

# On the Design of Maximally Sparse Beamforming Arrays

Richard M. Leahy, *Member, IEEE*, and Brian D. Jeffs, *Member, IEEE*

**Abstract**—The problem of designing nonuniformly spaced arrays is formulated as one of constrained optimization in which the cost function is chosen to select the array with the minimum number of elements. The response of the array is controlled by a set of inequality point response constraints. It is shown that a suitable cost function for this problem is the  $l_p$  quasi-norm for  $0 < p < 1$  and that there exists a number  $p_1$  such that for all  $0 < p < p_1$  the resulting array is maximally sparse. Furthermore it is shown that a solution to the problem lies at an extreme point of the simplex formed by the point constraints. A simplex search algorithm is described which will converge to a local minimum of the cost function on this simplex. The algorithm is illustrated in application to the design of sparse linear and planar arrays.

## I. INTRODUCTION

THE problem of designing antenna arrays in which the elements are not equally spaced has been studied widely over the last 30 years. A prime motivation for this work is the potential for improved resolution or reduced sidelobe levels compared to an equispaced array with the same number of elements. In this paper we describe a new method for designing sparse, unequally spaced arrays. This result is achieved by optimizing a nonlinear cost function over a set of linear inequality response constraints to select the minimum number of array elements, their locations and the gain factors necessary to meet a given response criterion. The use of an optimization theoretic approach is made possible by a simplex search algorithm based on an  $l_p$  quasi-norm which serves as a sparseness objective function.

Consider a  $P$  element arbitrarily shaped three-dimensional (3-D) array of matched omnidirectional sensors located at positions  $\mathbf{z}_i$ ,  $i = 1, 2, \dots, P$ . For a far-field, narrow-band source, with wavenumber  $k = 2\pi/\lambda$ , impinging on the array from direction  $(\theta, \psi)$  the array factor or beamformer spatial response is given by

$$F(\theta, \psi) = \sum_{i=1}^P A_i \exp \{jk(\mathbf{z}_i \cdot \mathbf{s}_{\theta, \psi})\} \quad (1)$$

where  $A_i$ ,  $i = 1, 2, \dots, N$ , are a set of complex weights, the unit vector  $\mathbf{s}_{\theta, \psi} = (\sin \psi \cos \theta, \sin \psi \sin \theta, \cos \psi)$  and  $(\cdot)$  denotes the vector inner product.

Manuscript received August 10, 1990; revised January 4, 1991.

R. M. Leahy is with the Signal and Image Processing Institute, Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, CA 90089-2564.

B. D. Jeffs is with the Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT 84602.  
IEEE Log Number 9100288.

The problem addressed in this paper is that of finding an optimal set of array weights,  $A_i$ , such that 1) the spatial response,  $F(\theta, \psi)$ , is held within a specified tolerance limit of a desired response,  $F_D(\theta, \psi)$ , and 2) the maximum possible number of weights,  $A_i$ , are zero valued. By removing the zero weighted elements from the array we then produce a maximally sparse array for the given set of initial positions,  $\mathbf{z}_i$ , and the desired response.

It will be shown that this problem is related to the nonlinear program

$$\min_{\mathbf{x} \geq 0} g(\mathbf{x}) = \sum_{i=1}^N |a_i|^p \quad \text{subject to } \mathbf{H}\mathbf{x} = \mathbf{b}$$

for some  $0 < p < 1$ , and a locally optimal solution may be found using the algorithm presented in Section IV.

The design of maximally sparse arrays belongs to the class of related nonuniformly spaced array problems. It has long been known that element placements and shadings exist that exhibit one or more of the following advantages over uniform arrays.

- 1) Significantly increased resolution (decreased main lobe width) when compared to a uniform array of the same number of elements [1]. The array aperture may be increased using nonuniformly spaced elements without generating the grating lobes found with evenly spaced elements at greater than  $\lambda/2$  separation.
- 2) Sidelobe level control by element placement in lieu of element shading [2]. For example, in applications where both maximum energy output and sidelobe control are desired for a transmitting array, each element may be operated at maximum power with the sidelobes controlled by element placement rather than "tapering" the array with reduced power in the outer elements.
- 3) Placement and shading of elements when physical constraints restrict potential sites [3]. For example, conformal arrays which must follow the surface of a supporting body, such as a ship's hull, and arrays in which minimum spacing is limited by mutual coupling or element size. In this case the most favorable element locations are often nonintuitive and unaided placement can lead to using far more elements than necessary.

To place the algorithm described in this paper in context we will briefly review the alternative approaches to the design of nonuniformly spaced arrays. However, it is not our intention

here to give a comprehensive review of the many fine articles on this subject. For example, see [1]–[14], [16].

Many early attempts at designing nonuniform arrays were based on series expansions of the array factor. Equating the truncated series with the desired response, one may then solve for element gain [4] or location [5]. Others have addressed the problem using spatial tapering of the array elements [1], [2] to reduce sidelobe levels. While these methods were successful in achieving nonuniform designs which exhibit favorable properties in comparison to equispaced arrays, the optimality of the resulting designs is doubtful [6]. An alternative approach based on random element placement was investigated by Lo [7] and others [8]. The array patterns are selected as samples from a probability distribution which exhibit optimal statistical properties. However, any particular realization of the array will not in general share these properties. Furthermore, this technique is generally applied to uniformly illuminated arrays and is not easily modified to allow variable illumination.

