ON LEAST SQUARES CRITERIA FOR ESTIMATION OF SIGNAL PARAMETERS AND NOISE PARAMETERS IN WAVEFIELDS

D. Kraus and J.F. Böhme

Lehrstuhl für Signaltheorie, Ruhr-Universität Bochum, 4630 Bochum, FRG

1. INTRODUCTION

The source location estimation problem in the presence of partly unknown noise fields is investigated. In eigenstructure methods or parametric methods as maximum likelihood, the additive noise structure is usually assumed to be known, for instance to be sensor noise, i.e. of equal power and uncorrelated from sensor to sensor. The aim is to find a suitable estimate not requiring this knowledge. In sonar, noise structures can be complicated and unknown. The use of a wrong noise model can result in a break down of the estimate.
Certain knowledge about the structure of noise enables to estimate the other part together with the signal parameters. Recently some authors, cf. [1], [2], [3] and [4], adapt eigenstructure methods for different noise models. In continuation of [5], we investigate the conditional maximum likelihood estimate (CMLE), the conditional marginal maximum likelihood estimate (CMMLE) and three different least squares fits (LSF's). One of the LSF's is a slight modification of an estimate proposed in [3] that supplies asymptotic efficient estimates. In opposite to [3], we use an unbiased estimate of the inverse spectral density matrix obtained by an appropriate scaling of the inverse sample spectral density matrix. In addition, we develop a two stages estimate (TSE) combining CMLE and CMMLE for the location parameters estimation and the noise spectral parameters estimation, respectively.
An outline of the paper follows. In section 2, the data model and the parameter structure are introduced. The criterion of the CMLE is developed in section 3. The estimation of variance components via CMMLE is investigated in section 4. The LSF's and the TSE are described in section 5 and section 6, respectively. In section 7, results of numerical experiments are presented. We conclude with some remarks.

2. DATA MODEL

A conventional model is used. Sources m = 1,...,M generate signals that are transmitted by a wavefield. The wavefield has known properties of propagation except for some parameters. As in sonar, the outputs of the sensors n = 1,...,N are Fourier transformed with a smooth, normalized window of length T. For a frequency ω of interest, we get data $\mathbf{X}^k(\omega) = (X_1^k(\omega),...,X_N^k(\omega))'$ of k = 1,...,K successive pieces of sensor outputs, similar to the radar baseband data. Correspondingly, $\mathbf{S}^k(\omega) = (S_1^k(\omega),...,S_M^k(\omega))'$ denotes the Fourier transformed signals received at the origin. The array output is assumed to be a zero-mean stationary vector process. The propagation-reception conditions for signals can be described by a (NxM) matrix $H(\omega)$ with the elements $H(\omega)_{nm} = \exp(j\omega\tau_{nm})$ (n=1,...,N; m=1,...,M), where τ_{nm} is the time delay of the m-th signal in the n-th sensor. The columns of $H(\omega)$ are known as the steering vectors $\mathbf{d}_m$ (m=1,...,M). At the sensors, the signals are measured additively disturbed by noise. The (NxN) spectral matrix $C_x(\omega)$ of the array output is assumed to be

$$C_x(\omega) = H(\omega)C_s(\omega)H(\omega)^* + C_u(\omega), \tag{1}$$

where $C_s(\omega)$ is the (MxM) spectral matrix of the signals, $C_u(\omega)$ is the (NxN) spectral matrix of noise and * denotes the hermitian operation. Let us apply the well known asymptotic properties of the Fourier transformed array output for a large window length T:

i) $X^1(\omega), \dots, X^K(\omega)$ are independent and identically complex-normally distributed random vectors with zero-mean and covariance matrix $C_x(\omega)$ as in (1).

ii) Given $S^1(\omega), \dots, S^K(\omega)$, the $X^1(\omega), \dots, X^K(\omega)$ are independent and identically complex-normally distributed random vectors with mean $H(\omega)S^k(\omega)$ and covariance matrix $C_u(\omega)$.

