

(12)

AD-A146 594

Technical Report
654

Performance of Bayes-Optimal Angle-of-Arrival Estimators

F.M. White

13 August 1984

Lincoln Laboratory

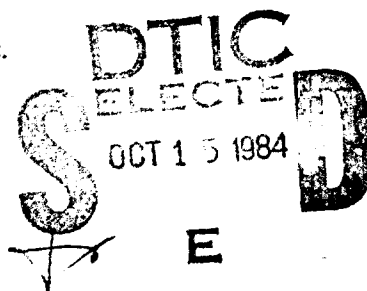
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LEXINGTON, MASSACHUSETTS



Prepared for the Department of Defense
under Electronic Systems Division Contract F19628-80-C-0002.

Approved for public release; distribution unlimited.



DTIC FILE COPY

84 10 11 001

**MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY**

**PERFORMANCE OF BAYES-OPTIMAL
ANGLE-OF-ARRIVAL ESTIMATORS**

F.M. WHITE

Group 44

TECHNICAL REPORT 654

13 AUGUST 1984

Approved for public release; distribution unlimited.

LEXINGTON

MASSACHUSETTS

ABSTRACT

The angle-of-arrival estimation problem for waves incident upon a sensor array was examined through a Monte Carlo evaluation of the performance of the Bayes-optimal MAP (maximum a posteriori) and MMSE (minimum mean square error) estimators. The case of two independent wave emitters of known powers as well as a multiple look, Gaussian signal in Gaussian noise statistical model were assumed. The Cramer-Rao bound on the estimator's rms error was computed for comparison.

The evaluation proceeded with the computation of MAP and MMSE angle estimates for 1000 random samples of array outputs and the accumulation of their rms errors. The probability of detecting both emitters with the optimal detector was also accumulated. This was done for .1, .03, and .01 beamwidths emitter separations and a range of signal-to-noise ratios (SNRs). The accuracy of the computations was assured through a simple finite grid approximation for the estimates, with no convergence problems, and through the evaluation of statistical confidence intervals for the Monte Carlo data.

The results of the evaluation indicated that the Cramer-Rao bound was achievable by both the MAP and MMSE estimators over a wide range of SNR provided a few as 10 looks had been taken. In general, the bound was achieved wherever both signals were detectable. These results were surprising since the bound exhibited unusual behavior; for example, in one SNR region, the bound showed smaller rms errors for more closely-spaced emitters.

Additional results included properties of the a posteriori probability density and an analytical computation of the performance of the known angles-of-arrival optimal detector.



Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

CONTENTS

Abstract	111
1.0. INTRODUCTION	1
1.1. AOA Estimation with Sensor Arrays	1
1.2. Statistical Model	5
1.3. Cramer-Rao Bound	9
1.4. Optimal Estimation and Detection	14
1.4.1. Estimation	14
1.4.2. Detection	17
1.5. Previous Work - Present Contribution	18
2.0. THE MAP AND MMSE ESTIMATORS AND THE OPTIMAL DETECTOR	20
2.1. MAP and MMSE Estimators	20
2.2. Optimal Detector	22
2.3. Examples of Properties of $p(\underline{\theta} \hat{R})$	24
3.0. MONTE CARLO PERFORMANCE EVALUATION	31
3.1. Computational Procedure	31
3.1.1. RMS Error Calculation	31
3.1.2. Probability of Detection Computation	35
3.1.3. Bias Computation	37
3.2. Results	37
3.2.1. RMS Error Results	38
3.2.2. Probability of Detection Results	41
3.2.3. Bias Results	47
4.0. CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH	51
ACKNOWLEDGEMENTS	53
REFERENCES	54
APPENDIX A - CONFIDENCE INTERVALS FOR THE MONTE CARLO RESULTS	57
APPENDIX B - THE KNOWN ANGLE-OF-ARRIVAL DETECTOR	61

1.0 INTRODUCTION

The estimation of the angles-of-arrival of wavefronts incident upon a sensor array is a well-known and important problem, occurring in such varied fields as geophysics, oceanography, emitter location, etc. For small angular separations of the waves or for small array apertures, estimation is difficult. As more and more performance is desired of such estimators, as the cost of physically large sensor arrays increases, and as the cost of real-time digital computation drops, the possibility of statistically optimum estimation becomes increasingly interesting and practical.

The goal of this report is not to develop a practical optimal estimator but rather to compute the performance of the optimal estimators and compare this performance to the Cramer-Rao bound. More specifically, assuming a Gaussian signal model with two incident waves of known powers, the rms error of the Bayes-optimal MAP (maximum a posteriori) and MMSE (minimum mean square error) estimators will be compared via a Monte Carlo simulation, and compared to the known powers Cramer-Rao bound. These rms errors will provide an absolute lower bound on the attainable rms error in angle-of-arrival estimation and will determine the tightness of the Cramer-Rao bound for this problem.

1.1 AOA Estimation with Sensor Arrays

The angle-of-arrival (AOA) estimation problem in its most general form consists of the estimation of the wavevectors $\underline{k}_i$ of say D incident waves based on the outputs of N arbitrarily placed wavefield sensors. Each sensor output is some known function of the field in its vicinity.

In this report, we specialize the problem to a simpler though common case. Later, we assume only two waves ($D=2$) in order to make computations feasible, but will keep D arbitrary whenever possible. Assume that all of the D waves are monochromatic plane waves with wave-vectors $\underline{k}_i$, complex amplitudes a_i at the origin ($\underline{r} = 0$), and frequency ω_0 . The total field at position $\underline{r}$ is then

$$E(\underline{r}) = \sum_{i=1}^D \operatorname{Re} \{ a_i e^{j\omega_0 t} e^{j\mathbf{k}_i \cdot \underline{r}} \} \quad (1.1)$$

by superposition (linear medium). Assume that the sensors are isotropic with coherent detectors so that the output of a sensor at location $\underline{r}$ is

$$x(\underline{r}) = \sum_{i=1}^D a_i e^{j\mathbf{k}_i \cdot \underline{r}} \quad (1.2)$$

and that the sensors are equally spaced along the x axis with spacing $\lambda/2$, λ = wavelength, so that the position $\underline{r}_n$ of the nth sensor is

$$\underline{r}_n = \begin{pmatrix} \frac{n\lambda}{2} \\ 0 \\ 0 \end{pmatrix}, \quad n = 1, \dots, N \quad (1.3)$$

Finally, assume that all the wavevectors lie in the upper x-y half-plane, thus

$$\mathbf{k}_i = \frac{2\pi}{\lambda} \begin{pmatrix} \sin\phi_i \\ \cos\phi_i \\ 0 \end{pmatrix}, \quad -\frac{\pi}{2} \leq \phi_i \leq \frac{\pi}{2}.$$

where ϕ_i is the angle of arrival of the ith wave with respect to the y axis (see Fig. 1.1). This is necessary to make the array output unique for unique directions. Arbitrary wavevectors could be estimated by merely replacing the array along the y and z axes, however.

Plugging these assumptions in Eq. (1.2), we have that the output of sensor n is

$$x_n = \sum_{i=1}^D a_i e^{j\mathbf{k}_i \cdot \underline{r}_n} = \sum_{i=1}^D a_i e^{j\pi n \sin\phi_i} \quad (1.4)$$

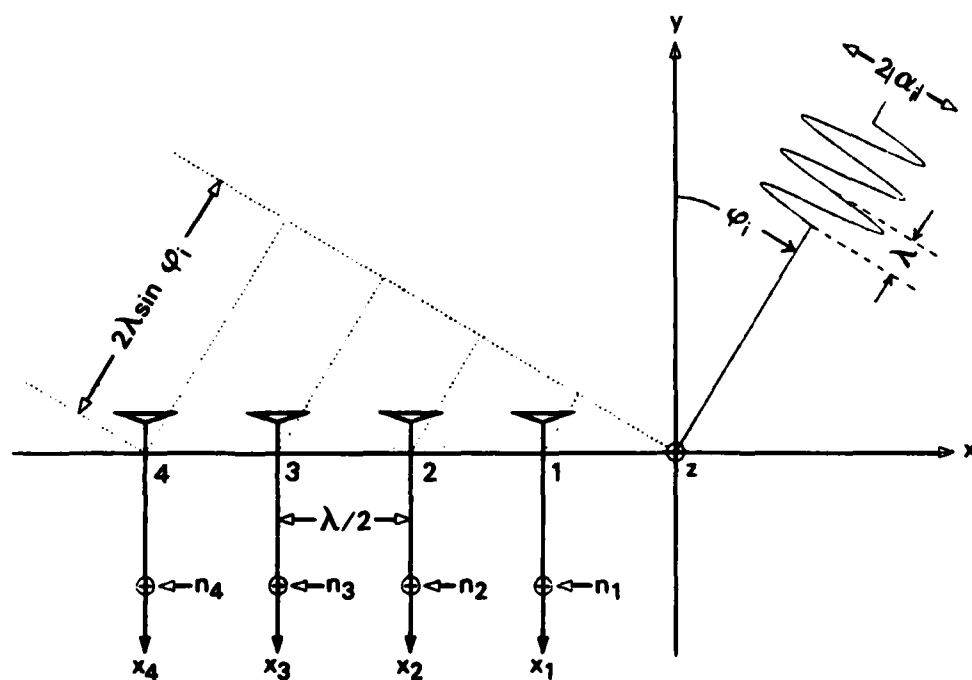


Fig. 1.1. Array geometry.

We can rewrite (1.4) concisely by using vector notation. Define

$$\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \quad \text{"received signal vector"}$$

$$\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_D \end{pmatrix} \quad \text{"signal amplitudes at origin vector"}$$

$$\underline{v}(\theta) = \begin{pmatrix} e^{j\pi\theta} \\ \vdots \\ e^{jn\pi\theta} \end{pmatrix} \quad \text{"direction vector"}$$

$$\theta_i = \sin \phi_i, \quad i=1, \dots, D \quad \text{"directions-of-arrival"}$$

$$\underline{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_D \end{pmatrix}$$

$$V = (\underline{v}(\theta_1) \quad \dots \quad \underline{v}(\theta_D)) \quad \text{"direction matrix."}$$

Then (1.4) becomes

$$\underline{x} = V \underline{a} \quad . \quad (1.5)$$

We see that $\underline{x}$ is a linear combination of "discrete-space" sinusoids, with spatial frequencies $f_1 = 1/2$ $\theta_1 = 1/2 \sin \phi_1$. Thus, this angle estimation problem is equivalent to spectral estimation for uniformly sampled time series, to within the non-linear transformation $f = 1/2 \sin \phi$. In this report, for the sake of mathematical simplicity, we will estimate the θ_1 's rather than the ϕ_1 's, thus the relation to time series frequency is a linear one, $f = \theta/2$. Our results are therefore directly applicable to time series spectral estimation (for one look, see next section).

1.2 Statistical Model

Any real sensor array is corrupted by noise. Thermal noise in the receivers, for example, is inevitable. The sources, additionally, are best modelled as noisy due to propagation environment effects, etc. For the usual physical and analytical reasons, these noises will be taken as Gaussian, specifically zero-mean circular complex Gaussian*.

Accordingly, our model for the actual array output is now

$$\underline{X} = V \underline{a} + \underline{n} \quad (1.6)$$

*z is circular complex Gaussian, CN (m, σ^2), if $\text{Re } z$ and $\text{Im } z$ are jointly Gaussian with

$$E (\text{Re } z) = \text{Re } m$$

$$E (\text{Im } z) = \text{Im } m$$

$$\text{cov} (\text{Re } z, \text{Im } z) = \begin{pmatrix} \frac{1}{2} \sigma^2 & 0 \\ 0 & \frac{1}{2} \sigma^2 \end{pmatrix}$$

where

$$\underline{n} = \begin{pmatrix} n_1 \\ \vdots \\ n_N \end{pmatrix}$$

complex Gaussian sensor noise with

$$E(\underline{n}) = \underline{0}$$

zero mean,

$$E(\underline{n} \underline{n}^H) = I$$

unit power, and independent from sensor to sensor, and

$$\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_D \end{pmatrix}$$

complex Gaussian signal amplitudes at origin with

$$E(\underline{a}) = \underline{0}$$

zero mean and covariance

$$E(\underline{a} \underline{a}^H) = P$$

, the "signal-in-space" covariance matrix.

We assume also that the sensor noise is independent of the signal amplitude, i.e.,

$$E(\underline{a} \underline{n}^H) = [0]$$

which implies that the received signal covariance R is

$$\begin{aligned}
R &\triangleq E(\underline{x} \underline{x}^H) = E(V \underline{a} \underline{a}^H V^H) + E(V \underline{a} \underline{n}^H) + E(\underline{n} \underline{a}^H V^H) + E(\underline{n} \underline{n}^H) \\
&= V E(\underline{a} \underline{a}^H) V^H + 0 + 0 + E(\underline{n} \underline{n}^H) \\
&= VPV^H + I
\end{aligned} \tag{1.7}$$

We also assume that by sampling the array output at a sufficiently large time interval (greater than the signal-in-space's (finite) correlation time), we have access to L statistically independent samples of the random vector $\underline{x}$,

$$X = \{\underline{x}_i\}_{i=1}^L, \quad L = \text{number of "looks" at array.}$$

Note that this is where the array processing problem differs from the time series problem, where L is one though N may be large.

The estimation problem is to determine the true directions of arrival θ^* , given the observations X and knowing apriori the powers P , the number of signals D , the number of looks L , and the array direction vectors $\underline{v}(\theta)$ for all θ . This is a classical parameters-in-the-covariance, zero mean complex Gaussian estimation problem, described, for example, in [15].

For such a problem, there is a simple sufficient statistic, the sample covariance matrix, which obviates the need to save all the data vectors $\underline{x}_i$, $i=1, \dots, L$. This can be seen from the probability densities (pdf's): since $\underline{x}$ is complex Gaussian with zero mean and covariance R , its pdf is

$$p(\underline{x}) = \frac{1}{|\pi R|} e^{-\underline{x}^H R^{-1} \underline{x}}$$

Since the $\underline{x}_i$ vectors are all independent and identically distributed, their joint pdf is

$$p(X) = \prod_{i=1}^L \frac{1}{|\pi R|} e^{-\underline{x}_i^H R^{-1} \underline{x}_i}$$

$$\begin{aligned}
&= \frac{1}{|\pi R|^L} e^{-\sum_{i=1}^L \underline{x}_i^H R^{-1} \underline{x}_i} \\
&= \frac{1}{|\pi R|^L} e^{-\sum_{i=1}^L \text{tr } \underline{x}_i^H R^{-1} \underline{x}_i} \\
&= \frac{1}{|\pi R|^L} e^{-\text{tr } R^{-1} \sum_{i=1}^L \underline{x}_i \underline{x}_i^H}
\end{aligned}$$

where the last step follows through the use of the trace identity

$$\text{tr}(AB) = \text{tr}(BA) \quad .$$

Thus, if we define

$$\hat{R} \triangleq \frac{1}{L} \sum_{i=1}^L \underline{x}_i \underline{x}_i^H$$

to be the sample covariance matrix, then

$$p(X) = \frac{1}{|\pi R|^L} e^{-L \text{tr } R^{-1} \hat{R}} \quad , \quad (1.8)$$

which depends on the data only through R . Thus, R is a sufficient statistic for the collection X . This fact is especially useful for Monte Carlo simulations since random samples of R can be generated directly with no need to generate the potentially large amount of data ($L > 1000$ for some applications) in X . All these facts are well known [15].

By way of notation, we will often write the expression $p(X)$ equivalently as $p(X|\underline{\theta})$ or $p(\hat{R}|\underline{\theta})$.

1.3 Cramer-Rao Bound

An effort to determine the best that one can estimate the parameters $\underline{\theta}^*$ leads naturally to the Cramer-Rao bound. The Cramer-Rao bound [15] is a well-known lower bound on the mean-square error of any unbiased estimator. Here, we assume that $\underline{\theta}^*$ is a non-random, though unknown, vector parameter, which leads to a more useful bound, although the random $\underline{\theta}^*$ case will be considered in following sections.

The bound is expressed in terms of the log likelihood function $\lambda(\underline{\theta})$, where

$$\lambda(\underline{\theta}) \triangleq \ln p(X|\underline{\theta})$$

and the Fisher information matrix F , where

$$F_{ij} \triangleq E_{p(X|\underline{\theta})} \left[\frac{\partial \lambda(\underline{\theta})}{\partial \theta_i} \cdot \frac{\partial \lambda(\underline{\theta})}{\partial \theta_j} \right], \quad 1 \leq i, j \leq D.$$

Then if $\hat{\underline{\theta}}$ is any unbiased estimate of $\underline{\theta}^*$, i.e.,

$$E(\hat{\underline{\theta}}) = \underline{\theta}^*,$$

the mean square error of the estimate, Λ_e , given by

$$\Lambda_e \triangleq E((\hat{\underline{\theta}} - \underline{\theta}^*) (\hat{\underline{\theta}} - \underline{\theta}^*)^T)$$

is bounded below through

$$\Lambda_e \geq F^{-1} ,$$

where the matrix inequality means that $\Lambda_e - F^{-1}$ is positive semidefinite. In particular, since a positive semidefinite matrix must have non-negative diagonal elements, this implies

$$\sigma_i^2 \triangleq E(\hat{\theta}_i - \theta_i^*)^2 \geq (F^{-1})_{ii} , \quad i=1 \dots D$$

For our estimation problem, λ is given from Eq. (1.8) by

$$\lambda = \ln p(X|\underline{\theta}) = -L (\ln |\pi R| + \text{tr } R^{-1} \hat{R}) .$$

After much algebra, detailed in [6], the bound emerges in the form

$$\Lambda_e \geq F^{-1} \quad \text{where}$$

$$F = 2 L \text{ Re } \{ (P-Q) \times (\hat{V}^H \hat{V} - \hat{W}^H \hat{Q} \hat{W})^T + \hat{Q} \hat{W} \times (\hat{Q} \hat{W})^T \} \quad (1.9)$$

$$A \times B \triangleq \text{matrix direct product, } (A \times B)_{ij} = A_{ij} B_{ij}$$

$$W \triangleq V^H V$$

$$V \triangleq \begin{pmatrix} \underline{v}(\theta_1) & \dots & \underline{v}(\theta_D) \end{pmatrix}$$

$$\dot{\underline{v}}(\theta) \triangleq \frac{\partial \underline{v}(\theta)}{\partial \theta}$$

$$\dot{V} \triangleq \begin{pmatrix} \dot{\underline{v}}(\theta_1) & \dots & \dot{\underline{v}}(\theta_D) \end{pmatrix}$$

$$\dot{W} \triangleq V^H \dot{V}$$

$$O \triangleq (W + P^{-1})^{-1}$$

Notice the appearance of $\dot{\underline{v}}(\theta)$, which is the key dependence of the observed data on small changes in the desired parameter θ , i.e., for small emitter separations. Also notice that the number of looks L enters only as a scale factor.

