed in Section 3 by a description of a proposed new algorithm, Balanced Data Focusing (BDF), which overcomes some of these difficulties. Detailed descriptions of various aspects of the BDF algorithm are discussed in Sections 4 and 5. In Section 6, the results of computer simulations are used to compare the performance of the BDF method to the DML method, as well as to the MUSIC algorithm (an algorithm which has received considerable attention in the open literature). The conclusions are presented in Section 7.

2.0   THE DETERMINISTIC MAXIMUM LIKELIHOOD METHOD

Assuming that $x_n(t)$ has the same form as $y_n(t)$ with the addition of white Gaussian noise which is uncorrelated from sensor to sensor, then for K data samples the best fit between the model and data is achieved when [6] the mean squared error (MSE) given by

$$MSE = \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} |x_n(t_k) - y_n(t_k)|^2 \tag{3}$$

is minimized. This is called the deterministic maximum likelihood (DML) approach since information about the statistics of the model parameters is not incorporated into the fitting procedure i.e. the statistics are assumed to be unknown.

For the purposes of this reports, it is also convenient to represent the above expression in vector form as

$$MSE = \sum_{k=0}^{K-1} (\mathbf{x}(t_k) - \mathbf{y}(t_k))^H (\mathbf{x}(t_k) - \mathbf{y}(t_k)) \tag{4}$$

where $\mathbf{x}(t_k)$ and $\mathbf{y}(t_k)$ are $N \times 1$ elements vectors whose elements are given by the measured data $x_0(t_k), x_1(t_k), ..., x_{N-1}(t_k)$ and the model data $y_0(t_k), y_1(t_k), ..., y_{N-1}(t_k)$ respectively, and the superscript H represents the conjugate transpose operation. The vectors $\mathbf{x}(t_k)$ and $\mathbf{y}(t_k)$ represent samples of the measured and modelled data taken at time, t_k, and are referred to as snapshots.

The main difficulty with determining the model parameters which minimize the

2

MSE is that no closed form solution exists. There are, however, closed form solutions for the complex amplitudes (represented by $c_m(t_k)$) given that the values for the bearings ϕ_1, $\phi_2, \ldots, \phi_M$ [6] are known. For example, the signal model data can be represented as

$$\mathbf{y}(t_k) = \mathbf{Sc}(t_k) \tag{5}$$

where $\mathbf{S}$ is the $N \times M$ normalized signal matrix given by

$$\mathbf{S} = \begin{bmatrix} 1, & 1, & 1, & \cdots, & 1 \\ e^{-j\omega_1 d}, & e^{-j\omega_1 d}, & e^{-j\omega_1 d}, & \cdots, & e^{-j\omega_1 d} \\ e^{-j\omega_1 2d}, & e^{-j\omega_2 2d}, & e^{-j\omega_3 2d}, & \cdots, & e^{-j\omega_M 2d} \\ \vdots, & \vdots, & \vdots, & & \vdots \\ e^{-j\omega_1 (N-1)d}, & e^{-j\omega_2 (N-1)d}, & e^{-j\omega_3 (N-1)d}, & \cdots, & e^{-j\omega_M (N-1)d} \end{bmatrix} \tag{6}$$

and $\mathbf{c}(t_k)$ is the $M \times 1$ vector of complex amplitudes where

$$\mathbf{c}(t_k) = \begin{bmatrix} c_1(t_k) \\ c_2(t_k) \\ c_3(t_k) \\ \vdots \\ c_M(t_k) \end{bmatrix} \tag{7}$$

The MSE can be rewritten as,

$$MSE = \sum_{k=0}^{K-1} (\mathbf{x}(t_k) - \mathbf{Sc}(t_k))^H (\mathbf{x}(t_k) - \mathbf{Sc}(t_k)) \tag{8}$$

Minimizing this expression with respect to $\mathbf{c}(t_k)$ and solving leads to the result [6]

$$\mathbf{c}(t_k) = (\mathbf{S}^H \mathbf{S})^{-1} \mathbf{S}^H) \mathbf{x}(t_k) \tag{9}$$

Plugging this result back into equation (8) yields

$$MSE = \sum_{k=0}^{K-1} \mathbf{x}(t_k)^H (\mathbf{I} - \mathbf{S}(\mathbf{S}^H \mathbf{S})^{-1} \mathbf{S}^H) \mathbf{x}(t_k) \tag{10}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix.

The most obvious advantage of equation (10) is that it simplifies the problem

from a search for the complex amplitudes and bearings ($3M$ real parameters) to a search for bearings only (M real parameters). If the initial estimate of the bearings can be made with reasonable accuracy, several good methods exist which can fine-tune the estimates until the MSE is minimized [6] - [9]. The main danger of these methods is that if the initial values are improperly chosen, they will converge on solutions which are not optimum (i.e. a local minimum of the MSE function instead of the global minimum).

The selection of initial values is typically performed using an estimator which requires no initial values but is less accurate than the DML approach. At low signal-to-noise ratios this can cause problems when the initial value estimates become so inaccurate that the search method used to determine the DML estimates is no longer guaranteed to find the global minimum. The results are therefore unreliable under these conditions.

To overcome the initial value problem, an exhaustive search of all possible solutions can be performed, or more efficient approaches such as the Simulated Annealing algorithm [10] can be used to overcome the initial value problem. Unfortunately these methods achieve better reliability by sacrificing convergence speed and are generally too slow for most practical applications.

3.0 THE BALANCED DATA FOCUSING ALGORITHM

To overcome the initial value problem without sacrificing convergence speed a new approach, the BDF algorithm, is proposed. To simplify the following discussion, the single snapshot case is discussed here — extensions to multiple snapshots are presented in Section 5. Dropping the argument t_k for convenience, the equation for MSE simplifies to

$$MSE = (\mathbf{x} - \mathbf{y})^H (\mathbf{x} - \mathbf{y}) \tag{11}$$

In the BDF algorithm, the initial value problem is overcome by adjusting or "focusing" the data instead of adjusting the signal model. The purpose is to focus the data until it represent a valid form of the model data, i.e. defining S as the set containing all possible values $\mathbf{y}$, the focusing procedure continues until $\mathbf{x}' \in S$ where $\mathbf{x}'$ is the focused data.

A simple approach to performing the focusing is to use a gradient descent technique. Assuming there is an error function $\varepsilon(\mathbf{x})$ that provides an estimate of the least

4

square distance from $\mathbf{x}$ to S, then the focusing procedure can be defined by

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mu \mathbf{g}_i \tag{12}$$

where $\mathbf{x}_i$ represents the data after the ith iteration, $\mathbf{x}_0 = \mathbf{x}$ the initial value , μ the step size, and $\mathbf{g}_i$ is the gradient defined by,

$$\mathbf{g}_i = \nabla_\mathbf{x} \varepsilon(\mathbf{x}_i) \tag{13}$$

The procedure terminates when $|\varepsilon(\mathbf{x}_i)|$ is less than a predetermined tolerance value.

Since the set S contains an infinite number of solutions, there will be an infinite number of possible solutions for the focusing procedure. The ideal choice is the solution which lies closest to $\mathbf{x}_0$ in the least squares sense. Therefore the development of the error function $\varepsilon(\mathbf{x})$ is critical to the success of the BDF algorithm and is discussed in Section 3.1. The gradient is derived in Section 3.2. Once the algorithm has converged, any number of DF algorithms [1]-[5] can be used to extract the bearings from $\mathbf{x}_i$. However, since the calculation of $\varepsilon(\mathbf{x})$ involves the estimation of the autoregressive filter coefficients of the underlying signal model in $\mathbf{x}$ (see Section 4), the bearing estimates can also be determined by rooting the filter coefficients once the algorithm has converged. The step size μ also plays a role in the accuracy and convergence time of the algorithm and its selection is discussed in Section 3.3.

3.1 The Error Function

The definition of the signal model data was given in equation (1). An alternate expression is the autoregressive signal model given by

$$y_n = \sum_{m=1}^{M} a_m^* y_{n-m} \tag{14}$$

for $M \leq n < N$, where the filter coefficients are defined by [16]

$$
\begin{aligned}
a_1 &= e^{j\omega_1 d} + e^{j\omega_2 d} + ... + e^{j\omega_M d} \\
a_2 &= -e^{j(\omega_1+\omega_2)d} - e^{j(\omega_2+\omega_3)d} - ... - e^{j(\omega_{M-1}+\omega_M)d} \\
a_3 &= e^{j(\omega_1+\omega_2+\omega_3)d} + e^{j(\omega_2+\omega_3+\omega_4)d} + ... + e^{j(\omega_{M-2}+\omega_{M-1}+\omega_M)d} \\
&\vdots
\end{aligned} \tag{15}
$$

$$a_M = (-1)^{M+1} e^{j(\omega_1 + \omega_2 + \omega_3 + \dots + \omega_M)d}$$

Equation (14) can be rewritten as

$$\sum_{m=0}^{M} b_m^* y_{n-M+m} = 0 \tag{16}$$

where

$$b_m = \frac{a_{M-m}}{\sqrt{a_M}} \tag{17}$$

and

$$b_M = \frac{-1}{\sqrt{a_M}} \tag{18}$$

It is relatively straightforward to show that the filter coefficients are conjugate symmetric, i.e. $b_m = b_{M-m}^*$.

The filter

$$\mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix} \tag{19}$$

is a spatial moving average filter with M nulls corresponding to the directions of the M signals in the model data. Applying this filter to the measured data $\mathbf{x}$, the signals will be nulled out and the resultant output will be noise only. That is,

$$\sum_{m=0}^{M} b_m^* x_{n-m} = \eta_n \tag{20}$$

where η_n is the noise or error output. Applying this to all the data (i.e. letting $n = M, M+1, \dots, N-1$) then

$$\mathbf{B}^H \mathbf{x} = \mathbf{n} \tag{21}$$

6

where

$$\mathbf{B} = \begin{bmatrix} b_0 & & & 0 \\ \vdots & \ddots & & \\ b_M & & & \\ & & b_0 & \\ & \ddots & & \vdots \\ 0 & & & b_M \end{bmatrix}.$$ (22)

and $\mathbf{n}$ is an $(N - M) \times 1$ vector with elements $\eta_M, \eta_{M+1}, ..., \eta_{N-1}$.

Since the error output of the filter will be a function of the signal-to-noise ratio, then an appropriate error function for the focusing procedure discussed in Section 2.0 is given by the sum of the square errors. Using equation (21) this can be expressed as

$$\varepsilon(\mathbf{x}) = \mathbf{n}^H \mathbf{n} = \mathbf{x}^H \mathbf{B} \mathbf{B}^H \mathbf{x}$$ (23)

The filter coefficients $b_0, b_1, ..., b_M$ could be generated using equations (16), (17), and (18), however this provides no advantage since initial value estimates are still required for the bearings $\phi_1, \phi_2, ..., \phi_M$. Instead, the filter coefficients are estimated from the data itself. The methods by which this is done are discussed in some detail in Section 4.

3.2 The Gradient

The gradient was previously defined in equation (13). Using the expression developed for $\varepsilon(\mathbf{x})$ (equation (23)) and performing the gradient operation,

$$\begin{aligned} \mathbf{g}_i &= \nabla_{\mathbf{x}} \varepsilon(\mathbf{x}_i) \\ &= \nabla_{\mathbf{x}} (\mathbf{x}_i^H \mathbf{B} \mathbf{B}^H \mathbf{x}_i) \\ &= 2\mathbf{B} \mathbf{B}^H \mathbf{x}_i \end{aligned}$$ (24)

3.3 The Step Size μ

An appropriate step size can be determined using a first order Taylor series

7

expansion of $\varepsilon(\mathbf{x}_i)$ to get

$$\varepsilon(\mathbf{x}_{i+1}) \approx \varepsilon(\mathbf{x}_i) + \left(\frac{1}{2}\mathbf{g}_i^H(\mathbf{x}_{i+1} - \mathbf{x}_i) + \frac{1}{2}(\mathbf{x}_{i+1} - \mathbf{x}_i)^H \mathbf{g}_i\right) \qquad (25)$$

Substituting the right side of equation (12) for $\mathbf{x}_{i+1}$ and solving for μ yields

$$\mu \approx \frac{\varepsilon(\mathbf{x}_i) - \varepsilon(\mathbf{x}_{i+1})}{\mathbf{g}_i^H \mathbf{g}_i} \qquad (26)$$

To minimize convergence time $\varepsilon(\mathbf{x}_{i+1}) = 0$ is chosen which gives

$$\mu \approx \frac{\varepsilon(\mathbf{x}_i))}{\mathbf{g}_i^H \mathbf{g}_i} \qquad (27)$$

Ideally, with this choice the focusing procedure converges in one step. However, since equation (26) is only an approximation a smaller step size is required. Accordingly,

$$\mu = \frac{\mu_0 \varepsilon(\mathbf{x}_i)}{\mathbf{g}_i^H \mathbf{g}_i} \qquad (28)$$

where $0 < \mu_0 \leq 1$. Through empirical testing it has been found by the author that $\mu_0 = 0.3$ provides a good compromise between convergence speed and accuracy.

Using this last result the expression for updating the data (equation (12)) becomes

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \frac{\mu_0 \varepsilon(\mathbf{x}_i)\mathbf{g}_i}{\mathbf{g}_i^H \mathbf{g}_i} \qquad (29)$$

4.0 GENERATING THE FILTER COEFFICIENTS

For the purposes of this analysis Equation (23) can be rewritten as

$$\sigma^2 = \mathbf{b}^H \mathbf{X}_f \mathbf{X}_f^H \mathbf{b} \qquad (30)$$

where $\sigma^2 = \varepsilon(\mathbf{x})$ and $\mathbf{X}_f$ is the forward data matrix defined by

$$\mathbf{X}_f = \begin{bmatrix} x_0, & x_1, & x_2, & \ldots, & x_{N-M-1} \\ x_1, & x_2, & x_3, & \ldots, & x_{N-M} \\ x_2, & x_3, & x_4, & \ldots, & x_{N-M+1} \\ \vdots & \vdots & \vdots & & \vdots \\ x_M, & x_{M+1}, & x_{M+2}, & \ldots, & x_{N-1} \end{bmatrix} \tag{31}$$

The goal here is to estimate the $M + 1$ filter coefficients, $b_0, b_1, b_2, ..., b_M$ from the data instead of being required to explicitly specify $\phi_1, ..., \phi_M$. This can be done by minimizing σ^2 with respect to $\mathbf{b}$ subject to some constraints.

An obvious constraint on $\mathbf{b}$ is that it must be conjugate symmetric, since filters generated from the model signals have this property. An important feature of the conjugate symmetry property is that it only occurs for

(i) complex sinusoidal signals of the form (see also equation (1))

$$c_m(t_k)\rho^{-n}e^{-j\omega_m nd} \tag{32}$$

(ii) dampled complex sinusoidal signal pairs of the form

$$a_m(t_k)\rho^{-n}e^{-j\omega_m nd} + b_m(t_k)\rho^n e^{j\omega_m nd} \tag{33}$$

or combinations of the two-signal models where $a_m(t_k)$ and $b_m(t_k)$ are the complex amplitudes and ρ is the real-valued damping factor. The conjugate symmetry constraint therefore ensures that the bearing estimates derived from $\mathbf{b}$ will be based on solutions which correspond to case (i) or (ii). Other researchers [11] have noted that the inclusion of case (ii) does not significantly degrade estimator performance. Results supporting this assertion are given in Section 5.