Now, we have to specify the parameters. Because we have fixed the frequency ω, we omit its notation in the sequel. The wave parameters are described by the vector ξ, and we can write $H = H(\xi)$. For spheric waves, ξ summarizes, e.g., bearings and ranges of the sources. The spectral parameters of the signals are given by the entries of C_s. The spectral matrix C_u is defined by

$$C_u = \sum_{i=0}^{L} \nu_i J_i = \nu_0\left[I + \sum_{i=1}^{L} \mu_i J_i\right] = \nu_0 C_v, \text{ with } J_0 = I, \tag{2}$$

where $\nu = (\nu_0,\dots,\nu_L)'$ are the noise spectral parameters and $\mu = (\mu_1,\dots,\mu_L)'$ with $\mu_i = \nu_i/\nu_0$ for i=1,...,L. The J_i are supposed to be known nonnegatively hermitian matrices for angle band-limited noise of different directions and widths. For example, assuming farfield noise in the plane, known angle spectra $C_i(\alpha)$ and a half wavelength equispaced line array, the entries of J_i can be

$$J_{i\,nm} = \int_{-\pi}^{\pi} \exp[(n-m)\pi \sin\alpha]\, C_i(\alpha)\, d\alpha/(2\pi). \tag{3}$$

Thus, for i) the parameters of $C_x = C_x(\theta)$ are $\theta = (\mathrm{col}(C_s)',\zeta')'$ with $\zeta = (\xi',\nu')'$. For ii), the components of S^k are unknown. The S^k can be interpreted as parameters, and we get $\vartheta = (S^{1'},\dots,S^{K'},\xi',\nu_0,\mu')'$. Generally, the number M of sources is unknown in addition. We assume to know M with M<N in this study.

3. CONDITIONAL MAXIMUM LIKELIHOOD ESTIMATES

Properties i) and ii) permit to formulate approximate maximum likelihood estimates. The application of ii) characterizing a conditional distribution leads to the conditional log-likelihood function

$$l(\vartheta) = -N \log\nu_0 - \log\det C_v - \frac{1}{\nu_0 K}\sum_{k=1}^{K}(X^k - HS^k)^* C_v^{-1}(X^k - HS^k) \tag{4}$$

that has to be maximized over the parameters ϑ. We first optimize $l(\vartheta)$ over S^k and ν_0 while the parameters ξ and μ are fixed. The necessary conditions for S^k and ν_0 result in

$$\bar{S}^k = (H^* C_v^{-1} H)^{-1} H^* C_v^{-1} X^k \Big|_{\xi,\mu}, \qquad \bar{\nu}_0 = \mathrm{tr}[C_v^{-1}(I - P_v)\hat{C}_x]/N \Big|_{\xi}, \tag{5}$$

where $P_v = H(H^*C_v^{-1}H)^{-1}H^*C_v^{-1}$ and the data are collected to

$$\hat{C}_x = \frac{1}{K}\sum_{k=1}^{K} X^k X^{k*}. \quad (6)$$

An explicit solving for μ seems to be not feasible. Therefore, using (5) in (4) we get CMLE's by minimizing the criterion

$$q(\xi,\mu) = \frac{1}{N}\, \mathrm{logdet} C_v + \mathrm{logtr}[C_v^{-1}(I-P_v)\hat{C}_x], \quad (7)$$

iteratively over the wave parameters ξ and the parameters μ.

