High-Resolution Signal and Noise Field Estimation Using the $L1$ (Least Absolute Values) Norm

CEDRIC A. ZALA, IAN BARRODALE, AND J. S. KENNEDY

(Invited Paper)

Abstract—**In this paper a new method for obtaining a quantitative estimate of an acoustic field consisting of a set of discrete sources and background noise is described. The method is based on the $L1$ (least absolute values) norm solution to an underdetermined system of linear equations defining the Fourier transform of the signal series. Implementations of the method with either equality or inequality constraints are presented and discussed. The much faster and more compact equality constraint version with a provision for modeling the noise field is recommended in practice. Experience with real data has shown the necessity of correcting for an observed Gaussian decay on the covariances. A simple means of estimating this effect and taking it into account in the signal estimation procedure is discussed, and the implications of this effect in high-resolution beamforming are considered. The effectiveness and versatility of the $L1$ method indicate that it has a useful role in high-resolution signal estimation.**

I. Introduction

A CENTRAL problem in array signal processing is the estimation of a set of discrete signals in the presence of background noise. In the classical approach to this problem, conventional beamforming is applied to a set of covariances derived from the outputs of an array of sensors to yield an estimate of the strength of the acoustic field in the direction of the beam. Since the method of conventional beamforming is based on the Fourier spatial power spectrum of the array, the resolution of discrete signals by this procedure is limited by the aperture of the array. Because of this, closely spaced discrete arrivals will be detected only as a single broadened peak in the beamforming output. In addition, a strong source may mask a weaker arrival of interest by sidelobe leakage. Thus it is desirable to consider alternatives to conventional beamforming which have higher resolution and lower levels of sidelobe interference.

The problem of obtaining direction estimates of a noise field is mathematically equivalent to estimating a spatial spectrum for the array. In recent years, many alternatives to the classical Fourier methods for spectral estimation have been developed [1]. The application of several of these methods to direction estimation, including maximum likelihood, linear prediction, eigenvector and maximum entropy algorithms, has been discussed by Johnson [2], who showed that each approach is the solution of a particular constrained optimization problem. Although these methods are often capable of extremely high resolution, they can suffer from certain serious practical drawbacks, such as difficulties in dealing with other than white-noise backgrounds, a requirement to choose the model order (i.e., assume a set number of signals), a decreased stability relative to conventional beamforming, and the absence of quantitative information in the noise field estimate (i.e., peak heights which do not correspond linearly to signal amplitudes). It is thus desirable to consider other high-resolution methods for spectral estimation which may reduce or eliminate these difficulties and which may be adapted to certain characteristics of real data. In general we seek a method of obtaining a quantitative estimate of a signal field consisting of relatively few plane-wave signals in a smoothly varying noise background.

The fundamental problem in high-resolution spectral estimation is nonuniqueness. In theory, an infinite number of acoustic fields exist which will account exactly for the given observations. The origin of this problem is the finite aperture and number of elements of the array, which result in the data being band limited at the sampling densities required to achieve high resolution. Any "solution" with appropriate characteristics in the nonzero spectral region of the data will then be consistent with the data, no matter what its spectral properties outside this region. Many of these solutions will be physically impossible and most will be unreasonable. To achieve meaningful high-resolution results, it is therefore necessary to introduce assumptions or constraints (maximum entropy, for example) so as to obtain a unique and physically acceptable solution.

High-resolution bearing estimation may also be formulated as a deconvolution problem. For a linear equispaced array, the output of conventional beamforming is the convolution of the unknown spatial signal field with a band-limited beam function. The deconvolution problem with band-limited data is notoriously ill-posed: a unique solution cannot be obtained in the absence of further constraints.

One method of performing the deconvolution is frequency-domain division of the beamforming output by the beam function for the spectral region where the beam function is nonzero. However, the resulting values are equal, within a scaling factor, to the original set of covariances and thus this simple approach to deconvolution does not yield an increase in resolution.

Manuscript received April 1, 1986; revised September 16, 1986.

C. A. Zala and I. Barrodale are with the Department of Computer Science, University of Victoria, Victoria, B.C., V8W 2Y2, Canada.

J. S. Kennedy is with the Defence Research Establishment Pacific, F.M.O., Victoria, B.C., V0S 1B0, Canada.

IEEE Log Number 8714345.

0364-9059/87/0100-0253$01.00

An acceptable method for high-resolution signal estimation in this application, i.e., for a relatively small number of discrete signals on a background of isotropic noise, should ideally be constrained so as to have the following properties:

a) it should produce a sparse, quantitative, and accurate estimate of the signal series—preferably, the simplest which can account for the data;
b) it should be stable to white noise and able to model or accommodate isotropic spherical and cylindrical noise;
c) it should produce only physically reasonable results (e.g., nonnegative signal intensities);
d) it should not require assumptions about model order or number of signals; and
e) it should be computationally efficient.

Several deconvolution-based methods for signal estimation have been examined recently [3]. A method originally proposed by Levy and Fullagar [4], which is based on the $L1$ (least absolute values) norm, showed very promising results and satisfied most of the above specifications. This $L1$ norm-based procedure has also been successfully applied to spectral estimation of harmonic processes [5], [6] and "superresolution" of one-dimensional images [7], [8]. In the present paper the background and implementation of this approach is presented, using an efficient and compact algorithm which has been developed recently [9]. Modifications to the method are introduced to model white, spherical, and/or cylindrical noise (or any "noise" with a defined covariance) along with discrete signals. Some observations for real data indicated the frequent presence of a decay in the covariances. A simple correction to model and overcome this effect is presented.

II. Background to the $L1$ Norm Method

Consider a set of M omnidirectional sensors comprising a linear array with constant interelement separation d. Then for a particular frequency and assuming stationarity, the $M \times M$ covariance matrix $\underset{\sim}{R}$ is defined by

$$\underset{\sim}{R} = E[\underset{\sim}{X}\underset{\sim}{X}^*] \tag{1}$$

where $\underset{\sim}{X}$ is a complex M-element column vector consisting of the Fourier coefficients for the sensor outputs at the frequency of interest, and $E[\ ]$ denotes the expectation operator.

If the measured field consists of uncorrelated plane-wave signals and isotropic noise, $\underset{\sim}{R}$ may be modeled as

$$\underset{\sim}{R} = \underset{\sim}{Q} + \sum_{j=1}^{NS} p_j \underset{\sim}{v}_j \underset{\sim}{v}_j^* \tag{2}$$

where

$\underset{\sim}{Q}$ covariance matrix of the noise,
NS number of discrete signals,
p_j power of the jth signal,
$\underset{\sim}{v}_j$ direction vector for the jth signal, with elements defined by $v_{jm} = \exp(-i2\pi dm \cos\theta_j/\lambda) = \exp(-imkd \cos\theta_j)$,
θ_j arrival direction of the jth signal relative to the array axis,
λ wavelength corresponding to the particular frequency, and
k $= 2\pi/\lambda$ is the wavenumber.

The output of conventional beamforming applied to $\underset{\sim}{R}$ may be defined as

$$B(\omega_n) = \underset{\sim}{d}^*(\omega_n)\underset{\sim}{R}\underset{\sim}{d}(\omega_n) \tag{3}$$

where $\underset{\sim}{d}$ is a look-direction vector related to a physical direction ϕ_n by $d_m(\omega_n) = \exp(-im\omega_n)$ for $m = 0, 1, \cdots, M-1$ and $\omega_n = 2\pi d \cos\phi_n/\lambda$. In many applications, including the present one, it is convenient to perform beamforming such that ω consists of a set of Fourier frequencies, i.e., $\omega_n = 2\pi n/N$, $n = 0, 1, \cdots, N-1$, and N is the overall number of beams formed. Under these conditions the beam function is fixed and independent of arrival direction. The beamforming output is the convolution of this beam function with the measured field and it corresponds to a spectral estimate of the covariances throughout the entire range of Fourier frequencies. (Note, however, that for $d/\lambda < 0.5$ some beams will not correspond to any physical angles, while for $d/\lambda > 0.5$ some physical angles will be aliased.) The correspondence between the physical angles and Fourier angles is illustrated clearly in all figures displaying the results of beamforming and high-resolution signal estimation.