Based on the conjugate symmetry property of $\mathbf{b}$, equation (30) can also be written as,

$$\sigma^2 = \mathbf{b}^H \mathbf{X}_b \mathbf{X}_b^H \mathbf{b} \tag{34}$$

where $\mathbf{X}_b$ is the backward data matrix defined by

$$
\mathbf{X}_b = \begin{bmatrix}
x^*_{N-1}, & x^*_{N-2}, & x^*_{N-3}, & \cdots, & x^*_M \\
x^*_{N-2}, & x^*_{N-3}, & x^*_{N-4}, & \cdots, & x^*_{M-1} \\
x^*_{N-3}, & x^*_{N-4}, & x^*_{N-5}, & \cdots, & x^*_{M-2} \\
\vdots & \vdots & \vdots & & \vdots \\
x^*_{N-M-1}, & x^*_{N-M-2}, & x^*_{N-M-3}, & \cdots, & x^*_0
\end{bmatrix}. \tag{35}
$$

The equivalence of equations (30) and (35) was derived assuming $\mathbf{b}$ was conjugate symmetric, however this equivalence also holds for any choice of $\mathbf{b}$ as long as there exists some complex constant α such that $|\alpha| = 1$ and the vector $\alpha\mathbf{b}$ is conjugate symmetric. Since the constant multiplier α has no effect on the filter nulls (and therefore the estimated signal bearings), then if $\mathbf{b}$ is not conjugate symmetric but a value for α exists, it is assumed that $\mathbf{b}$ is multiplied by α to make it conjugate symmetric.

A second constraint that is required is a nontriviality constraint. The purpose is to avoid the solution $b_0 = b_1 = \ldots = b_M = 0$ which is obviously meaningless. Accordingly, two such constraints have been investigated in this report, namely

$$
\mathbf{b}^H\mathbf{b} = 1 \tag{36}
$$

which leads to an eigenanalysis based approach, and

$$
\mathbf{b}^H\mathbf{u}_k = 1 \tag{37}
$$

which leads to a linear prediction/interpolation approach. For the latter approach k represents the row in $\mathbf{X}_f$ or $\mathbf{X}_b$ to be predicted/interpolated ($0 \le k \le M$) and $\mathbf{u}_k$ is an $M + 1$ column vector of zeros except for the kth element which is $u_k = 1$.

The choice of the nontriviality constraint influences the performance of the BDF algorithm. The purpose of investigating two different choices is to illustrate the differences. Neither choice was selected based on optimality criteria, rather they were selected because they are commonly used constraints in DF estimation theory. The next two sections describe how these constraints are incorporated with the conjugate symmetry constraint into a solution for the minimization of σ^2.

10

4.1 Eigenvector Approach

For the eigen analysis approach, advantage is taken of the equivalence between equations (30) and (34) to recast the minimization as

$$2\sigma^2 = \mathbf{b}^H \mathbf{X}_{fb} \mathbf{X}_{fb}^H \mathbf{b} \tag{38}$$

where $\mathbf{X}_{fb}$ the forward-backward data matrix is defined by

$$\mathbf{X}_{fb} = [\mathbf{X}_f, \mathbf{X}_b] \tag{39}$$

Minimizing this expression subject to the constraint given in (36) and temporarily ignoring the conjugate symmetry constraint leads to the result (also known as Pisarenko Harmonic Decomposition [12])

$$\mathbf{X}_{fb} \mathbf{X}_{fb}^H \mathbf{b} = 2\sigma^2 \mathbf{b}, \tag{40}$$

where vector $\mathbf{b}$ is an orthonormal eigenvector of the matrix $\mathbf{X}_{fb} \mathbf{X}_{fb}^H$ and $2\sigma^2$ is the corresponding eigenvalue. Since the solution is not unique, $\mathbf{b}$ is chosen as the eigenvector corresponding to the smallest eigenvalue of $\mathbf{R}_M$. Eigen analysis or singular value decomposition techniques can be used to determine the eigenvectors and eigenvalues.

Although no obvious conjugate symmetry constraint has been applied, an eigenvector of the desired form corresponding to the eigenvalue $2\sigma^2$ will always exist. For example, the matrix

$$\mathbf{R}_{fb} = \mathbf{X}_{fb} \mathbf{X}_{fb}^H \tag{41}$$

has the structure

$$\mathbf{R}_{fb} = \begin{bmatrix} r_{00}, & r_{10}^*, & r_{20}^*, & \cdots, & r_{M,0}^* \\ r_{10}, & r_{11}, & r_{21}^*, & \cdots, & r_{M-1,0}^* \\ r_{20}, & r_{21}, & r_{22}^*, & \cdots, & r_{M-2,0}^* \\ \vdots & \vdots & \vdots & & \vdots \\ r_{M-1,0}, & r_{M-2,1}, & r_{M-3,1}, & \cdots, & r_{10}^* \\ r_{M,0}, & r_{M-1,0}, & r_{M-2,0}, & \cdots, & r_{00} \end{bmatrix} \tag{42}$$

The matrix $\mathbf{R}_{fb}$ has the properties that it is both Hermitian symmetric ($r_{ij} = r_{ji}^*$) and persymmetric ($r_{ij} = r_{M-i,M-j}^*$). Assuming that $\mathbf{b}$ is a valid solution to equation (40) but is not conjugate symmetric (nor can it be made conjugate symmetry through the multiplication of a complex constant) then it is apparent from the form of $\mathbf{R}_{fb}$ that $\tilde{\mathbf{b}}$

must also be a valid solution where

$$\tilde{\mathbf{b}} = \begin{bmatrix} b_M^* \\ b_{M-1}^* \\ b_{M-2}^* \\ \vdots \\ b_0^* \end{bmatrix} \qquad (43)$$

The new vector $\tilde{\mathbf{b}}$ will also be an orthonormal eigenvector and $\mathbf{b}^H\tilde{\mathbf{b}} = 0$. Using these two eigenvector solutions it will always be possible to construct a new eigenvector solution $\mathbf{b}'$ which is conjugate symmetric, for example,

$$\mathbf{b}' = \frac{1}{\sqrt{2}}(\mathbf{b} + \tilde{\mathbf{b}}). \qquad (44)$$

4.2   The Balance Linear Interpolator

The solution to the equation (30) using the linear prediction/interpolation non-triviality constraint (equation (36)) is given by

$$\mathbf{X}_f\mathbf{X}_f^H\mathbf{b} = \sigma^2\mathbf{u}_m \qquad (45)$$

where $b_m = 1$, $\mathbf{u}_m$ is an $(M+1) \times 1$ vector with elements $u_0 = u_1 = ... = u_M = 0$ except $u_m = 1$, and σ^2 is the prediction error given by

$$\sigma^2 = \mathbf{b}^H\mathbf{X}_f\mathbf{X}_f^H\mathbf{b} \qquad (46)$$

Since the multiplication of the filter coefficients by a constant phase term has no effect on the location of the filter roots or prediction error, an alternate form of equation (45) which more easily allows the conjugate symmetry constraint to be imposed is given by

$$\mathbf{X}_f\mathbf{X}_f^H\mathbf{b} = c\sigma^2\mathbf{u}_m \qquad (47)$$

where c is an arbitrary complex term such that $|c| = 1$, and now $b_m = c$.

The conjugate symmetry constraint can be imposed by noting that if $\mathbf{b}$ has

12

conjugate symmetry, it also satisfies the relationship

$$\mathbf{X}_b \mathbf{X}_b^H \mathbf{b} = c^* \sigma^2 \mathbf{u}_{M-m} \tag{48}$$

where $\mathbf{X}_b$ is the backward data matrix defined by

$$\mathbf{X}_b = \begin{bmatrix} x_{N-1}^*, & x_{N-2}^*, & x_{N-3}^*, & \cdots, & x_M^* \\ x_{N-2}^*, & x_{N-3}^*, & x_{N-4}^*, & \cdots, & x_{M-1}^* \\ x_{N-3}^*, & x_{N-4}^*, & x_{N-5}^*, & \cdots, & x_{M-2}^* \\ \vdots & \vdots & \vdots & & \vdots \\ x_{N-M-1}^*, & x_{N-M-2}^*, & x_{N-M-3}^*, & \cdots, & x_0^* \end{bmatrix}. \tag{49}$$

Therefore if $\mathbf{b}$ is conjugate symmetric it must simultaneously satisfy the both equations (47) and (48). This can be forced by adding the two expressions to get

$$\mathbf{X}_f \mathbf{X}_f^H \mathbf{b} + \mathbf{X}_f \mathbf{X}_f^H \mathbf{b} = \mathbf{R}_{fb} \mathbf{b} = 2\sigma^2 (c\mathbf{u}_m + c^* \mathbf{u}_{M-m}) \tag{50}$$

where $\mathbf{R}_{fb}$ was defined previously (equation (42)). In this new expression, c is no longer arbitrary as will be seen shortly.

Before discussing the actual solution of $\mathbf{b}$, σ^2, and c, it is useful to define the matrix

$$\mathbf{Q} = \mathbf{R}_{fb}^{-1} \tag{51}$$

where the elements of $\mathbf{Q}$ are given by q_{ij} for $i, j = 0, 1, ..., M$. It has been assumed here that $\mathbf{R}_{fb}$ is invertible. The matrix $\mathbf{Q}$ has the properties that it is both Hermitian symmetric ($q_{ij} = q_{ji}^*$) and persymmetric ($q_{ij} = q_{M-i,M-j}^*$) since $\mathbf{R}_{fb}$ also has these properties [12]. These properties will be used to make simplifications in the following analysis.

From this definition the filter coefficients can be expressed as

$$\mathbf{b} = 2\sigma^2 \mathbf{Q}(c\mathbf{u}_m + c^* \mathbf{u}_{M-m}) \tag{52}$$

Using the fact that $b_m = c$ then $\mathbf{u}_m^H \mathbf{b} = c$ and

$$\sigma^2 = \frac{1}{2\mathbf{u}_m^H \mathbf{Q}(|c|^2 \mathbf{u}_m + (c^*)^2 \mathbf{u}_{M-m})} \tag{53}$$

13

In terms of the elemental values of $\mathbf{Q}$,

$$\sigma^2 = \frac{1}{2(q_{mm} + (c^*)^2 q_{m,M-m})} \tag{54}$$

Noting that by Hermitian symmetry q_{mm} must be real, then the choice of c that minimizes σ^2 occurs when

$$(c^*)^2 q_{m,M-m} = |q_{m,M-m}| \tag{55}$$

Therefore

$$c = \sqrt{\frac{q_{m,M-m}}{|q_{m,M-m}|}} \tag{56}$$

Summarizing these results and using the persymmetry property of $\mathbf{Q}$ to express the results in terms of the column elements $q_{0m}, q_{1m}, ..., q_{Mm}$ then

$$c = \sqrt{\frac{q^*_{M-m,m}}{|q_{M-m,m}|}} \tag{57}$$

$$\sigma^2 = \frac{1}{2(q_{mm} + |q_{M-m,m}|)} \tag{58}$$

and

$$b_i = 2\sigma^2(cq_{im} + c^* q^*_{M-i,m}) \quad \text{for} \quad i = 0, 1, ..., M \tag{59}$$

5.0 EXTENSION TO THE MULTIPLE SNAPSHOT CASE

Although single snapshot processing can be useful in some situations, in many real world systems multiple snapshot data is available. Utilizing such data can greatly enhance estimator performance. The extension of the BDF method to multiple snapshots is relatively straightforward.

Reintroducing the argument t_k to represent a quantity based on the sensor data measured at time instance t_k, then for the multiple snapshot case the sensor data $\mathbf{x}(t_0)$, $\mathbf{x}(t_1)$, ..., $\mathbf{x}(t_{K-1})$ will be available for processing. For convenience the sensor data matrix is defined as

$$\mathbf{X} = [\mathbf{x}(t_0), \mathbf{x}(t_1), ..., \mathbf{x}(t_{K-1})] \tag{60}$$

and the error function is then represented by $\varepsilon(\mathbf{X})$. The forward-backward data matrix

14

in this case is also redefined as

$$\mathbf{X}_{fb} = [\mathbf{X}_{fb}(t_0), \mathbf{X}_{fb}(t_1), \ldots, \mathbf{X}_{fb}(t_{K-1})] \tag{61}$$

where $\mathbf{X}_{fb}(t_k)$ is the forward-backward data matrix formed from a single data snapshot $\mathbf{x}(t_k)$ using equations (31), (35), and (39).

The determination of the filter tap weights proceeds in the same manner described in Section 4 using the new definition of $\mathbf{X}_{fb}$. The gradient vector given in equation (24) is easily modified to become the gradient matrix equation

$$\mathbf{G} = 2\mathbf{B}^H\mathbf{B}\mathbf{X} \tag{62}$$

Finally, the expression for updating the data (equation (29)) becomes

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \frac{\mu_0\varepsilon(\mathbf{X}_i)\mathbf{G}_i}{\text{trace}\{\mathbf{G}_i^H\mathbf{G}_i\}} \tag{63}$$

From this derivation it can be seen that, compared to the single snapshot case, increasing the number of snapshots to K increases the amount of processing by a factor of K. For a large number of snapshots the processing requirements may become prohibitive. These requirements can be eased somewhat by noting that the MSE for the DML approach given in equation (10) can be rewritten as

$$MSE = \text{trace}\{\mathbf{X}\mathbf{X}^H(\mathbf{I} - \mathbf{S}(\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H)\} \tag{64}$$

By inspection, it is clear that the data matrix $\mathbf{X}$ can be replaced by an equivalent matrix $\mathbf{Z}$ as long as

$$\mathbf{Z}\mathbf{Z}^H = \mathbf{X}\mathbf{X}^H \tag{65}$$

For the BDF algorithm, this means that the columns of the matrix $\mathbf{Z}$ can be processed in place of $\mathbf{x}(t_0)$, $\mathbf{x}(t_1)$, ..., $\mathbf{x}(t_{K-1})$. If the choice of $\mathbf{Z}$ is restricted to square matrices, then the maximum effective number of snapshots that need to be processed by the BDF algorithm is limited to N, which results in a considerable savings in processing when $K > N$.

One approach to computing the matrix $\mathbf{Z}$ is to perform an eigen decomposition

on $\mathbf{XX}^H$ to get

$$\mathbf{XX}^H = \sum_{i=0}^{N-1} \lambda_i \mathbf{v}_i \mathbf{v}_i^H, \tag{66}$$

where the eigenvalues are ordered so that $\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{N-1}$ and the corresponding eigenvectors $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{N-1}$ form an orthonormal basis set. The columns of $\mathbf{Z}$ are then given by

$$\mathbf{z}_i = \sqrt{\lambda_i} \mathbf{v}_i \quad \text{for } i = 0, 1, 2, \ldots, N-1 \tag{67}$$

Although there are much more efficient approaches to generating $\mathbf{Z}$ (e.g. Cholesky or QR decomposition), the eigen decomposition approach provides a simple method to introduce further improvements. In other research the observation has been made that the eigenvectors $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{M-1}$ are mainly a function of the signal(s) present (i.e. signal subspace) and the eigenvectors $\mathbf{v}_M, \mathbf{v}_{M+1}, \ldots, \mathbf{v}_{N-1}$ (i.e. noise subspace) are mainly a function of noise, e.g. [4] and [5]. Therefore the noise subspace eigenvectors can be ignored and the columns of $\mathbf{Z}$ computed according to

$$\mathbf{z}_i = \sqrt{\lambda_i - \lambda_\eta} \mathbf{v}_i \quad \text{for } i = 0, 1, 2, \ldots, M-1 \tag{68}$$

where λ_η compensates for the cumulative effects of noise in the main diagonal of the matrix $\mathbf{XX}^H$ and is estimated using,

$$\lambda_\eta = \frac{1}{N-M} \sum_{i=M}^{N-1} \lambda_i. \tag{69}$$

A secondary advantage of rejecting the noise subspace eigenvectors is that the number of input vectors is further reduced from N to M.