4. CONDITIONAL MARGINAL MAXIMUM LIKELIHOOD ESTIMATES

That CMLE's of variance components can be heavy biased is well known in statistics. To overcome this problem, already mentioned in [5], the less biased CMMLE is introduced. The CMMLE is performed by maximizing the conditional likelihood function based on the transformed data TX^k over the noise spectral parameters v, where T is any matrix orthogonal to H.

$$L(v) = -\, \mathrm{logdet}(T'C_uT) - \mathrm{tr}\{T(T'C_uT)^{-1}T'\hat{C}_x\} \quad (8)$$

Differentiation of (8) with respect to v and after some algebra, the CMMLE's are obtained by solving the nonlinear equation system

$$\mathrm{tr}\{C_u^{-1}(I-P_u)J_i\} = \mathrm{tr}\{C_u^{-1}(I-P_u)J_i\, C_u^{-1}(I-P_u)\hat{C}_x\}, \quad i=0,...L \quad (9)$$

which is independent of the chosen T. To solve (9) several techniques are known in numerical mathematics. For the case that one of the J_i is equal to the unit matrix, a special iteration procedure

$$v_i^{k+1} = v_i^k\, \mathrm{tr}\{C_u^{-1}(I-P_u)J_i\, C_u^{-1}(I-P_u)\hat{C}_x\}/\mathrm{tr}\{C_u^{-1}(I-P_u)J_i\}\Big|_{v_i^k} \quad (10)$$

has been proposed in [6]. This procedure has the interesting property that, if the inital v_i^0 are nonnegative and the J_i are nonnegative definite, the v_i^k stay nonnegative.

5. LEAST SQUARES FITS

A least squares fit of the model (1) to the sample spectral matrix $\hat{C}_x$ can be performed by a variety of different criteria. In this contribution, we present the following three,

$$q^a(\theta) = \mathrm{tr}[(C_x-\hat{C}_x)^2], \quad (11a)$$

$$q^b(\zeta) = \mathrm{tr}\{[(I-P)(\hat{C}_x-C_u)]^2\}, \quad (11b)$$

$$q^c(\theta) = \mathrm{tr}[(C_x\check{C}_x^{-1}-I)^2], \quad (11c)$$

where $P = H(H^*H)^{-1}H^*$ and $\check{C}_x^{-1} = (1-N/K)\,\hat{C}_x^{-1}$ with $E\check{C}_x^{-1} = C_x^{-1}$.
The idea of constructing LSF's is as follows. We first try to optimize $q^a(\theta)$, $q^b(\zeta)$ or $q^c(\theta)$ over the spectral parameters where the location parameters are fixed. If we get explicit solutions, we put them into the criterion and minimize over the location parameters.
Minimization of $q^a(\theta)$, $q^b(\zeta)$ and $q^c(\theta)$ over the spectral parameters without restrictions yields the explicit solutions, respectively,

ON LEAST SQUARES CRITERIA

$$\tilde{C}^a_s = (H^*H)^{-1}H^*(\hat{C}_x - C_u)H(H^*H)^{-1}\Big|_{\xi,v}, \quad \tilde{v}^a = A^{-1}b, \tag{12a}$$

$$A_{ij} = tr[(J_i - PJ_iP)J_j], \quad b_i = tr[(\hat{C}_x - P\hat{C}_xP)J_i],$$

$$\tilde{v}^b = A^{-1}b, \tag{12b}$$

$$A_{ij} = tr[(I-P)J_i(I-P)J_j], \quad b_i = tr[(I-P)J_i(I-P)\hat{C}_x],$$

$$\tilde{C}^c_s = (H^*\check{C}_x^{-1}H)^{-1}H^*\check{C}_x^{-1}(\check{C}_x - C_u)\check{C}_x^{-1}H(H^*\check{C}_x^{-1}H)^{-1}\Big|_{\xi,v}, \quad \tilde{v}^c = A^{-1}b, \tag{12c}$$

$$A_{ij} = tr[(J_i - P_xJ_iP_x^*)\check{C}_x^{-1}J_j\check{C}_x^{-1}], \quad b_i = tr[(I-P_x)J_i\check{C}_x^{-1}],$$

where $P_x = H(H^*\check{C}_x^{-1}H)^{-1}H^*\check{C}_x^{-1}$. After replacing the spectral parmeters in (11a) to (11c) by their corresponding estimates, we find estimates of the location parameters by minimizing