If the array covariances are modeled as consisting purely of a set of discrete plane-wave signals, then the relation between the two may be expressed in terms of $2M-1$ equations expressing an N-element Fourier transform evaluated at the first M Fourier frequencies as follows:

$$\left.\begin{aligned} \frac{1}{N}\sum_{k=0}^{N-1} s_k \cos(k\omega_j) &= \mathrm{Re}(r_j), \qquad 0 \le j \le M-1, \\ -\frac{1}{N}\sum_{k=0}^{N-1} s_k \sin(k\omega_j) &= \mathrm{Im}(r_j), \qquad 1 \le j \le M-1, \end{aligned}\right\} \tag{4}$$

where

M number of sensors in the array,
N number of points in the Fourier transform of s ($N \ge 2M$),
s unknown spatial signal series for which a solution is sought,
ω_j Fourier frequency defined by $2\pi j/N$, and
r_j complex covariance between sensors at a separation of j units.

This is an underdetermined system of linear equations for $N \ge 2M$, since r_j is only defined at the lowest M frequencies. It should also be noted that the imaginary part of the equation at $\omega_j = 0$ has been dropped since all the terms are zero.

A unique solution to (4) may be obtained by specifying that s be minimized in some norm. The $L2$ (least squares) solution may be shown to be equal to the inverse Fourier transform of $\underset{\sim}{r}$ as follows. Let $\underset{\sim}{F}$ represent the $M \times N$ matrix with elements

$F_{jk} = \exp[-i2\pi(j-1)(k-1)/N]$ for $1 \le j \le M$ and $1 \le k \le N$. Then (4) may be written in matrix form as

$$\underset{\sim}{F}\underset{\sim}{s} = \underset{\sim}{r}. \tag{5}$$

The solution which minimizes the $L2$ norm of $\underset{\sim}{s}$ is defined by $\underset{\sim}{s}_{(L2)} = \underset{\sim}{F}^*(\underset{\sim}{F}\underset{\sim}{F}^*)^{-1}\underset{\sim}{r}$. Since the rows of $\underset{\sim}{F}$ are orthogonal, $(\underset{\sim}{F}\underset{\sim}{F}^*)^{-1} = \underset{\sim}{I}^{-1} = \underset{\sim}{I}$, so that

$$\underset{\sim}{s}_{(L2)} = \underset{\sim}{F}^*\underset{\sim}{r} \tag{6}$$

i.e., the inverse Fourier transform of the covariance series r. This estimate of s is therefore band limited and consequently smooth and continuous, so it does not yield a high-resolution result.

The $L1$ (least absolute values) norm solution to any underdetermined set of equations, however, has the desirable property that the number of nonzero elements of the solution will be equal at most to the number of equations. In this case, the number of nonzero elements in s would be at most $2M - 1$. This will be the case no matter how many points N are considered for the purpose of resolving s (provided, of course, $N \ge 2M$). The use of the $L1$ norm therefore provides a means of obtaining a sparse quantitative spike series whose transform is exactly equal to the covariance within the spectral range for which the covariances are defined. In contrast to the $L2$ solution, the $L1$ solution for s is not band limited, i.e., its spectrum is not constrained to be zero outside the range for which r is defined. The $L1$ procedure is thus a form of spectral extrapolation technique, capable of yielding broadband high-resolution estimates from band-limited input.

It should be emphasized that the $L1$ norm method will always yield a high-resolution estimate of a signal series, whether the underlying series in fact consists of a few discrete signals or not, just as the $L2$ norm will always yield a smooth estimate. The methods thus differ greatly in their representation of the available information, and the choice of which procedure is appropriate will depend on the assumptions about the underlying structure of the data. For example, a flat signal field due to pure white noise may be perfectly modeled by a uniformly spaced series of discrete signals, even though this is not the physical case. However, while it is shown below how to adapt the $L1$ procedure to include the effects of white or other model noise while at the same time achieving high resolution of a signal series, such a modification applied to the $L2$ approach will still yield a low-resolution estimate for the signal series. It is unclear whether this joint signal–noise estimation may be realized with other high-resolution spectral estimation techniques. Thus the $L1$ procedure may possess a substantial advantage over existing methods for the resolution of signals in the presence of noise.

III. Implementation of the $L1$ Norm Method

Two approaches to implementing the $L1$ deconvolution method were developed and compared, one involving equality constraints and the other involving inequality constraints. For the equality constraint version, a new numerical algorithm for the $L1$ solution of an underdetermined system of equations was devised [9]. For the inequality constraint version, an existing general-purpose $L1$ algorithm [10] was employed.

A. Equality Constraints

The equality constraint version involves an $m \times n$ matrix $\underset{\sim}{A}$ (where $m = 2M - 1$ and $n = N$) and m-element vector $\underset{\sim}{b}$ for the matrix equation $\underset{\sim}{A}\underset{\sim}{s} = \underset{\sim}{b}$ used as input to the algorithm:

$$\overset{m}{\updownarrow}\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \cos\omega & \cos 2\omega & \cdots & \cos((n-1)\omega) \\ 0 & -\sin\omega & -\sin 2\omega & \cdots & -\sin((n-1)\omega) \\ 1 & \cos 2\omega & \cos 4\omega & \cdots & \cos(2(n-1)\omega) \\ 0 & -\sin 2\omega & -\sin 4\omega & \cdots & -\sin(2(n-1)\omega) \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & \cos\left(\left(\frac{m}{2}-1\right)\omega\right) & \cos\left(2\left(\frac{m}{2}-1\right)\omega\right) & \cdots & \cos\left(\left(\frac{m}{2}-1\right)(n-1)\omega\right) \\ 0 & -\sin\left(\left(\frac{m}{2}-1\right)\omega\right) & -\sin\left(2\left(\frac{m}{2}-1\right)\omega\right) & \cdots & -\sin\left(\left(\frac{m}{2}-1\right)(n-1)\omega\right) \end{bmatrix}$$

$$\longleftarrow n \longrightarrow$$

$$\begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ s_{n-1} \end{bmatrix} = \begin{bmatrix} r_0 \\ \mathrm{Re}\,(r_1) \\ \mathrm{Im}\,(r_1) \\ \mathrm{Re}\,(r_2) \\ \mathrm{Im}\,(r_2) \\ \cdot \\ \cdot \\ \cdot \\ \mathrm{Re}\,(r_{m/2-1}) \\ \mathrm{Im}\,(r_{m/2-1}) \end{bmatrix}.$$

Here each consecutive pair of rows expresses the n-element Fourier transform at a Fourier frequency ω_j, except the first row, for which frequency ($\omega_0 = 0$) the transform is real-valued only. Thus there is a maximum of $2M - 1$ rows in the matrix $\underset{\sim}{A}$. It is frequently necessary to use fewer than this number of rows in practice, because of the effects of uncertainties in the data, which generally become more pronounced at higher Fourier frequencies. These uncertainties may arise from several sources. Since r_j is generally computed by averaging the covariances for $M - j$ pairs of hydrophones, it is evident that the error in estimating r_j will increase with j. Also, the presence of a Gaussian decay in the covariances (Section V) reduces their amplitudes. It is possible to correct for this effect by estimating the parameter σ which defines the decay, but there will still be an uncertainty in the estimate for σ. This uncertainty will result in an increased relative error in the evaluation of the Gaussian decay as the Fourier frequency increases.

The issue of how to allow for uncertainties in the data has been considered by Levy and Fullagar [4]. Both empirical and statistically based procedures for limiting the number of equations to those with acceptable signal-to-noise ratios were suggested. It was found useful in the present study to be able to limit empirically the number of rows based on the product of the (real-valued) spectrum of the beam function and the value of the Gaussian function. The form of the spectrum of the beam function is $1 - j/M$, and that of the Gaussian is $\exp(-j^2/(2\sigma^2))$ for $j = 0, 1, \cdots, M - 1$. This formulation takes both the above sources of uncertainty into account. Thus rows would only be included in the matrix for those Fourier frequencies at which the product of these two functions was greater than some specified value. For the types of uncertainties in our data, a value in the range 0.05–0.10 was found to give reasonable results in practice. The final system of equations built up in this way was then solved, using the new underdetermined $L1$ algorithm [9].

B. Inequality Constraints

In the inequality constraint version of the method, the equations expressing the Fourier transform of the signal series become

$$\left.\begin{aligned} \left| \frac{1}{N}\sum_{k=0}^{N-1} s_k \cos(k\omega_j) - \mathrm{Re}(r_j) \right| &\le \epsilon_j, \qquad 0 \le j \le M-1 \\ \left| -\frac{1}{N}\sum_{k=0}^{N-1} s_k \sin(k\omega_j) - \mathrm{Im}(r_j) \right| &\le \epsilon_j, \qquad 1 \le j \le M-1 \end{aligned}\right\} \tag{7}$$

where ϵ_j are nonnegative Fourier frequency-dependent tolerances.