6.0   SIMULATION RESULTS

To investigate the performance of the BDF approach, computer simulations were run based on a linear N element sensor array with uniform half-wavelength spacing. Noise was additive white Gaussian noise with variance η^2 at each sensor output. The noise was also considered to be isotropic in nature and uncorrelated from sensor to sensor. The signal-to-noise ratio (SNR) is defined here as the ratio of a signal with unity power (i.e.

$c_m(t) = 1$ in equation (1)) to the noise power, or

$$SNR = -10\log(\eta^2) \quad \text{dB}. \tag{70}$$

In the simulations used to generate the following results all signals were constant amplitude and unity power (unless otherwise stated). For uncorrelated signals, the signal phase for each signal was independently and randomly chosen between $-\pi$ and π radians for each snapshot. For fully correlated signals, signal phases were randomly chosen between $-\pi$ and π radians for each snapshot, but the relative phases between signals was kept constant. Each trial consisted of processing K snapshots of sensor data to generate a single estimate of the signal bearings. Successive trials consisted of using the same signal data but increasing the contribution of the noise data in 1 dB steps until the entire range of interest was covered. Then new signal and noise data was generated and the process repeated L times.

For comparison purposes three different methods were used to estimate the signal bearings for each trial. These three methods were the BDF method, the DML method, and root-MUSIC [15]. The root-MUSIC method was included since it has been extensively reported in the open literature and become a standard by which other estimators are often compared.

For the BDF method the value of the step adjustment parameter μ_0 in equation (63) was chosen to be 0.3. It was empirically found that values greater than this degraded accuracy, while smaller values increased the computation time without significantly increasing accuracy. In the event that an iteration resulted in a greater error (i.e., $\varepsilon(\mathbf{X}_{i+1}) > \varepsilon(\mathbf{X}_i)$), μ was decreased by 30% and the iteration repeated. Iterations in the algorithm were stopped when the value of $\varepsilon(\mathbf{X}_i) \leq 10^{-14}\text{trace}\{\mathbf{X}^H\mathbf{X}\}$, which typically occurred between 50 and 200 iterations. This value was more than sufficient to ensure accuracy and could probably have been relaxed in most cases.

For the DML method an exhaustive search was performed by varying each angle in 1 degree increments to locate minimums in the MSE function, followed by a fine 0.1 degree search around each minimum to determine the global minimum. The Alternating Projection Maximization algorithm [7] was then used to further fine-tune the results.

For the root-Music estimator the estimated $(p+1) \times (p+1)$ covariance matrix was computed using equations (31), (35), (39), and (41) where the parameter p was used in place of M. For multiple snapshot processing with uncorrelated signals, the value of

p was chosen to be $N - 1$ (instead of M as for the BDF method). For single snapshot processing, or fully correlated signals, the value of p was chosen to optimize accuracy and threshold performance for each particular signal environment (i.e. spatial smoothing was used).

Accuracy is defined here as

$$Accuracy = \sqrt{\frac{1}{ML} \sum_{m=1}^{M} \sum_{l=1}^{K} (\hat{\phi}_{ml} - \phi_m)^2}, \tag{71}$$

where L represents the number of trials and $\hat{\phi}_{ml}$ is the estimate of the mth bearing for the lth trial.

6.1 Comparing Estimates from Different Methods

Figure 1 compares the performance of the BDF method with the DML and root-MUSIC methods over 100 trials with single snapshot processing ($K = 1$), $N = 8$ sensors, a signal-to-noise ratio of 30 dB, and two signals with bearings of 40 and 120 degrees. For the root-MUSIC method the value $p = 5$ was used since this not only optimizes accuracy but maximizes the number of signal bearings that can be estimated by this method using a single snapshot. In this example the BDF method produces nearly identical results to the DML method. The root-MUSIC estimates are different and slightly poorer (in this example by a factor of 1.4).

6.2 Damped Solutions

In deriving an error function for the BDF method, the signal model also incorporated undesired damped sinusoid pairs of the form represented by equation (33). The undesired solutions introduce a mechanism for error in the BDF algorithm. For example, against two signals with different bearings the DML method will estimate two distinct bearings. In the BDF method, if equation (33) is chosen as the best model for the data, then the estimated bearings for the two signals will be identical. This is obviously not correct.

In practice the selection of the undesired forms are far less common in the solution than the desired signals, and when they do occur, do not significantly degrade the

18
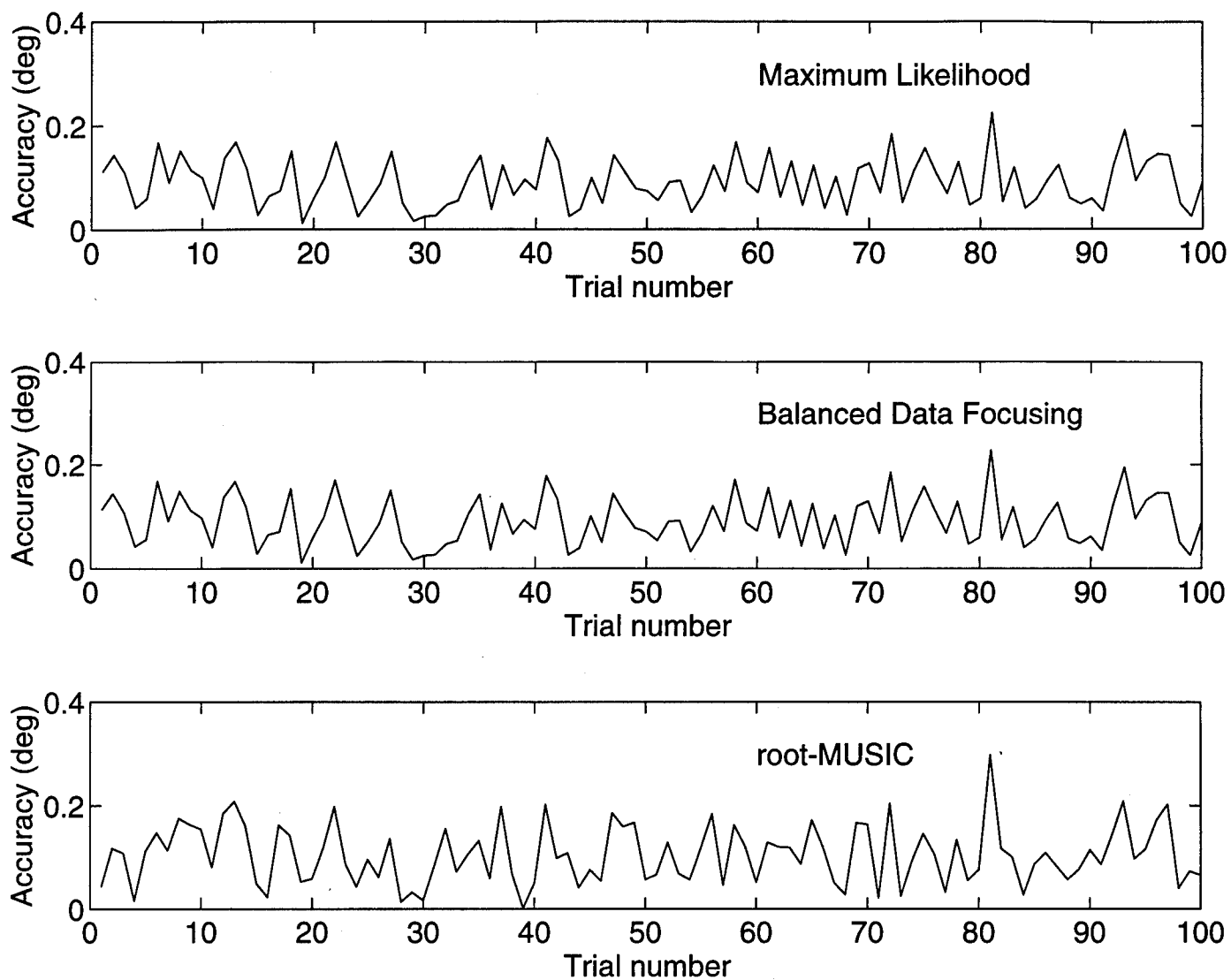
Figure 1: Estimator bearing errors for a series of 100 trials

Table 1: Damped sinusoid solutions

| BDF Signal Pair		DML Estimate	
ρ	ϕ	ϕ_1	ϕ_2
0.862	67.61	67.58	67.66
0.873	68.14	68.00	68.15
0.886	67.00	66.96	67.08
0.895	66.92	66.88	67.00
0.914	67.52	67.47	67.56
0.937	67.65	67.60	67.68
0.940	67.67	67.66	67.73
0.946	66.85	66.81	66.90
0.955	67.47	67.41	67.49
0.963	67.61	67.58	67.65

results. To illustrate this a simulation experiment was run with two signals at bearings of 65 and 70 degrees and a signal-to-noise ratio of 23 dB. Using single snapshot data ($K = 1$) estimates of the signal bearings were made using both the DML and BDF methods.

Out of 500 trials, damped sinusoid pair estimates occurred 3 times for the BDF method. The estimated bearing ϕ and corresponding damping factor ϕ for the first 10 of these undesired solutions is shown in Table 1. The corresponding DML estimates, ϕ_1 and ϕ_2, are also shown in the table.

It is readily apparent from the results shown in Table 1 that when pairs of undesired signal types are selected by the BDF method, the DML solution has almost converged to a single bearing. The result is, that compared to the DML method, the bearing accuracy of the BDF method is not significantly degraded.

6.3 Comparison of the EV and BLI Approaches

Figure 2 shows the performance of the BDF method under difficult conditions (8 sensors, single snapshot processing, and signals at closely space bearings of 69 and 70 degrees) for two different approaches to filter estimation, namely, the eigenvector approach (EV) and the balanced linear interpolator (BLI). For this and the following simulations $L = 500$ trials (see equation (71)). At high signal-to-noise ratios both approaches perform equally well. At lower signal-to-noise levels, a point is reached where performance degrades
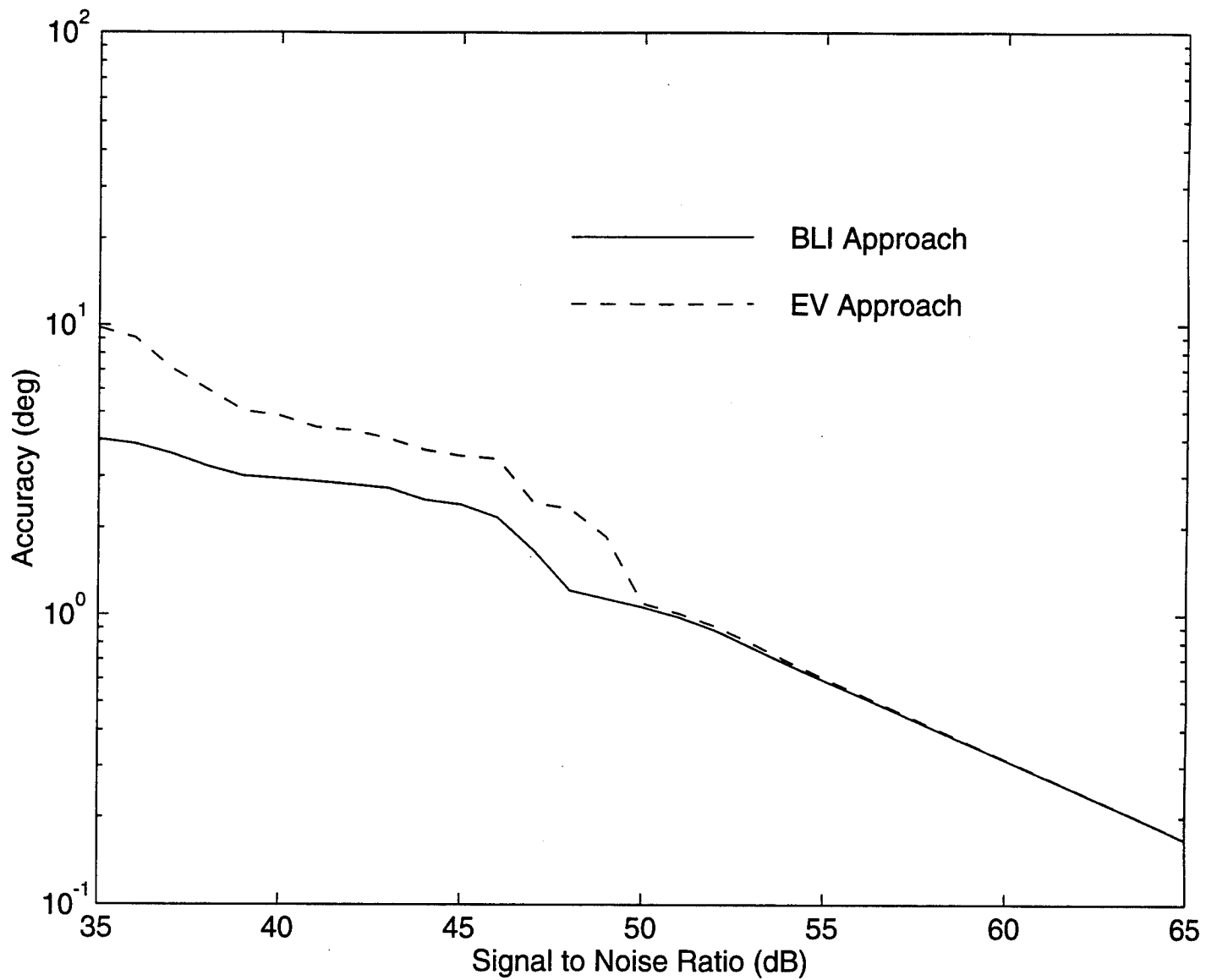
Figure 2: Estimator bearing accuracy for different focusing schemes ($\phi = 69$ and 70 degrees, $N = 8$, $K = 1$)

21

at a significantly higher rate than would be predicted from performance at higher signal-to-noise ratios. Defining this as the threshold point, then in this example, threshold occurs at 51 dB for the EV approach, and 47 dB for the BLI approach. In comparing the BLI approach to the EV approach, the BLI clearly outperforms the EV approach. This is a fortuitous result since the BLI approach is computationally less intensive. Given these advantages, the following simulations were performed using the balanced linear interpolator (as described in Section 4.2) to generate the filter coefficients.