Due to the complexity of (1.9), the bounds' behavior was examined numerically to generate the curves in Fig. 1.2. Here, we assume the case for which performance evaluations were made in Chapter 3, namely two uncorrelated

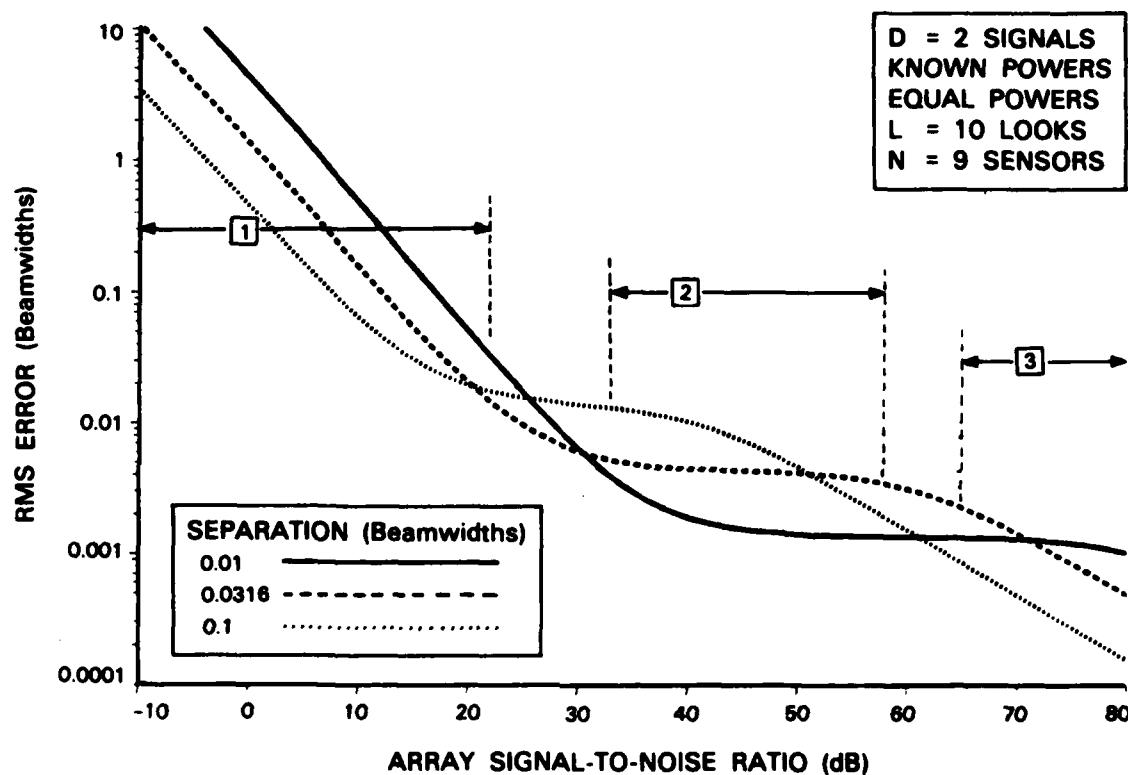


Fig. 1.2. Known powers Cramer-Rao bound.

signals of equal power. The rms error bound on θ_1 , in beamwidths (one beamwidth = $2/N$), is plotted versus the array signal-to-noise (SNR) Np , where

$$P = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}$$

for various values of the emitter separation $\Delta\theta$,

$$\Delta\theta = \theta_1^* - \theta_2^*$$

This is sufficient since, by symmetry, the bound on θ_1 equals the bound on θ_2 , and since by the isotropy of the sensors, the bound does not depend on the absolute position of the emitters θ_1 and θ_2 , but only on their separation, $\Delta\theta$.

We see that the bound has an unusual behavior. In region 2 of SNR, as shown in Fig. 1.2, the rms error is larger for larger emitter separations, which is saying that widely spaced emitters are more difficult to estimate than closely spaced. This is counter to intuition and also in contrast to the bound in the low SNR and high SNR regions 1 and 3, respectively.

Also in region 2, we see a plateau in the bound. This is saying that once a certain rms error is attained (namely, $\sigma_1 \approx \Delta\theta/3$ at $L = 10$ looks), a very large step in SNR is required to decrease the error further. Both of these effects would be of great concern in the AOA estimation problem if they were "real." But the curves are only lower bounds on the rms error, hence if they are not tight bounds, the actual achievable rms error could behave differently. For example, since it is known that for high SNR the Cramer-Rao bound is always achievable [15], the actual error might follow straight lines coinciding with the bounds' high SNR asymptotes.

Additionally, the curves are bounds only on the estimation error and presume that it is known apriori that there are two signals present, or that two signals have been detected. Where the rms error predicted by the bound is

much larger than the separation (region 1 in the figure), one would not expect both signals to be detectable.

These questions can be resolved by computing the actual achievable rms error and the actual achievable signal detectability (probability of detection), namely, the performances of the optimal estimator and detector.

1.4 Optimal Estimation and Detection

1.4.1 Estimation

The obvious way to verify the tightness of the Cramer-Rao bounds of the previous section are through the evaluation of the rms errors of the statistically optimal estimators ML (maximum likelihood), MAP (maximum a posteriori), and MMSE (minimum mean square error).

The ML estimate $\hat{\theta}_{ML}$ is defined to be the θ which maximizes the likelihood function λ :

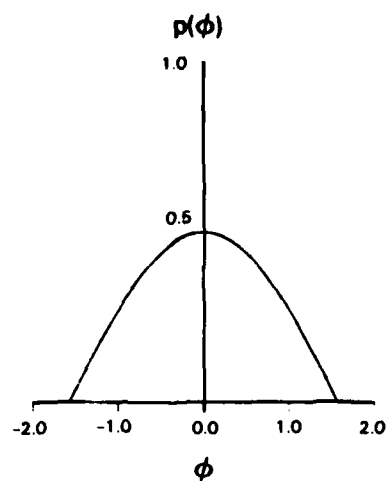
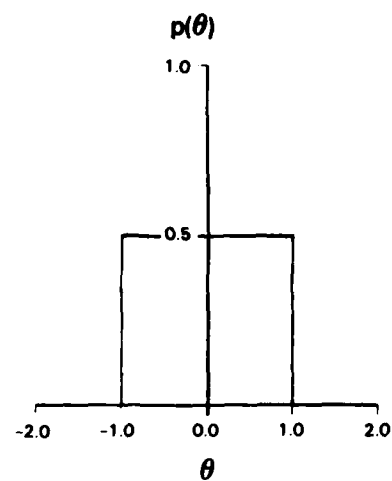
$$\hat{\theta}_{ML} \triangleq \underset{\theta}{\operatorname{argmax}} \lambda(\theta) .$$

It has the important property that if any unbiased estimate achieves the Cramer-Rao bound, then the ML estimate does if it is also unbiased.

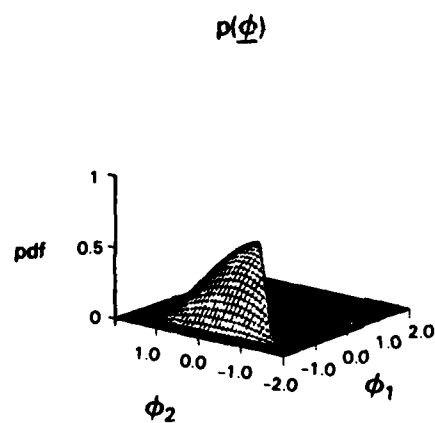
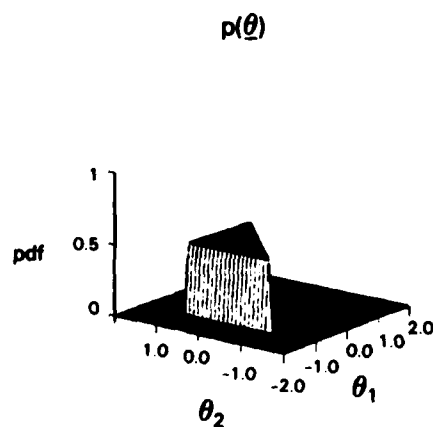
To define the MAP and MMSE estimators, we must assume that θ^* is a random unknown parameter with some probability density $p(\theta)$. Here, we choose $p(\theta)$ to be a uniform density over all possible θ values. Since $\theta_1 = \sin \phi_1$, we have $-1 \leq \theta_1 \leq 1$ $i=1,2$, and since the labelling of "1" and "2" is arbitrary because the signals are symmetric (we assume equal power here), we also have, without loss of generality, that $\theta_2 \geq \theta_1$. Thus, $p(\theta)$ is a uniform density on the triangle with $p(\theta) = 1/2$ (on the left in Fig. 1.3).

Notice that this does imply a nonuniform density for $\delta = \sin^{-1} \theta$, which is seen by the change of variables formula from probability theory to be

$$p(\phi) = \frac{1}{2} \cos \phi_1 \cos \phi_2 \quad , \quad -\frac{\pi}{2} \leq \phi_1 \leq \phi_2 \leq \frac{\pi}{2}$$



(a)



(b)

Fig. 1.3. (a) One-dimensional prior densities. (b) Two-dimensional prior densities.

(on right in Fig. 1.3). This does not concern us here but would affect estimators of $\underline{\phi}$ directly.

With $p(\underline{\theta})$ in hand, we now have the aposteriori density of $\underline{\theta}$, $p(\underline{\theta}|\hat{R})$, which is given by Bayes' rule:

$$p(\underline{\theta}|\hat{R}) = \frac{p(\hat{R}|\underline{\theta}) p(\underline{\theta})}{p(\hat{R})} .$$

Now the MAP estimator $\hat{\underline{\theta}}_{\text{MAP}}$ is the maximum of this density

$$\hat{\underline{\theta}}_{\text{MAP}} \triangleq \underset{\underline{\theta}}{\operatorname{argmax}} p(\underline{\theta}|\hat{R}) .$$

Because of our choice of a uniform prior on $\underline{\theta}$, the MAP estimate is identical to the ML estimate.

$$\hat{\underline{\theta}}_{\text{MAP}} = \hat{\underline{\theta}}_{\text{ML}} ,$$

so we will, for convenience, consider only $\hat{\underline{\theta}}_{\text{MAP}}$.

The MMSE estimate $\hat{\underline{\theta}}_{\text{MMSE}}$ is the mean of the aposteriori density,

$$\hat{\underline{\theta}}_{\text{MMSE}} \triangleq E(\underline{\theta}|\hat{R}) = \int \int d\underline{\theta} \underline{\theta} p(\underline{\theta}|\hat{R}) ,$$

and has the property that it minimizes the total squared error, averaged over the aposteriori density $p(\underline{\theta})$, i.e.,

$$\min E(\operatorname{tr} \Lambda_e) = \min \int \int d\underline{\theta} \operatorname{tr} \Lambda_e p(\underline{\theta}) .$$

This is not the mean square error Λ_e defined previously, since Λ_e did not include the average over the prior $p(\underline{\theta})$. The MMSE estimator is still well-defined, however, and could be better than MAP.

The reader may be disturbed by our changing from nonrandom to random $\underline{\theta}^*$. In all of the work which follows, $\underline{\theta}^*$ is actually assumed fixed and nonrandom (though unknown to the estimator) because to average the statistics over many samples of $\underline{\theta}^*$ from the density $p(\underline{\theta})$ would wash out effects of interest, for example, the dependence of the error on $\underline{\theta}^*$. However, assuming random $\underline{\theta}^*$ allows the definition and computation of the conceptually appealing density of $p(\underline{\theta}|\hat{R})$, which describes exactly what is known about $\underline{\theta}$ given the observation $\hat{R}$. By choosing a uniform prior, we make the random and nonrandom $\underline{\theta}^*$ viewpoints coincide (e.g., $\hat{\theta}_{ML} = \hat{\theta}_{MAP}$).

1.4.2 Detection

In order to determine the regions where the Cramer-Rao bound is meaningful, it is of interest to examine the regions where the two signals can be detected. We assume that the presence of some signal is easy to detect so that the problem is to distinguish one signal from two closely spaced ones.

Let $p(D)$ be the apriori probability of the number of signals; then the minimum probability of error estimate of the number of signals, D , is

$$\hat{D} = \underset{D}{\operatorname{argmax}} p(D|\hat{R}) ,$$

where $p(D|\hat{R})$ is the aposteriori density of D given $\hat{R}$, derived from Bayes' rule as

$$p(D|\hat{R}) = \frac{p(\hat{R}|D) p(D)}{p(\hat{R})}$$

and the probability of error (P_E) is

$$P_E \triangleq \text{prob} (\hat{D} \neq D^*) ,$$

where D^* is the true number of signals.

This detector is closely related to the optimal estimators, as will be seen in Chapter 2.

1.5 Previous Work - Present Contribution

Research relating to optimal angles-of-arrival estimation has been carried out since the early 1960's. None of the published results solve the problem for our particular combination of assumptions, however. Most of the work assumed the easier deterministic model, which implied also that just one look was considered. Sklar and Schweppe [12] and Pollon [10] computed the Cramer-Rao bounds for this problem, for various array configurations and two signals. Young [17] assumed a prior density on θ as we have and derived the form of an analog signal processor for the MMSE estimator. Ksienski and McGhee [8] considered the ML estimator for two signals and computed it approximately by maximization over a finite 10 by 10 grid of directions. Pollon and Lank [11] ran an analog simulation of the two signal ML estimator for a certain circular array.

Gallop and Nolte [4] and Hodgkiss and Nolte [7] examined the one-signal case only, considered mainly the detection problem. Rife and Boorstyn [13, 14] simulated the one- and two-signal estimators and compared them to the Cramer-Rao bound, but approximated ML by a simple discrete Fourier transform.

The multiple look, Gaussian model problem is the more difficult one and has been little examined. El-Behery and Macphie [3] investigated the one-signal ML estimator, which reduces to a discrete Fourier transform. They performed an extensive Monte Carlo Simulation as we will here and found attainment of the Cramer-Rao over a surprising range of SNR. Lainiotis [9]

considered the Gaussian model for Bayes' estimation in a control theory context. The estimator was computed by a maximization over only 10 grid points, however, one of which was the true parameter value. Such a choice probably explains the performance observed. A control theory or state variable approach with finite grid approximations is also considered by Bucy and Senne [2], in the context of the phase demodulation problem, and with a thorough Monte Carlo performance evaluation.

The contribution of this report consists principally of the accurate computation of performance of the MAP and MMSE estimators for Gaussian two-signal, known powers problem. The calculation is accurate enough to distinguish the unusual features of the Cramer-Rao bound mentioned earlier.

Additionally, the performance of the optimal detector is computed and compared to an ad-hoc approximation based on the Cramer-Rao bound, and to a rigorous known-directions bound.

In Chapter 2, a more detailed derivation of the form of the MAP and MMSE estimators and the optimal detector for our problem is given. The chapter closed with some example plots of $p(\underline{\theta}|\hat{R})$ for two emitters. These plots provide insight into the nature of the estimation errors.

In Chapter 3, the results of the performance evaluation are presented. First, error sources in the computer calculation are quantified and seen to be acceptable. Then the results of the calculation are plotted and compared to the Cramer-Rao bound and the P_D bounds. The bounds' behavior is seen to indeed be real.

In Chapter 4, the consequences of the tightness of the bound are discussed, along with suggestions for further research.

The appendices contain derivations of the confidence intervals for the Chapter 3 results and of the known-directions P_D bound.

2.0 THE MAP AND MMSE ESTIMATORS AND THE OPTIMAL DETECTOR

In this chapter, the equations for the MAP and MMSE estimators and the optimal detector are derived in detail. Although the derivations are classical [15], some simplifications specific to the two-emitter problem are applied. The equations center on the expression for the aposteriori density, $p(\underline{\theta}|\hat{R})$, and the chapter is concluded with example plots of $p(\underline{\theta}|\hat{R})$ illustrating some properties which have been derived or observed.

Since $p(\underline{\theta}|\hat{R})$ embodies all the information about $\underline{\theta}$ in $\hat{R}$, it is an important function itself in the angle estimation problem.