These tolerances may be set to accommodate the uncertainties noted above (where they would increase with j) and also to allow for other sources of error in the covariances. An example of the latter might be a smoothly varying background noise field of unknown covariance, where the uncertainties in r_j would increase with decreasing j.

The inequality constraint version was implemented using an existing algorithm for the $L1$ solution to a system of equations subject to linear constraints [10]. Although this algorithm does not lend itself to an efficient implementation for this problem, it does provide a means of evaluating the inequality constraint approach. The adaptation of the problem to this algorithm involved setting up an $(n + m) \times n$ matrix $\underset{\sim}{B}$ (where $m = 4M - 2$ and $n = N$) and vector $\underset{\sim}{c}$ with the following structure:

$$\underset{\sim}{B} = \begin{bmatrix} \underset{\sim}{I}_{n\times n} \\ \hline 1 & \cos\omega & \cos 2\omega & \cdots & \cos(n-1)\omega \\ -1 & -\cos\omega & -\cos 2\omega & \cdots & -\cos(n-1)\omega \\ 0 & -\sin\omega & -\sin 2\omega & \cdots & -\sin(n-1)\omega \\ 0 & \sin\omega & \sin 2\omega & \cdots & \sin(n-1)\omega \\ 1 & \cos 2\omega & \cos 4\omega & \cdots & \cos 2(n-1)\omega \\ -1 & -\cos 2\omega & -\cos 4\omega & \cdots & -\cos 2(n-1)\omega \\ 0 & -\sin 2\omega & -\sin 4\omega & \cdots & -\sin 2(n-1)\omega \\ 0 & \sin 2\omega & \sin 4\omega & \cdots & \sin 2(n-1)\omega \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \cos m\omega/4 & \cos 2m\omega/4 & \cdots & \cos(n-1)m\omega/4 \\ -1 & -\cos m\omega/4 & -\cos 2m\omega/4 & \cdots & -\cos(n-1)m\omega/4 \\ 0 & -\sin m\omega/4 & -\sin 2m\omega/4 & \cdots & -\sin(n-1)m\omega/4 \\ 0 & \sin m\omega/4 & \sin 2m\omega/4 & \cdots & \sin(n-1)m\omega/4 \end{bmatrix} \quad \underset{\sim}{c} = \begin{bmatrix} \underset{\sim}{0}_{n\times 1} \\ \hline \mathrm{Re}(r_1)+\epsilon_1 \\ -\mathrm{Re}(r_1)+\epsilon_1 \\ \mathrm{Im}(r_1)+\epsilon_1 \\ -\mathrm{Im}(r_1)+\epsilon_1 \\ \mathrm{Re}(r_2)+\epsilon_2 \\ -\mathrm{Re}(r_2)+\epsilon_2 \\ \mathrm{Im}(r_2)+\epsilon_2 \\ -\mathrm{Im}(r_2)+\epsilon_2 \\ \vdots \\ \mathrm{Re}(r_{m/4})+\epsilon_{m/4} \\ -\mathrm{Re}(r_{m/4})+\epsilon_{m/4} \\ \mathrm{Im}(r_{m/4})+\epsilon_{m/4} \\ -\mathrm{Im}(r_{m/4})+\epsilon_{m/4} \end{bmatrix}$$

(m rows, n columns)

In this formulation the minimization of the $L1$ norm of the spike series is achieved by setting up the identity matrix and zero vector above in the first n rows of $\underset{\sim}{B}$ and $\underset{\sim}{c}$, respectively. The inequality constraints are then specified in the remaining

m rows, each Fourier frequency generating four rows. The constraint at $\omega = 0$ (and involving r_0) is omitted here, as it was found that an effective way of dealing with white noise (the covariance of which is nonzero only at r_0) was to use the option available in this algorithm to specify nonnegativity constraints on the solution, while dropping the constraint at $\omega = 0$. Thus the algorithm would automatically adjust the baseline upwards as far as possible (i.e., to the white-noise level) without introducing negative signal powers into the solution.

In using this version, the tolerances ϵ_j must be specified according to some scheme appropriate to the uncertainties expected in the data. Again, Levy and Fullagar [4] discuss empirical and statistical methods for setting these tolerances. We devised separate empirical methods of defining ϵ_j for the two types of uncertainties discussed above. For the uncertainties which increased with ω_j, ϵ_j were set according to

$$\epsilon_j = r_0 c_1 / [(1 - j/M) \exp(-j^2/(2\sigma^2))]$$

where c_1 is some small positive constant. Only for values of j for which $\epsilon_j \leq r_0$ would the constraints at the corresponding Fourier frequencies be included in the matrix. For accommodating uncertainties in the covariances which decrease with increasing j (due, for example, to non-white background noise) ϵ_j were set according to $\epsilon_j = r_0 c_2 M/j$, where c_2 is another small positive constant.

C. Comparison of Algorithms Using Synthetic Data

When the full set of equations was used and, in the inequality constraint version, all tolerances were set to zero, both algorithms produced identical results for noise-free synthetic data. Fig. 1 shows an example of these results for a 32-element linear equispaced array with a 0.4 wavelength spacing in a signal field consisting of the 20 synthetic signals given in Table I. In every case when there were two signals under a single peak these were resolved by the algorithm. In addition, several small signals whose presence was not apparent from the beamforming output were successfully resolved. The resulting estimate is quantitative and highly accurate and demonstrates the potential of the method.

It is noted here that the estimate of the signal series given by either algorithm almost always consists of sets of adjacent pairs of nonzero signals separated by a series of zeros. This is to be expected for two reasons. First, an array of M sensors can detect at most $M - 1$ discrete signals, but will give rise to $2M - 1$ equations (in the equality constraint version). The observed grouping of the output in pairs will result in a maximum of $M - 1$ pairs (i.e., discrete signals), corresponding to the maximum number of signals. Second, it has been observed with synthetic data that a signal which corresponds exactly to a grid point will usually be estimated by a single spike at that point, but one occurring between two points will produce a pair of nonzero values at two adjacent points, the sum of which is equal to the true amplitude. For greater clarity, in all illustrations in this paper the amplitudes of pairs of adjacent spikes separated from other spikes have been combined and the sum rather than the individual amplitudes is shown.

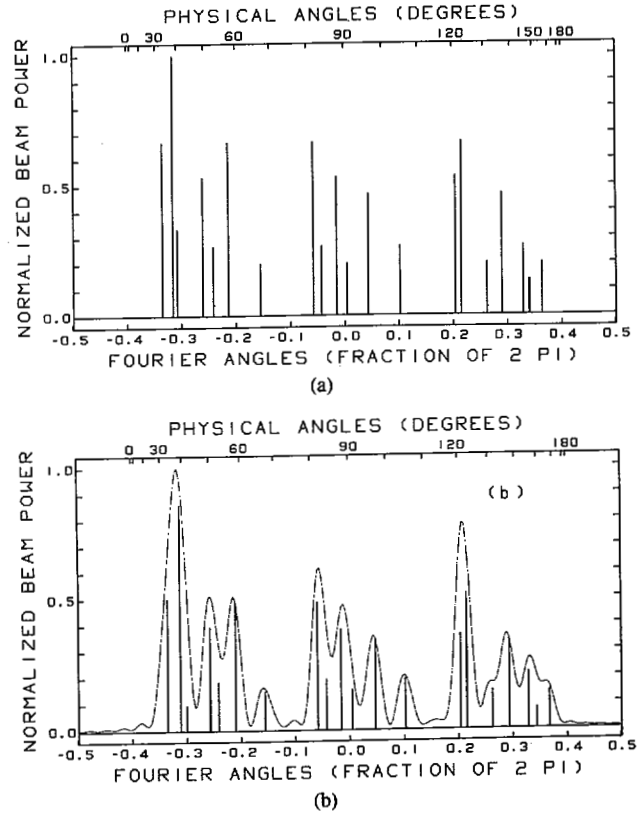


Fig. 1. (a) Synthetic signal field used in testing the algorithms; the corresponding numerical values are given in Table I. (b) Output from a 256-beam conventional beamformer (continuous curve) and the high-resolution signal estimate given by the $L1$ signal estimation procedure for a noise-free 32×32 covariance matrix generated using the synthetic series in (a). In this and subsequent plots of this form, the vertical axis is in linear units and the output has been normalized. The horizontal axis is linear in $\cos\theta$, where θ is the arrival direction relative to the array axis; under these conditions the beam function is independent of θ. The signals in the output of the $L1$ algorithms usually occurred as pairs of immediately adjacent spikes (see text). For greater clarity, a single spike corresponding to the sum of each pair is shown in the plots.