More closely related to our technique are the various optimization schemes which have been proposed. The general form involves a solution of the problem:

$$\min_{\alpha \in \Omega, z \in Z} f(F(\theta, \psi)_{\alpha, z}, \alpha, z)$$

where  $\alpha$  denotes the vector of complex weights,  $z$  the vector of element positions,  $\Omega$  and  $Z$  the respective feasible sets and  $f$  is some function of the array response parameterized in terms of  $\alpha$  and  $z$ . Typically  $f$  is either an  $l_2$  or  $l_\infty$  measure on the difference between the desired and actual array factor summed over a finite set of point responses. Direct minimization of this functional over all possible values of  $\alpha$  and  $z$  is infeasible. Consequently, the set of potential element positions  $Z$  is normally restricted to a finite set of points.

Skolnik *et al.* [9] use dynamic programming to search a grid of potential element placements for a uniformly illuminated, linear symmetric array with the objective of minimizing peak sidelobe level. This work was extended by Bratkovic [10] who used a combination of analytical and dynamic programming techniques, with the constraint of monotonically increasing element spacing away from the array center, to achieve the minimax solution with guaranteed optimality. Redlich [11] and others describe algorithms for the design of nonuniform arrays which minimize the  $l_2$  error between the desired and actual array factors, using the derivative of the cost function. The major deficiency of these and the many related optimization techniques is that, with a few exceptions, the number of elements in the array is fixed. Thus, they may not be used directly for the design of optimally sparse arrays.

Mitrou [12] and Kishi *et al.* [13] have investigated minimum squared error array synthesis by adding one element at a time and reoptimizing the weights and/or positions with each iteration. By repositioning each element and reweighting at each iteration, Mitrou's technique is guaranteed to converge to a local minimum of the cost as a function of element position, but no global measure of sparseness is applied. A problem with this approach is that the  $l_2$  cost function used here does not permit precise control of maximum sidelobe level or main lobe beamwidth. In many practical data inde-

pendent beamforming design problems these are the parameters of greatest interest, and  $l_\infty$  optimization is used to reduce the peak sidelobe level [15], [30]. Peak sidelobe level control is particularly important in sparse array design because of the potential for narrow, large, grating lobes when spacing is greater than one half-wavelength [16]. An alternative approach due to Jarske *et al.* [14] begins with a uniform array and with each iteration removes the pair of elements which after weighting optimization would leave the lowest peak sidelobe level. In this case, however, one must optimize the response for each potential configuration and the method becomes impractical for large arrays.

In summary, while a large number of techniques exist for the design of nonuniformly spaced arrays, we have found no computationally feasible method for designing optimally sparse arrays in which both the element location and gain may be simultaneously varied to meet a desired response. In Sections II, III, and IV we present a problem formulation and propose an algorithm which minimizes a cost function that directly reflects the degree of sparseness of the array. The relationship between the desired and actual response is controlled by a set of point constraints.

## II. PROBLEM FORMULATION

The response of the array given in (1) can be controlled by choosing the order  $P$  and the pairs  $(A_i, z_i)$ ,  $i = 1, 2, \dots, P$  to meet a given desired response  $F_D(\theta, \psi)$  within a specified tolerance  $\pm \epsilon(\theta, \psi)$ . These parameters remain as variables in the method described here, however we adopt the following design constraints in order to achieve computational tractability.

1) The response is controlled using a finite set of point constraints:

$$\begin{aligned} |(F(\theta_l, \psi_m) - F_D(\theta_l, \psi_m))| &\leq \epsilon(\theta_l, \psi_m), \\ l &= 1, 2, \dots, M_1, \quad M = M_1 M_2, \\ m &= 1, 2, \dots, M_2. \end{aligned} \quad (2)$$

Placement and density of these point constraints is arbitrary. However, a sufficiently dense grid of response samples should be used in the critical regions to control sidelobe leakage between points. Some techniques used for frequency constraint points in FIR filter design are applicable here [17].

2) The element locations are confined to a large finite set of candidate locations  $z_i$ ,  $i = 1, 2, \dots, K$  where  $K$  is chosen to be considerably larger than the expected order of the sparse array. For the remainder of this paper,  $K$  is assumed even with  $N = K/2$ . The extension to  $K$  odd is accomplished trivially by representing the center element as two co-located elements.

3) The element locations are constrained to be symmetric about their center point and the associated weights to be conjugate symmetric. This ensures  $F(\theta, \psi)$  will be real thus simplifying evaluation of (2).<sup>1</sup>

$$\begin{aligned} z_i &= -z_{2N-i+1}, \quad i = 1, 2, \dots, N \\ A_i &= A_{2N-i+1}^*, \quad i = 1, 2, \dots, N. \end{aligned} \quad (3)$$

<sup>1</sup>  $A^*$  denotes the complex conjugate of  $A$ .

4) The array is presteered by element phasing for maximum response in a desired direction,  $\mathbf{s}_0$ . This leaves only a single real unknown weight,  $a_i$ , at each element:

$$\mathbf{A}_i = a_i \exp \{ -jk(\mathbf{z}_i \cdot \mathbf{s}_0) \}, \quad i = 1, 2, \dots, N \quad (4)$$

where  $a_i$  are a set of real coefficients.