2.1 MAP and MMSE Estimators

As mentioned in Chapter 1, the MAP and MMSE estimates of the direction parameters are given in terms of the aposteriori density $p(\underline{\theta}|\hat{R})$ by

$$\hat{\underline{\theta}}_{\text{MAP}} = \underset{\underline{\theta}}{\operatorname{argmax}} p(\underline{\theta}|\hat{R})$$

$$\hat{\underline{\theta}}_{\text{MMSE}} = \int \int d\underline{\theta} \underline{\theta} p(\underline{\theta}|\hat{R}) \quad .$$

The form of $p(\underline{\theta}|\hat{R})$ now follows our Gaussian assumptions. We have first (from Eq. (1.8)) that $p(\hat{R}|\underline{\theta})$ is a multivariate Gaussian density with zero mean and covariance $R(\underline{\theta})$:

$$p(\hat{R}|\underline{\theta}) = |\pi R(\underline{\theta})|^{-L} e^{-L \operatorname{tr} R(\underline{\theta})^{-1} \hat{R}}^*$$

To define the aposteriori density, we must use the prior density of $\underline{\theta}$, $p(\underline{\theta})$, defined in Chapter 1. Then, by Bayes' rule,

*The notation here is loose since this is actually $p(X|\underline{\theta})$ is expressed in terms of $\hat{R}$. $\hat{R}$ itself is Complex Wishart [5]. This detail does not effect $p(\underline{\theta}|\hat{R})$.

$$p(\underline{\theta}|\hat{R}) = \frac{p(\hat{R}|\underline{\theta}) p(\underline{\theta})}{p(\hat{R})}$$

Since $p(\hat{R})$ is independent of $\underline{\theta}$, it is just a normalizing constant for the density. Likewise, we chose a uniform prior density on $\underline{\theta}$ (in Chapter 1), so that $p(\underline{\theta})$ is also a constant. Therefore,

$$p(\underline{\theta}|\hat{R}) = K p(\hat{R}|\underline{\theta})$$

where K depends on $\hat{R}$ but not on $\underline{\theta}$.

We now absorb more constants by using a well-known matrix inversion identity,

$$R(\underline{\theta})^{-1} \equiv (I + VPV^H)^{-1} = I - VQV^H$$

$$\text{where } Q \triangleq (P^{-1} + W)^{-1}$$

$$W \triangleq V^H V$$

(same definitions as used for Cramer-Rao bound in Section 1.3), and get

$$\begin{aligned} p(\underline{\theta}|\hat{R}) &= K |\pi R(\underline{\theta})|^{-L} e^{-L \operatorname{tr} (I - VQV^H) \hat{R}} \\ &= K \pi^{-NL} |R(\underline{\theta})|^{-L} e^{-L \operatorname{tr} (\hat{R} - VQV^H \hat{R})} \\ &= K' |R(\underline{\theta})|^{-L} e^{L \operatorname{tr} QV^H \hat{R} V} \end{aligned}$$

where $K' = K \pi^{-NL} e^{-L \operatorname{tr} \hat{R}}$ is once again independent of $\underline{\theta}$.

Finally, the determinant can be reduced since we assume only two emitters by

$$\begin{aligned} |R(\underline{\theta})| &= |VPV^H + I| = (1 + \lambda_1)(1 + \lambda_2) \\ &= 1 + \lambda_1 + \lambda_2 + \lambda_1 \lambda_2 \end{aligned}$$

where λ_1 and λ_2 are the two non-zero eigenvalues of VPV^H . By a theorem from linear algebra, λ_1 and λ_2 are also the only two eigenvalues of $PV^H V$. Thus,

$$\lambda_1 + \lambda_2 = \text{tr}(PV^H V) = \text{tr}(PW)$$

and

$$\lambda_1 \lambda_2 = |PV^H V| = |PW|$$

implying

$$|R(\underline{\theta})| = 1 + \text{tr}(PW) + |PW|$$

which is convenient since PW is only a 2 by 2 matrix. Therefore, we now have

$$p(\underline{\theta}|R) = \frac{K' e^{L \text{tr} QV^H R V}}{(1 + \text{tr}(PW) + |PW|)^L} \quad (2.1)$$

This expression, unfortunately, cannot be solved analytically for its maximum or mean, so numerical procedures must be used to compute $\hat{\theta}_{\text{MAP}}$ and $\hat{\theta}_{\text{MMSE}}$. For our two-emitter known power case, $\underline{\theta}$ is only two-dimensional, hence, such numerical procedures are reasonable. Of course, this was the principal reason for limiting the investigation to this case. The procedures used are described in Chapter 3.

2.2 Optimal Detector

The optimal detector, as also covered in the first chapter, is given by

$$\hat{D} = \max_D p(D|\hat{R})$$

and is the minimum probability-of-error (P_E) estimate of the number of signals present. Since D must be determined without knowing the signal directions, this is the composite hypothesis testing problem as defined in [15], with $\underline{\theta}$ as the nuisance parameter. The test is seen to be simply related to $p(\underline{\theta}|\hat{R})$ of the previous section.

We assume as in Chapter 1 that the presence of some signal is easy to detect, therefore we need only distinguish $D=1$ and $D=2$. Also, we assume that these cases are equally likely apriori, thus the apriori density of D is $p(D=1) = p(D=2) = 1/2$. Finally, we must specify the known power if one signal is present. The physically meaningful choice for this power is $p_1 + p_2$ where p_1 and p_2 are the powers when two signals are present. This choice also should be the most difficult detection problem since otherwise one and two signals could be distinguished on the basis of total signal power alone.

Now, by Bayes' rule,

$$p(D|\hat{R}) = \frac{p(\hat{R}|D) p(D)}{p(\hat{R})} \quad D = 1, 2$$

$$= K p(\hat{R}|D)$$

since $p(D)$ and $p(\hat{R})$ are each independent of D . The constant K here is unrelated to the K of the previous section.

Now, $p(\hat{R}|D)$ is obtained from $p(\hat{R}|D, \underline{\theta})$ by

$$p(\hat{R}|D) = \int d\underline{\theta} p(\hat{R}|D, \underline{\theta}) p(\underline{\theta}) d\underline{\theta} \quad , \quad (2.2)$$

where the integral is one dimensional for $D=1$ and two dimensional for $D=2$. For $D=2$, $p(\hat{R}|D, \underline{\theta})$ is what we actually meant by $p(\hat{R}|\underline{\theta})$ in the previous section, where two-dimensionality was implicit. Furthermore, $p(\hat{R}|D=1, \underline{\theta})$ is obtained by evaluating the former along the diagonal $\theta_1=\theta_2$,

$$p(\hat{R}|D=1, \underline{\theta}) = p(\hat{R}|D=2, \underline{\theta})|_{\underline{\theta} = \begin{pmatrix} \theta \\ \theta \end{pmatrix}} \quad ,$$

which simply says that one signal in direction θ with power $p_1 + p_2$ is indistinguishable from two signals at the same direction θ with powers p_1 and p_2 . Also, we take the apriori density of θ for $D=1$, $p(\theta)$, to be uniform on $-1 \leq \theta \leq 1$, hence $p(\theta) = 1/2$. Therefore, plugging the above and Eq. (2.1) into (2.2), we have

$$p(D|\hat{R}) = \begin{cases} K' \int_{-1}^1 d\theta p(\theta) |R(\theta)|^{-L} e^{L \text{tr} QV(\theta)^H \hat{R} V(\theta)}, & D=1 \\ K' \iint d\theta p(\theta) |R(\theta)|^{-L} e^{L \text{tr} QV(\theta)^H \hat{R} V(\theta)}, & D=2 \end{cases} \quad (2.3)$$

where K' is chosen such that $p(D=1|\hat{R}) + p(D=2|\hat{R}) = 1$ and where the determinant has been left unexpanded for conciseness. Thus, $\hat{D}$ is 1 or 2 according to whether the 1 dimensional or 2 dimensional integral of $p(\theta|\hat{R})$ is larger.

The integrals are, as usual, intractable and must be computed numerically. Integration procedures used will be described in Section 3.2.

2.3 Examples of Properties of $p(\theta|\hat{R})$

The aposteriori density reveals precisely how much information about θ is contained in R and is the function in terms of which the optimal estimators, MAP, MMSE, and detector can be expressed. Therefore, before examining the statistics of the estimators in the next chapter, it is interesting to see just what it looks like.

First, we see (Fig. 2.1) that for equal powers p_1 and p_2 , the pdf is symmetric about the diagonal $\theta_1 = \theta_2$. This clearly must be true, since the numbering of the angles as "1" or "2" is arbitrary. The positions of $\hat{\theta}_{\text{MAP}}$ and $\hat{\theta}_{\text{MMSE}}$ are also shown in the figure. Note that $\hat{\theta}_{\text{MMSE}}$ in particular must be computed by integrating only over the triangle, since the mean of the density over the square (due to the symmetry) will always be on the diagonal.

Another important feature we see is that extraneous maxima tend to be in the form of ridges along the lines $\theta_1 \approx \theta_1^*$ and $\theta_2^* \approx \theta_2$, i.e., one

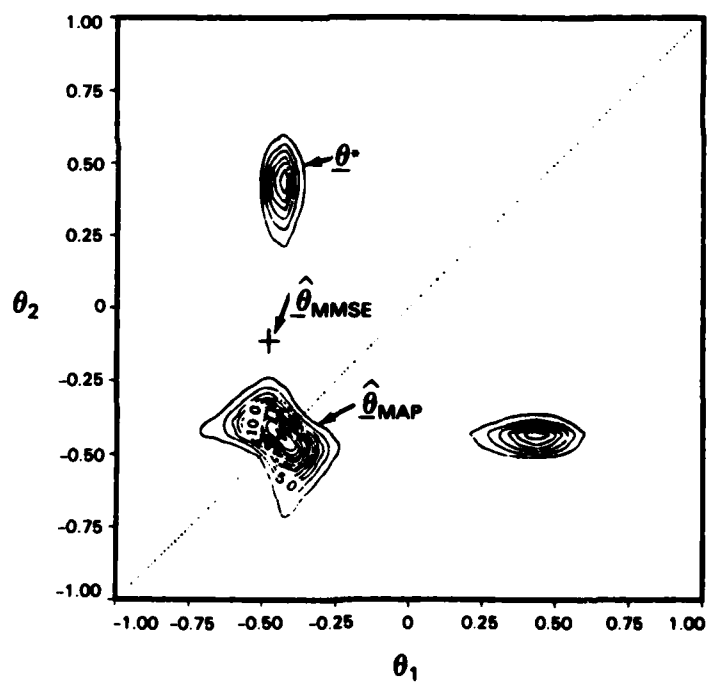
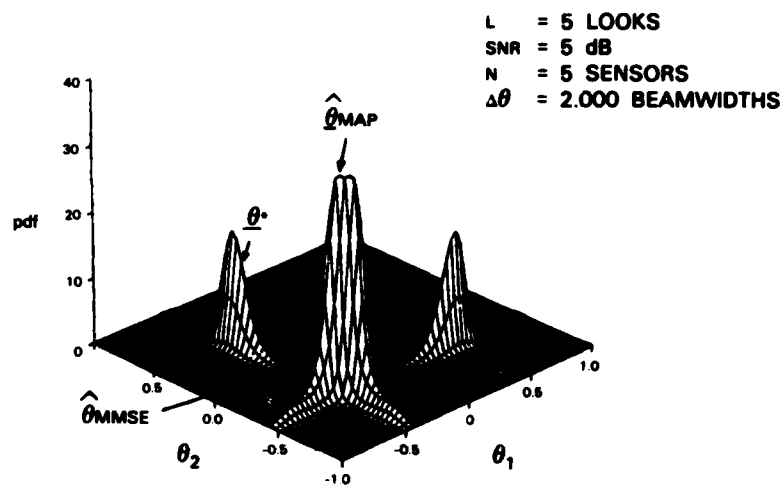


Fig. 2.1. Example of $p(\underline{\theta}|\hat{R})$ for equal powers.

angle correct and the other arbitrary. The ridges tend to disappear from the pdf as the number of looks increases (Fig. 2.2a) but, as can be seen from the log pdf, which is the same as the log-likelihood function (Fig. 2.2b), are always present to some extent. This property could confuse a gradient search type algorithm for finding the $\hat{\theta}_{\text{MAP}}$.

If $p_1 \neq p_2$, there is an ordering on the signals, say "1" is the larger one. Hence, the pdf is not symmetric and should have a peak only on the correct triangle (Fig. 2.3). Notice that the peak is no longer approximately circular. In fact, it is about 10 times broader in the θ_2 direction as the θ_1 direction, which reflects the greater uncertainty in the smaller signal's location θ_2 . Of course, if p_1 is almost equal to p_2 , the estimator will have difficulty distinguishing which is actually larger, hence there will tend to be peaks on each triangle until many looks are taken (Fig. 2.4).

Finally, we note that once the aposteriori density is essentially unimodel on the triangle, it can be well approximated by a Gaussian mound. In fact, it is close to a Gaussian with mean given by the true signal locations θ^* and variances given by the Cramer-Rao bound (Fig. 2.5). Notice that the grid has been shrunk in this figure to a small region about the peak.

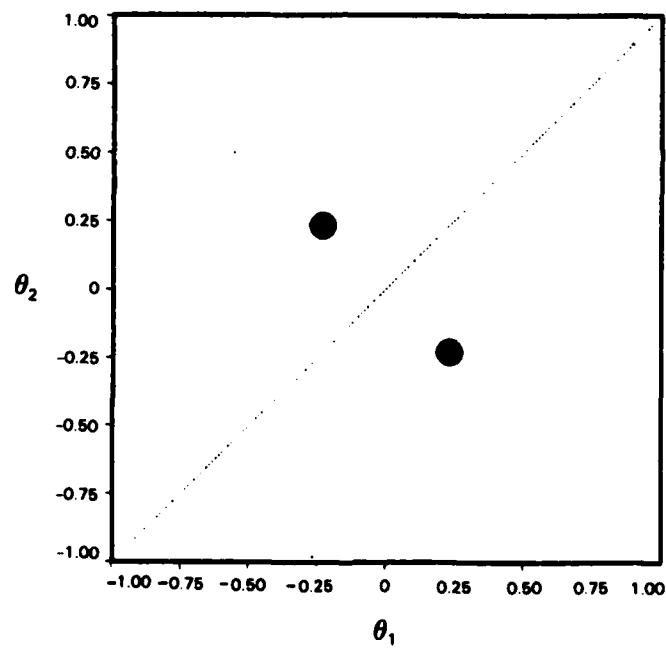
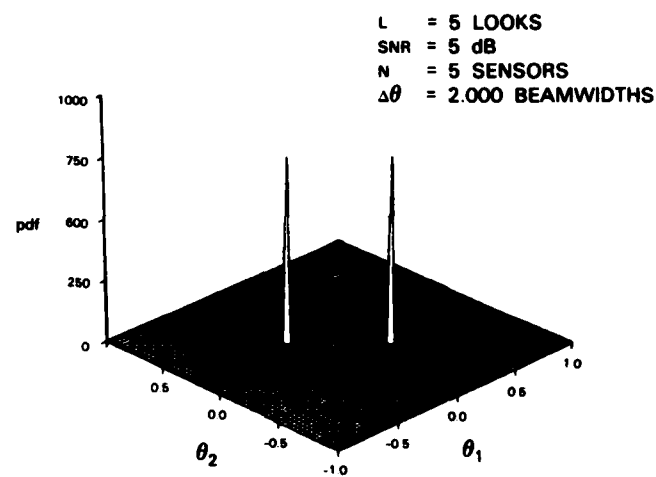


Fig. 2.2a. $p(\underline{\theta}|\hat{R})$ for additional looks.

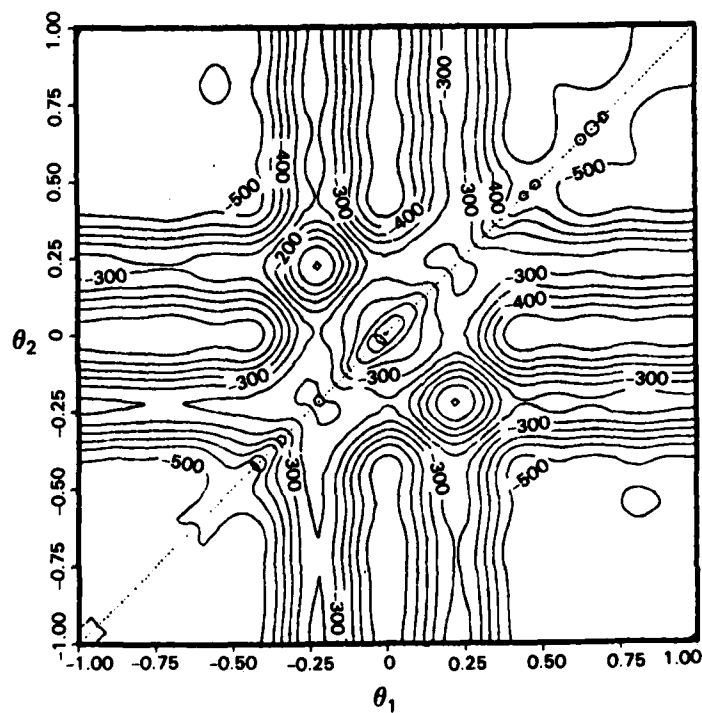
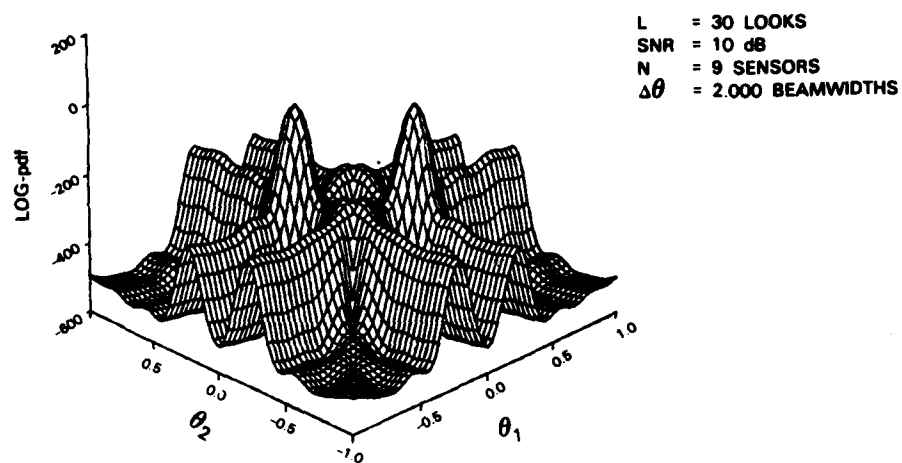


Fig. 2.2b. Log-pdf illustrating ridges.