Ideally, for maximum accuracy in resolution, the widest possible spectral bandwidth should be employed. This would correspond to using as many equations as possible and, for the inequality constraint version, setting the tolerances to low values. As uncertainties in the data often prevent this goal from being achieved, a reduction in the number of rows or an increase in the tolerances is necessary, though at the expense of accuracy in resolution. Fig. 2 illustrates for the equality constraint version the effects on resolution of decreasing the number of rows based on the decay of the beam-function spectrum $(1 - j/M)$ to the level indicated. Fig. 3 shows the effect for the inequality constraint version of varying c_1 to increase the tolerances ϵ_j based on $\epsilon_j = r_0 c_1/(1 - j/M)$. In both cases a progressive loss of accuracy is observed, as expected, although the results are still of high accuracy for decay levels of up to 0.10 (for the equality version) or values for c_1 of up to 0.05 (for the inequality version).

The main advantage of the equality constraint version of the algorithm is its relative speed and compactness (each more than a factor of ten) compared to the current implementation of the inequality constraint version. The inequality constraint approach was originally developed to allow for the effects of

TABLE I
DESCRIPTION OF THE 20-ELEMENT SYNTHETIC SIGNAL SERIES
Fourier angles are computed assuming a sensor separation of 0.4 wavelength.

Signal Number	Physical Angle θ (degrees)	Fourier Angle (radians/2π)	Power
1	32.54	-.3372	1.0
2	37.81	-.3160	1.5
3	39.91	-.3068	0.5
4	49.36	-.2605	0.8
5	52.55	-.2432	0.4
6	57.65	-.2140	1.0
7	66.86	-.1572	0.3
8	81.55	-.0588	1.0
9	83.56	-.0449	0.4
10	87.76	-.0156	0.8
11	90.28	.0020	0.3
12	96.25	.0435	0.7
13	104.59	.1008	0.4
14	120.26	.2016	0.8
15	122.28	.2136	1.0
16	130.54	.2600	0.3
17	136.38	.2895	0.7
18	144.69	.3264	0.4
19	148.54	.3412	0.2
20	155.37	.3636	0.3

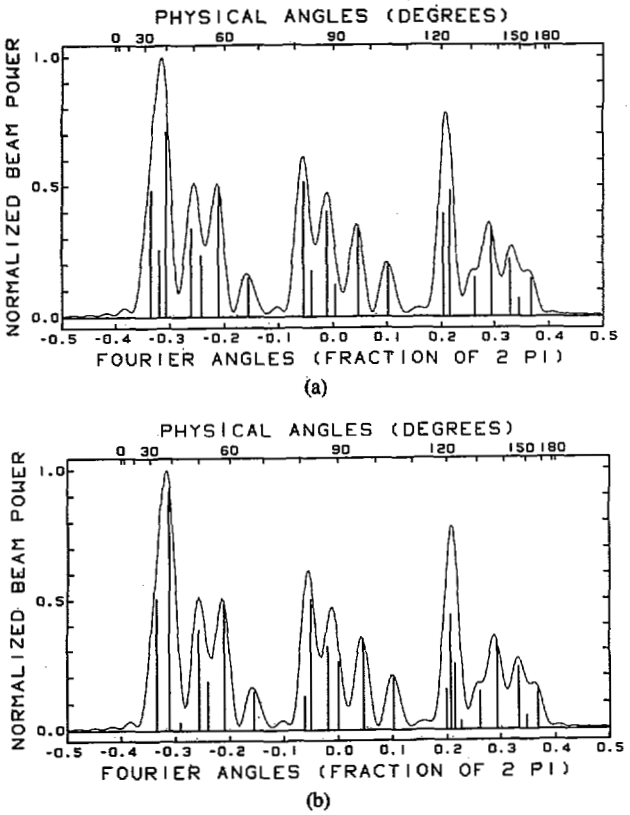


Fig. 2. Effect on signal resolution of reducing the number of rows (equations) corresponding to higher Fourier frequencies in the equality constraint version of the $L1$ signal estimation procedure. (a) Rows retained for which $1 - j/M > 0.1$. (b) Rows retained for which $1 - j/M > 0.2$.

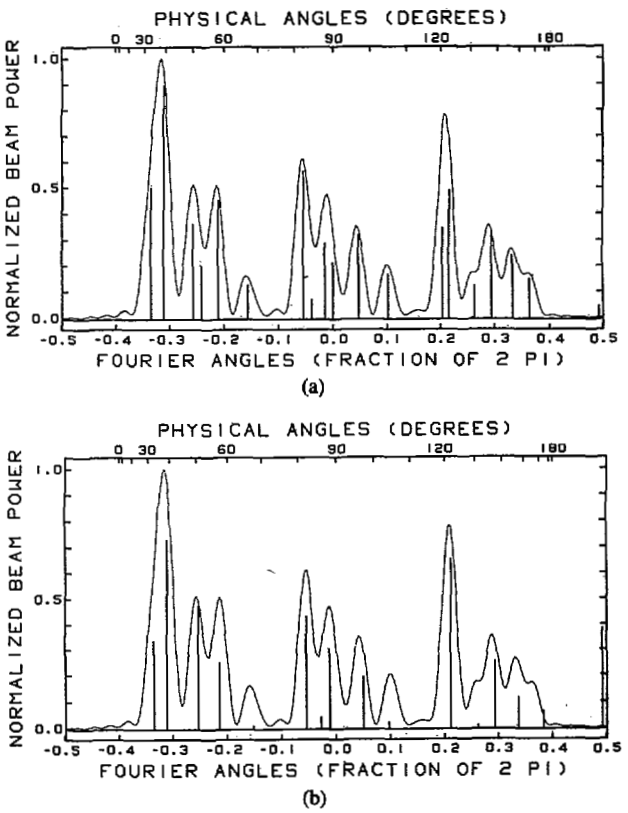


Fig. 3. Effect on signal resolution of increasing the tolerances ϵ_j ($=r_0 c_1/(1 - j/M)$) in the inequality constraint version of the $L1$ signal estimation procedure. (a) $c_1 = 0.02$. (b) $c_1 = 0.15$.

non-white noise and errors in the covariances and is of particular benefit when the spectral properties of these are not well defined. However, with the development, described below, of methods for including explicit noise models in the equality constraint version and for correcting for Gaussian decays in the covariance, the equality constraint version was found to be much more effective in practice.

IV. Adaptation to Isotropic Noise Fields

To be applied successfully to the analysis of real data, a high-resolution algorithm should be able to discriminate between discrete plane wave signals and isotropic or smoothly varying background noise fields. This may be done in the present case by considering the covariances due to certain types of model noise and incorporating their descriptions into the model.

A. Algorithmic Modification

The covariance at a sensor separation of j units due to a signal from direction θ has the form

$$r_j = p \exp[-ij2\pi d \cos\theta/\lambda]. \quad (8)$$

It is complex valued with constant modulus for all separations. The covariances q_j due to isotropic noise fields are real valued only and have a form which depends on the model for the field as follows:

a) white (uncorrelated) noise: $q_j^w = n_w\delta(j)$;

b) spherical noise: $q_j^s = n_s$ sinc (jkd); and

c) cylindrical noise: $q_j^c = n_c J_0(jkd)$ (in plane of array axis)

where n_w, n_s, and n_c are constants equal to the total power of the respective noise fields. The beamforming outputs for spherical and cylindrical noise are shown in Fig. 4 for typical conditions.