Substituting (3) and (4) in (1):

$$F(\theta, \psi) = \sum_{i=1}^N 2a_i \cos [k(\mathbf{z}_i \cdot (\mathbf{s}_{\theta, \psi} - \mathbf{s}_0))]. \quad (5)$$

Let  $\mathbf{a}^T = (a_1, a_2, \dots, a_N) \in \mathbf{R}^N$  denote the unknown weight vector. The constraints in (2) may then be written in matrix vector form:

$$|\mathbf{F}\mathbf{a} - \mathbf{f}_D| \leq \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon}, \mathbf{f}_D \in \mathbf{R}^M, \mathbf{a} \in \mathbf{R}^N, \mathbf{F} \in \mathbf{R}^{M \times N} \quad (6)$$

where  $\mathbf{f}_D = [F_d(\theta_1, \psi_1) \cdots F_d(\theta_M, \psi_M)]^T$  is the desired response vector,  $\boldsymbol{\epsilon} = [\epsilon(\theta_1, \psi_1) \cdots \epsilon(\theta_M, \psi_M)]^T$  is the tolerance vector, and the matrix  $\mathbf{F}$  has elements  $F(m, n) = 2 \cos [k(\mathbf{z}_n \cdot (\mathbf{s}_{\theta, \psi_m} - \mathbf{s}_0))]$ ,  $m = 1, 2, \dots, M$ ,  $n = 1, 2, \dots, N$ , where  $\mathbf{s}_{\theta, \psi_m}$  denotes the direction vector for the  $m$ th point constraint.

The set  $\Omega$  of vectors  $\mathbf{a}$  which satisfies the constraints in (6) form a potentially large set of feasible solutions. To achieve the desired sparseness in the design we adopt a cost function which favors sparse arrays and optimize this function over the feasible set  $\Omega$ . The appropriate cost function for this problem is the sum of indicator functions on each element of the weight vector,  $a_i$ ,  $i = 1, \dots, N$ , since minimization of this functional will result in the minimum number of nonzero weights. Since any zero weight implies that the corresponding array element is redundant, this optimization yields a maximally sparse array.

The problem may now be precisely stated as the following nonlinear program:

$$\min_{\mathbf{a}} f(\mathbf{a}) = \sum_{i=1}^N 1(a_i), \quad \text{subject to } \mathbf{a} \in \Omega \quad (7)$$

where  $1(a_i)$  denotes the indicator function on  $a_i$  defined as

$$1(a_i) = \begin{cases} 0, & \text{if } a_i = 0 \\ 1, & \text{otherwise.} \end{cases}$$

It should be noted that additional linear constraints other than directional response limits may be added to (6) without changing the form of the optimization problem. For example, we have applied a nonnegativity constraint,  $a_i \geq 0$ ,  $i = 1, 2, \dots, N$  to improve stability of the solution, and a constraint on the maximum allowable weight magnitudes,  $|a_i| \leq c$ ,  $i = 1, 2, \dots, N$ , to control the array response to independent element noise.

### III. MAXIMALLY SPARSE OPTIMIZATION

The nonconvex optimization problem (7) is unattractive as the basis of an algorithm because it is particularly difficult to solve. The function  $f(\mathbf{a})$  is discontinuous and has zero gradient except at the discontinuities. The discrete nature of the indicator function yields no information as to potential direction of movement from one point in the solution space to one

of lower cost since  $f(\mathbf{a})$  is constant until a solution with a different number of zero terms is encountered. Furthermore, accurate computational evaluation of the functional is difficult due to finite machine precision as elements of  $\mathbf{a}$  approach zero. In this section we consider the relationship between maximally sparse optimization and generalized  $l_p$  optimization for  $0 < p < 1$  and show the conditions under which a solution to the computationally simpler  $l_p$  optimization problem leads to a solution of (7).

Consider the generalized linearly constrained  $l_p$  optimization problem:

$$\min_{\mathbf{a}} g(\mathbf{a}) = \sum_{i=1}^N |a_i|^p, \quad \text{subject to } \mathbf{a} \in \Omega \quad (8)$$

for values of  $p$  in the range ( $0 < p < 1$ ). To show how this problem is related to (7) consider the unit ball surfaces in  $\mathbf{R}^2$  for the quasi-norm  $\|\mathbf{a}\|_p = [\sum_{i=1}^N |a_i|^p]^{1/p} = g(\mathbf{a})^{1/p}$  for values of  $p$  in the range  $0 \leq p \leq \infty$  as illustrated in Fig. 1. For  $p \geq 1$  we have the conventional  $l_p$  norm, which is a convex functional and obeys the triangle inequality. Since the linear constraints in (2) form a convex set it is well known that any local minimum of  $\|\mathbf{a}\|_p$  satisfying the constraints is a global optimum. Many efficient algorithms exist for solving such problems [18], [19]. Of particular interest are the cases for values of  $p = 1, 2$ , and  $\infty$ , which form the basis of many widely used optimization procedures. However, the resulting solutions for these do not achieve the "sparse" results of interest in this paper.