$$q^a(\theta) = tr\{[P(\hat{C}_x - \tilde{C}^a_u)P - (\hat{C}_x - \tilde{C}^a_u)]^2\}, \tag{13a}$$

$$q^b(\zeta) = tr\{[(I-P)(\hat{C}_x - \tilde{C}^b_u)]^2\}, \tag{13b}$$

$$q^c(\theta) = tr\{[(P_x(\check{C}_x - \tilde{C}^c_u)P_x - (\check{C}_x - \tilde{C}^c_u))\check{C}_x^{-1}]^2\}, \tag{13c}$$

iteratively, where $\tilde{C}^a_u$ is given by (2) if v is replaced by $\tilde{v}^a$ etc. The minimization of the criteria (13a) to (13c) as well as the minimization of criterion (7) require global searches and local optimization techniques.

6. TWO STAGES ESTIMATES

The goal to separat the wave parameters estimation from the spectral parameters estimation fails for the CMLE. In this case, a reduction of the computational burden can be achieved by a two-stages method different from the method described above. In [5] the TSE as an alternative to the CMLE has been proposed by heuritic arguments. Because both, CMLE and CMMLE belong to the same family of estimators an even better motivated TSE can be constructed by combining CMLE and CMMLE for source location estimation and noise spectral parameters estimation, respectively.

The procedure is as follows. Initially, the bearings and ranges are estimated by a simple procedure introduced in [7]. With these estimates and appropriate initial values of v, a CMMLE of the noise spectral parameters is achieved by (10). The resulting estimate $\tilde{C}_v$ of C_v is used for a CMLE of the bearings and ranges that minimizes

$$q(\xi) = tr[\tilde{C}_v^{-1}(I-P_v)\hat{C}_x], \tag{14}$$

if C_v^{-1} is replaced by $\tilde{C}_v^{-1}$ in (7). The spectral parameters could then be estimated again etc. These two stages allow a separated determination of wave parameters and spectral parameters as well as a recursive procedure. The minimization of criterion (14) is carried out similar to that indicated in the previous section. The convergence speed can be enhanced when performing only one iteration for each recursion.

7. NUMERICAL EXPERIMENTS

We investigate the precision, the stability and the common behaviour of the TSE in

comparison with the CMLE using the exact noise structure as well as the LSF's in comparison with the CMLE in the presence of sensor noise only. Several numerical experiments were performed, especially when resolution problems are expected.
Model (1) is used for a line array of 15 sensors spaced by a half wavelength in the plane. The matrices J_i of the noise model are determined by (3) with the angle spectra

$$C_i(\alpha) = \begin{cases} \pi^2/\gamma_i \cos[\pi/\gamma_i(\alpha-\varphi_i)], & \varphi_i - \gamma_i/2 \le \alpha \le \varphi_i + \gamma_i/2 \\ 0 & \text{otherwise} \end{cases},$$

where φ_i, γ_i denote the centre and the extent of the i-th noise profile, respectively.
In the experiments presented, three sources located approximately broadside generate uncorrelated signals, cf. figs. 1a to 3a. Unknown wave parameters are the bearings β_m and the ranges ρ_m (m=1,2,3). The following noise models,

1) sensor noise,
2) three-parameter noise model with $v_0=v_1=v_2$, cf. fig. 2a,
3) four-parameter noise model with $v_0=v_1=v_2=v_3$, cf. fig. 3a,

are used. The $\hat{C}_x$ are complex-Wishart distributed with K=20 degrees of freedom, except for the LSF (8b) with K=50. For each experiment, 2048 pseudo-random matrices and also all estimates are computed using scoring. In figs. 1 to 3 scatter diagrams show the results of the bearing estimates for the noise models, respectively. Crosses indicate the exact signal parameters. The signal-to-noise ratio (SNR) of a source is defined by $SNR_m = 10\log(NCs_{mm}/trC_u)$. Space limitation does not allow to present the range estimates.