The equality constraint matrix $\underset{\sim}{A}$ and the vector of unknowns $\underset{\sim}{s}$ (but not the right-hand side $\underset{\sim}{b}$) may be modified as follows to include these relationships, where $\underset{\sim}{A}'$ and $\underset{\sim}{s}'$ are the modified arrays:

$$\underset{\sim}{A}' = \left[\begin{array}{c|ccc} & 1 & 1 & 1 \\ & 0 & \text{sinc } (kd) & J_0(kd) \\ \underset{\sim}{A} & 0 & 0 & 0 \\ & 0 & \text{sinc } (2kd) & J_0(2kd) \\ & 0 & 0 & 0 \\ & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot \end{array}\right].$$

$$\underset{\sim}{s}'^T = [\ \underset{\sim}{s}^T \ | \ n_w \quad n_s \quad n_c\ ]$$

This formulation then provides the potential for appropriate modeling of an acoustic field consisting of discrete signals and any combination of the above types of isotropic noise. However, since it is the ($L1$) norm of the vector $\underset{\sim}{s}'$ which is being minimized, it is necessary to provide a means of biasing the solution to include a nonzero value for such noise when it is present. In the absence of such biasing, the norms of the solutions will be the same, whether signals or noise was brought in to model the part of the covariances due to noise. This biasing may be accomplished easily by scaling the columns of $\underset{\sim}{A}'$ corresponding to noise by a factor slightly greater than unity. The final estimates for the power of the noise components are then obtained by multiplying the values found for n_w, n_s, and n_c by this scaling factor. Tests with synthetic and real data have established that any value between 1.1 and 1.5 is an appropriate scaling factor.

This approach may be generalized to include any noise field for which the covariances can be defined. Thus the potential exists to adapt the $L1$ method to a wide variety of noise environments, including background noise fields with measured rather than modeled covariances. These covariances need not necessarily be real valued only; they could also contain imaginary components and thus be used to model anisotropic background noise.

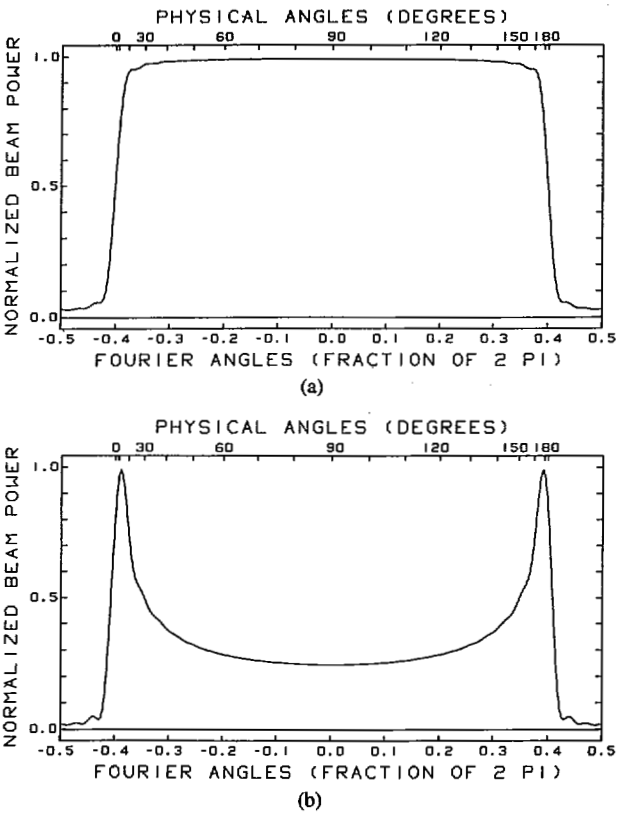


Fig. 4. Beamforming output for (a) spherical and (b) cylindrical noise models.

B. Performance

The above method was tested on synthetic data consisting of the 20-signal field described above and containing various combinations of isotropic white, spherical, and cylindrical noise. Representative results of these tests are shown in Fig. 5. The method yielded an excellent estimate of the signal field and accurately identified the type and powers of the model noise fields present. Having this capability is especially important in the case of cylindrical noise, which is concentrated into two signal-like peaks near the ends of the physical region of the beamforming output.

The effects of failing to model these types of noise when they are in fact present are shown in Fig. 6. Even for the simple case of white noise only, a series of incorrect discrete signals is brought in and there is a large loss of accuracy in the modeling of the true signals. As expected, for substantial levels of all three noise models, the results are so inaccurate as to be of little value. The necessity of taking these effects into account in real data analysis is thus clear.

The inequality constraint version was also applied to this purpose, where the tolerances were set according to $\epsilon_j = r_0 c_2 M/j$ (Section II-B), but the noise was not specifically modeled. The parameter c_2 was varied and the results compared with the true signal series. The best estimate found (for the case where all three types of noise were present) is shown in Fig. 7; it is considerably less accurate than that given

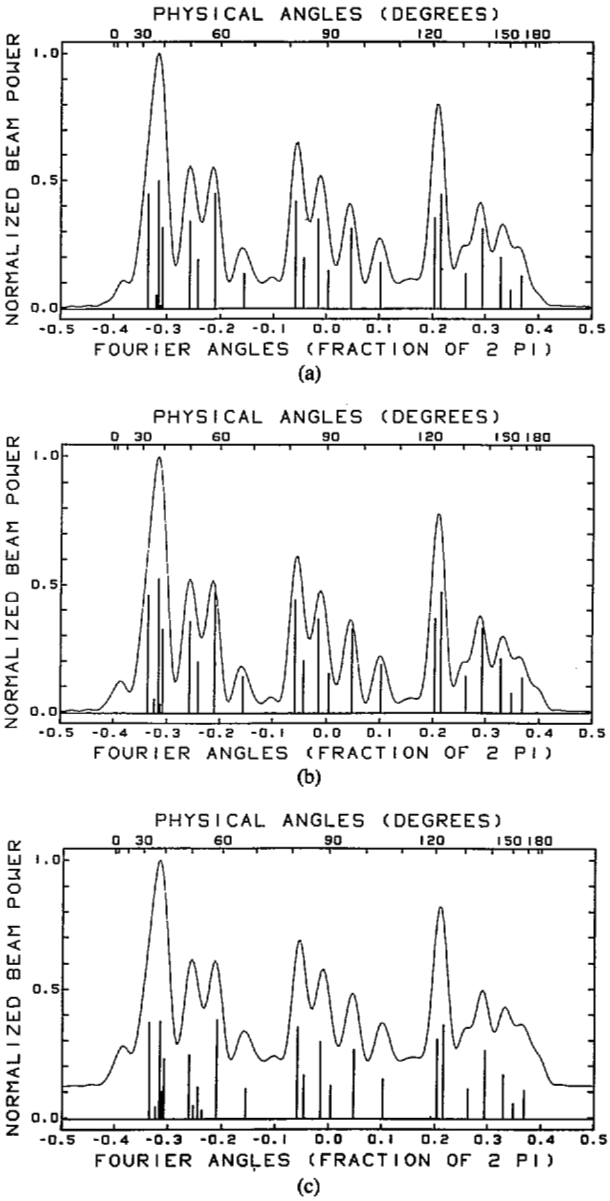


Fig. 5. Signal estimates given by the equality constraint version of the $L1$ signal and noise estimation procedure in the presence of model noise. Noise powers are expressed in the same relative units as the signal powers in Table I. (a) Spherical noise of power 4.883 units added: the output noise estimate was white noise 0, spherical noise 4.896 and cylindrical noise 0.008. (b) Cylindrical noise of power 1.953 added: the output noise estimate was white noise 0, spherical noise 0, and cylindrical noise 1.873. (c) White noise of power 9.766, spherical noise of power 4.883, and cylindrical noise of power 1.953 added: the output noise estimate was white noise 9.986, spherical noise 4.903, and cylindrical noise 2.018.

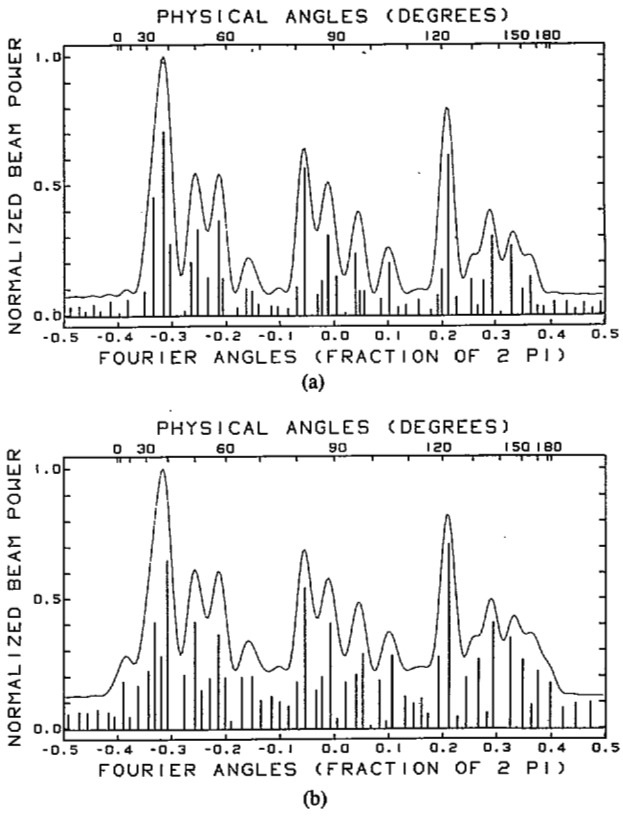


Fig. 6. Best signal estimates in the presence of model noise given by the $L1$ inequality constraint version without provision for noise estimation. (a) White noise of power 9.766 added. (b) White noise of power 9.766, spherical noise of power 4.883, and cylindrical noise of power 2.018 added.