For  $0 < p < 1$ ,  $l_p$  is only a quasi-norm [20], since the triangle inequality does not hold, and in fact the inequality is reversed for positive  $a_i$ . Over  $\mathbf{R}^N$ ,  $\|\mathbf{a}\|_p$  is neither convex nor concave, containing many strong local minima and presenting a difficult optimization problem. Large values of  $p$  result in smooth solutions, however, as  $p \rightarrow 0$  the solutions tend to become more "spiked," or sparse [21]. The reason for this can be seen in Fig. 1. As  $p \rightarrow 0$ , the curves in Fig. 1 approach the  $x_1, x_2$  axes, on which the unit ball lies for  $f(\mathbf{a})$  in (7). We observe that

$$\lim_{p \rightarrow 0} |x|^p = \begin{cases} 1, & \text{for } x \neq 0 \\ 0, & \text{for } x = 0. \end{cases}$$

This suggests that we may identify minimum order optimization as a special case of generalized  $l_p$  optimization:

$$\lim_{p \rightarrow 0} g(\mathbf{a}) = \sum_{i=1}^N |a_i|^p = \sum_{i=1}^N 1(a_i) = f(\mathbf{a}). \quad (9)$$

The utility of this observation is that for  $p > 0$ ,  $g(\mathbf{a})$  eliminates some of the handicaps of  $f(\mathbf{a})$ ;  $g(\mathbf{a})$  is continuous everywhere and differentiable except at the axes.

If, however, we must allow  $p \rightarrow 0$  before (8) leads to a solution of (7), then we cannot benefit from the practical advantages of  $g(\mathbf{a})$  mentioned above. Theorem 4 (see Appendix) provides justification for minimum order optimization based on minimizing  $g(\mathbf{a})$  by demonstrating that for a bounded basic feasible solution set there exists a finite  $p_1 > 0$  such that for all  $0 < p < p_1$ , any solution to (8) is a solution to (7). The optimization problem in (8) therefore defines a

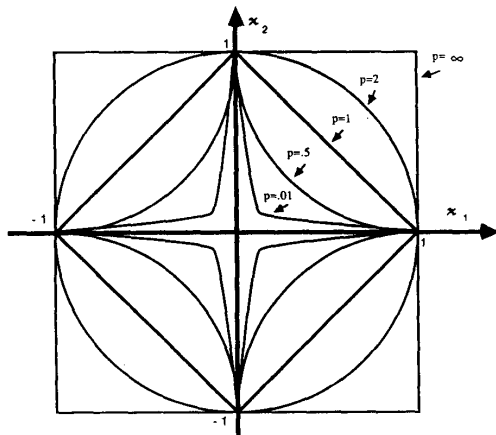


Fig. 1. Unit balls in  $R^2$  for the  $l_p$  quasinorm for various values of  $p$ . Note that as  $p \rightarrow 0$ , the unit ball collapses onto the  $x_1$  and  $x_2$  axes.

class of problems, indexed by  $p$ , whose global solutions are increasingly sparse as  $p$  decreases, until  $p < p_1$ , at which point an optimally sparse solution is given. For  $p \geq 1$  the optimal  $\mathbf{x}$  changes continuously as a function of  $p$ , but as  $p$  varies over the range  $0 < p < 1$  only a finite number of unique globally optimum solutions are encountered [22]. A given  $\mathbf{a}_{opt}$  will remain optimal over a range of  $p$  values, and as  $p$  decreases, we step from one solution to another in a discrete fashion [22].

It should be noted that since solutions to (7) are not necessarily unique, and lacking any justification for accepting one over another, we are satisfied with any algorithm which will select one from the optimal set. Theorem 4 proves that solutions to (8), for  $p < p_1$ , form a subset of solutions to (7), so we accept any  $l_p$  optimum provided  $0 < p < p_1$ . In order to improve the computational stability of an algorithm, we wish to use the largest value of  $p$  which reasonably approximates  $f(\mathbf{a})$ . The  $p_1$  as computed in Appendix I is a conservative upper bound, and we have found that in practice a much larger value may often be used.

Fig. 2 is an illustrative example of the effect of changing  $p$  on the minimum  $l_p$  solution to the following set of underdetermined linear equality constraints:

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0.2 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (10)$$

The solution space to this system is given by the line passing through the points  $(a_1, a_2, a_3) = (1, 0, 1)$  and  $(0, 10, 0)$ . The optimal solutions for various  $p$  lie along this line and have been computed and tabulated in Table I. It is interesting to note that for this example the bound  $p_1$  from Theorem 4 is correctly predicted with  $\Gamma = 10$ ,  $\Psi = 1$ , and  $r = 1$  to be  $p_1 = \log[(r + 1)/r] / \log[\Gamma/\Psi] = 0.3012$  since for  $0 < p < 0.3012$  we obtain the minimum-order solution.

Although to the best of our knowledge the above approach to minimum order optimization has not been previously suggested, a number of investigators have studied related problems. In [22] Barrowdale discusses the behavior of lin-

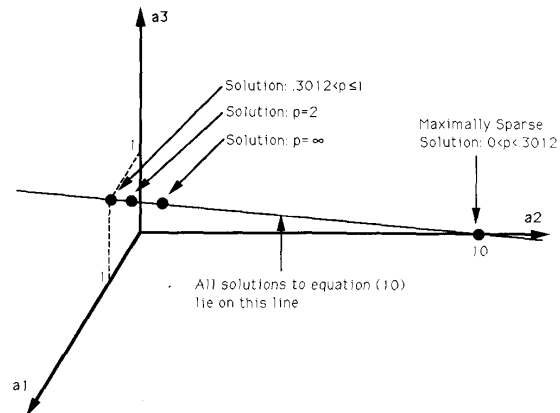


Fig. 2. An example of the effect of varying  $p$  in linearly constrained  $l_p$  optimization for the linear system defined in (10). All feasible solutions to the constraints in (10) lie on a line through  $R^3$ ; varying the value of  $p$  in the cost function selects different optimal solutions lying on this line. Optimal solutions for various values of  $p$  are given in Table I.