8. CONCLUDING REMARKS

The estimates developed in this paper have been empirically investigated by a multitude of numerical experiments, especially when resolution problems are expected. The numerical experiments indicate that, although the LSF (11c) supplies asymptotic efficient estimates, the finite behaviour is even worse in comparison to CMLE and the other LSF's. Compared with the CMLE, the LSF (11c) uses approximately a 2.5 times higher number of degrees of freedom for the same accuracy. The accuracy and stability of the TSE combining CMLE and CMMLE are satisfactory and depend slightly on the parameter number of the noise model. The TSE well approximates the CMLE for known noise structure.

9. REFERENCES

[1] Koop L. and Bienvenu G.: Multiple detection using eigenvalues when the noise spatial coherency is partially unknown, in Urban H. ed.: Adaptive Methods in Acoustics, Reidel Publ. Comp., Dordrecht Holland, pp. 385-391, 1985.

[2] Paulraj A. and Kailath T.: Eigenstructure methods for direction of arrival estimation in the presence of unknown noise fields, IEEE Trans. ASSP, Vol.34, pp. 13-20, 1986.

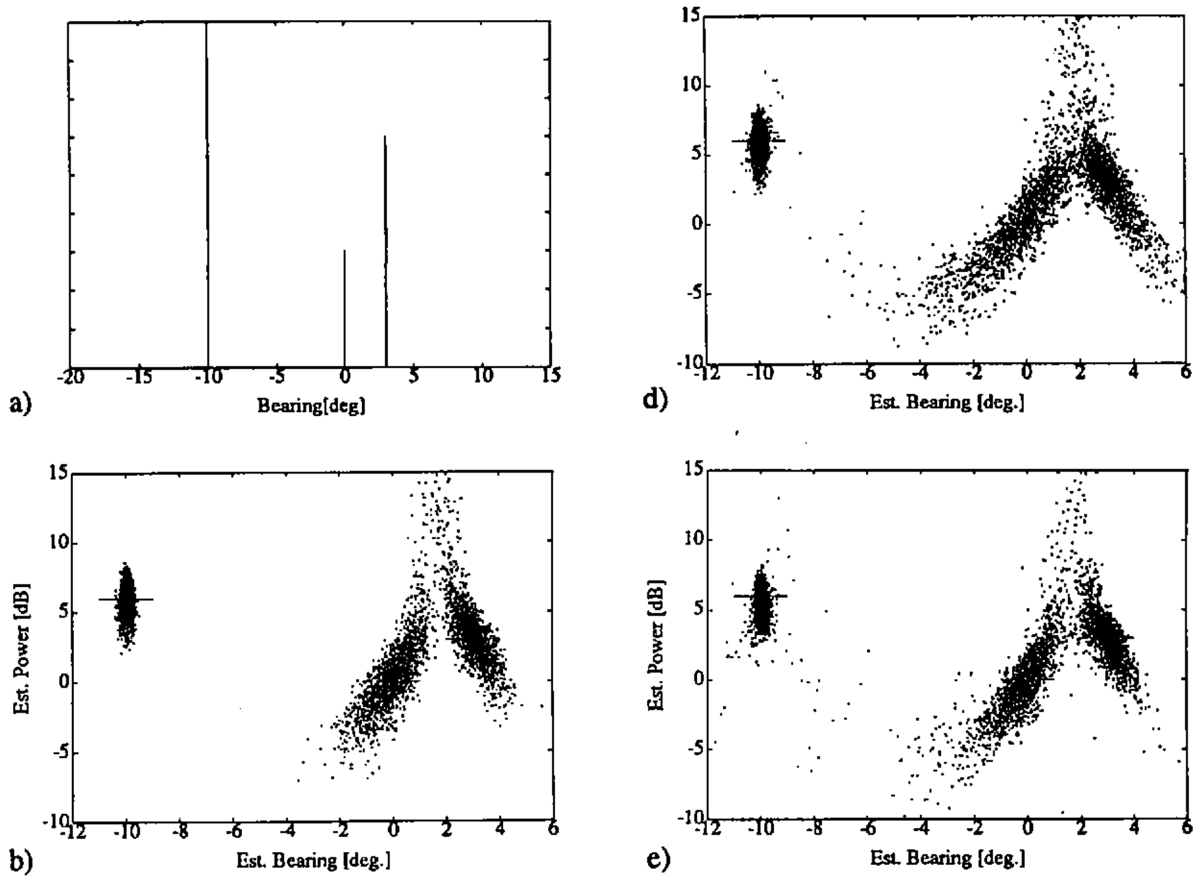
[3] Forster P. and Vezzosi G.: Approximate ML after spheroidal beamforming in the narrow band and wide band cases, Proc. EUSIPCO, Grenoble, pp. 307-310, 1988.