142953-R

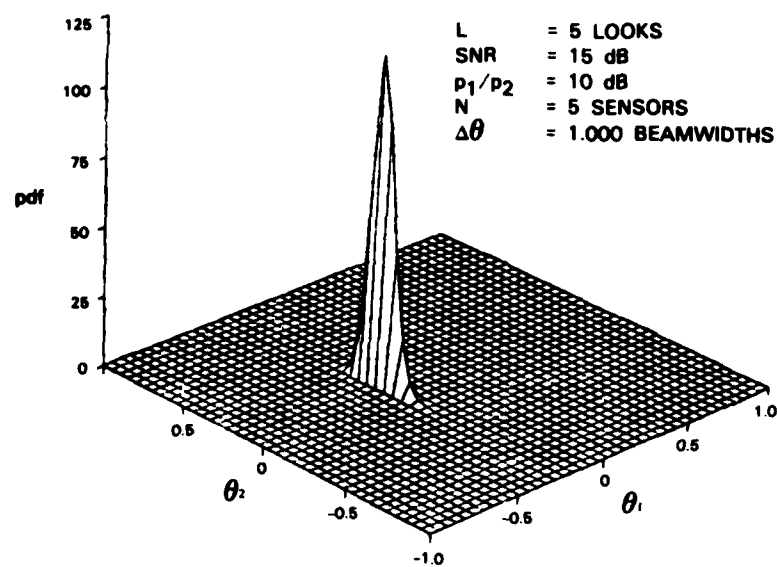


Fig. 2.3. Example of $p(\theta|\hat{R})$ for unequal powers.

141871-R

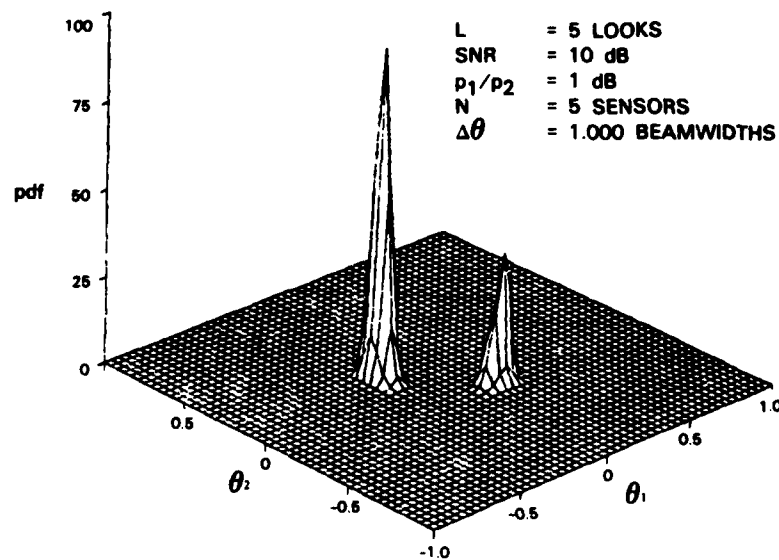


Fig. 2.4. Example of $p(\theta|\hat{R})$ for almost equal powers.

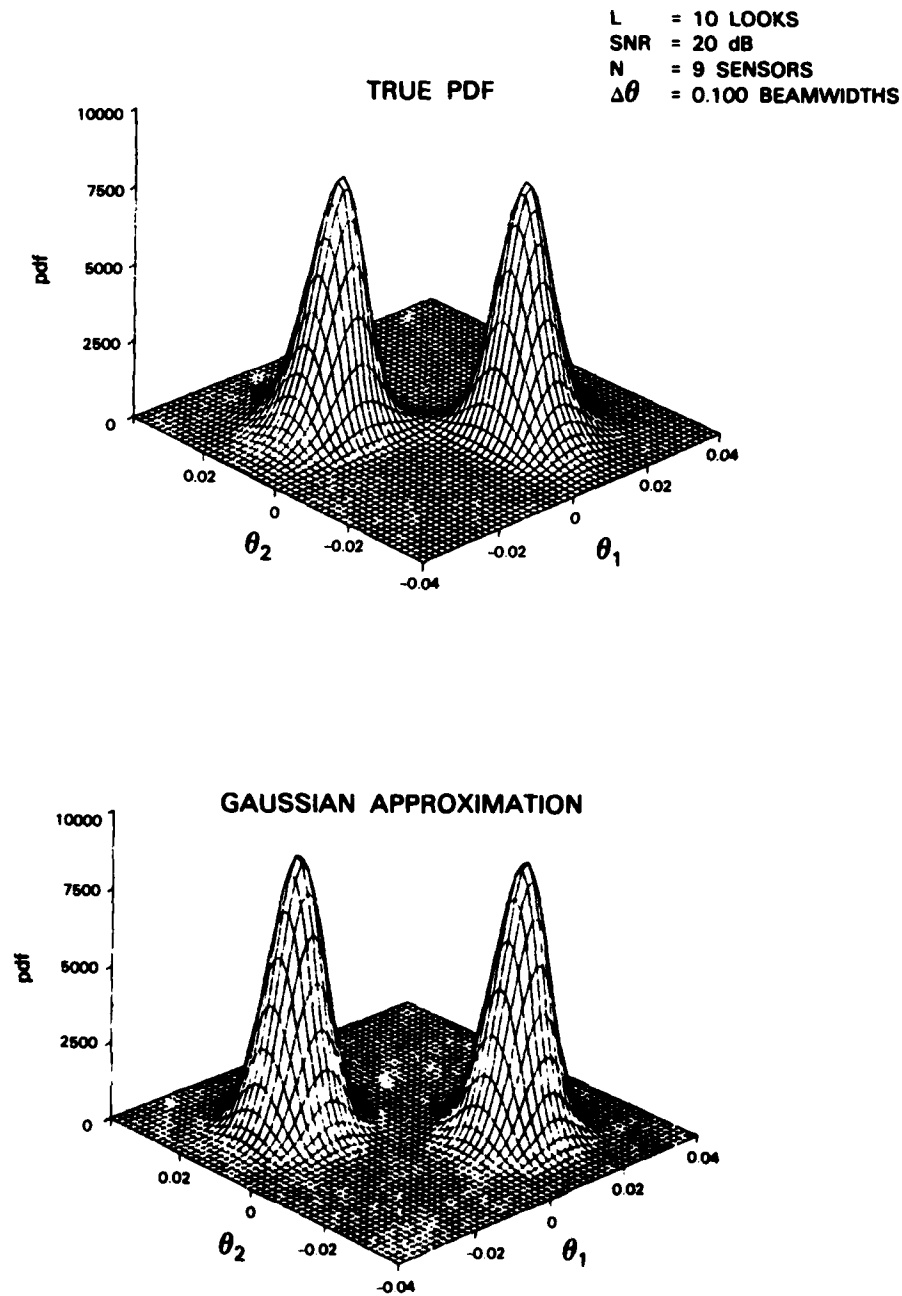


Fig. 2.5. Comparison of $p(\underline{\theta}|\hat{R})$ to Gaussian pdf with Cramer-Rao bound as its covariance.

3.0 MONTE CARLO PERFORMANCE EVALUATION

In this chapter, we present the results of the Monte Carlo evaluation of the performance of the MAP and MMSE estimators (their rms errors) and of the optimal detector (its probability of detection, P_D). The computational procedure and its sources of error are first described. The errors arise from two sources, the piecewise constant approximation to $p(\underline{\theta}|\hat{K})$ and the finite number of Monte Carlo trials performed. The results of the evaluation are presented next and compared to the Cramer-Rao bound. The main result is that the rms error of the estimators achieves the Cramer-Rao bound wherever the P_D is usefully high and deviates from the bound only for low P_D , where the estimates become biased. This result is surprising due to the bound's unusual behavior.

3.1 Computational Procedure

Simple approximations were used in the computations of performance to make the program fast and to make the errors easy to understand and quantify. The following sections describe first the rms error calculation and then the P_D and bias calculations.

3.1.1 RMS Error Calculation

The conditional error covariance of an estimator $\hat{\underline{\theta}} = \hat{\underline{\theta}}(\hat{K})$ is, by definition,

$$\Lambda_e \triangleq E((\hat{\underline{\theta}} - \underline{\theta}^*)(\hat{\underline{\theta}} - \underline{\theta}^*)^T | \underline{\theta}^*) = \int d\hat{K} p(\hat{K} | \underline{\theta}^*) (\hat{\underline{\theta}} - \underline{\theta}^*)(\hat{\underline{\theta}} - \underline{\theta}^*)^T \quad (3.1)$$

where $\underline{\theta}^*$ is the true theta value. The estimator's rms error on the i^{th} direction is the square root of the i^{th} diagonal element of Λ_e ,

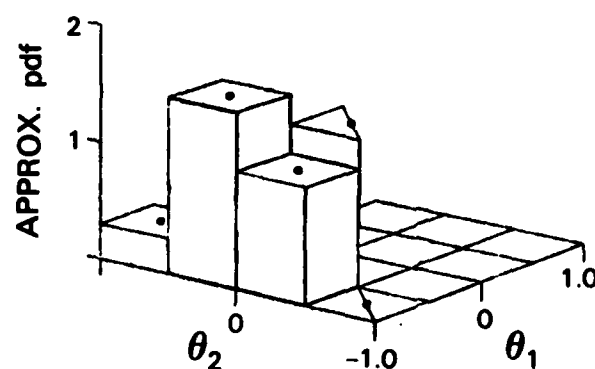
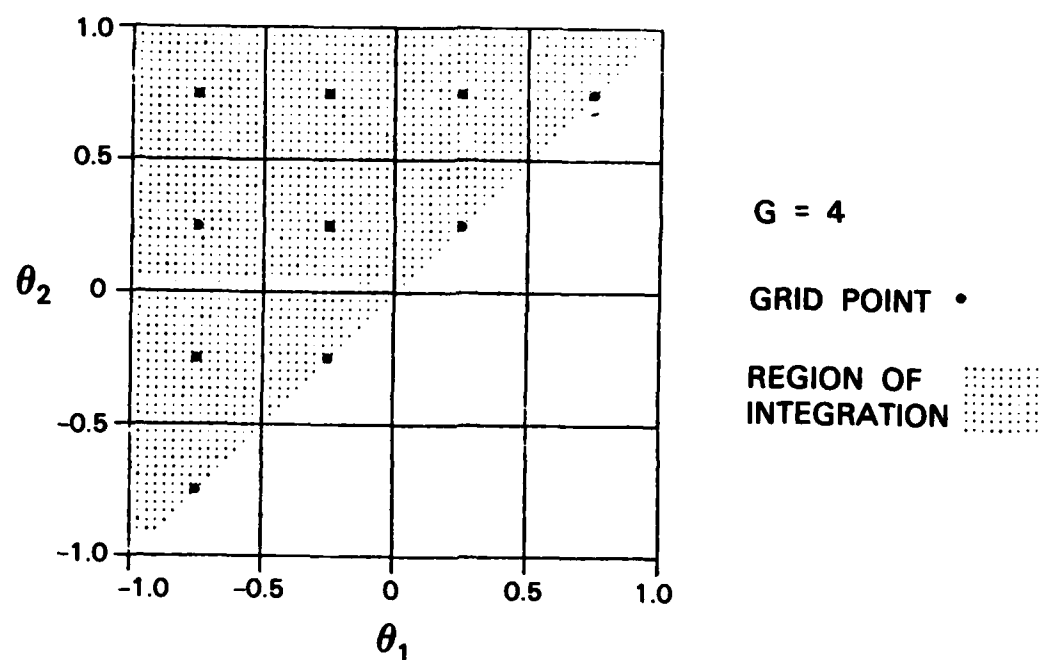
$$\sigma_i = \sqrt{(\Lambda_e)_{ii}} \quad .$$

In order to numerically compute Λ_e from (3.1), we must make basically two different approximations. One is the approximate computation of $\hat{\theta}$ given an accurate $\hat{R}$ and the other is the approximate integration with respect to $\hat{R}$ over its density. The approximations used and their errors will now be described.

Given an $\hat{R}$, the two estimates $\hat{\theta}_{MAP}$ and $\hat{\theta}_{MMSE}$ are computed as follows. First, $p(\underline{\theta}|\hat{R})/K'$ (Eq. 2.1) is evaluated over a fixed finite grid of $\underline{\theta}$ values, say G by G , uniformly spaced over the square (see Fig. 3.1). Since for high SNR or large numbers of looks $p(\underline{\theta}|\hat{R})$ is zero everywhere except in a small region around $\underline{\theta}^*$, the square is taken to be the smallest one which safely encloses the pdf's support. This is actually $[-1, 1] \times [-1, 1]$ only for the low SNR cases examined. Then this is numerically integrated over the upper triangle to find K' . $\hat{\theta}_{MMSE}$ is now the numerical integral of $\underline{\theta}p(\underline{\theta}|\hat{R})$ over the triangle and $\hat{\theta}_{MAP}$ is the location of the grid point with the largest value of $p(\underline{\theta}|\hat{R})$. The integrals are taken such that they are exact if $p(\underline{\theta}|\hat{R})$ is assumed to be piecewise constant, i.e., constant on each grid cell, as shown in Fig. 3.1. This implies that the computation is a simple summation plus a few correction terms and that the only error source is the degree to which $p(\underline{\theta}|\hat{R})$ is not piecewise constant. The error will go to monotonically zero as G is increased since this approximation converges uniformly on the square to $p(\underline{\theta}|\hat{R})$ [2].

In fact, a plot was made of the integrals versus G to determine how fine a grid was required (Fig. 3.2). For typically smooth true pdfs, $G \geq 50$ was adequate for an accuracy of about 2% of the true value. Errors in $\hat{\theta}_{MMSE}$ and $\hat{\theta}_{MAP}$ were not surprisingly of the order of the grid cell size. Making these errors about 2% places them below the errors due to the second source, the integration with respect to $\hat{R}$.

This integration, involving as it does the complicated function $p(\hat{R}|\underline{\theta}^*)$, is hopeless to perform analytically. It is also of such dimensionality (e.g., 81-dimensional for nine antennas) that the only tractable numerical integration method is the Monte Carlo method [16].



APPROXIMATE pdf = TRUE pdf AT GRID POINTS

Fig. 3.1. Piecewise-constant approximation.

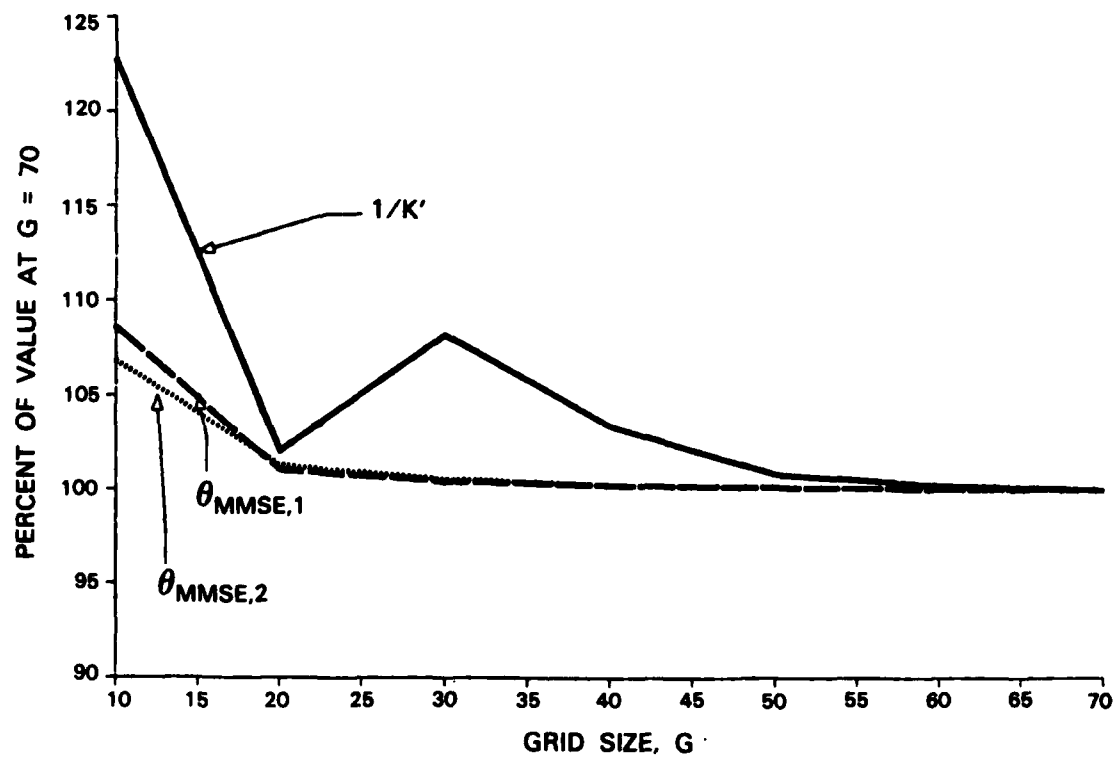


Fig. 3.2. Integrals versus grid size, G.