TABLE I  
OPTIMAL SOLUTIONS TO THE  $l_p$  OPTIMIZATION PROBLEM FOR THE LINEAR CONSTRAINTS IN (10). NOTE THAT FOR  $p > 1$ , THE OPTIMAL SOLUTION CHANGES CONTINUOUSLY WITH  $p$  BUT FOR  $0 < p < 1$  A SOLUTION REMAINS OPTIMAL FOR A FINITE RANGE OF  $p$

$p$ values	solution type	$x_1$	$x_2$	$x_3$
$p = \infty$	min $l_\infty$	0.91	0.91	0.91
$p = 2$	min $l_2$	0.98	0.20	0.98
$0.3012 \leq p \leq 1$	min $l_1$	1.00	0.00	1.00
$0 < p < 0.3012$	max sparse	0.00	10.00	0.00

early constrained  $l_p$  optimization for  $0 < p < 1$  and demonstrates the stepwise fashion in which the optimal solutions change with  $p$ . Pham and De Figueiredo discuss optimal estimation of sparse random distributions using  $l_p$  quasi-norm minimization [23]. As we shall see in Section IV, problem (8) is closely related to the problem of concave minimization over linear constraints. A number of techniques have been proposed for global optimization for problems of this type based on collapsing polytopes and branch and bound procedures [24]–[26]. These methods, however, make a number of assumptions not appropriate to our problem. In addition, due to the use of multiple nested linear programming subproblems, are probably computationally infeasible for the type of problem considered in this paper.

#### IV. A SIMPLEX SEARCH ALGORITHM

Having established the connection between maximally sparse optimization and  $l_p$  optimization, we now proceed to develop an algorithm for solving these problems. In Section IV-A three theorems are given which show that an optimal solution to (8) may be found using a simplex search. A suitable algorithm for solving (8) is therefore related to the well-known simplex method for linear programming (LP) [19]. The new algorithm differs from LP since the cost is nonlinear and must therefore be specifically computed rather than using a reduced cost row in the simplex tableau.

### A. Theorems

To facilitate the solution of (8) we borrow a well-known result from linear programming, namely that any set of linear inequality constraints in which the solution vector may contain both negative and positive values may be rewritten as a set of linear equality constraints, with a nonnegative solution vector, by the introduction of slack and surplus variables and doubling the length of the solution vector [19]. As shown in Theorems 1 and 2, the following two problems are equivalent for  $0 < p < 1$ :

$$\min_{\mathbf{a}} g(\mathbf{a}) = \sum_{i=1}^N |a_i|^p, \quad \text{subject to } |\mathbf{F}\mathbf{a} - \mathbf{f}_D| \leq \epsilon, \\ \epsilon, \mathbf{f}_D \in \mathbf{R}^M, \mathbf{a} \in \mathbf{R}^N, \mathbf{F} \in \mathbf{R}^{M \times N}, 0 < p < 1, \quad (11)$$

and

$$\min_{\mathbf{x}} h(\mathbf{x}) = \sum_{i=1}^{2N} (x_i)^p, \quad \text{subject to } \mathbf{H}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \\ \mathbf{H} = \begin{bmatrix} \mathbf{F} & -\mathbf{F} & \mathbf{I} & \mathbf{0} \\ \mathbf{F} & -\mathbf{F} & \mathbf{0} & -\mathbf{I} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} \mathbf{a}^+ \\ \mathbf{a}^- \\ \mathbf{s}^+ \\ \mathbf{s}^- \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} \mathbf{f}_D + \epsilon \\ \mathbf{f}_D - \epsilon \end{bmatrix}. \quad (12) \\ \mathbf{a}^+, \mathbf{a}^- \in \mathbf{R}^N, \quad \mathbf{s}^+, \mathbf{s}^- \in \mathbf{R}^M, \quad \mathbf{x} \in \mathbf{R}^{2N}, \quad \mathbf{b} \in \mathbf{R}^{2M}, \\ N' = 2N + 2M, \quad M' = 2M$$

and  $\mathbf{I}$  is the  $M \times M$  identity matrix.

The following statement and theorems show the equivalence between problems (11) and (12) and hence that an optimal solution to (11) can be found from an optimal solution to (12). These results are also the basis for the algorithm described in Section IV-B.

The following definitions are commonly used in the development of the simplex method for linear programming:

*Definition:* A *feasible solution* to problems (11) or (12) is any solution vector which satisfies the constraints.

*Definition:* A *basic feasible solution* to (12) is any  $\mathbf{x}$  which is feasible and has at most  $2M$  nonzero components. The following statement follows from the construction of (12).

*Statement:* If  $\mathbf{x}_0^T = [\mathbf{a}_0^{+T}, \mathbf{a}_0^{-T}, \mathbf{s}_0^{+T}, \mathbf{s}_0^{-T}]$  is a feasible solution to (12) then  $\mathbf{a}_0 = (\mathbf{a}_0^+ - \mathbf{a}_0^-)$  is a feasible solution to (11). Furthermore, if  $\mathbf{a}_0$  is a feasible solution to (11) then there exists an  $\mathbf{x}_0^T = [\mathbf{a}_0^{+T}, \mathbf{a}_0^{-T}, \mathbf{s}_0^{+T}, \mathbf{s}_0^{-T}]$ , such that  $(\mathbf{a}_0^+ - \mathbf{a}_0^-) = \mathbf{a}_0$  and  $\mathbf{a}_0^{+T} \mathbf{a}_0^- = 0$ , which is a feasible solution to (12).