6.4 Single Snapshot Estimator Performance

Figures 3 and 4 show the accuracy of the BDF, DML, and root-MUSIC methods for a single-signal environment and a two-signal environment, respectively, using $N = 8$ sensors. Not surprisingly, all three estimators perform almost equally well (within 1 dB) in the single-signal case. The performance advantage of the DML or BDF estimators compared to root-MUSIC becomes apparent when multiple signal environments are considered. In Figure 4, for example, the DML and BDF estimators have better accuracy than root-MUSIC for signal-to-noise ratios above threshold. The values of p used for root-MUSIC in Figures 3 and 3 were $p = 7$ and $p = 5$ respectively.

Figures 5 and 6 show two examples of single snapshot processing with $N = 8$ sensors where two signals are closely spaced ($p = 5$ for root-MUSIC). In both cases the threshold points for the BDF and DML methods occur at signal-to-noise ratios up to 6 dB lower than root-MUSIC. The apparently superior performance of the BDF method compared to the DML method for signal-to-noise ratios below the threshold point for each estimator is not important, since the bearing estimates in this region are unreliable and often meaningless (random). Therefore, there is relatively little to be gained by comparing the accuracies of different estimators for signal-to-noise ratios below the threshold point. The superior performance of root-MUSIC above threshold in Figure 6 is surprising since using a value of $p < N - 1$ effectively reduces the array aperture. This usually results in poorer accuracy, as in Figure 5, however in Figure 6, the opposite is true.

Figures 7 and 8 show two more examples of single snapshot processing with $N = 4$ and $N = 16$ sensors respectively (with $p = 2$ and $p = 10$ respectively for root-MUSIC). In these two examples the threshold points for all three estimators is identical. Above threshold the BDF and DML methods have superior accuracy to root-MUSIC equivalent to a 2 dB increase in the SNR.
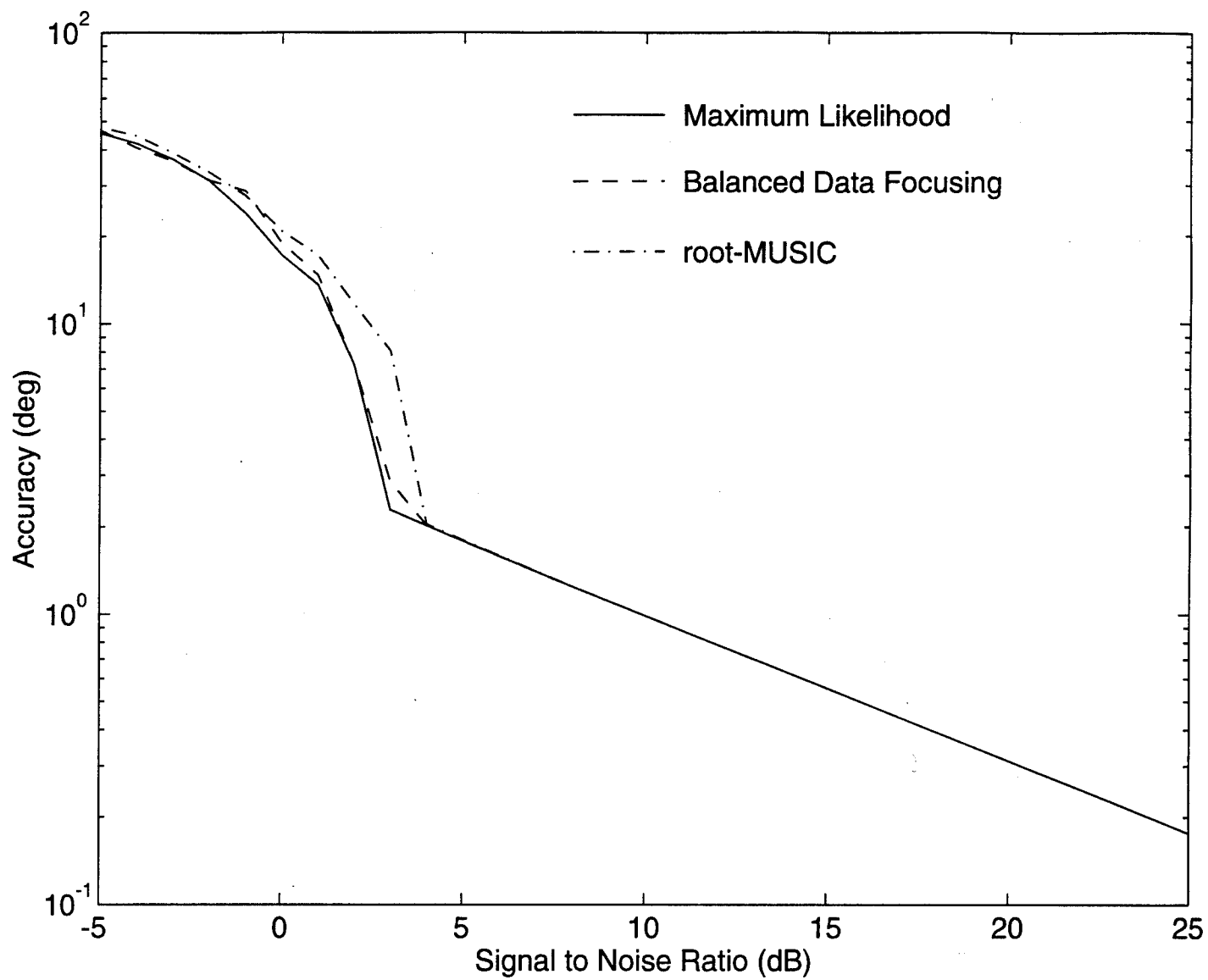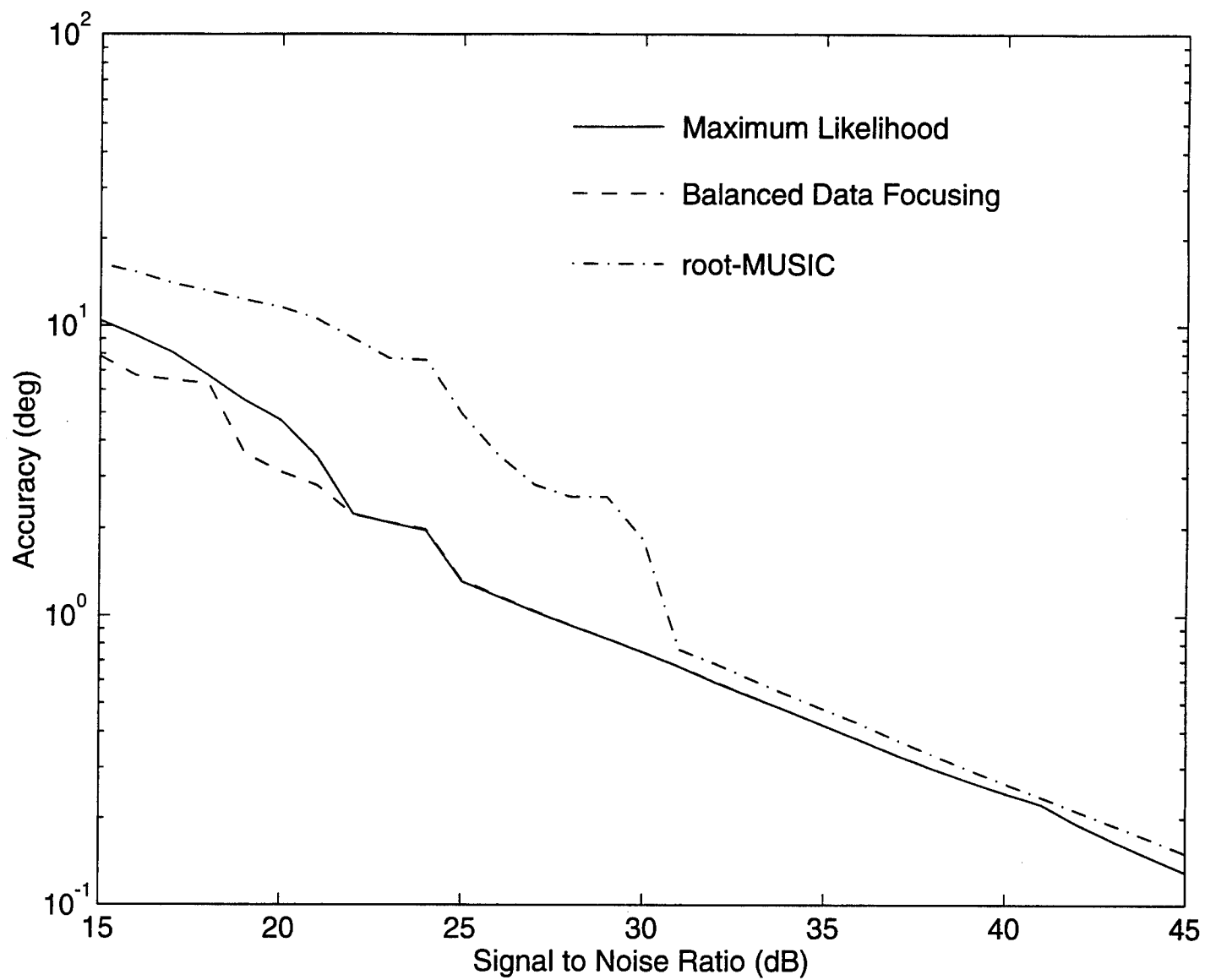
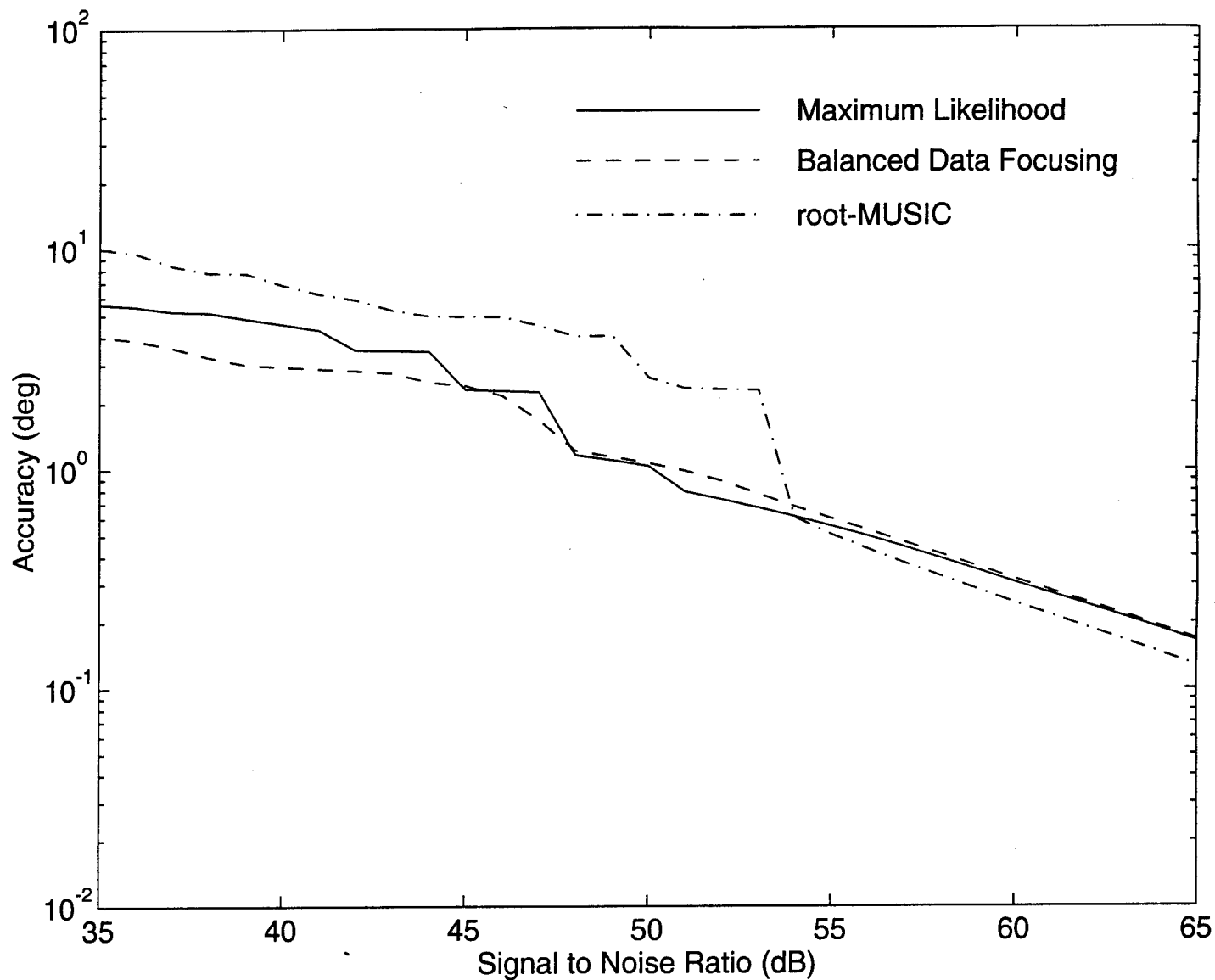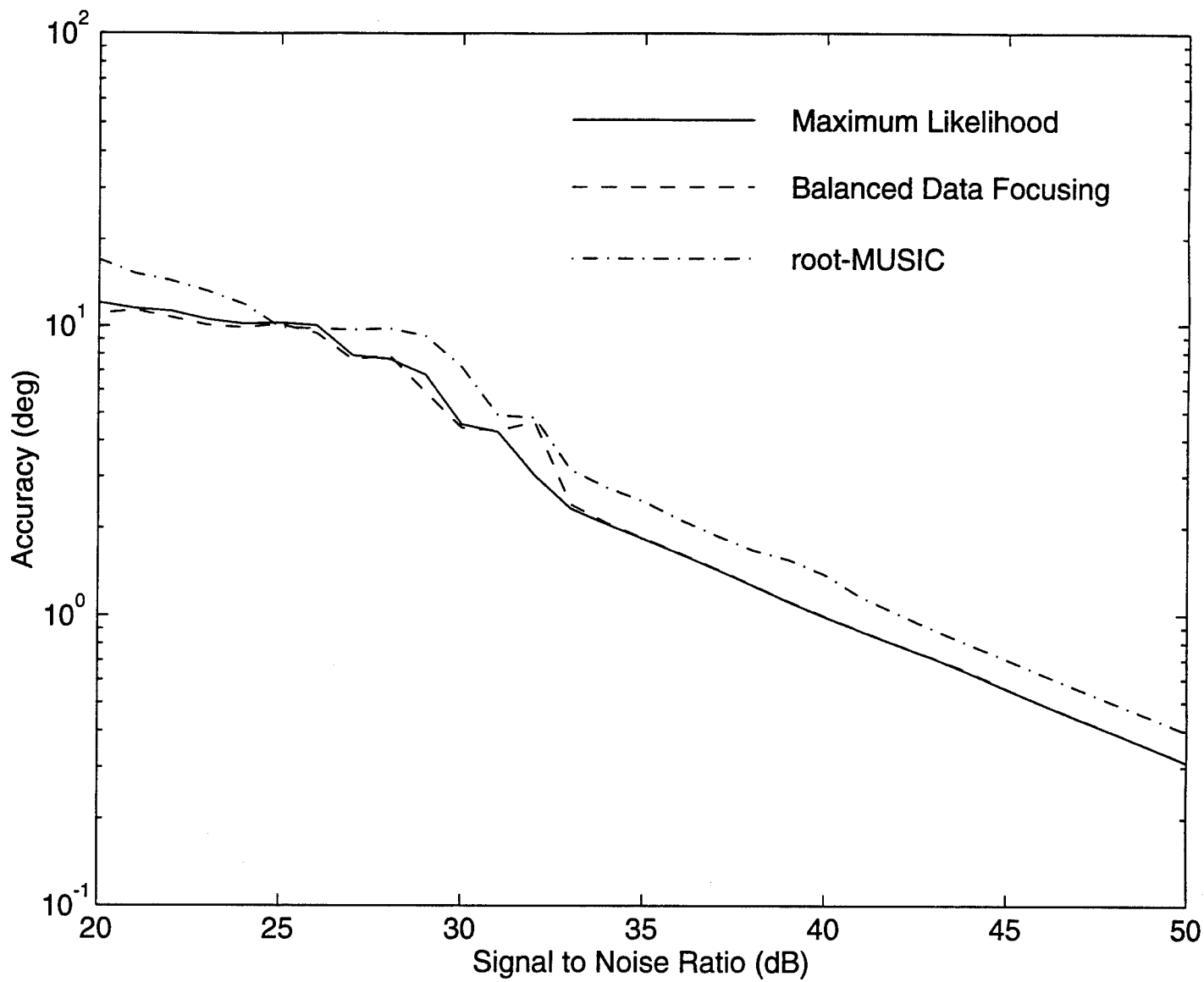**Figure 3**: Estimator bearing accuracy versus SNR ($\phi = 40$ degrees, $N = 8$, $K = 1$)
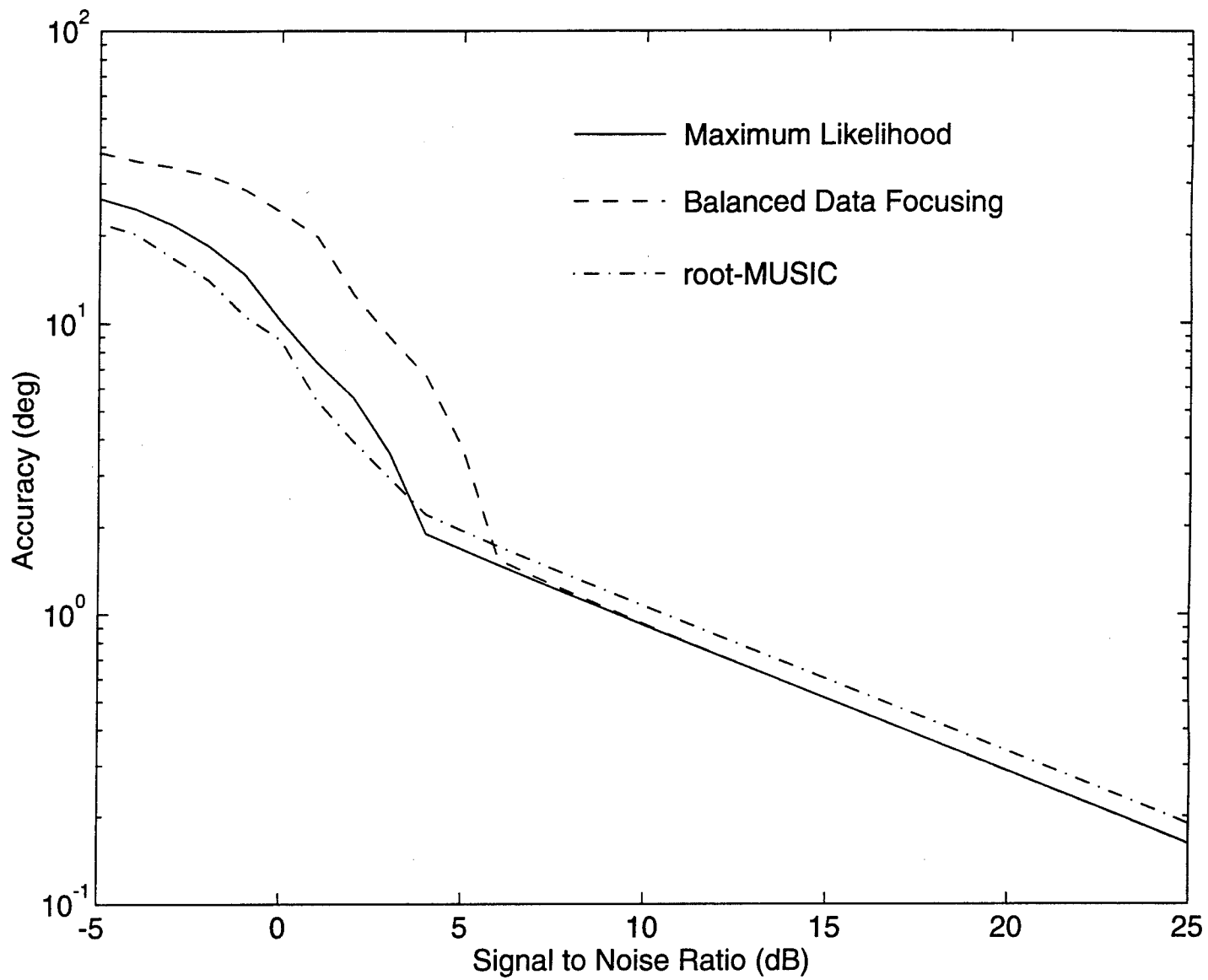
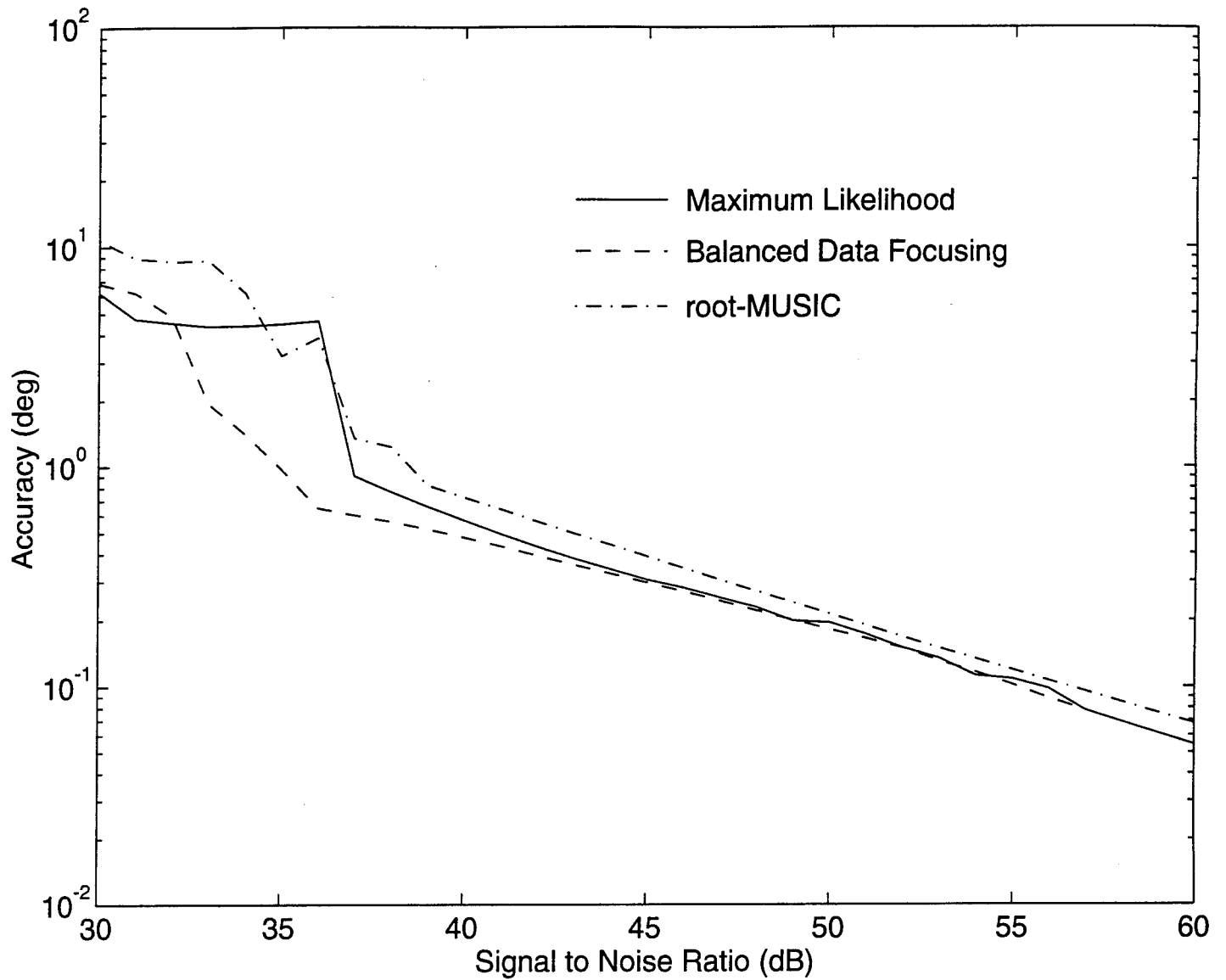**Figure 4**: Estimator bearing accuracy versus SNR ($\phi = 40$ and 120 degrees, $N = 8$, $K = 1$)