141874-R

Monte Carlo integration in this case amounts to the simulation of the estimators, computation of the errors and squared errors, and accumulation of their sample means, i.e.,

$$\hat{\Lambda}_e = \frac{1}{M} \sum_{i=1}^M (\hat{\theta}(\hat{R}_i) - \underline{\theta}^*) (\hat{\theta}(\hat{R}_i) - \underline{\theta}^*)^T$$

where M is the number of statistically independent Monte Carlo trials, $\hat{R}_i$ is the sample covariance chosen from the population $p(\hat{R}|\underline{\theta}^*)$ on the i th trial, and $\hat{\Lambda}_e$ is the estimated error covariance.

Now $\hat{\Lambda}_e$ is a random variable with some unknown probability density. Its expected value is the true Λ_e and its variance is determined by its pdf. If we assume that its pdf is Gaussian, which will be approximately true for large M , then its variance can be computed and used to compute the number of trials required for a specified accuracy. Choosing a 99% confidence interval width of ± 2 dB as adequate for distinguishing the interesting features of the bounds, we found by making trial runs that about 1000 Monte Carlo trials were required (see Appendix A).

Since the rms errors for each direction θ_1 and θ_2 are equal by symmetry, the final estimate of the rms error for either direction was taken to be

$$\hat{\sigma} = \frac{\hat{\sigma}_1 + \hat{\sigma}_2}{2}$$

where

$$\hat{\sigma}_1 = \sqrt{(\hat{\Lambda}_e)_{11}}$$

3.1.2 Probability of Detection Computation

The probability of detection (P_D) for the optimal detector is given by

$$P_D = \int_0^\infty d\ell \, p(\ell|\theta^*) \quad (3.2)$$

where

$$\ell = \int_{-1}^1 \int_{\theta_1}^1 d\theta_1 d\theta_2 p(\underline{\theta}|\hat{R}) - \int_{-1}^1 d\theta p(\underline{\theta}|\hat{R}) \quad (3.3)$$

and $p(\ell|\underline{\theta}^*)$ is ℓ 's probability density. So P_D is the probability that the integral over the triangle is greater than the integral along the diagonal.

Unfortunately, $p(\ell|\underline{\theta}^*)$ is not tractable to compute analytically and probably not easy to integrate as required in (3.2). So P_D must be computed by a Monte Carlo simulation*.

We define the 0-1 random variable

$$\ell' = \begin{cases} 0, & \ell < 0 \\ 1, & \ell > 0 \end{cases} \quad (3.4)$$

i.e., $\ell' = 1$ iff two signals are detected. Then P_D is the expected value of ℓ' ,

$$P_D = E(\ell') .$$

Therefore, P_D can be approximated by

$$\hat{P}_D = \frac{1}{M} \sum_{i=1}^M \ell'_i \quad (3.5)$$

where M is the number of statistically independent Monte Carlo trials, as before, and ℓ'_i is a sample of ℓ' from the population of $p(\ell'|\underline{\theta}^*)$ on the i th trial. ℓ'_i is generated, of course, by computing $p(\underline{\theta}|\hat{R}_i)$, finding its 2-D and 1-D integrals as in (3.3), and using the definition (3.4).

If it is assumed that the detector knows $\underline{\theta}^$, then this P_D can be computed and serves as a convenient upper bound for the present one (see Appendix B).

The approximations made here are in the piecewise constant approximation for $p(\underline{\theta}|\hat{R})$ in the integrals and in the finite variance of the random variable $\hat{P}_D$. Assuming the grid has already been chosen fine enough to make the integrals for the rms error calculation in the previous section sufficiently accurate, only the Monte Carlo error remains.

As explained in Appendix A, this error is quantified by the observation that $\hat{P}_D$ has approximately a Gaussian pdf. Using the 1000 trials required for the rms error calculation, we find in the Appendix that the 99% confidence interval for P_D is about $\pm .05$.

3.1.3 Bias Computation

The bias b of an estimate $\underline{\theta}$ is defined as

$$b = E(\underline{\hat{\theta}} - \underline{\theta}^*)$$

and was computed for MMSE and MAP exactly the same way as the rms error and the probability of detection. Given M sample angle estimates, the estimate for the bias was

$$\hat{b} = \frac{1}{M} \sum_{i=1}^M (\hat{\underline{\theta}}_i - \underline{\theta}^*)$$

Confidence intervals for b follow easily and are derived in Appendix A.

3.2 Results

A computer simulation program was written which performed the Monte Carlo evaluation of rms errors, probability of detection, and bias of the optimal estimators (MAP and MMSE). The results of running that program for various array signal-to-noise ratios, emitter separations, and numbers of looks are presented in this section. In every case, two equal power, uncorrelated emitters and a uniform linear array with nine or five elements were assumed. The results indicate that rms errors and detection thresholds corresponding to the Cramer-Rao bound are in fact achievable given as few as 10 looks at the array.

Incidentally, the running time for the program for 9 elements and a 50 by 50 grid was about 0.19 sec/trial on an Amdahl 470 and 3.1 sec/trial on a VAX 11/750.

3.2.1 RMS Error Results

Figures 3.3a through 3.3c show the computed rms errors of the MAP and MMSE estimators versus SNR for emitter separations of .1, $.01/\sqrt{10}$, and .01 beamwidths, respectively, and all at 10 looks. Figure 3.3d is an overlay of 3.3a, b, and c. The array for these cases has nine elements implying a beamwidth of $2/9 \approx .222$. For an emitter separation of $\Delta\theta$, the true signal locations were taken to be

$$\theta_1^* = -\frac{1}{2} \Delta\theta \quad ,$$

$$\theta_2^* = \frac{1}{2} \Delta\theta \quad ,$$

i.e., centered on the array boresight.

Below about 5 dB in SNR, we see from the figure that the rms error is independent of the true signal separation. In fact, it is independent of almost everything in this region since the aposteriori pdf is approximately uniform over the triangle, on the average, implying that the estimates vary approximately uniformly over the triangle. An exactly uniform pdf would imply an rms error of $4.5/\sqrt{3} \approx 2.6$ beamwidths, approximately that of MAP. MMSE is somewhat better since it tends to choose estimates in the center of the triangle given a uniform pdf, and this happens to be close to where the signals actually are.

Above about 20 dB is the more interesting region of SNR. First, we see that the MAP and MMSE estimator errors are identical. This is not surprising, since the aposteriori density is substantially unimodal and symmetric about

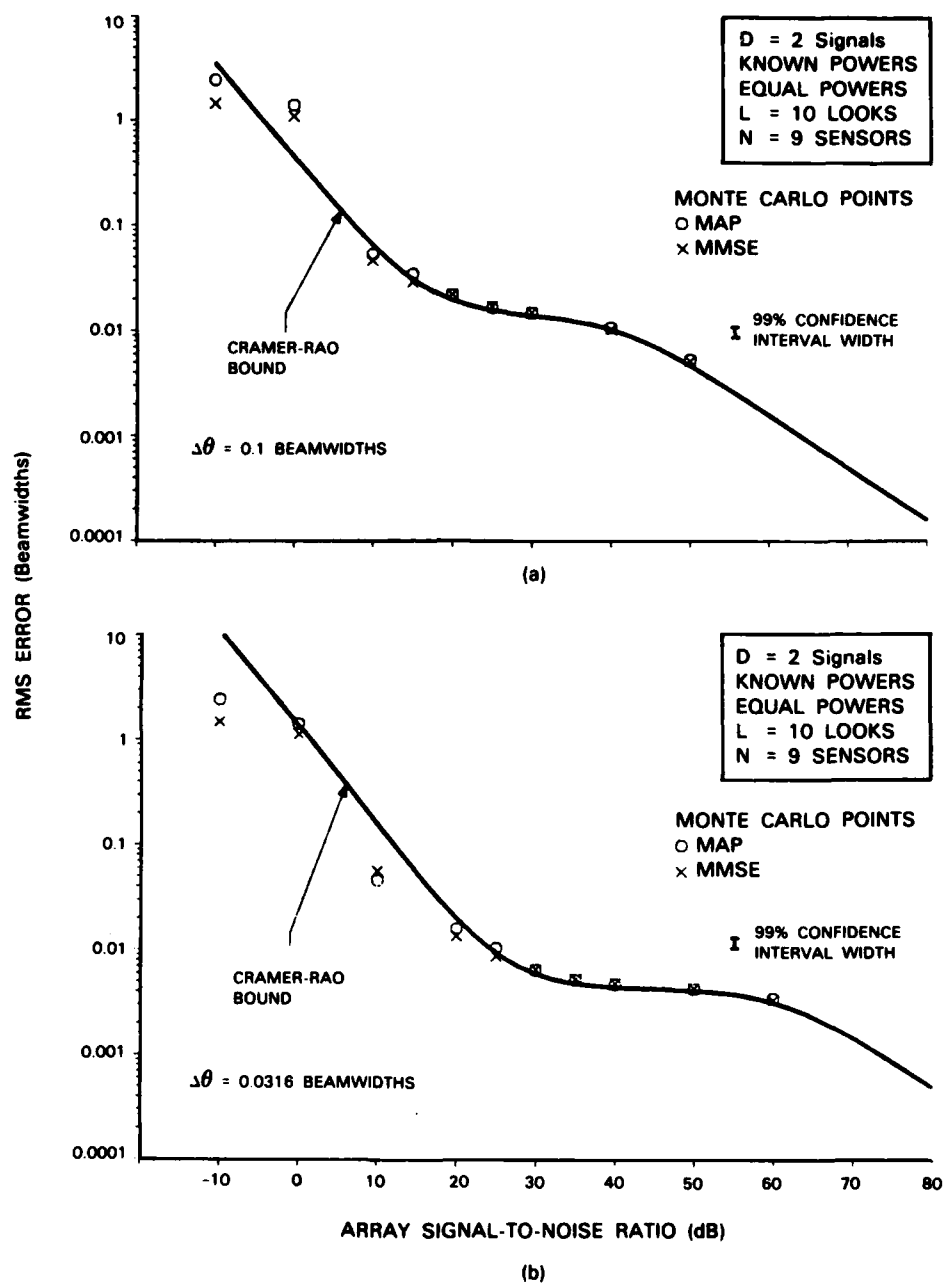


Fig. 3.3. Comparison of MAP and MMSE estimator accuracies with Cramer-Rao bound.

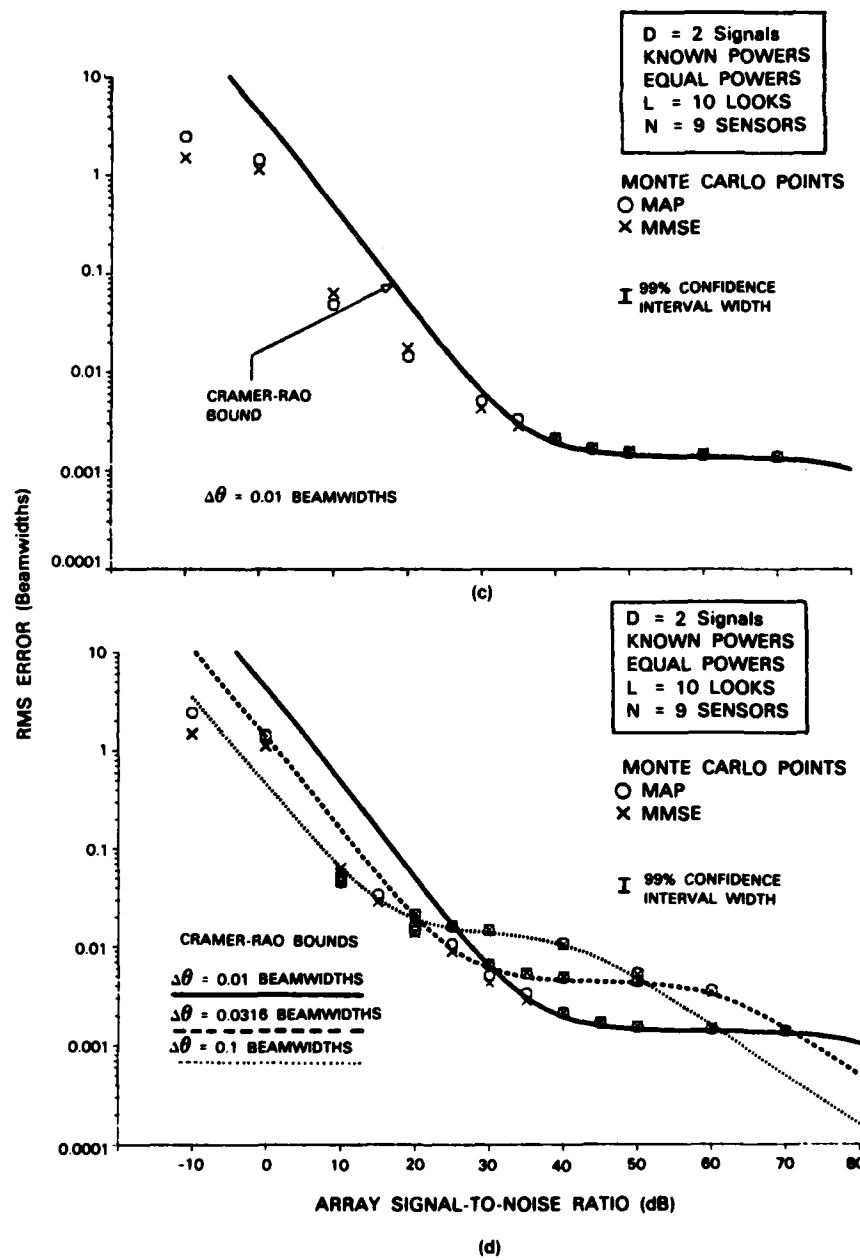


Fig. 3.3. Continued.

141876-R

its mode in this region, so that its mean and peak coincide. More significantly, we see that the rms error is always less than or equal to the Cramer-Rao bound (when less, the estimator was biased) and tracks the bound closely beyond its first "knee," which we will later see is for ($L=10$) at the detection threshold. At 40 dB SNR, for example, we see that the rms error is clearly (including the confidence interval) smaller for the smaller separations, verifying this unusual behavior of the bound noted in Chapter 1. Additionally, we see that the rms error is approximately constant as the SNR increases from, for example, 50 to 70 dB for the .01 beamwidths emitter separation, verifying the plateau in the bound. Therefore, these counterintuitive aspects of the known powers Cramer-Rao bound observed in Chapter 1 apply equally well to the behavior of the optimal estimators. It was not "just a bound" after all.

Of course, the preceding observations are for 10 looks, but this is a surprisingly small number: the convergence of the rms error to the bound as the number of looks increases is quite rapid. Figures 3.4a and 3.4b show this convergence for two selected SNRs and separations, 50 dB, .01 beamwidths, and 30 dB, .1 beamwidths, respectively. The bound is achieved to within the 99% confidence interval after about 10 looks and continues to track the bound for higher numbers of looks.

Expecting that the number of looks required to achieve the bound may be related to the number of antennas, a run was made at 50 dB, .01 beamwidths for only five antenna elements (Fig. 3.5). As can be seen from the figure, the threshold is about the same as for nine antennas. It could be that the threshold depends on the number of signal directions one is trying to estimate.

3.2.2 Probability of Detection Results

Figures 3.6 and 3.7 show the probability of detection computed versus SNR for the same three separations and the same number of looks as used previously. From these plots, we see the region of SNR for which the rms errors are of any practical interest.

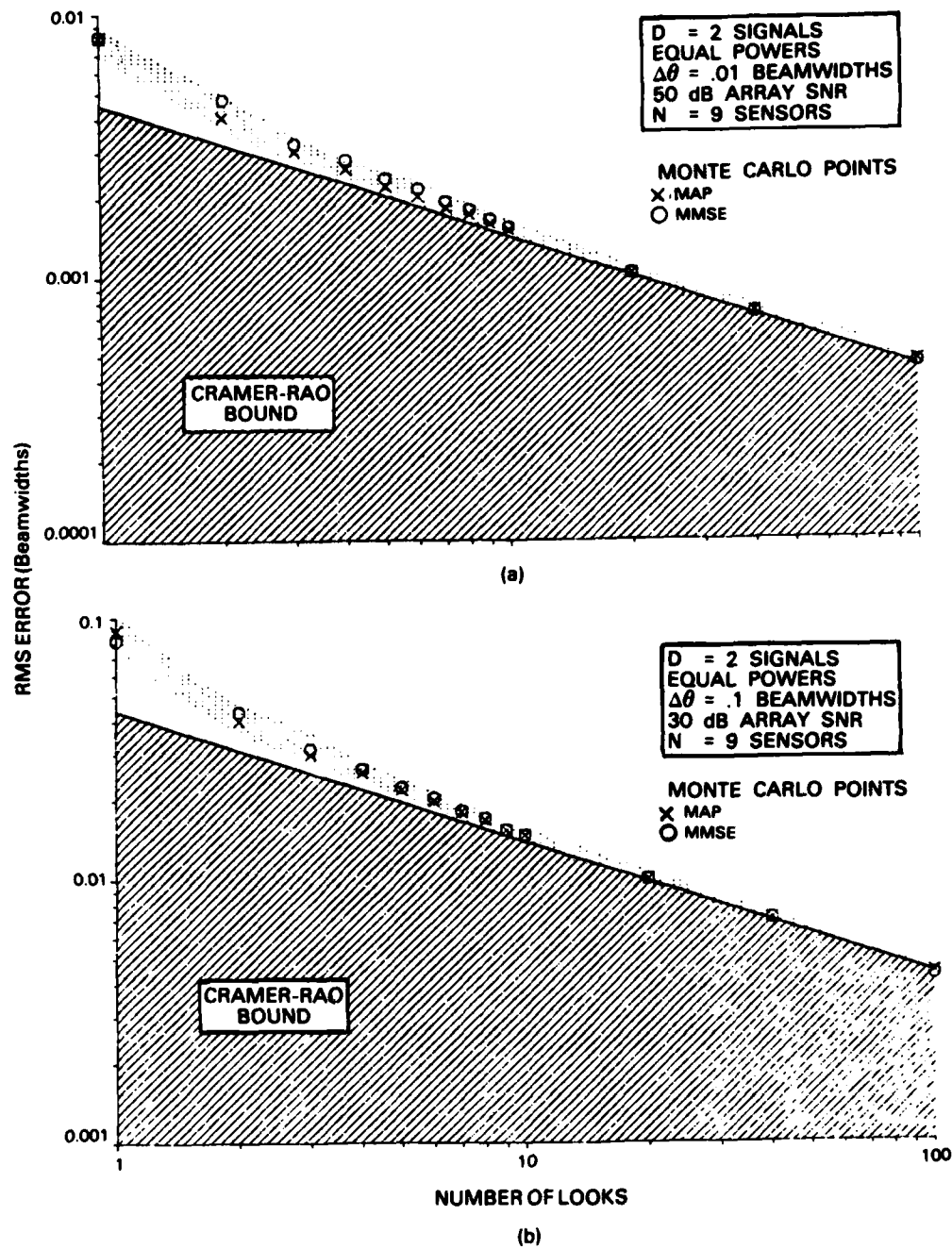


Fig. 3.4. Convergence of MAP and MMSE estimators with number of looks.