[4] Le Cadre J.P.: Parametric methods for spatial signal processing in the presence of unknown colored noise fields, IEEE Trans. ASSP, Vol. 37, pp. 965-983, 1989.

[5] Böhme J.F. and Kraus D.: On least squares methods for direction of arrival estimation in the presence of unknown noise fields, Proc. IEEE ICASSP, New York, pp. 2833-2836, 1988.

[6] Rao C.R. and Kleffe J.: Estimation of variance components, in Krishnaiah P.R. ed.: Handbooks of Statistics, Vol.1 , North-Holland Publ. Comp., Amsterdam, pp.1-40, 1980.

[7] Böhme J.F.: Source parameter estimation by approximate maximum likelihood and nonlinear regression, IEEE J. OE-10, pp. 206-212, 1985.



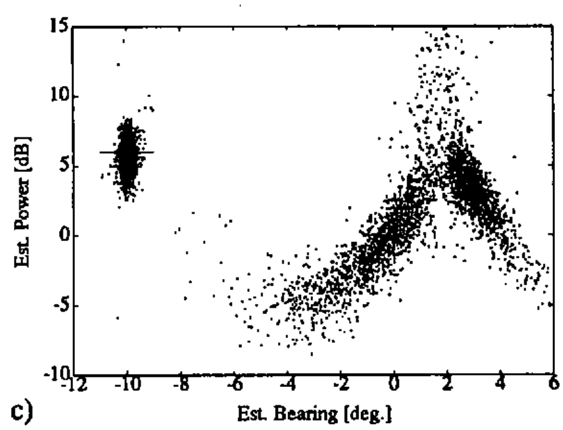
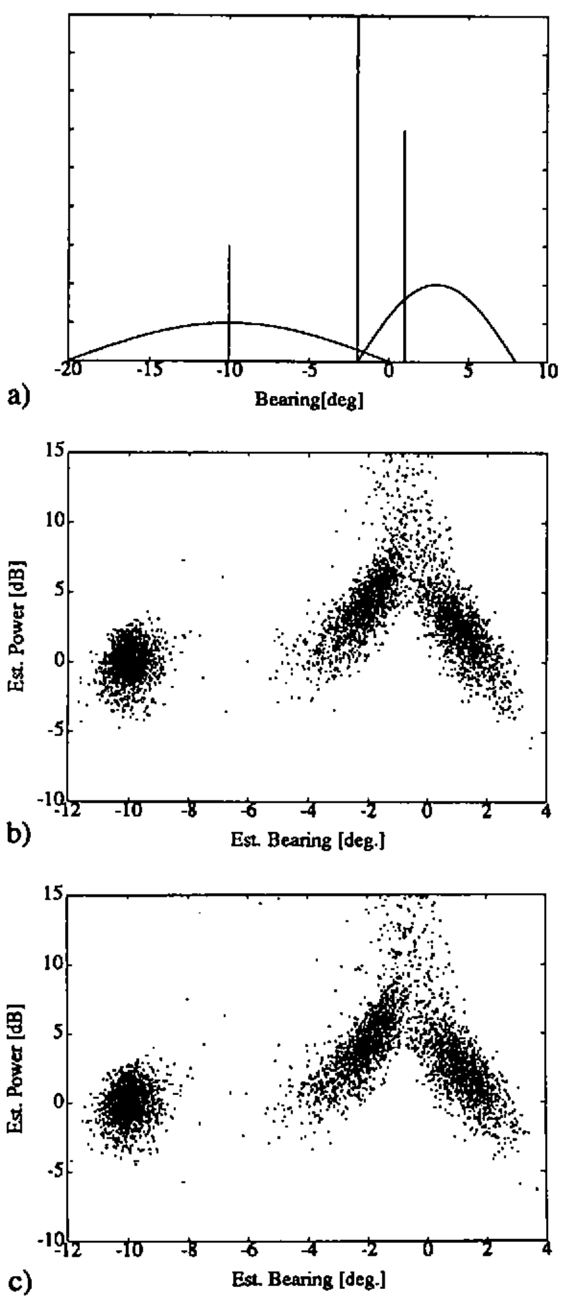
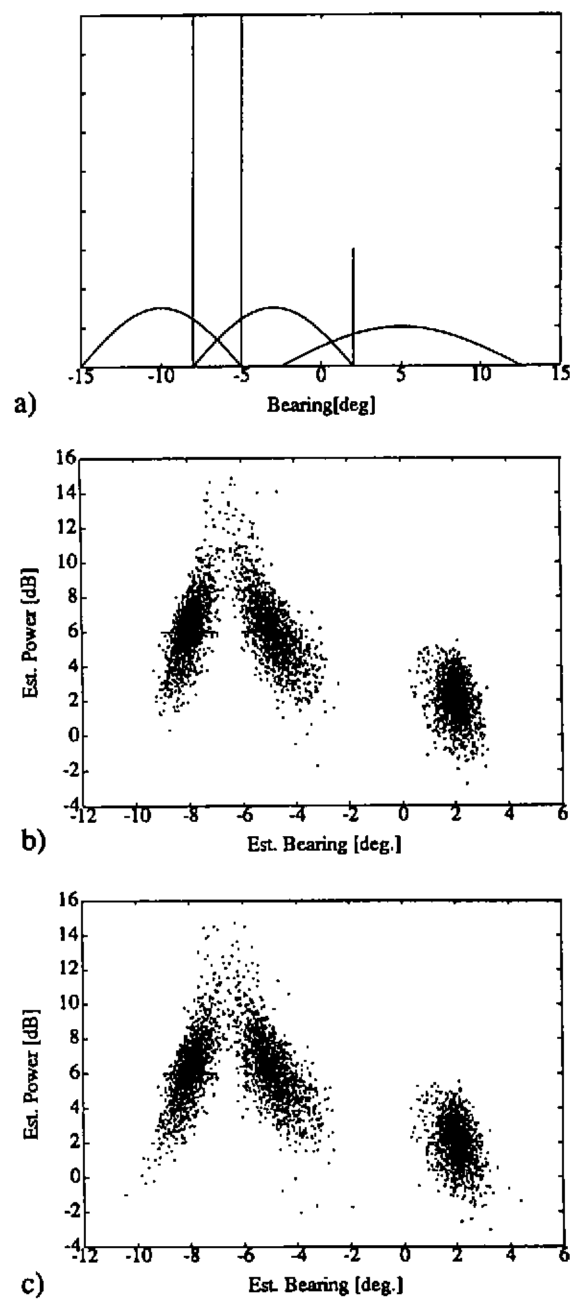


Fig. 1: Scatter diagrams of experiment 1, sensor noise only.
a) Signal configuration, qualitative.
b) CMLE, K=20 DOF.
c) LSF (11a), K=20 DOF.
d) LSF (11b), K=20 DOF.
e) LSF (11c), K=50 DOF.





<u>Fig. 2:</u> Scatter diagrams of experiment 2.

<u>Fig. 3:</u> Scatter diagrams of experiment 3.

a) Signal and noise configuration, qualitative.
b) CMLE using the exact noise structure, K=20 DOF.
c) TSE, K=20 DOF.