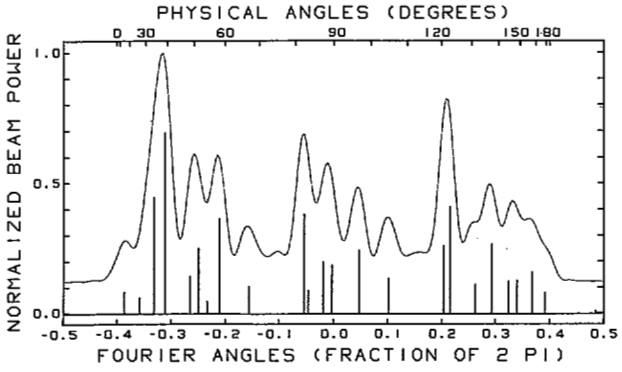


Fig. 7. Best signal estimate produced by the $L1$ inequality constraint version applied to the signal and noise field of Fig. 5(c) (white noise 9.766, spherical noise 4.883, and cylindrical noise 2.018).

by the modified equality constraint version. Because of this and its storage and speed limitations, the inequality constraint version is likely to be of practical use only for situations where the noise covariances do not correspond well with any model.

V. Correction of Decay on Covariances

In applying the $L1$ method to the covariance matrices of real data, it was observed that a much denser signal series than was physically reasonable would sometimes be produced. Closer inspection revealed that with these data sets the moduli of the covariances decayed with increasing sensor separation. For certain other real data sets and all synthetic data sets the moduli of the covariances did not decay noticeably. Typical examples of each of these two cases are shown in Fig. 8, for plots in the form of $\ln(|r_j|)$ versus j^2, which should be linear for a Gaussian model. In the second example (Fig. 8(b)) the covariances are adequately modeled by the Gaussian function r_j (observed) $= r_j$ (true)$\cdot \exp[-j^2/(2\sigma^2)]$. This distortion of covariances would explain the unreasonably dense signal series noted above.

It may be shown that under certain conditions phase fluctuation can introduce a Gaussian decay on covariances

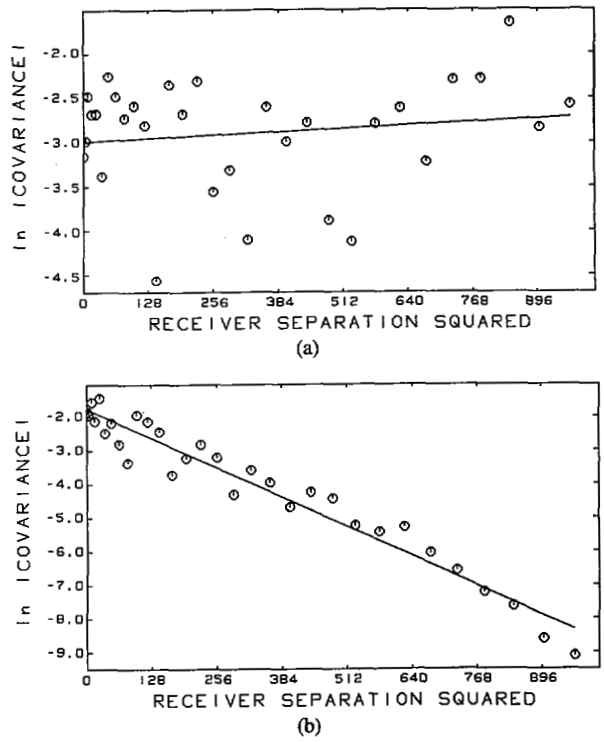


Fig. 8. Presence of a Gaussian decay of covariance in certain real data. (a) Covariance modulus versus the square of the receiver separation for a real data set not exhibiting a decay in covariance. (b) Corresponding plot for a data set where a substantial decay of covariance was present.

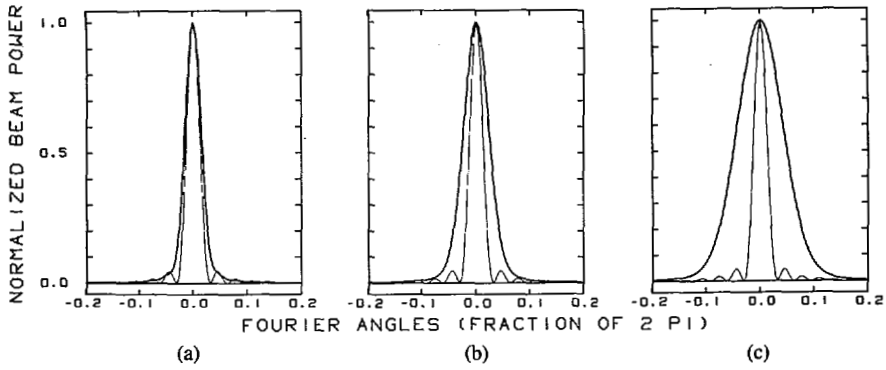


Fig. 9. Beam-function broadening and sidelobe smoothing in the presence of a Gaussian decay of covariance with receiver separation. Thin line: beam function in the absence of decay; bold line: smoothed and broadened beam function in the presence of a Gaussian decay with σ of (a) 16 sensor separation units, (b) 8 units, and (c) 4 units.

with increasing sensor separation, while not affecting the ratio of their real and imaginary parts. The term σ may thus be used to characterize the width of the Gaussian decay function in terms of units of sensor separation. It is noted that the assumption of amplitude rather than phase fluctuation does not lead to a decay, but simply to a mean value for each covariance.

The presence of such an effect will, if left uncorrected, have serious implications on high-resolution beamforming algorithms, which require that the modulus of the covariance for a given signal be independent of receiver separation. If input data are provided which contain a significant decay, then any algorithm, including the present $L1$ method, will yield a broadened estimate or multiple peaks when in fact there is only a single signal present. A second effect is that the amplitude of the conventional beamforming output will be underestimated.

Further consequences of the decay are an effective shortening of the array, a broadening of the beam function, and a smoothing of its sidelobes, resulting from the convolution of the ideal full aperture beam function with a Gaussian function. Fig. 9 illustrates this broadening, showing the ideal beam function and the effective beam function resulting from the presence of a Gaussian decay, for several values of σ, in the covariances. The necessity of correcting for the effect is evident in Fig. 10, which shows the signal series obtained by applying the $L1$ method to the uncorrected covariances due to

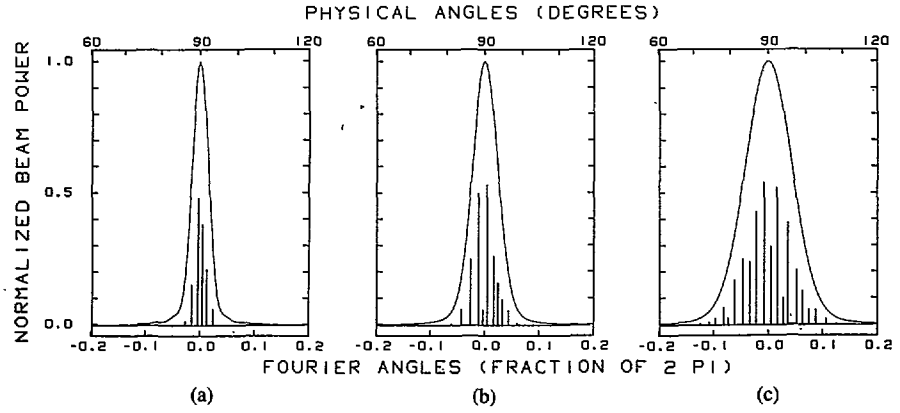


Fig. 10. Output of the $L1$ signal estimation method where a Gaussian decay on the covariances with σ of (a) 16 units, (b) 8 units, and (c) 4 units was present but not corrected for. A single signal is in fact present.

a single signal, under the same set of conditions as for Fig. 9. Several signals are required to model this effect, when in fact only a single signal is present.