141877-R

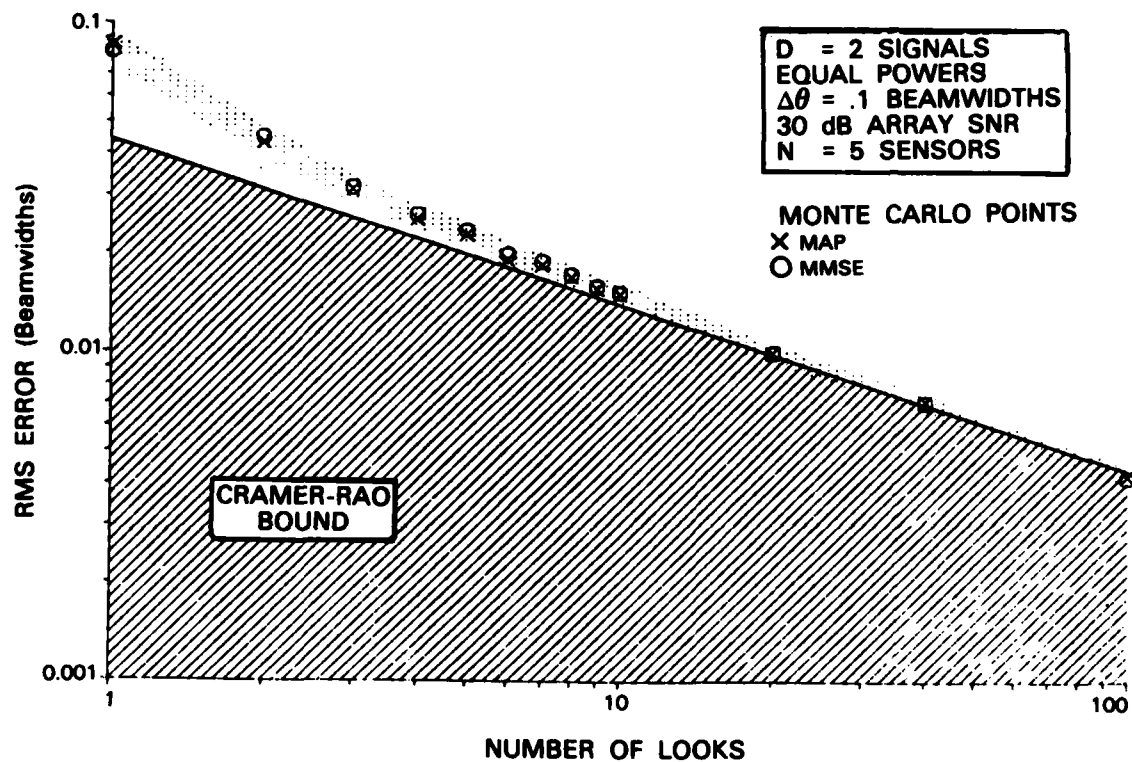


Fig. 3.5. Convergence of MAP and MMSE estimators with number of looks.

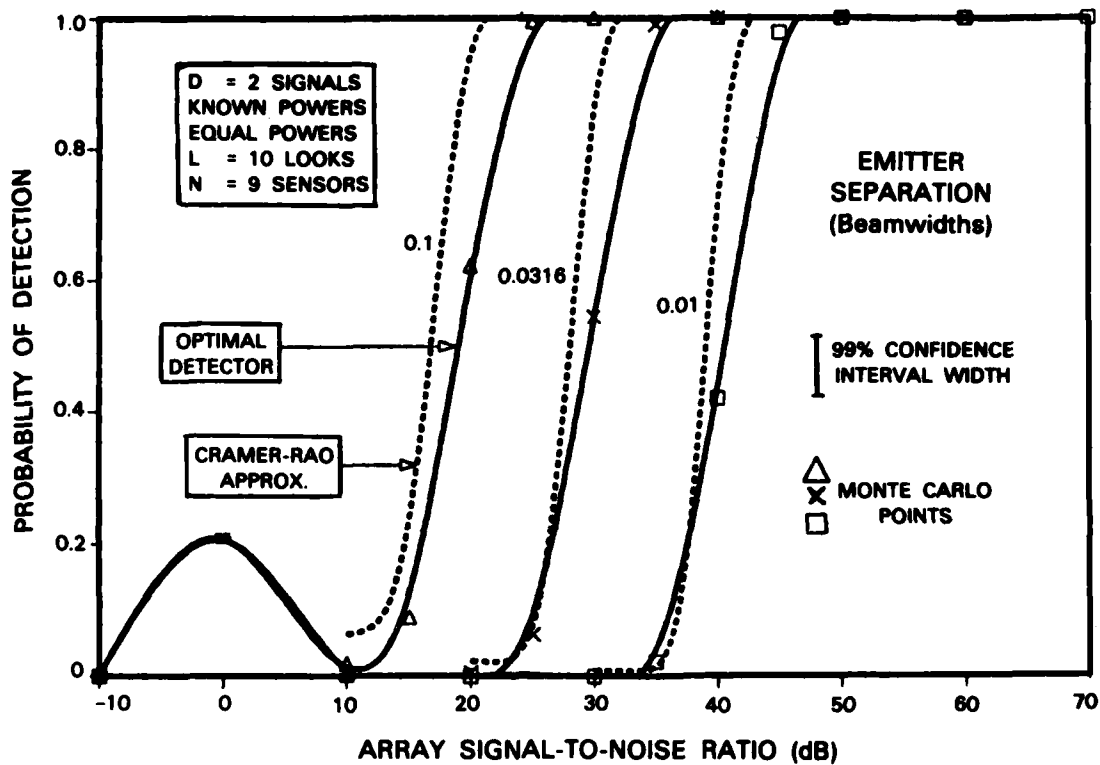


Fig. 3.6. Comparison of optimal detector performance to Cramer-Rao derived approximation.

141879-R

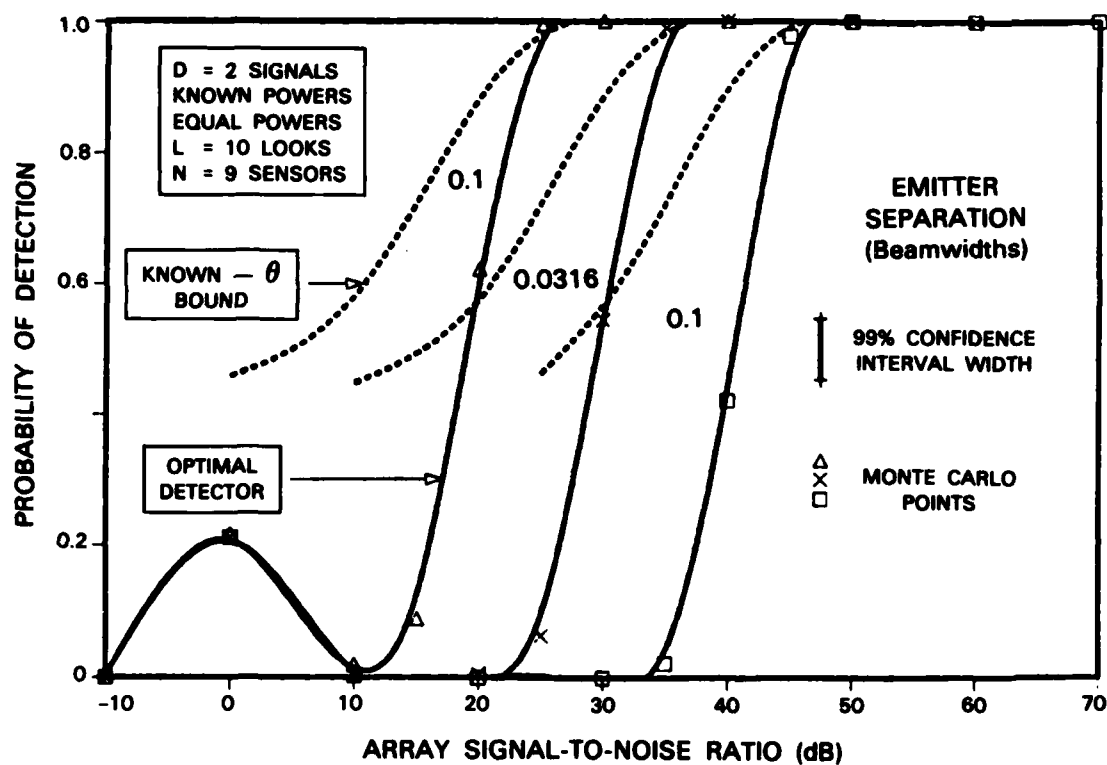


Fig. 3.7. Comparison of optimal Bayes detector performance to known- θ bound.

Below 10 dB in SNR, we see that the P_D does not depend on the true separation, as was the case for rms error. The detection here is probably no better than a random choice.

At higher SNR, we see sharply defined detection thresholds which, in fact, occur at the point where the rms error began to track the Cramer-Rao bound. If we take the threshold to be when $P_D \geq .5$, then a simple formula based on the figures for the threshold SNR versus separation is

$$\text{Detection Threshold SNR} \approx \frac{10}{(\Delta\theta)^2} \quad (3.6)$$

where $\Delta\theta$ = emitter separation in beamwidths, valid for $L = 10$ looks. This formula yields 20 dB, 30 dB, and 40 dB SNR for .1, .01 $\sqrt{10}$, and .01 beamwidths separation, respectively.

The dotted curves in Fig. 3.6 are an ad-hoc approximation to P_D based on the Cramer-Rao bound. They were generated by assuming that $p(\theta|\hat{R})$ was a Gaussian mound with mean given by θ^* and covariance given by the Cramer-Rao bound for the separation and SNR used. The integrals of the density over the upper triangle (call this I_2) and along the diagonal (I_1) were evaluated exactly as in Eq. (2.3). The approximation to P_D was then taken as

$$P_D \approx \frac{I_2}{I_2 + I_1} \quad .$$

Although this derivation is loose, here we see that it works quite well. This implies that the a posteriori density is fairly well approximated by a Gaussian mound with the Cramer-Rao bound as its covariance. This is also saying that the off-diagonal term of the Cramer-Rao bound is being achieved, since it corresponds to the orientation of the Gaussian mound, hence to the relative values of I_2 and I_1 and thus to the P_D . Such was expected from examining the picture of the pdf for various cases, as in Fig. 2.5.

In Fig. 3.7, the computed P_D 's are compared to the known $\underline{\theta}$ P_D mentioned in Chapter 2 and derived in Appendix B. It is indeed an upper bound, although not a very tight one. The known $\underline{\theta}$ curves also follow the $1/(\Delta\theta)^2$ dependence in Eq. (3.1) closely, indicating that knowing $\underline{\theta}$ is equivalent to a fixed gain in SNR, independent of separation.

In Fig. 3.8, P_D is plotted versus the number of looks, for the same two cases as in Fig. 3.4. The threshold here is at about five looks, which agrees fairly closely with the threshold observed for the rms error to achieve the bound.

3.2.3 Bias Results

In Figs. 3.9a through 3.9c, the bias of the MAP and MMSE estimators versus SNR is plotted as a percentage of the rms error. Below $\approx 20\%$, one can consider the estimator unbiased since the bias affects the rms error only by $(.2)^2 = 4\%$. From the plot, we see that the estimator is indeed biased, as it must be, in the low SNR region where its rms error was below the Cramer Rao bound, but becomes unbiased at about the detection threshold.

The confidence interval shown is so wide because it was taken as the maximum (most conservative) estimate, i.e.,

$$\text{upper limit} = 100\% \cdot \frac{\text{upper bias limit}}{\text{lower rms error limit}} \quad \text{and}$$

$$\text{lower limit} = 100\% \cdot \frac{\text{lower bias limit}}{\text{upper rms error limit}} \quad .$$

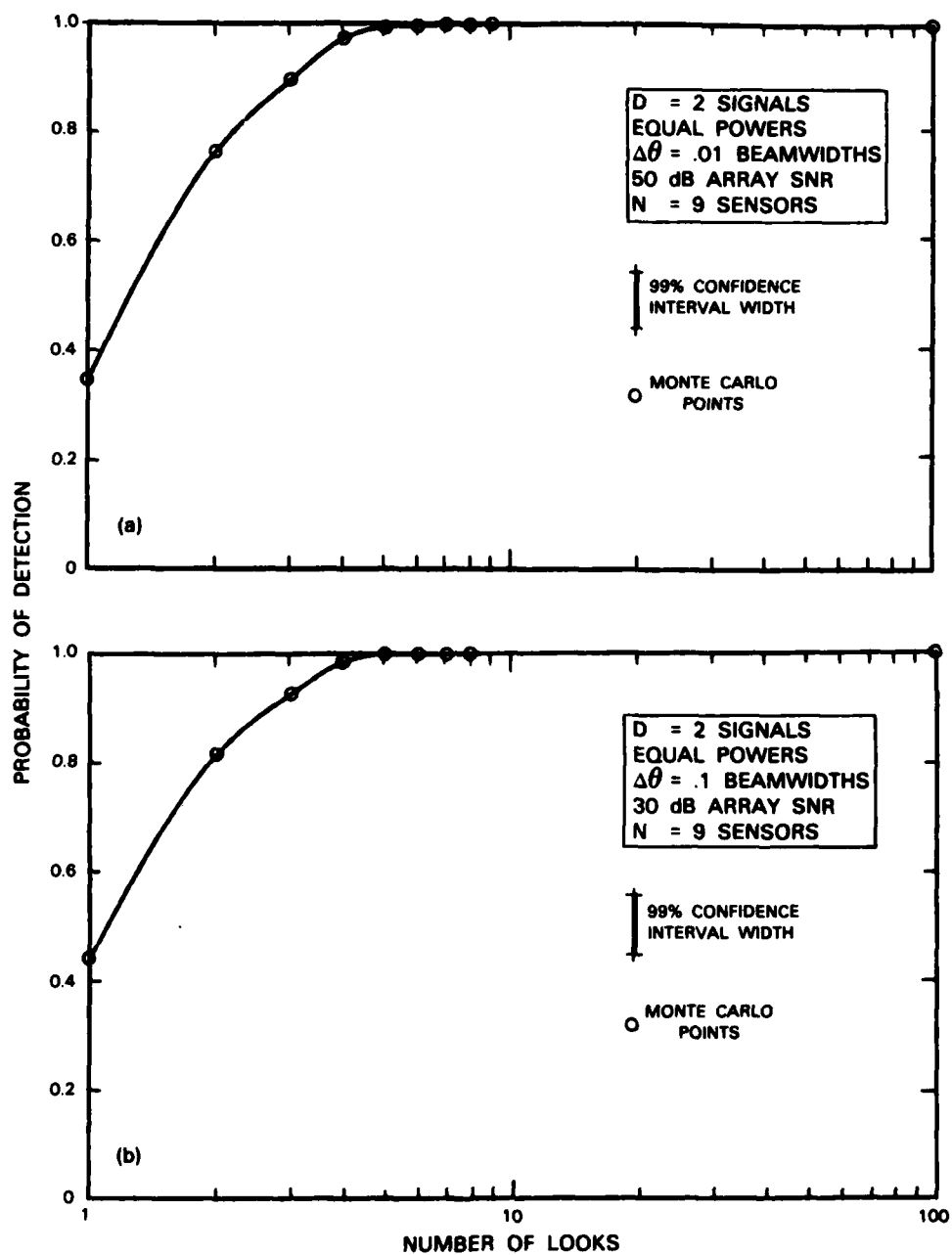


Fig. 3.8. Performance of optimal Bayes detector versus number of looks.