24

Figure 5: Estimator bearing accuracy versus SNR ($\phi = 65$ and 70 degrees, $N = 8$, $K = 1$)
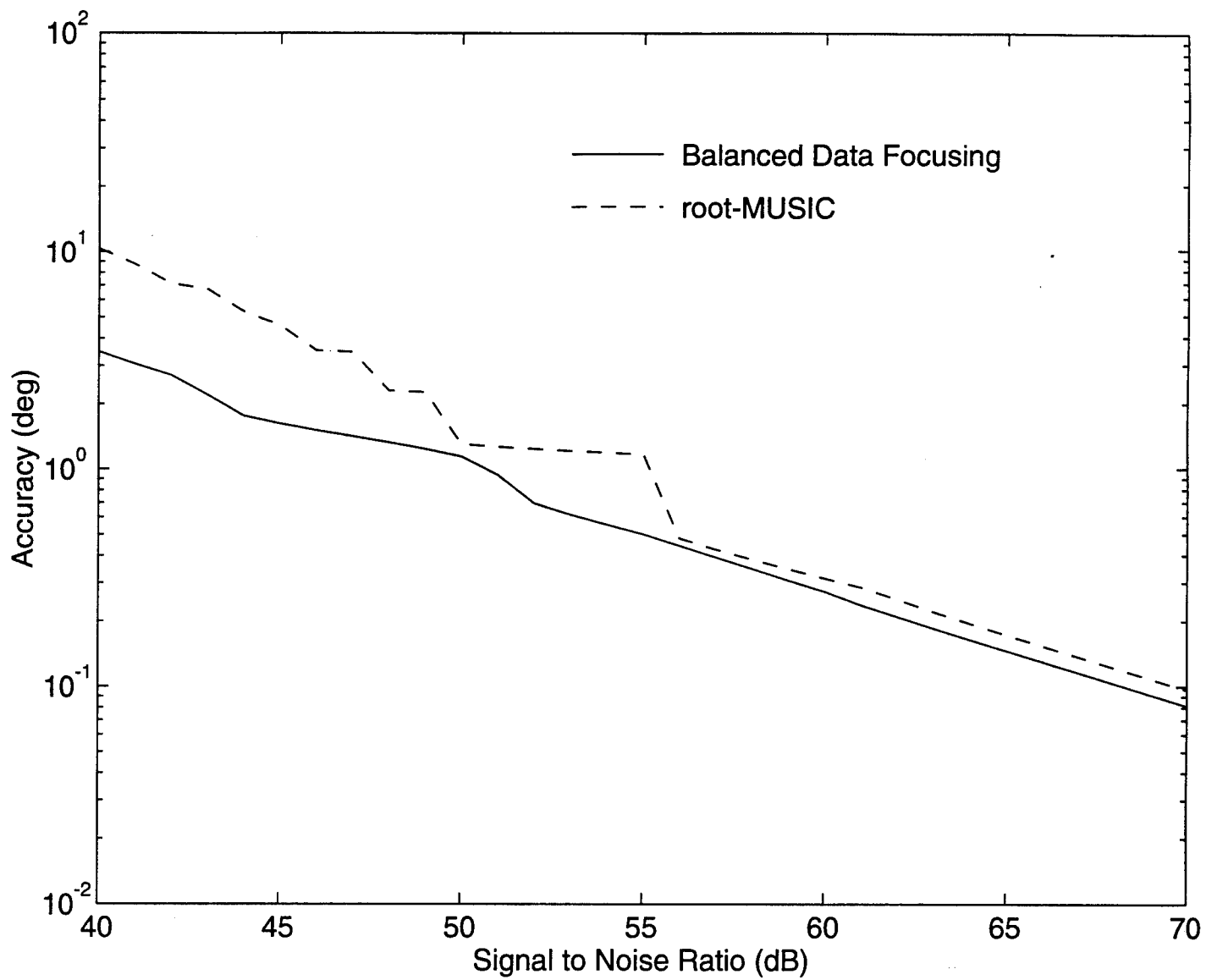
Figure 6: Estimator bearing accuracy versus SNR ($\phi = 69$ and 70 degrees, $N = 8$, $K = 1$)

Figure 7: Estimator bearing accuracy versus SNR ($\phi = 45$ and 55 degrees, $N = 4$, $K = 1$)

27

Figure 8: Estimator bearing accuracy versus SNR ($\phi = 29$ and 31 degrees, $N = 16$, $K = 1$)

28

Figure 9: Estimator bearing accuracy versus SNR (ϕ = 65, 70, and 75 degrees, N = 8, K = 1)

Figure 9 shows one final single snapshot processing example for 8 sensors and three signals ($p = 5$ for root-MUSIC). Note that the DML results are not shown here due to the computational difficulties of performing an exhaustive three dimensional search. Analysis of estimator performance on a trial by trial basis for SNR = 47 dB and SNR = 48 dB has shown that for this example threshold performance of the DML and BDF methods are within 1 dB. Above threshold their accuracies are the same.

## 6.5 Multiple Snapshot Estimator Performance

The performance of all three estimators in the last five examples is quite poor. Better results can be achieved if more snapshots are available for each trial. Figure 10 shows the performance of the BDF method using $K = 100$ snapshots per trial for various types of noise preprocessing (see Section 5.0) and assuming the signals are uncorrelated. These include no preprocessing, full processing using equation (66), and partial processing using equation (66) but including all eigenvectors (i.e. $i = 0, 1, 2, \ldots, N - 1$ which is equivalent to subtracting the diagonal noise matrix $\lambda_\eta \mathbf{I}$ from $\mathbf{R}_N$). From the results, performance with full noise processing is the best, so that in the following examples full noise processing for the BDF method is performed.

6.5.1 Uncorrelated Signals

Figure 11 shows the same signal environment as in Figure 10 but comparing the BDF, DML, and root-MUSIC approaches. Again both the DML and BDF methods outperform root-MUSIC, despite the fact that under these conditions root-MUSIC can utilize the full array (i.e. no spatial smoothing, or $p = 7$).