Proofs of the following theorems are given in the Appendix.

*Theorem 1:* If a feasible solution to (12) exists then a basic feasible solution exists.

*Theorem 2:* If there exists a globally optimal feasible solution to (12) then there exists a basic feasible solution which is globally optimal.

*Theorem 3:* If  $\mathbf{x}_0^T = (\mathbf{a}_0^{+T}, \mathbf{a}_0^{-T}, \mathbf{s}_0^{+T}, \mathbf{s}_0^{-T})$  is a globally optimal basic feasible solution to (12) then  $\mathbf{a}_0 = (\mathbf{a}_0^+ - \mathbf{a}_0^-)$  is a globally optimal solution to (11).

Theorem 3 shows that the the solution to (11) may be

found from the solution to (12). Theorems 1 and 2 show that the search for a global optimum may be confined to the set of basic feasible solutions of (12). Since for linear equality constraints the basic feasible solutions correspond to the extreme points or vertices of a convex polytope [19], it then follows that a global optimum may be found using a finite search of these extreme points. Furthermore, the simplex method for linear programming gives us a well-known technique for searching these points using successive pivoting operations. The method differs from the linear programming simplex algorithm only in the form of the cost function, which in our case is nonlinear.

In summary, Theorems 1, 2, and 3 show that a global minimum of the  $l_p$  minimization problem (11) may be found by a simplex search of the basic feasible solutions associated with the linear constraint equation in (12). Theorem 4 shows that there exists a value of  $p$ ,  $0 < p < 1$ , such that a global minimum of the  $l_p$  problem (11) is also a global minimum of the sparse optimization problem (7), and hence a maximally sparse array meeting the desired response constraints may be found by a simplex search algorithm which solves (12).

### B. A Simplex Algorithm

The algorithm described in this section is a modification of the LP simplex algorithm. Emphasis is placed on the differences in the two algorithms. For a comprehensive description of the LP algorithm see [19], [28].

As in linear programming, we begin by forming a tableau for ease of notation and computation.  $\mathbf{H}$  may be partitioned as

$$[\mathbf{A} | \mathbf{D}] = \mathbf{H}, \quad \text{where } \mathbf{A} \in \mathbf{R}^{2M \times 2M}, \quad (13)$$

Multiplying by  $\mathbf{A}^\dagger$ , the left pseudo-inverse of  $\mathbf{A}$ , leads directly to the basic solution  $\mathbf{x}_B$

$$[\mathbf{I} | \mathbf{A}^\dagger \mathbf{D}] \mathbf{x} = \mathbf{A}^\dagger \mathbf{b}, \quad \mathbf{x}_B = [\mathbf{A}^\dagger \mathbf{b}, 0, \dots, 0]^T. \quad (14)$$

The tableau is then formed by augmenting the matrix with the right-hand side

$$\mathbf{Y} = [\mathbf{I} | \mathbf{A}^\dagger \mathbf{D} | \mathbf{A}^\dagger \mathbf{b}]. \quad (15)$$

To compute (15) a standard linear programming phase one procedure [19] may be used to find a nonnegative initial basic solution and corresponding tableau.

The variables  $x_i$  associated with columns of  $\mathbf{A}$  are termed basic variables and  $\mathbf{A}$  the basis. As in the standard LP simplex algorithm, an adjacent tableau is formed by moving (pivoting) one variable out of the basis and one nonbasic variable into the basis by swapping a column in  $\mathbf{A}$  with one in  $\mathbf{D}$ , adjusting  $\mathbf{x}$  indices, and updating  $\mathbf{A}^\dagger$ .

For the  $l_p$  simplex search algorithm, the reduced cost row of the LP tableau does not appear in  $\mathbf{Y}$  due to the nonlinear cost functional which must be explicitly computed at each potential adjacency before the best pivot is identified.

For the  $l_p$  simplex algorithm, we may view the set of tableaus as a connected graph with a node for each tableau satisfying (14) and (15).

Let  $\mathbf{S}$  = the set of all basic feasible solutions (BFS) to (12)

Let  $\mathbf{T}$  = the set of all tableaus,  $\mathbf{Y}^i$ , associated with  $\mathbf{S}$ .



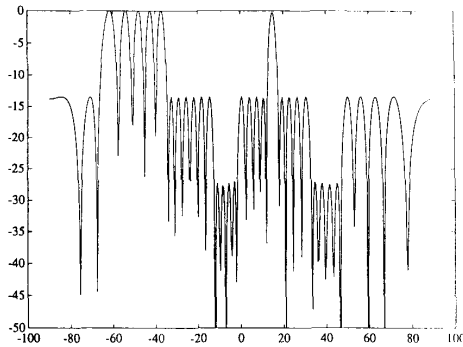


Fig. 3. Response of linear 35 element, equispaced array with constraints defined in (17) designed using the Dolph-Chebyshev technique.