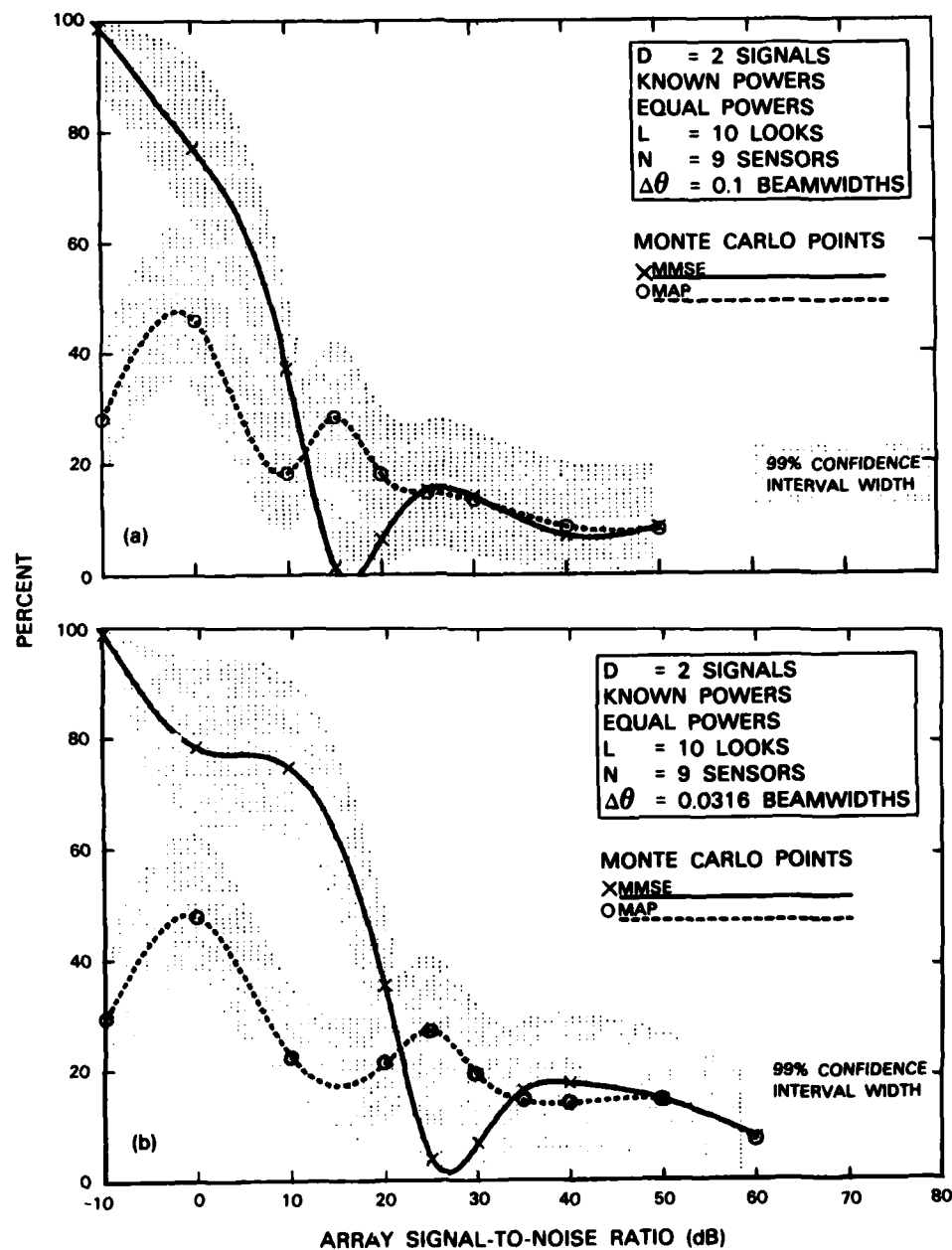


Fig. 3.9. Biases as percentage of rms error.

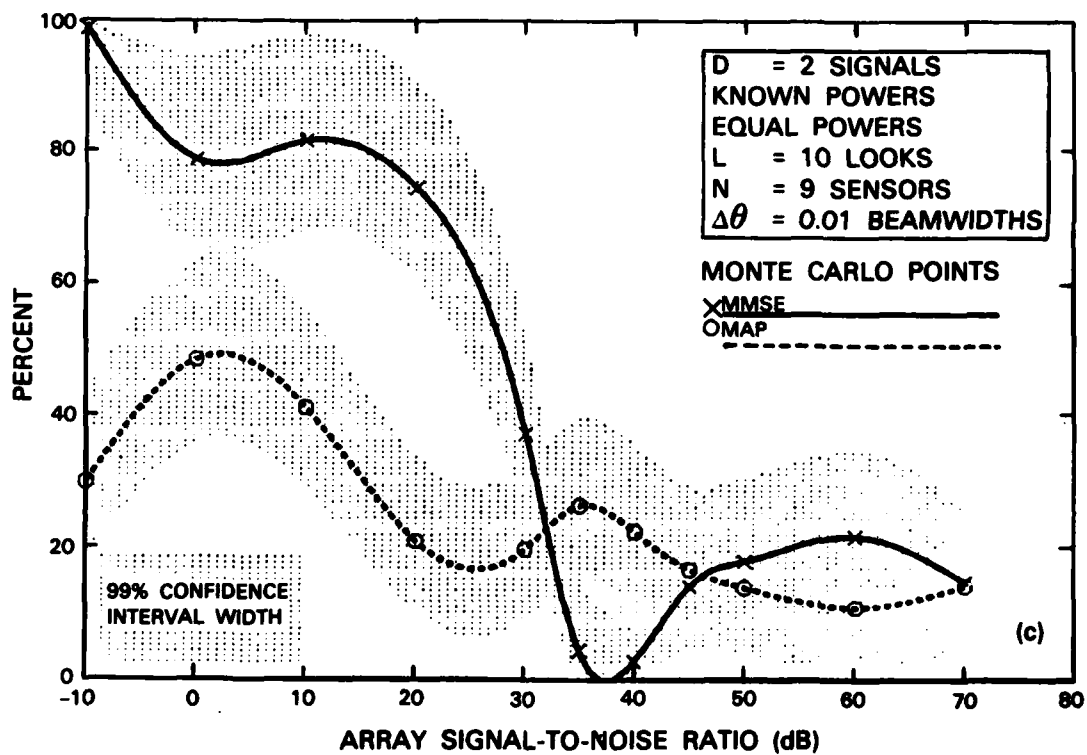


Fig. 3.9. Continued.

141803-R

4.0 CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

The most important result of this research is the presentation of the agreement between the rms errors of the optimal estimators and the known powers, Gaussian signal Cramer-Rao bound. The two unusual features of the bound, that the rms error is constant over a range of SNR and is smaller for larger emitter separations over about the same SNR range, have been verified. The verification was made rigorous through the use of a straightforward finite-grid computation for the MAP and MMSE estimates and through the evaluation of statistical confidence intervals for Monte Carlo results.

As a consequence of this agreement, the Cramer-Rao bound can be now used with some confidence to predict the ultimate performance of angle-of-arrival estimators for various combinations of assumptions. This is valuable since the bound is considerably easier to compute than the optimal estimator performance. One application of this was the ad-hoc formula for the P_D derived from the Cramer-Rao bound in Chapter 3, which agreed well with the P_D computed by the simulation. Without running additional simulations, the bound can be used for cases of different numbers of sensors, looks, or different array configurations from those run in this research to predict performance. The unknown powers Cramer-Rao bound may even be used, with some confidence, a case for which the simulation is considerably more difficult, requiring a 4-dimensional grid.

Although it was the non-random parameters Cramer-Rao bound which was actually verified, a random parameter (Bayesian) approach to the overall estimation problem was used in order to define the MMSE estimator, for one, and more importantly the aposteriori density of the directions-of-arrival $p(\theta|K)$. The properties of this function are important since they could lead to efficient procedures for computing the optimal estimators and to insights into the sources of estimation errors.

While the known powers problem is certainly of theoretical interest, isolating as it does the difficulty in angle estimation alone, it might not be of great practical utility. Further research should therefore center on the unknown-powers problem. This research could be approached in several ways.

A straightforward extension to a 4-dimensional grid search is, of course, not computationally feasible. One could test the importance of knowing the powers by providing the known powers estimator with incorrect powers, which would determine how finely tuned the estimator is to the particular signal model assumptions. If it is not very sensitive to misinformation about powers, then one would expect completely unknown powers to cause little additional degradation.

Extensions of this research to even more general cases, such as > 2 signals or to 2-dimensional angles parameters, await sophisticated but provably accurate procedures to search a many-dimensional space. Our restriction to just two known power signals was made specifically to avoid all questions of convergence which typically arise with such procedures. Hopefully, the results presented here will be of some help in the general problem.

The plots of $p(\underline{\theta}|\hat{R})$, for example, provide some insight into the animal which must be searched. The fact that $p(\underline{\theta}|\hat{R})$ seems to approach a Gaussian mound with the Cramer-Rao bound as its covariance, if quantified, could be used.

ACKNOWLEDGEMENTS

I must thank Ken Senne and Dick Bucy for suggesting the present problem, rescuing me from a previous intractable one, and for guidance throughout the research. Thanks to Jim McClellan for many helpful discussions and much patience as my advisor. Thanks also to Russ Johnson for technical suggestions and moral support and to Diane Young for excellent typing on a recalcitrant word processor. Finally, I thank Lincoln Laboratory for its generous support over my years as a research assistant there.

REFERENCES

1. R. S. Bucy, C. Hecht, and K. D. Senne, An Engineer's Guide to Building Nonlinear Filters, Vol. 1, Frank J. Seiler Research Laboratory (AFSC), USAF Academy, Colorado, 1972.
2. R. S. Bucy and K. D. Senne, "Nonlinear Filtering Algorithms for Vector Processing Machines," Comp. & Maths. with Appls. 6, 317 (1980).
3. I. N. El-Behery and R. H. Macphie, "Maximum-Likelihood Estimation of Source Parameters from Time-Sampled Outputs of a Linear Array," Journal of the Acoustical Society of America 61, 125 (1977).
4. M. A. Gallop and L. W. Nolte, "Bayesian Detection of Targets of Unknown Location," IEEE Trans. Aerospace and Electronic Systems AES-10, 429 (1974).
5. N. R. Goodman, "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)," Ann. Mathematics and Statistics 34, 152 (1963).
6. D. F. DeLong, "Multiple Signal Direction Finding with Thinned Linear Arrays," Project Report TST-68, Lincoln Laboratory, M.I.T. (13 April 1983), DTIC AD-A128924.
7. W. S. Hodgkiss and L. W. Nolte, "Bayes Optimum and Maximum-Likelihood Approaches in the Array Processing Problem," IEEE Trans. Aerospace and Electronic System, AES-11, 913 (1975).
8. A. A. Ksienski and R. M. McGhee, "A Decision Theoretic Approach to the Angular Resolution and Parameter Estimation for Multiple Targets," IEEE Trans. Aerospace and Electronic Systems AES-4, 443 (1968).
9. D. G. Lainiotis, "Optimal Adaptive Estimation: Structure and Parameter Adaptation," IEEE Trans. Automatic Control AC-16, 160 (1971).
10. G. E. Pollon, "On the Angular Resolution of Multiple Targets," IEEE Trans. Aerospace and Electronic Systems AES-4, 145 (1967).
11. G. E. Pollon and G. W. Lank, "Angular Tracking of Two Closely Spaced Radar Targets," IEEE Trans. Aerospace and Electronic Systems AES-4, 145 (1967).

REFERENCES (continued)

12. J. R. Sklar and F. C. Schweppe, "On the Angular Resolution of Multiple Targets," *Proc. IEEE* 52, 1044 (1964).
13. D. C. Rife and R. R. Boorstyn, "Single Tone Parameter Estimation from Discrete-Time Observations," *IEEE Trans. Inform. Theory* IT-20, 591 (1974).
14. D. C. Rife and R. R. Boorstyn, "Multiple Tone Parameter Estimation from Discrete-Time Observations," *BSTJ* 55, 1398 (1976).
15. H. L. Van Trees, Detection, Estimation, and Modulation Theory, Part 1 (Wiley, New York, 1968).
16. S. J. Yakowitz, Computational Probability and Simulation, Chapter 6 (Addison Wesley, Reading, 1977).
17. G. O. Young, "Optimum Space-Time Signal Processing and Parameter Estimation," *IEEE Trans. Aerospace and Electronic Systems* AES 4, 334 (1968).
18. S. S. Wilks, Mathematical Statistics (Wiley, New York, 1962).

APPENDIX A CONFIDENCE INTERVALS FOR THE MONTE CARLO RESULTS

In this appendix, we derive the confidence intervals for the estimates of bias, mean square error, and probability of detection shown in the figures of Chapter 3. Much of this appendix is abstracted from Chapter 4 of [1].

Let α be the true value of some parameter to be estimated and let $\hat{\alpha}$ be its estimate. Then a $100\gamma\%$ confidence interval for α is a pair of random variables $l(\hat{\alpha})$ and $u(\hat{\alpha})$ such that

$$\text{prob} (l(\hat{\alpha}) \leq \alpha \leq u(\hat{\alpha})) = \gamma$$

[18, p. 365]. Here we choose the symmetric interval where l and u are of the form

$$l(\hat{\alpha}) = \hat{\alpha} - k$$

$$u(\hat{\alpha}) = \hat{\alpha} + k \quad ,$$

thus,

$$\text{prob}(|\hat{\alpha} - \alpha| \leq k) = \gamma \quad .$$

To compute k , it suffices to know the pdf of $z = \hat{\alpha} - \alpha$, say $p(z)$. This implies that k is given by

$$\int_{-k}^k dz p(z) = \gamma \quad .$$

In each of our cases, we are estimating the mean $\bar{x}$ of a random variable x given M independent samples $\{x_1, \dots, x_M\}$ of that variable. Hence, the best estimate of x is

$$\hat{x} = \frac{1}{M} \sum x_i \quad .$$

If we assume now that M is so large that $\hat{\bar{x}}$ is approximately Gaussian, then $\hat{\bar{x}} - \bar{x}$ is Gaussian with mean, variance given by

$$E(\hat{\bar{x}} - \bar{x}) = \frac{1}{M} \sum E(x_i) - \bar{x} = 0$$

$$\sigma^2 = \text{var}(\hat{\bar{x}}) = \frac{E(\bar{x}^2) - \bar{x}^2}{M} = \frac{\text{var}(x)}{M}$$

and we can easily derive a symmetric confidence interval. In particular, the 99.74% confidence interval is given by

$$|\hat{\bar{x}} - \bar{x}| \leq 3\sigma \quad .$$

Unfortunately, although the pdf of $\hat{\bar{x}}$ may be very nearly Gaussian, the pdf of x is unknown so σ^2 is also unknown. Here, we get an approximate confidence interval by using the estimate of σ^2

$$\sigma^2 \approx \frac{1}{M} \sum x_i^2 - \hat{\bar{x}}^2$$

which should be a good approximation for large M .

For bias b and mean square error s , we are estimating the means of

$$x = \hat{\theta} - \theta^*$$

and

$$x = (\hat{\theta} - \theta^*)^2$$

respectively. Hence, given M sample angle estimates $\{\hat{\theta}_1, \dots, \hat{\theta}_M\}$, we compute the estimates

$$\hat{b} = \frac{1}{M} \sum (\hat{\theta}_i - \theta^*)$$

$$\hat{s} = \frac{1}{M} \sum (\hat{\theta}_i - \theta^*)^2$$

and the fourth moment

$$\hat{u}_4 = \frac{1}{M} \sum (\hat{\theta}_i - \theta^*)^4$$

implying that the approximate 99.74% confidence intervals as derived above are

$$\begin{aligned} |\hat{b} - b| &\leq 3\sqrt{\hat{s} - b^2} \\ |\hat{s} - s| &\leq 3\sqrt{\hat{u}_4 - s^2} \equiv f \end{aligned}$$

The interval for the rms error, $\sqrt{s}$, was derived from the interval for s as

$$\sqrt{\hat{s} - f} \leq \sqrt{s} \leq \sqrt{\hat{s} + f}$$

On the plots, all of the intervals (for each SNR and separation) were approximately the same so the largest one was taken as the interval for all the points. By making runs for various values of M and computing the confidence intervals, $M=1000$ trials was found adequate for a ± 2 dB interval for rms error.

For the probability of detection, we can be a little more precise since we are estimating the mean of a binomial random variable

$$x = \begin{cases} 0, & \text{no detection, probability } 1 - P_D \\ 1, & \text{detection, probability } P_D \end{cases}$$

and can compute the worst case $\text{var}(x)$:

$$\text{var}(x) = E(x^2) - E(x)^2 = P_D - P_D^2 = P_D(1 - P_D)$$

$$\leq \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \quad \text{for all } 0 \leq P_D \leq 1$$

implying

$$\text{var}(\hat{P}_D) \leq \frac{1}{4M}$$

Therefore, the worst case 99.74% confidence interval for P_D is (still assuming P_D is approximately Gaussian)

$$|\hat{P}_D - P_D| \leq \frac{3}{2\sqrt{M}}$$

which, at $M = 1000$ is

$$|\hat{P}_D - P_D| \leq .047 \quad ,$$

as shown in the P_D plots of Chapter 3.

APPENDIX B

THE KNOWN ANGLE-OF-ARRIVAL DETECTOR

The probability of detection (P_D) of the optimal Bayes detector described in Chapter 2 can be upper bounded by the P_D of a detector which knows a priori the angles of arrival $\underline{\theta}$. This latter P_D , unlike the former, may be computed analytically and hence is a convenient bound. In this appendix, we derive in detail the P_D of the known $\underline{\theta}$ bound, which is seen to be the solution of the general Gaussian hypothesis testing problem in [15].