We corrected for this Gaussian decay by simply estimating σ using a least squares linear fit to the data in the form $\ln |r_j|$ as a function of j^2. To reduce possible artifacts this fit did not include the covariance for $j = 0$, at which the spectral power for noise is largely concentrated. Confidence limits for the slope were computed and the Gaussian decay taken to be present only if the slope was nonzero with the required degree of confidence (e.g., 90 percent). The covariances were then corrected by division by $\exp(-j^2/(2\sigma^2))$.

The results of neglecting and correcting for Gaussian decay for a real data set are compared in Fig. 11. If the effect is neglected, a solution with many signals spaced at close intervals is obtained (Fig. 11(a)). If the Gaussian decay is estimated ($\sigma = 8.1$) and the correction is performed over the entire set of covariances without regard to the uncertainties in the data, a dense meaningless estimate is obtained (Fig. 11(b)). However, if the effect is corrected for, using the criterion of Section III-A to limit the number of rows in the matrix, a sparse physically reasonable estimate is obtained (Fig. 11(c)).

VI. Performance of $L1$ Method Using Real Data

A standard procedure was adopted for applying the $L1$ underdetermined method to high-resolution beamforming. Covariance matrices at each of several frequencies were computed from the output of a 32-element linear equispaced array. A value of σ for a Gaussian decay on the covariances was estimated, and when necessary, a correction was applied to the covariances. The equality constraint version of the $L1$ method was employed, with 256 points in the signal field and the provision for modeling white, spherical, and/or cylindrical noise outlined in Section IV.

The results of the $L1$ high-resolution algorithm for four sample data sets (with pairs of adjacent spikes combined) are shown in Fig. 12. The output of conventional beamforming applied to the data is plotted on the same graphs for comparison. Note that the discrete signal amplitudes are substantially greater than the output from conventional beamforming when a Gaussian decay is present (Fig. 12 (b)–(d)). This is due to the effect of the decay of covariance when

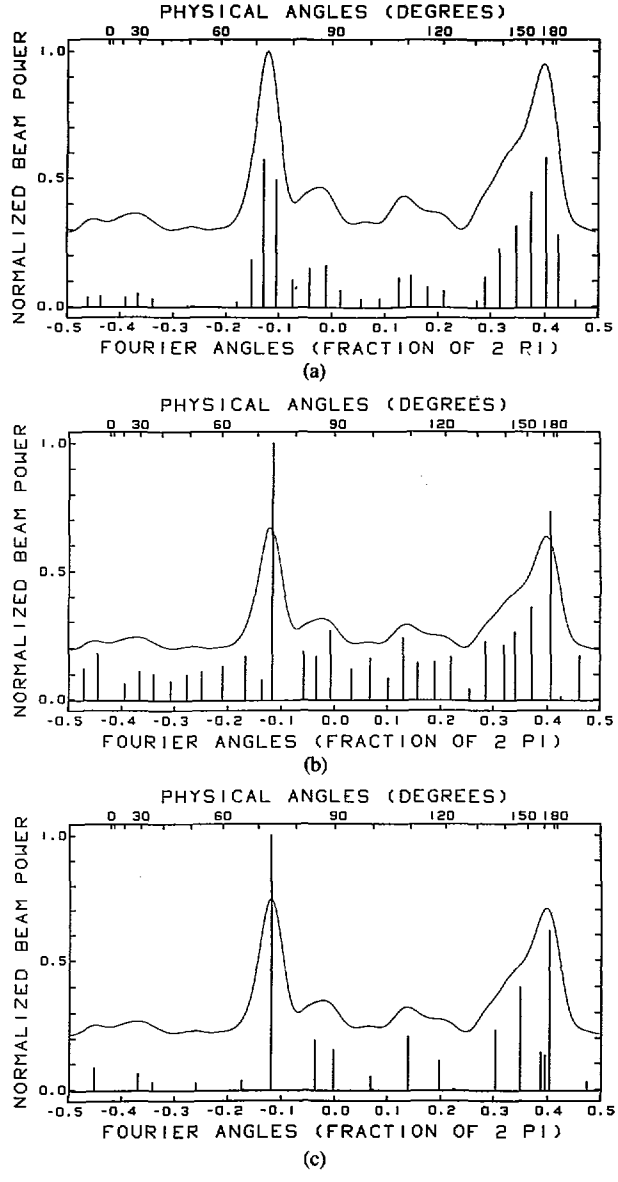


Fig. 11. Signal estimate given by the $L1$ signal and noise estimation procedure for a real data set exhibiting decay in the covariances. (a) No correction for decay applied and rows (equations) retained for which $1 - j/M > 0.05$. (b) Gaussian decay estimated ($\sigma = 8.65$) and corrected, but all rows 0 to $M - 1$ retained. (c) Gaussian decay corrected for and only rows retained for which $\exp(-j^2/2\sigma^2)(1 - j/M) > 0.05$.

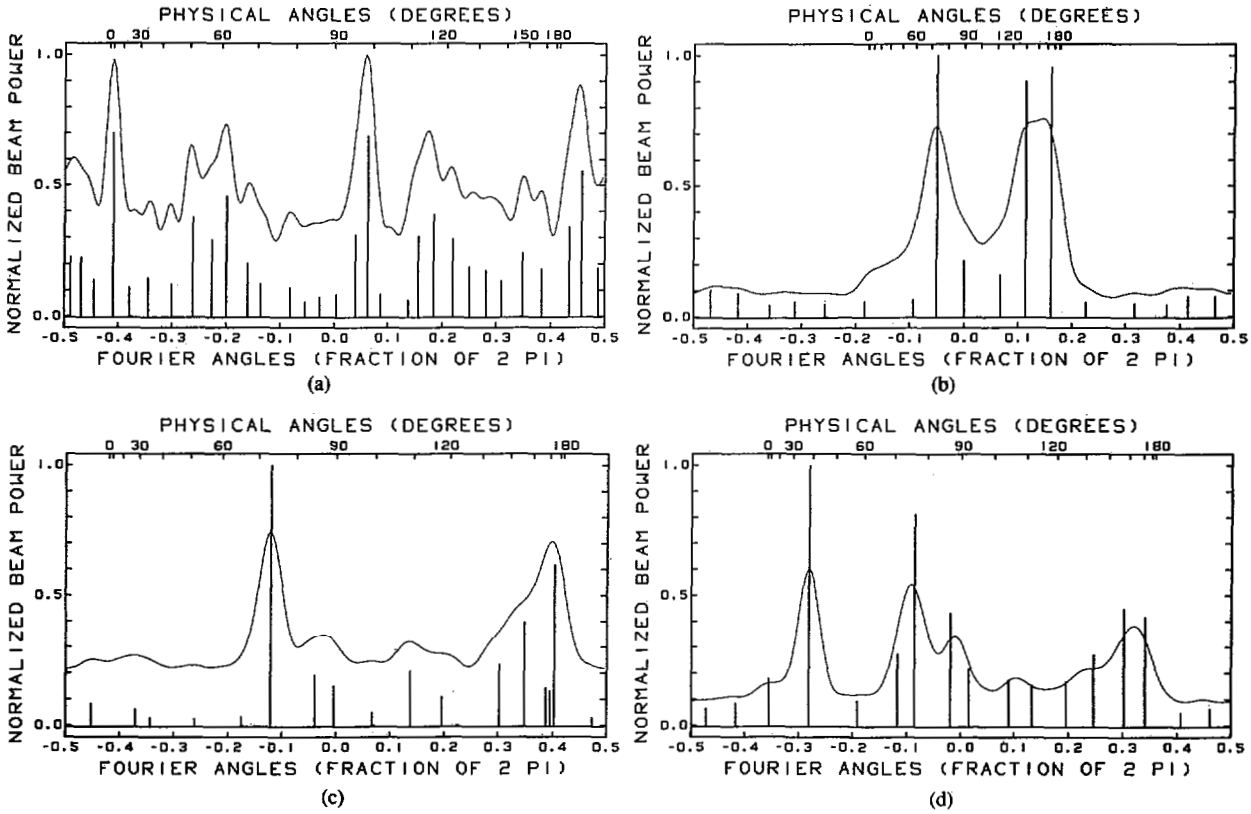


Fig. 12. Signal estimates given by the $L1$ signal and noise estimation method applied to several real data sets (32-element linear equispaced array). The noise powers below are expressed relative to the $L1$ norm of the output signal series. The data sets were tested for the presence of a Gaussian decay of covariance and corrected, and those rows for which the quantity exp $(-j^2/2\sigma^2)(1 - j/M)$ was greater than 0.05 were retained. (a) d/λ = 0.415. No Gaussian decay present; white noise power = 1.113, spherical noise power = 0, and cylindrical noise power = 0. (b) d/λ = 0.178. Gaussian σ = 8.33; white noise power = 0.460, spherical noise power = 0.386, and cylindrical noise power = 0. (c) d/λ = 0.294. Gaussian σ = 8.60; white noise power = 0.743, spherical noise power = 0.035, and cylindrical noise power = 0. (d) d/λ = 0.358. Gaussian σ = 8.55; white noise power = 0.388, spherical noise power = 0, and cylindrical noise power = 0.

conventional beamforming is applied to the uncorrected data. The levels (relative to the $L1$ norm of the signal series) of white, spherical, and cylindrical noise determined by the algorithm (using a scaling factor of 1.1) are given in the legend to Fig. 12.