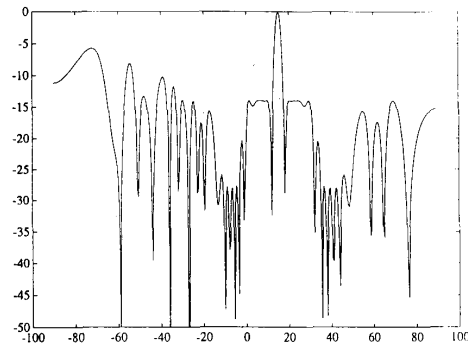


Fig. 4. Response of linear 26 element, nonequispaced array with constraints defined in (17) designed using the simplex search technique.

in Fig. 5. This result indicates the ability of the simplex method to design sparse arrays with fewer elements than an equivalent equispaced Dolph-Chebyshev array under the same constraints.

**B. Sparse Circular Arrays**

This example demonstrates the ability of the simplex method to design planar arrays. Consider the 60-element transparent concentric ring array of Fig. 6. We wish to form beams, steered horizontally, in the plane containing the array, i.e.,  $s = (\cos \theta, \sin \theta, 0)$ . This is similar to the configuration used by some "dipping" sonar systems which suspend a cylindrical ring array in the water from a helicopter and form horizontal search beams. Fig. 7 shows the beam response for the full array using unity magnitude shading, with complex phase shifting at each element to steer the maximum response angle (MRA) to a bearing of  $0^\circ$ . A sinusoidal signal at 1 kHz is assumed, which gives an average element to element spacing of just over  $\lambda/4$ .

The constraints for the simplex algorithm are that the main lobe width is the same as for the unity weighting of the full array and the maximum sidelobe level is allowed to be no higher than the first sidelobe in Fig. 7. Allowing some of the secondary sidelobes to come up to the level of the first introduces some degrees of freedom which are exploited by the algorithm to thin the array. Fig. 8 shows the remaining

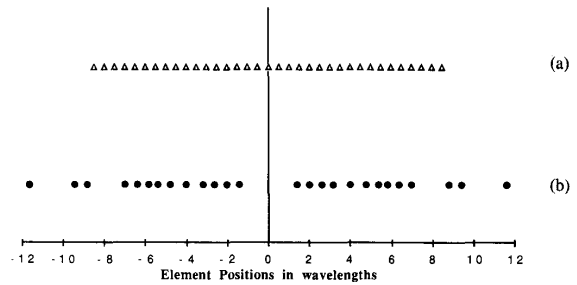


Fig. 5. Positions of the elements for (a) the equispaced array with response given in Fig. 3 and (b) the nonequispaced array with response given in Fig. 4.

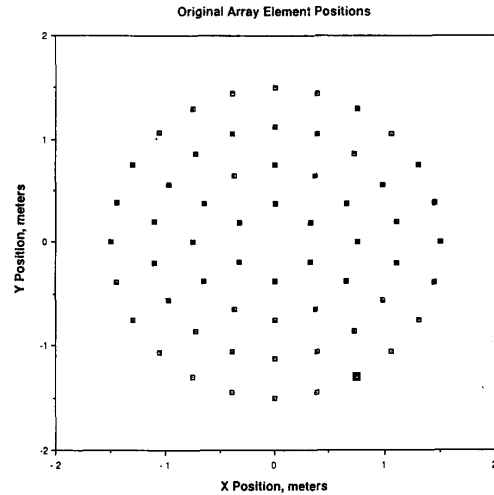


Fig. 6. Sixty element concentric ring array designed to operate at 1 KHz in seawater.

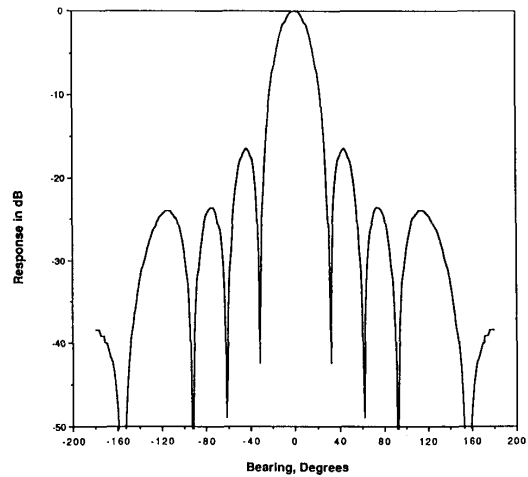


Fig. 7. In plane beam response for the array in Fig. 6 for unity magnitude shading and phase steering at bearing  $0^\circ$ .

elements of the array after thinning by the simplex search algorithm, and Fig. 9 shows the corresponding response pattern. Only 16 of the original elements are needed to maintain the original main lobe shape and maximum sidelobe level.

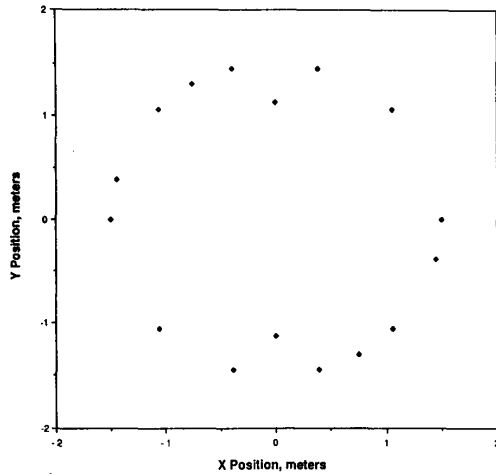


Fig. 8. Thinned array of 16 elements obtained using the simplex search on the configuration in Fig. 6 under the constraint that the mainlobe width and peak sidelobe level are no larger than for the full, unity weighted array response in Fig. 7.