By the phrase "known $\underline{\theta}$," we mean more precisely that if there are actually two signals present (hypothesis H_2), we know their angles

$$\underline{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix},$$

and if there is only one signal present (hypothesis H_1), we know its angle θ . Of course, we assume throughout this report that the powers of all the signals are also known, say p_1 and p_2 for two signals present, and p for one signal. Normally, for the problem to make physical sense, we assume

$$p = p_1 + p_2$$

(B.1)

$$\theta = \frac{p_1 \theta_1 + p_2 \theta_2}{p_1 + p_2} \quad (\text{center of power}).$$

Knowing the angles and the powers now means that we completely know the covariance of the observations $\{\underline{x}_i\}_{i=1}^L$ on each hypothesis. Therefore, the detection problem is to discern based on the L independent samples $\{\underline{x}_i\}_{i=1}^L$ whether

$$H_1: \underline{x}_1 \sim \text{CN}(\underline{0}, R_1)$$

or

$$H_2: \underline{x}_1 \sim \text{CN}(\underline{0}, R_2)$$

where

$$R_1 = I + p \underline{v}(\theta) \underline{v}(\theta)^H \quad (\text{B.2})$$

$$R_2 = I + V(\theta) P V(\theta)^H$$

$$P = \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix}$$

$\underline{v}(\theta)$ = direction vector

$$\underline{V}(\theta) = \begin{pmatrix} \underline{v}(\theta_1) & \underline{v}(\theta_2) \end{pmatrix}$$

$\text{CN}(\underline{m}, R)$ = complex normal pdf with mean $\underline{m}$ and covariance R .

This is the well-known equal mean, differing covariance Gaussian hypothesis testing problem [15, p. 113, case 1C]. The computation of its P_D is tedious, but straightforward.

The likelihood ratio test is

$$\frac{p(\{\underline{x}_i\}_{i=1}^L | H_2)}{p(\{\underline{x}_i\}_{i=1}^L | H_1)} \underset{H_1}{\overset{H_2}{\geq}} \gamma' \quad (\text{B.3})$$

where $\gamma' = 1$ for minimum probability error (P_E) detection assuming each hypothesis is equally likely ($P(H_1) = p(H_2) = 1/2$).

Plugging the complex normal distribution into (B.3) and using the independence of the $\underline{x}_i$'s, we get

$$\frac{|\pi R_1|^L}{|\pi R_2|^L} \exp \left(- \sum_{i=1}^L (\underline{x}_i^H R_2^{-1} \underline{x}_i - \underline{x}_i^H R_1^{-1} \underline{x}_i) \right) \gtrless \gamma'$$

or

$$\sum_{i=1}^L \underline{x}_i^H (R_1^{-1} - R_2^{-1}) \underline{x}_i \gtrless \ln \gamma' - L \ln \frac{|R_2|}{|R_1|}$$

or

$$L \operatorname{tr} ((R_1^{-1} - R_2^{-1}) \hat{R}) \gtrless \ln \gamma' + L \ln \frac{|R_2|}{|R_1|}$$

or, finally,

$$\operatorname{tr} ((R_1^{-1} - R_2^{-1}) \hat{R}) \gtrless \frac{\ln \gamma'}{L} + \ln \frac{|R_2|}{|R_1|} \quad (\text{B.4})$$

where

$$\hat{R} \triangleq \frac{1}{L} \sum_{i=1}^L \underline{x}_i \underline{x}_i^H$$

Let

$$\ell \triangleq \text{tr}((R_1^{-1} - R_2^{-1}) \hat{R})$$

be the statistic and

$$\gamma \triangleq \frac{\ln \gamma'}{L} + \ln \left| \frac{R_2}{R_1} \right|$$

be the threshold. Then the P_D of this test is

$$P_D = \text{prob}(\ell \geq \gamma) = \int_{\gamma}^{\infty} p_{\ell}(u|H_2) du$$

where $p_{\ell}(\cdot)$ is the pdf of ℓ , which will be obtained through the characteristic function of ℓ .

The characteristic function of ℓ ,

$$\phi_{\ell}(j\omega) \triangleq E(e^{j\omega\ell})$$

is given on hypothesis H_k by

$$\phi_{\ell}(j\omega|H_k) = \left| I - \frac{1}{L} j\omega R_k (R_1^{-1} - R_2^{-1}) \right|^{-L}$$

as derived in [5, Lemma 4.1]. Pulling out an $R_k^{1/2}$,

$$\phi_{\ell}(j\omega|H_k) = \left| I - \frac{1}{L} j\omega R_k^{1/2} (R_1^{-1} - R_2^{-1}) R_k^{1/2} \right|^L$$

Now let

$$U \Lambda U = R_k^{1/2} (R_1^{-1} - R_2^{-1}) R_k^{1/2} \quad (B.5)$$

be the eigenvalue decomposition of the Hermetian matrix on the right-hand side. Then,

$$\begin{aligned} \phi_L(j\omega | H_k) &= \left| I - \frac{j\omega}{L} U \Lambda U^H \right|^{-L} \\ &= \left| I - \frac{j\omega}{L} \Lambda \right|^{-L} \\ &= \det \begin{pmatrix} 1 - \frac{j\omega}{L} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & 1 - \frac{j\omega}{L} \lambda_n \end{pmatrix}^{-L} \\ \phi_L(j\omega | H_k) &= \frac{1}{\left(1 - \frac{j\omega \lambda_1}{L}\right)^L \cdot \cdot \cdot \left(1 - \frac{j\omega \lambda_n}{L}\right)^L} \end{aligned}$$

This is a difficult characteristic function to invert until we note that only three eigenvalues are non-zero since

$$\begin{aligned}
\text{rank } (R_k^{1/2} (R_1^{-1} - R_2^{-1}) R_k^{1/2}) &= \text{rank } (R_1^{-1} - R_2^{-1}) \\
&= \text{rank } (I - q \underline{v} \underline{v}^H - (I - v Q v^H)) \\
&= \text{rank } (v Q v^H - q \underline{v} \underline{v}^H) \\
&= 3 \text{ at most.}
\end{aligned}$$

Now

$$\phi_\ell(Ls | H_k) = \frac{(-1)^L}{\lambda_1^L \lambda_2^L \lambda_3^L (s - \lambda_1^{-1})^L (s - \lambda_2^{-1})^L (s - \lambda_3^{-1})^L}$$

which may be expanded into partial functions

$$\phi_\ell(Ls | H_k) = \frac{1}{\lambda_1^L \lambda_2^L \lambda_3^L} \sum_{j=1}^3 \sum_{i=1}^L \frac{a_{ij}}{(s - b_j)^i} \quad (\text{B.6})$$

where $b_j = 1/\lambda_j$ and the coefficients a_{ij} are given by

$$\begin{aligned}
a_{11} &= \frac{(-1)^L}{(L-1)!} \frac{d^{L-1}}{ds^{L-1}} [(s - b_2)^{-L} (s - b_3)^{-L}] \Big|_{s=b_1} \\
&= \frac{(-1)^L}{(L-1)!} \sum_{\ell=0}^{L-1} \binom{L-1}{\ell} (-1)^\ell \frac{(L+\ell-1)!}{(L-1)!} (b_1 - b_2)^{-(L+\ell)} \times \\
&\quad (-1)^{L-1-\ell} \frac{(L+L-1-\ell-1)!}{(L-1)!} (b_1 - b_3)^{-(L+L-1-\ell)}
\end{aligned}$$

$$= \frac{(-1)^{2L-1}}{(L-1)! (b_1 - b_2)^L (b_1 - b_3)^{2L-1}} \sum_{k=0}^{L-1} \left(\frac{b_1 - b_3}{b_1 - b_2} \right)^k \times$$

$$\frac{(L-1)! (L+\ell-1)! (2L-1-\ell-1)!}{\ell! (L-1-\ell)! (L-1)!^2}$$

or, finally,

$$a_{i1} = \frac{(b_3 - b_1)^i}{(b_1 - b_2)^L (b_3 - b_1)^{2L}} \sum_{\ell=0}^{L-1} u_{\ell} r^{\ell}$$

where

$$r = \frac{b_3 - b_1}{b_2 - b_1}, \quad u_0 = \frac{(2L-1-1)!}{(L-1)! (L-1)!} = \binom{2L-1-1}{L-1}$$

$$u_{\ell} = \frac{(L+\ell-1)! (2L-1-\ell-1)!}{\ell! (L-1-\ell)! (L-1)!^2} = u_{\ell-1} \frac{(L+\ell-1) (L-1-\ell+1)}{\ell (2L-1-\ell)}$$

and a_{12} and a_{13} are obtained symmetrically.

Given these a_{1j} , we now inverse transform Eq. (B.6):

$$p_{\ell}(\ell) = F\{\phi_{\ell}(j\omega | H_k)\}$$

$$= \frac{L}{(\lambda_1 \lambda_2 \lambda_3)^L} \sum_{j=1}^3 \sum_{i=1}^L a_{1j} F\{(j\omega - \lambda_j^{-1})^{-1}\} \Big|_{\ell=L \cdot \ell}$$

$$= \frac{L}{(\lambda_1 \lambda_2 \lambda_3)^L} \sum_{j=1}^3 \left\{ \begin{array}{l} u(\ell), \lambda_j > 0 \\ -u(-\ell), \lambda_j < 0 \end{array} \right\} e^{-L\ell/\lambda_j} \sum_{i=1}^L \frac{a_{1j} (L \cdot \ell)^{i-1}}{(i-1)!}$$

Therefore,

$$P_D = \int_{\gamma}^{\infty} dx p_{\ell}(x) = \frac{L}{(\lambda_1 \lambda_2 \lambda_3)^L} \sum_{j=1}^3 \sum_{i=1}^L \frac{a_{ij}}{(i-1)!} \int_{\gamma}^{\infty} dx \left\{ \frac{u(x)}{-u(-x)} \right\} (Lx)^{i-1} e^{-\frac{Lx}{\lambda_j}}$$

or

$$P_D = \frac{L}{(\lambda_1 \lambda_2 \lambda_3)^L} \sum_{j=1}^3 \sum_{i=1}^L a_{ij} \lambda_j^i I_{ij} \quad (B.7)$$

where

$$I_{ij} = \begin{cases} P(i, \frac{L\gamma}{\lambda_j}) & , & (\text{sgn } \lambda_j) & (\text{sgn } \gamma) \\ 0 & , & - & + \\ 1 & , & + & - \\ 1 - P(i, \frac{L\gamma}{\lambda_j}) & , & + & + \end{cases}$$

and $P(\cdot, \cdot)$ is the incomplete Gamma function

$$P(n, x) = \frac{1}{\Gamma(n)} \int_0^x dt t^{n-1} e^{-t}$$

Equation (B.7), of course, gives the probability of false alarm (P_F) also if the eigenvalues λ_j are those corresponding to R_1 instead of R_2 in Eq. (B.5).

A RATFOR program was written which computes P_D and P_F for arbitrary choices of θ_1 , θ_2 , p_1 , p_2 , and L from Eq. (B.7), and was used to generate the solid curves in Fig. B.1.

141884-R

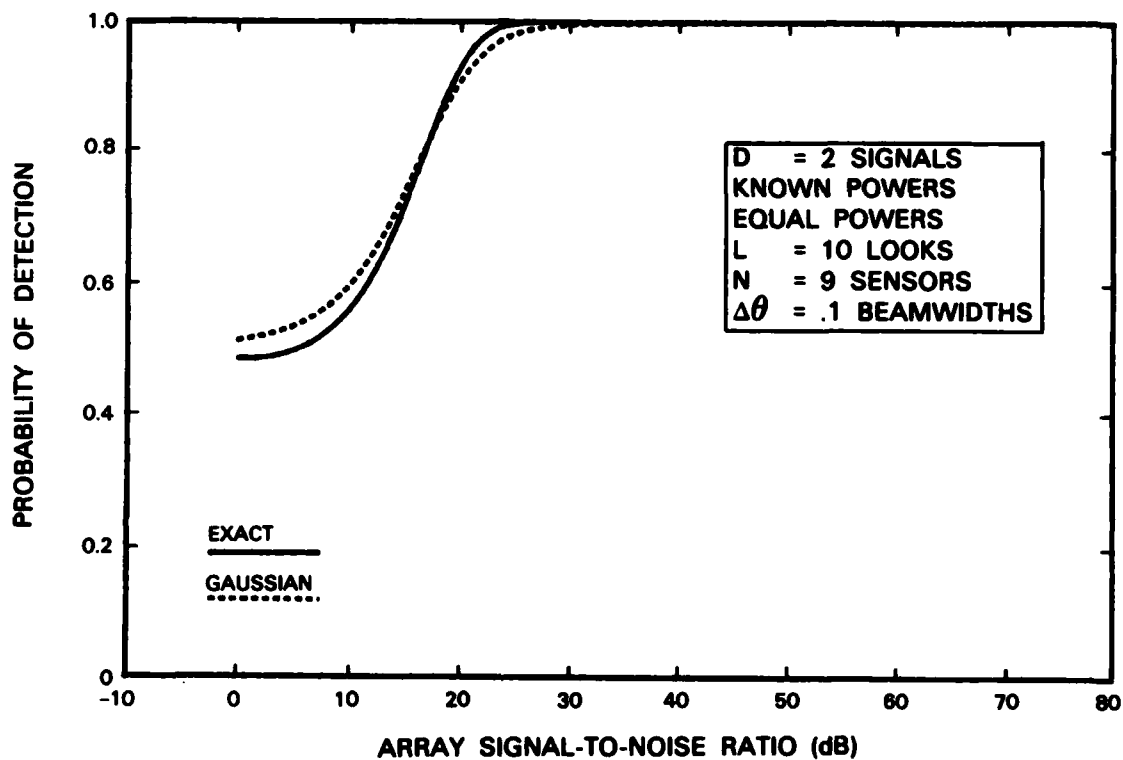


Fig. B.1. Known- θ probability of detection comparison of exact and Gaussian computations.

Unfortunately, there are numerical problems in computing the a_{ij} 's. Fortunately, the pdf of ℓ approaches Gaussian for large numbers of looks and is reasonably close to Gaussian at the point where the exact computation fails. Accordingly, the program optimally computes P_D and P_F based on a Gaussian approximation given by

$$P_d = \int_{\lambda}^{\infty} \frac{\lambda - \bar{\ell}}{\sqrt{2\pi} \sigma} e^{-(\lambda - \bar{\ell})^2 / 2\sigma^2}$$

$$= \begin{cases} \frac{1}{2} \operatorname{erf} \left(\frac{\lambda - \bar{\ell}}{\sqrt{2} \sigma} \right) & , \quad \lambda \leq \bar{\ell} \\ \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\bar{\ell} - \lambda}{\sqrt{2} \sigma} \right) \right) & , \quad \lambda \geq \bar{\ell} \end{cases}$$

where

$$\bar{\ell} = E(\ell) = \left. \frac{d \phi_{\ell}(s)}{ds} \right|_{s=0}$$

$$= \lambda_1 + \lambda_2 + \lambda_3$$

and

$$\sigma^2 = \operatorname{var}(\ell) = \left. \frac{d^2 \phi_{\ell}(s)}{ds^2} \right|_{s=0} - \bar{\ell}^2$$

$$= \frac{1}{L} (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)$$

This approximation appears as dotted curves in Fig. B.1.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ESD-TR-84-028	2. GOVT ACCESSION NO. AD-A146 594	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Performance of Bayes-Optimal Angle-of-Arrival Estimators		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER Technical Report 654
7. AUTHOR(s) Fredric M. White		8. CONTRACT OR GRANT NUMBER(s) F19628-80-C-0002
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lincoln Laboratory, M.I.T. P.O. Box 73 Lexington, MA 02173-0073		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Program Element No. 33401G
11. CONTROLLING OFFICE NAME AND ADDRESS Department of Defense The Pentagon Washington, DC 20301		12. REPORT DATE 13 August 1984
		13. NUMBER OF PAGES 78
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Electronic Systems Division Hanscom AFB, MA 01731		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES None		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
nonlinear parameter estimation nonlinear filtering Bayes-optimal estimation spectral estimators	angle-of-arrival estimators Monte Carlo evaluation techniques minimum mean-square error estimation maximum apriori probability estimation	Gaussian signal Gaussian noise Cramer-Rao bound
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)		
The angle-of-arrival estimation problem for waves incident upon a sensor array was examined through a Monte Carlo evaluation of the performance of the Bayes-optimal MAP (maximum a posteriori) and MMSE (minimum mean square error) estimators. The case of two independent wave emitters of known powers as well as a multiple look, Gaussian signal in Gaussian noise statistical model were assumed. The Cramer-Rao bound on the estimator's rms error was computed for comparison.		

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

28. ABSTRACT (Continued)

The evaluation proceeded with the computation of MAP and MMSE angle estimates for 1000 random samples of array outputs and the accumulation of their rms errors. The probability of detecting both emitters with the optimal detector was also accumulated. This was done for .1, .03, and .01 beamwidths emitter separations and a range of signal-to-noise ratios (SNRs). The accuracy of the computations was assured through a simple finite grid approximation for the estimates, with no convergence problems, and through the evaluation of statistical confidence intervals for the Monte Carlo data.

The results of the evaluation indicated that the Cramer-Rao bound was achievable by both the MAP and MMSE estimators over a wide range of SNR provided a few as 10 looks had been taken. In general, the bound was achieved wherever both signals were detectable. These results were surprising since the bound exhibited unusual behavior; for example, in one SNR region, the bound showed smaller rms errors for more closely-spaced emitters.

Additional results included properties of the a posteriori probability density and an analytical computation of the performance of the known angles-of-arrival optimal detector.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)