Figure 12 shows the same signal environment as in Figure 6 except where $K = 100$ snapshots were used for each trial. The relative performance of each of the estimators is the same as in Figure 12.

Figure 13 replots the performance of the BDF method from Figures 6 and 12 over the same range of SNR. This figure illustrates the improvements that result with multiple snapshot processing compared to single snapshot processing. In this example, increasing the number of snapshots K from 1 to 100 is equivalent to increasing the SNR by 40 dB.

30
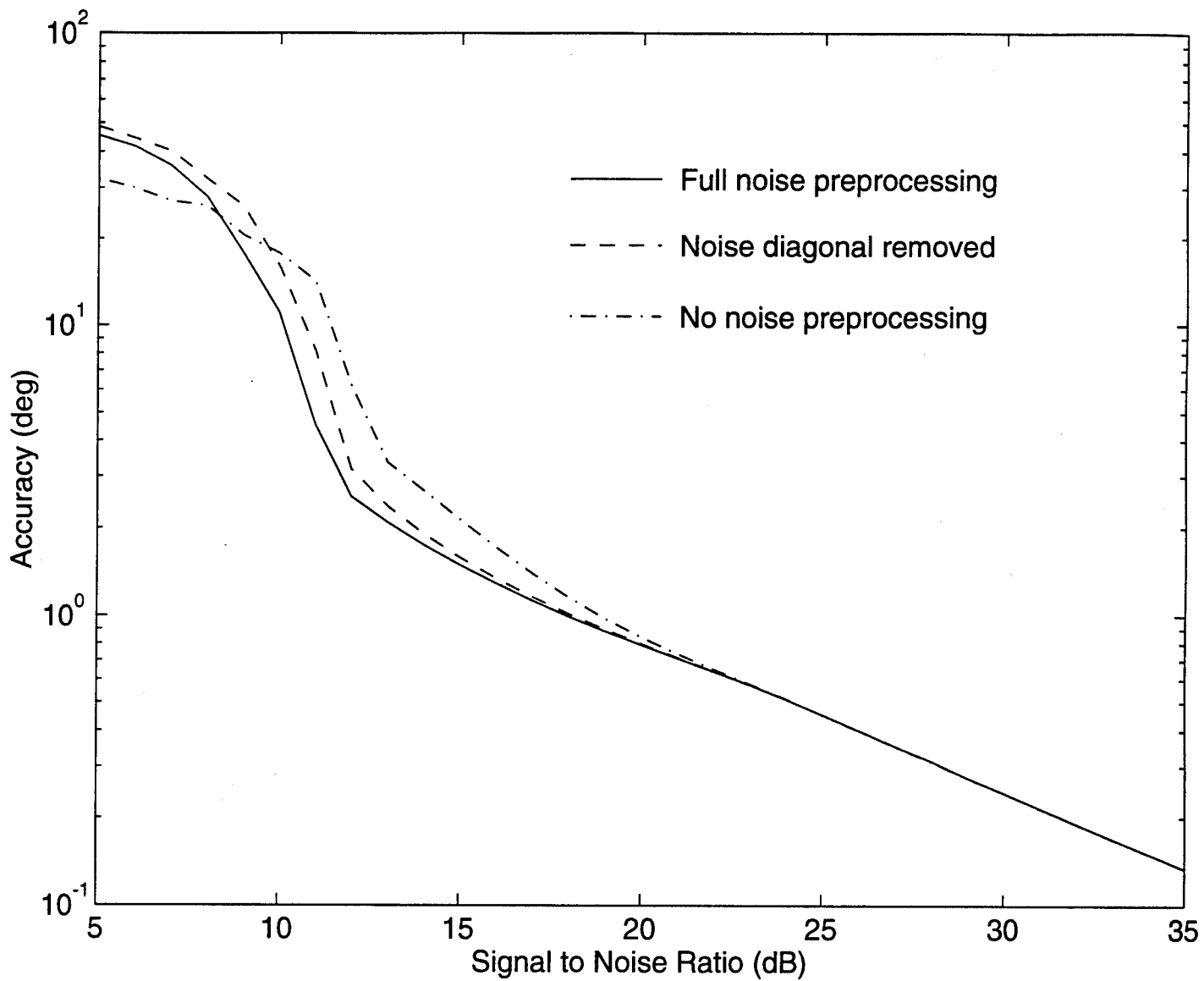
Figure 10: Comparison of Noise Preprocessing Schemes ($\phi = 29$ and 31 degrees, $N = 8$, $K = 100$)
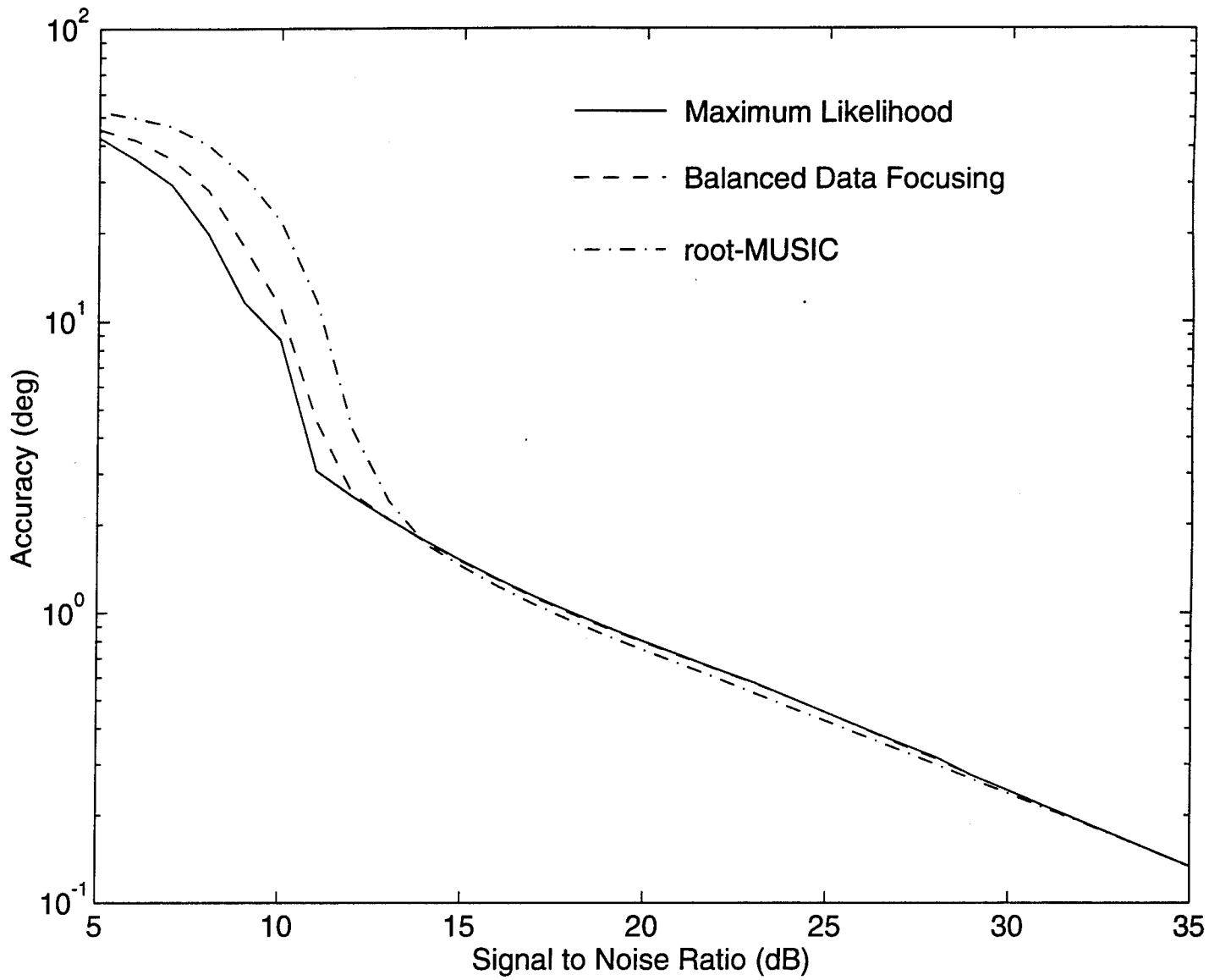
Figure 11: Estimator bearing accuracy versus SNR ($\phi = 29$ and 31 degrees, $N = 8$, $K = 100$)
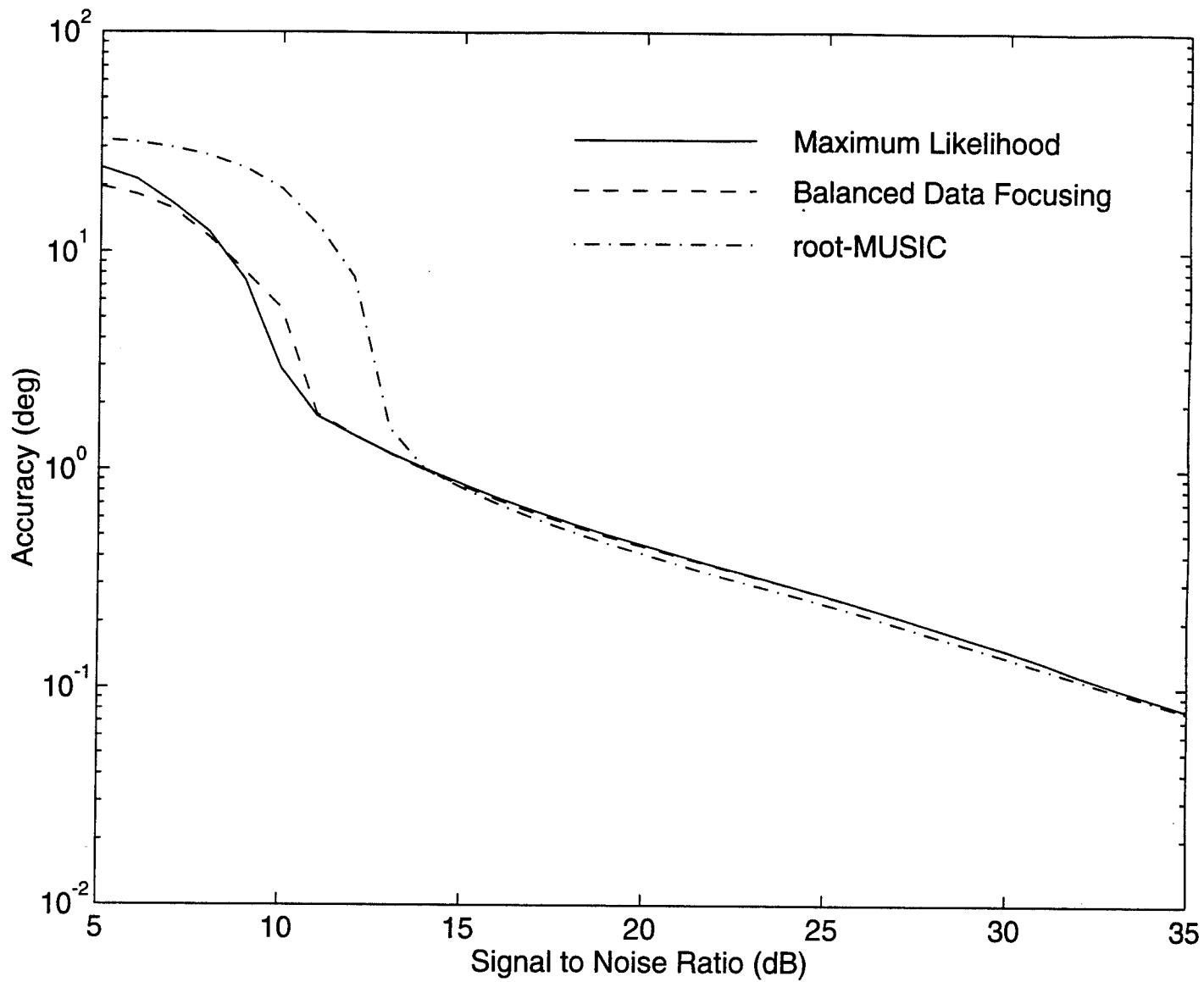
Figure 12: Estimator bearing accuracy versus SNR ($\phi = 69$ and 70 degrees, $N = 8$, $K = 100$)