The method produced physically reasonable high-resolution estimates of the measured field, with nonnegative amplitudes for all signal intensities being observed in practice under these conditions. This observation is especially important since the more efficient equality constraint algorithm used here does not provide for the specification of nonnegativity constraints on the solution. Most peaks in the beamforming output were resolved into either one or two discrete signals and in every data set a nonzero white-noise level was identified. In most data sets a nonzero value for either spherical or cylindrical noise was also observed.

As the threshold for limiting the number of rows in the matrix was increased (and the number of rows therefore decreased), several effects were observed. In general, progressively fewer spikes were extracted, as expected; spikes of lower amplitudes appeared to be lost preferentially. An increase in the relative power of the noise was also observed under these conditions. The representation of the discrete signals under a specific peak often changed as the number of rows was decreased, with several different distributions of signals accounting for the structure of the peak.

This decreased stability relative to conventional beamforming is common to all high-resolution methods of signal estimation. These methods are of necessity sensitive to small deviations of the covariances from their ideal values. The accuracy of the signal estimates produced by such methods depends largely on how closely the actual properties of the field resemble those of the model assumed by the methods. These assumptions usually involve stationarity, uncorrelated plane-wave signals, an upper limit to the number of signals, and well-defined signal and noise covariances. The current approach is also sensitive to deviations from these assumptions, but produces a signal estimate which takes into account background noise (by direct modeling), decaying covariance (by correction using a Gaussian model), and uncertainties in the covariances (by limiting the rows in the matrix to those with the smallest uncertainties).

The observed presence of small but significant spikes under relatively flat regions of the beamforming output raises the question of their physical validity. It is possible that they represent true signal arrivals, especially when they are seen to correspond to a shallow peak in the output of the conventional beamformer. An alternative explanation, however, is that they result from inexact modeling of the background noise field. In this connection, the algorithm has the advantage that it may easily be adapted to more sophisticated or realistic noise models, including anisotropic fields, for which the covariances contain imaginary as well as real components.

VII. Conclusions

The $L1$ high-resolution procedure appears to be an effective method for obtaining quantitative estimates for the discrete signal and smooth background noise components of an acoustic field as measured by a linear equispaced array. Signal and noise components may be modeled and estimated separately, based on their different covariances, and noise fields of any definable covariance are amenable to analysis by this procedure.

A decrease in the amplitudes of the covariances with increasing receiver separation is observed in some real data and may be modeled effectively by a Gaussian fit to the data. This decay must be corrected for if meaningful results are to be obtained from any high-resolution algorithm.

The equality constraint algorithm, with a provision for limiting the number of rows, is an effective means of obtaining a high-resolution estimate of an acoustic field. The inequality constraint version, while potentially better able to handle the types of uncertainties which may be present, would require a more efficient implementation to be of value in practical applications.

In tests with real data, the $L1$ equality constraint method produced quantitative and physically reasonable estimates of signal and noise fields and it is concluded that the procedure is a useful and potentially advantageous method for high-resolution acoustic field estimation.

Acknowledgment

The authors wish to thank Dr. F. Milinazzo and M. Greening for helpful discussions and criticism during the course of the work.

References

[1] S. M. Kay and S. L. Marple, Jr., "Spectrum analysis—A modern perspective," *Proc. IEEE*, vol. 69, pp. 1380–1419, 1981.

[2] D. H. Johnson, "The application of spectral estimation methods to bearing estimation problems," *Proc. IEEE*, vol. 70, pp. 1018–1028, 1982.

[3] C. A. Zala, I. Barrodale, and J. S. Kennedy, "Comparison of algorithms for high resolution deconvolution of array beamforming output," in *Proc. 14th Int. Symp. Acoustical Imaging*. New York: Plenum, 1985, pp. 699–702.

[4] S. Levy and P. K. Fullager, "Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution," *Geophysics*, vol. 46, pp. 1235–1243, 1981.

[5] S. Levy, C. Walker, T. J. Ulrych, and P. K. Fullagar, "A linear programming approach to the estimation of the power spectra of harmonic processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 675–679, 1982.

[6] G. Martinelli, G. Orlandi, and P. Burrascano, "Spectral estimation by repetitive L1-norm minimization," *Proc. IEEE*, vol. 74, pp. 523–524, 1986.

[7] R. J. Mammone and G. Eichmann, "Superresolving image restoration using linear programming," *Appl. Optics*, vol. 21, pp. 496–501, 1982.

[8] R. J. Mammone, "Spectral extrapolation of constrained signals," *J. Opt. Soc. Amer.*, vol. 73, pp. 1476–1480, 1983.

[9] I. Barrodale, M. Wilkie, and C. A. Zala, "A linear programming algorithm for solving underdetermined systems of equations in the L1 norm," Univ. Victoria, B.C., Canada, Tech. Rep. DCS-49-IR, 1985.

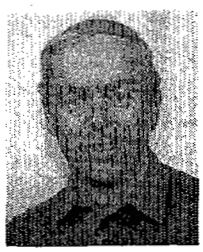
[10] I. Barrodale and F. K. D. Roberts, "Algorithm 552—Solution of the constrained L1 linear approximation problem [F4]," *ACM TOMS*, vol. 6, pp. 231–235, 1980.

★

Cedric A. Zala received the B.Sc. degree in biochemistry from the University of Victoria, Victoria, B.C., Canada, in 1971 and was awarded a Commonwealth Scholarship to study at the University of Manchester, Manchester, England, where he received the Ph.D. degree in biochemistry in 1974.

During several years of postdoctoral research into mechanisms of transmembrane glucose transport at the Jewish General Hospital in Montreal, Que., Canada, he became increasingly interested in computer applications, and in 1982 received a certificate in computer programming from McGill University, Montreal, Canada. Shortly afterward, he joined the staff of Barrodale Computing Services in Victoria, where he is engaged in signal processing software development. His special interest is in geophysical applications of the L_1 norm.

★



Ian Barrodale received the B.Sc. degree in mathematics from the University of Wales in 1960, the M.A. degree in mathematics from the University of British Columbia in 1965, and the Ph.D. degree in numerical analysis from the University of Liverpool, Liverpool, England, in 1967.

He is a Consultant and a part-time Professor at the University of Victoria, Victoria, B.C., Canada, where he has taught mathematics and computer science since 1961. He has held visiting positions as a Numerical Analyst at the Mathematics Research Center, Madison, WI, the Atomic Energy Research Establishment, Harwell, England, Liverpool University, and the Defense Research Establishment, Victoria. His research interests are in applied numerical analysis, operations research, and scientific programming. He has been an Editor of the *SIAM Journal on Numerical Analysis* and has served on advisory committees for the National Research Council and the Defense Research Board.

★

J. S. Kennedy received the B.E. degree in physics and the M.Sc. degree, in 1972 and 1975, respectively, both from the University of Saskatchewan, Saskatoon, Sask., Canada.

Since 1975, he has been working for the Canadian Department of National Defence at the Defence Research Establishment Pacific, Victoria, B.C., where he is currently Leader of the Ocean Acoustics Group. His primary interests are in the areas of signal and array processing for passive towed-array sonars, underwater ambient noise, and towed-array performance analysis.

Mr. Kennedy is a member of the Canadian Association of Physicists and the Acoustical Society of America.