Figure 13: BDF method bearing accuracy versus SNR for $K = 1$ and $K = 100$ ($\phi = 45$ and 55 degrees, $N = 4$)
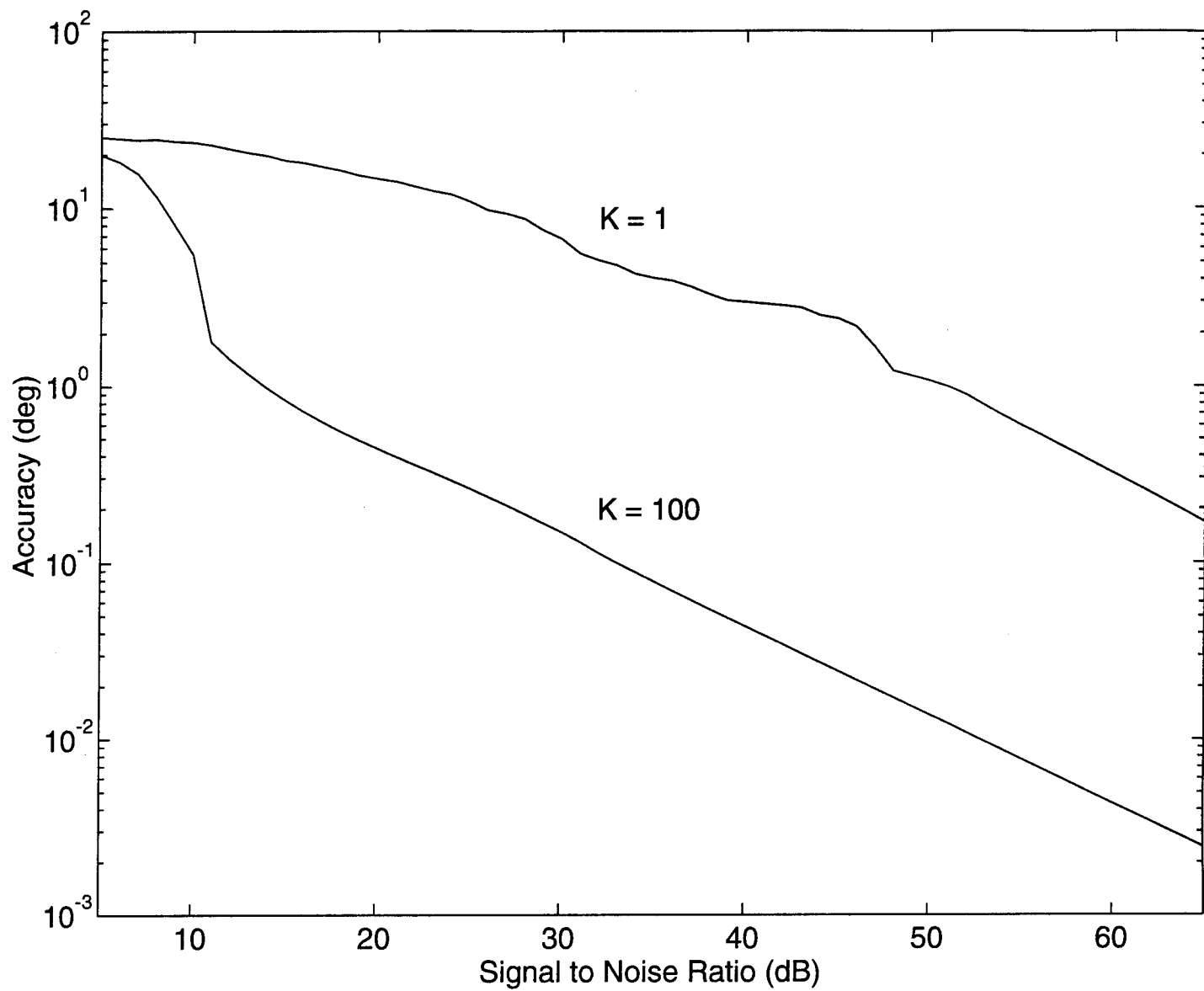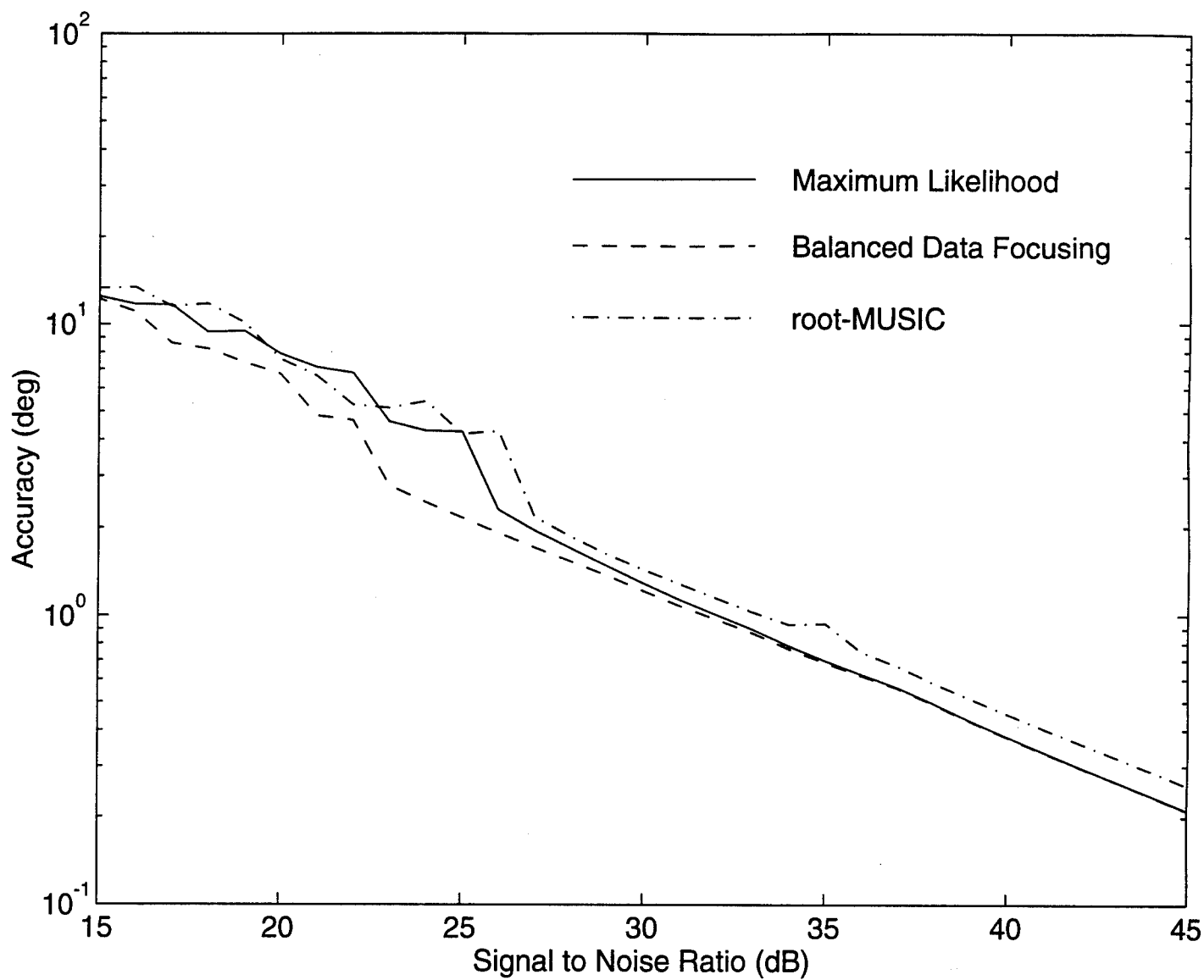
Figure 14: Estimator bearing accuracy versus SNR ($\phi = 45$ and 55 degrees, $N = 4$, $K = 10$)

35

Figure 15: Estimator bearing accuracy versus SNR ($\phi = 45, 47, 50, 60,$ and 120 degrees, $N = 10, K = 10$)

Figure 15 is a five-signal (uncorrelated) example comparing the BDF method to root-MUSIC ($p = 7$) for the case where the $K = 10$ snapshots and $N = 10$ sensors were used. Again due to the computational difficulties of implementing the DML method for more than two signals, its performance is not shown. Analysis of DML performance on a trial by trial basis for SNR's between 38 and 40 dB, and using various initial starting bearings, has determined that threshold for the DML method occurs at 39 dB or higher. Above threshold the DML and BDF methods were found to have the same accuracy.

6.5.2 Correlated Signals

Figures 14 and Figure 16 show examples of increasing the number of snapshots for the same signal environment used in Figure 7 and assuming the signals are fully correlated. In Figure 14 the threshold for the BDF method occurs at a lower SNR than either the DML or root-MUSIC methods. In Figure 16 the threshold for the BDF and root-MUSIC methods occur at a slightly lower SNR than the DML method. The superior threshold performance of the BDF method compared to the DML method can be explained by the fact that the BDF method appears to favour signal geometries with closely spaced signals. This assertion is supported by that the fact that in Figure 4 the threshold performance of the BDF method was slightly worse when the signals were widely spaced ($\phi = 4$ and 60 degrees). At higher signal-to-noise ratios both the BDF and DML methods have identical accuracies and outperform root-MUSIC due to the loss of effective aperture as a result of using spatial smoothing. Spatial smoothing was required for the root-MUSIC method due to the fact that the signals were fully correlated.

The advantage of increasing the number of snapshots K is also illustrated in Figure 17 which summarizes the results for the BDF method from Figures 7, 14, and 16. Although the improvement is smaller than for uncorrelated signals, it is still significant with an equivalent SNR improvement of approximately 20 dB for the case where $K = 100$ compared to the case where $K = 1$.

6.6 Unequal Signal Amplitudes

Figure 18 shows one final simulation example where the signal amplitudes were 1.0 and 0.1 corresponding to signal bearings of 95 and 105 degrees. The signals were assumed to be uncorrelated and $p = 5$ was used for the root-MUSIC method. Again
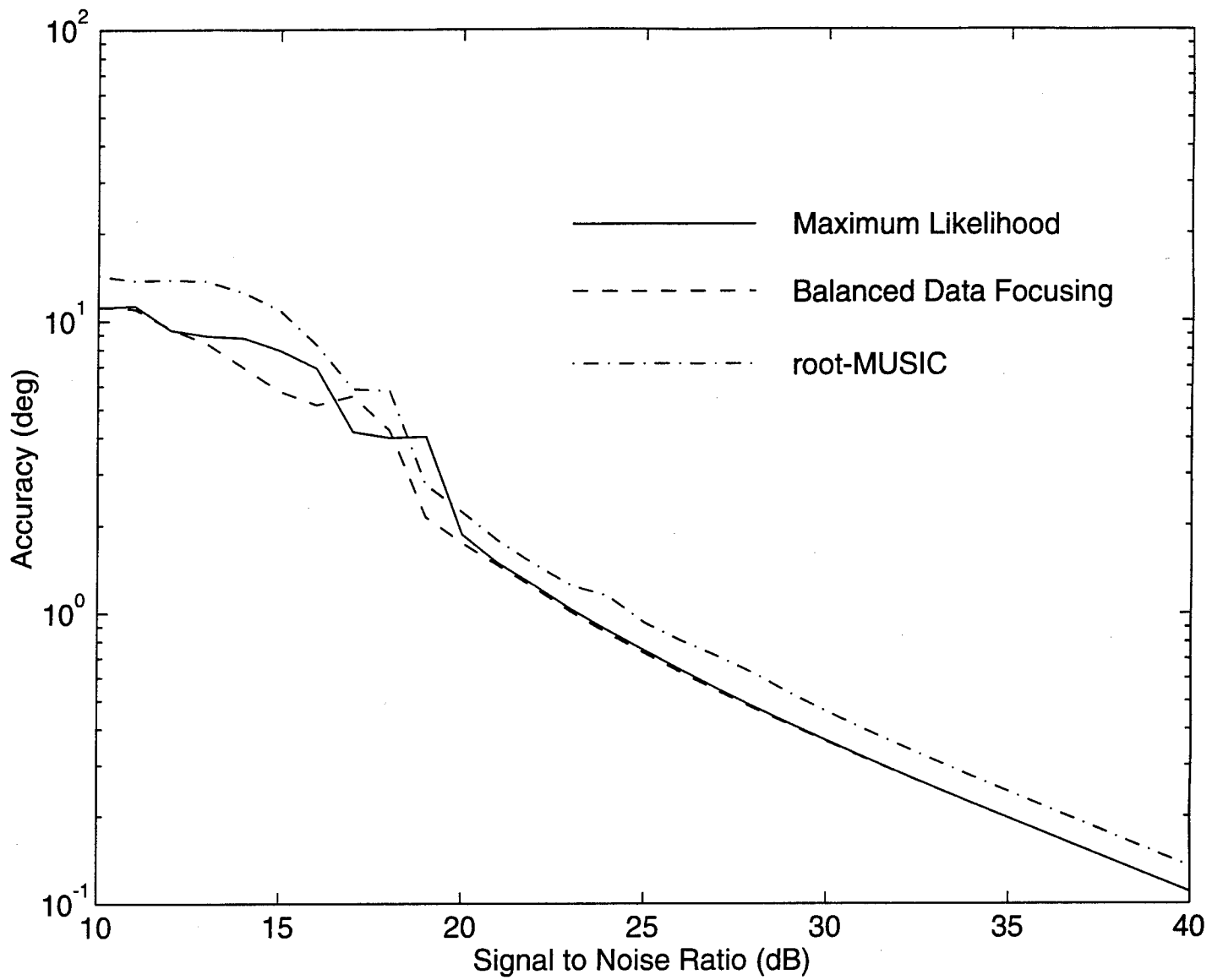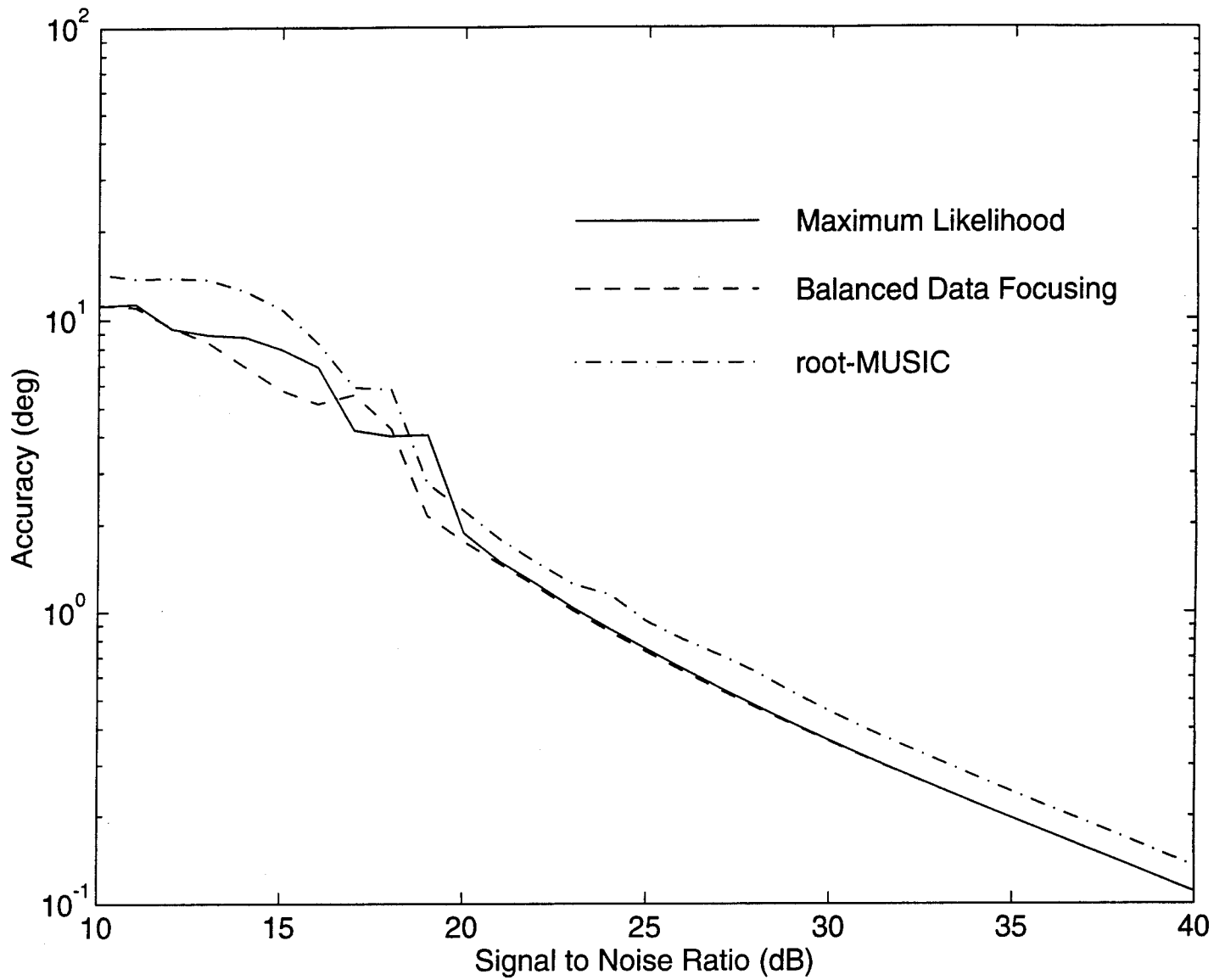
Figure 16: Estimator bearing accuracy versus SNR ($\phi = 45$ and 55 degrees, $N = 4$, $K = 100$)
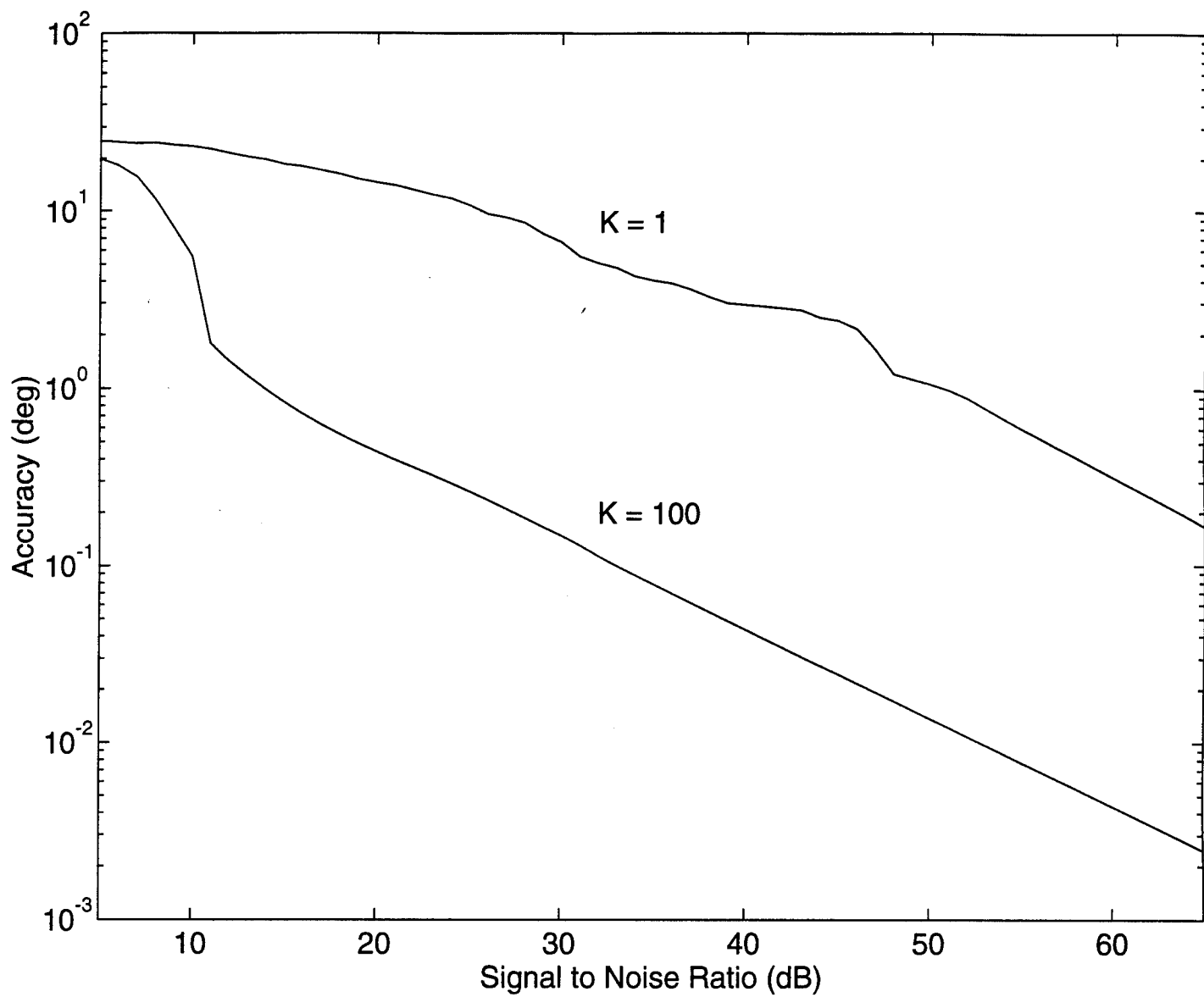
Figure 17: BDF method bearing accuracy versus SNR for $K = 1$, $K = 10$ and $K = 100$ ($\phi = 45$ and 55 degrees, $N = 4$)

Figure 18: Estimator bearing accuracy versus SNR for unequal signal amplitudes ($\phi = 95$ and 105 degrees, $N = 5$, $K = 100$)
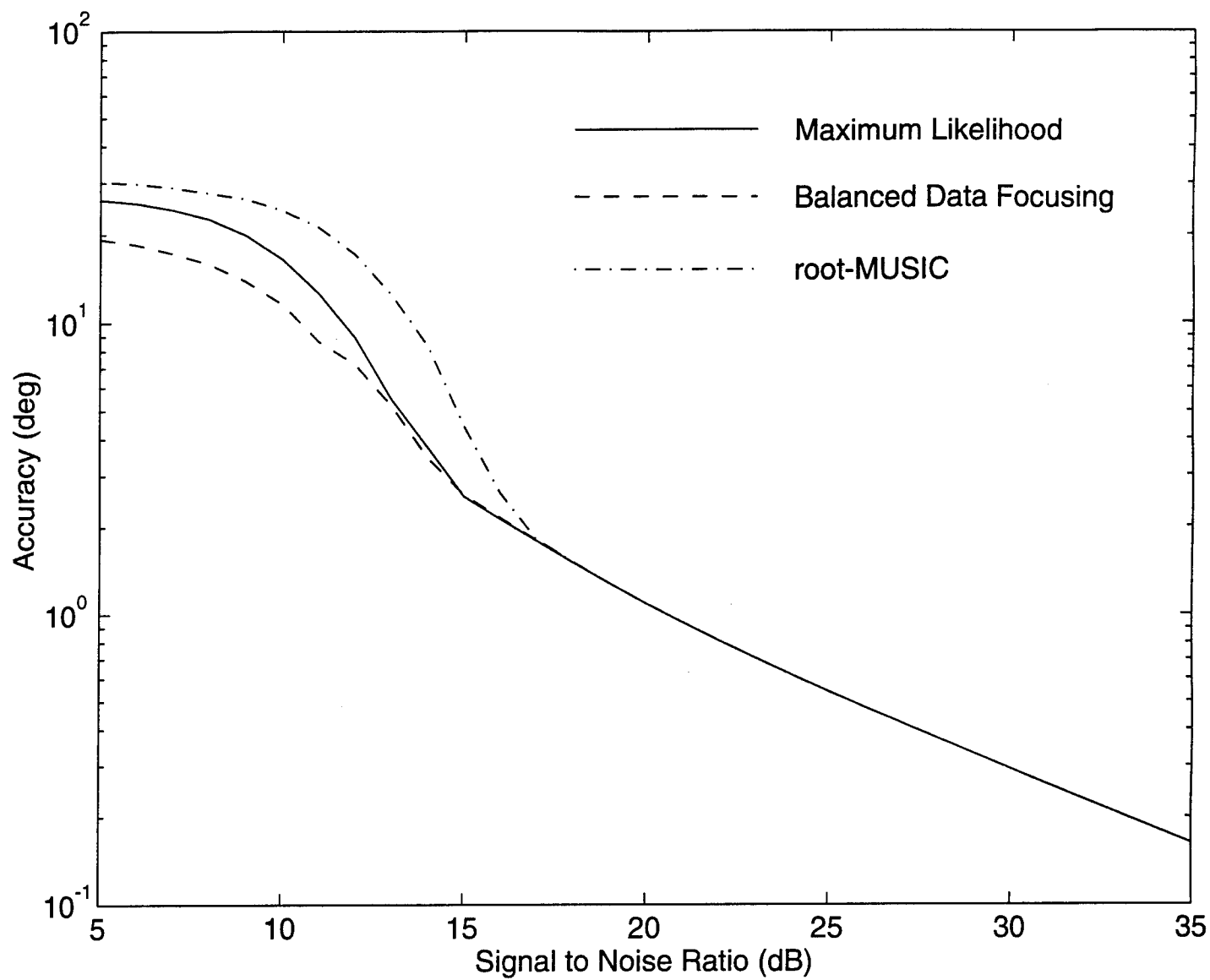
the BDF method performs as well as the DML method and better than the root-MUSIC method in terms of threshold. Above threshold all three estimators have identical accuracies.

6.7 Processing Speed

Although no claim is made that the algorithms tested were optimally coded, or for that matter, that the most appropriate search procedures were chosen, some mention of relative processing times is in order. Not surprisingly root-MUSIC was the fastest algorithm under all conditions. The BDF algorithm was found to be about an order of magnitude slower than root-MUSIC due to the slower convergence of the BDF algorithm (typically between 50 and 200 iterations). For the single snapshot case the convergence time of the BDF method was found to be relatively insensitive to the signal parameters including the number of signals, signal amplitudes, and bearing spacing. For the multiple snapshot case convergence time was slower for an increasing number of signals due to the increase in number of computations ($\propto M$) and a general increase in the number of iterations required.

As stated earlier, the DML method was implemented using an exhaustive search to approximately locate the global minimum of the DML error function (i.e. determine the initial bearing estimates), followed by a fine search to fine tune the results. Ignoring the estimation of the initial bearings for the moment, fine tuning was achieved using the APL-ML algorithm. This algorithm was chosen as it was found to be computationally the fastest of the algorithms surveyed ([6], [7], and [8]) for the case of two closely spaced signals. Generally, once properly initialized, the APM-ML algorithm converged in a slightly greater amount of time than the BDF method for the two-signal case, although occasionally it was many times slower. The APM-ML algorithm was significantly slower for cases involving three or more signals especially when they were closely spaced in bearing. Returning to the problem of estimating the initial bearings, the additional processing required adds to the total DML method processing time compared to simply fine tuning the results. In the case where an exhaustive search is used (as reported here), the additional time can be an order of magnitude or more greater than the time required for fine tuning. An alternate procedure for determining the initial bearings which is simple and fast was proposed in [7] . Although this procedure was reported to work well in the multiple snapshot case, for the single snapshot processing examples described in this paper, this procedure was not as effective.

7.0 CONCLUSIONS

In this paper a bearing estimation method for uniform linear arrays called Balanced Data Focusing was described which was shown to achieve maximum likelihood performance. The method uses an error function based on a linear interpolation filter which estimates the minimum mean squared error (MSE) of the deterministic maximum likelihood (DML) method. Using this error function, data input from the sensors is adjusted, or focused, until the minimum MSE is 0. The focused data can then be treated as noiseless data and any number of techniques used to extract the signal bearing information.

In simulations it was found that the BDF method performed as well as the DML method and outperformed root-MUSIC in terms of both accuracy and threshold performance, even under extreme conditions where signals were very closely spaced in azimuth. An important advantage of the BDF method compared to root-MUSIC, is that the BDF method can deal with both uncorrelated and fully correlated signals (or the equivalent single snapshot problem). An important advantage compared to the DML method is that no initial values are required. A poor choice of inital values for DML can lead to meaningless results.

A disadvantage of the BDF method is that it is slower than root-MUSIC, especially as the number of signals increases. Further research into better search algorithms (compared to the gradient descent technique used) would likely yield useful improvements in processing speed.

8.0 REFERENCES

[1] Burg, J.P., "Maximum Entropy Spectral Analysis", Ph.D. dissertation, Stanford University, Stanford, California, 1975.

[2] Gabriel, W.F., "Spectral Analysis and Adaptive Array Superresolution Techniques", *Proceedings of the IEEE*, vol. 68, no. 6, pp. 654-666, June 1980.

[3] Borgiotti, G.V., and Kaplan, L.J., "Superresolution of Uncorrelated Interference Sources by Using Adaptive Array Techniques", *IEEE transactions on Antennas and Propagation*, vol. 27, no. 6, pp. 842-845, November 1979.

[4] Schmidt, R.O., "Multiple Emitter Location and Signal Parameter Estimation", *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276-280, March 1986.

[5] Tufts, D., and Kumaresan, R., "Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood", *Proceedings of the IEEE*, vol. 70, no. 9, September 1982.

[6] Bresler, Y., and Macovski, A., "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 5, pp. 1081-1089, October 1986.

[7] Ziskind, I., and Wax, M., "Maximum Likelihood Estimation via the Alternating Projection Maximization Algorithm", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, ICASSP 87, pp. 2280-2283, 1987.

[8] Feder, M., and Weinstein, E., "Parameter Estimation of Superimposed Signals Using the EM Algorithm", *IEEE Transactions on Acoustics, Speech , and Signal Processing*, vol. 36, no. 4, pp. 477-489, April 1988.

[9] Stoica, P., Moses, R., Friedlander, B., and Soderstrom, T., "Maximum Likelihood Estimation of the Parameters of Multiple Sinusoids from Noisy Measurements", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 3, pp. 378-392, March 1989.

[10] Press, W., Flannery, B., Teukolsky, S., and Vetterling, W., *Numerical Recipes in C, The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1988.

[11] Stoica, P. and Sharman, K.C., "Maximum Likelihood Methods for Direction-of-Arrival Estimation", *IEEE Transactions on Acoustics, Speech , and Signal Processing*, vol. 38, no. 7, pp. 1132-1143, July 1990.

[12] Marple, S.L., *Digital Spectral Analysis with Applications*, Prentice-Hall, Alan V. Oppenheim, editor, Englewood Cliffs, New Jersey, 1987.

[13] Klema, V.C., and Laub, J.L., "The Singular Value Decomposition: Its Computation and Some Applications", *IEEE Transactions on Automatic Control*, vol. 25, no. 2, pp. 164-176, April 1980.

[14] Haykin, S., *Adaptive Filter Theory*, Prentice-Hall, Thomas Kailath, editor, Englewood Cliffs, New Jersey, 1986.

[15] Rao, B.D., and Hari, K., "Performance Analysis of Root-Music", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol.37, no. 2, pp. 1939-1949, December 1989.

[16] Read, W.J.L., "An Evaluation of Superresolution Methods for Tactical Radio Direction Finding", DREO Technical Report, No. 1091, October 1991.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.) DEFENCE RESEARCH ESTABLISHMENT OTTAWA NATIONAL DEFENCE SHIRLEYS BAY, OTTAWA, ONTARIO K1A 0Z4 CANADA	2. SECURITY CLASSIFICATION (overall security classification of the document. including special warning terms if applicable) UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C or U) in parentheses after the title.)

BALANCED DATA FOCUSING: DIRECTION OF ARRIVAL ESTIMATION WITH MAXIMUM LIKELIHOOD PERFORMANCE (U)

4. AUTHORS (Last name, first name, middle initial)

READ, W.J.L.

5. DATE OF PUBLICATION (month and year of publication of document) NOVEMBER 1994	6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.) 50	6b. NO. OF REFS (total cited in document) 16

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

DREO REPORT

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
NATIONAL DEFENCE
SHIRLEYS BAY, OTTAWA, ONTARIO K1A 0Z4 CANADA

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant) 041LX	9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)
10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.) DREO REPORT 1233	10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(XX) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availabilty (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

UNLIMITED

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

(U)   The performance of a direction of arrival estimation procedure at low signal-to-noise ratios and limited data samples is an important characteristic. The approach based on maximum likelihood (ML) estimation is considered to be among the best for this problem as long as the underlying signal model is properly chosen. Unfortunately, in most cases, there is no closed-form solution so fast search procedures are employed. Given no a priori knowledge, selecting the initial parameter values for these search procedures can be a difficult problem, especially under low signal-to-noise conditions. In this paper, a new method for uniform linear sensor arrays which overcomes the initial value problem is introduced. This method is called Balanced Data Focusing (BDF). Simulation results are included comparing the performance of this new method to that of the ML approach using the Alternating Projection Maximization search procedure and another popular estimation approach, the root-MUSIC method.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

DIRECTION FINDING
SUPERRESOLUTION
HIGH RESOLUTION
BEERING ESTIMATION
MAXIMUM LIKELIHOOD
MUSIC
NARROWBAND
LINEAR ARRAY
SIGNAL PROCESSING