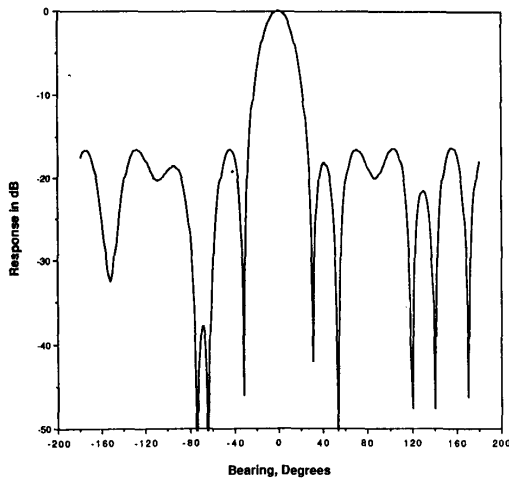


Fig. 9. Response of the thinned array of Fig. 8.

VI. CONCLUSION

The design of sparse beamforming arrays was formulated as a constrained optimization problem using a set of linear constraints on the array response and a cost function which favors sparse arrays. Theorems 1-4 show that under certain conditions the solution to this problem may be found using  $l_p$  optimization, for values of  $p$  in the range  $0 < p < 1$ . Using the properties of the  $l_p$  cost function, we show that the maximally sparse array (and its corresponding weight vector) must then lie at an extreme point of the simplex defined by the constraints, and thus the search may be restricted to a finite set of possible optimal solution vectors. An algorithm for finding a local minimum on the graph defined by the simplex tableaux and their adjacencies is described.

In Section V, examples of sparse linear and planar arrays obtained using the simplex method are given. These results

indicate that good practical results may be obtained using this approach. As described above, this method may also be applied to the design of 3-D or conformal arrays.

There are a number of limitations to this algorithm which must be discussed. Firstly, there is no guarantee of a global optimum since the algorithm will terminate on reaching a local minimum of the graph formed by the simplex tableaux and their adjacencies. The results in Section V indicate, however, that the local minima found by this algorithm are indeed sparse, and hence that a guaranteed global convergence is not essential for obtaining useful solutions. A second limitation of this method is that the array must be symmetric and the set of potential element locations is finite. Since there is no constraint on the number of candidate positions, the latter constraint is not too restrictive. Using a very large number of potential locations will however result in a large simplex tableau and hence computation time will increase. The requirement for symmetry in the array is necessary to reduce the problem to the estimation of a real set of coefficients. It is possible to modify the above formulation to incorporate complex variables, however, in this case the theorems given in the Appendix do not apply and hence the solution may not be found using the simplex method.

Application of this method is not restricted to the design of beamforming arrays. It can also be applied to any linearly constrained problem in which the dominant cost is the total number of "units" required in the solution vector. Thus we anticipate potential applications in operations research, the design of reduced-multiplier FIR filters and sparse deconvolution [31], [32].

APPENDIX

*Theorem 1:* If a feasible solution to (12) exists then a basic feasible solution exists.

*Proof of Theorem 1:* This theorem is well known and is the basis for the simplex algorithm for linear programming. A proof may be found in [19].  $\square$

*Theorem 2:* If there exists a globally optimal feasible solution to (12) then there exists a basic feasible solution which is globally optimal.

*Proof of Theorem 2:* It is easily shown that  $h(\mathbf{x})$  in (12) is a concave functional over the set  $S = \{\mathbf{x} \in R^{2N+2M}, \mathbf{x} \geq \mathbf{0}\}$  for any  $p \leq 1$ . It is well known that there exists an extreme point of the constraint set which is a global minimum of a concave functional over a set of linear constraints [19]. Since a feasible solution to a set of linear constraints is an extreme point iff it is basic, theorem 2 is proven.  $\square$

*Theorem 3:* If  $\mathbf{x}_0^T = (\mathbf{a}_0^{+T}, \mathbf{a}_0^{-T}, \mathbf{s}_0^{+T}, \mathbf{s}_0^{-T})$  is a globally optimal basic feasible solution to (12) then  $\mathbf{a}_0 = (\mathbf{a}_0^+ - \mathbf{a}_0^-)$  is a globally optimal solution to (11).

*Proof of Theorem 3:* Let  $\mathbf{x}_0^T = \{\mathbf{a}_0^{+T}, \mathbf{a}_0^{-T}, \mathbf{s}_0^{+T}, \mathbf{s}_0^{-T}\}$  be a globally optimal basic feasible solution to (12). Then:

$$h(\mathbf{x}_0) = \sum_{i=1}^N |a_0^+(i)|^p + |a_0^-(i)|^p. \quad (17)$$

If for any  $i \in \{1, 2, \dots, N\}$ ,  $a_0^+(i)a_0^-(i) \neq 0$  then replacing  $a_0^+(i)$  with  $\{a_0^+(i) - \min[a_0^+(i), a_0^-(i)]\}$



and

$$a_0^-(i) \text{ with } \{a_0^-(i) - \min[a_0^+(i), a_0^-(i)]\}$$

would produce a new basic feasible solution of lower cost. Since  $\mathbf{x}_0$  is a global optimum this is not possible, hence  $\mathbf{a}^{+T}\mathbf{a}^- = 0$ .

From Statement 1, if  $\tilde{\mathbf{x}}^T = \{\tilde{\mathbf{a}}^{+T}, \tilde{\mathbf{a}}^{-T}, \tilde{\mathbf{s}}^{+T}, \tilde{\mathbf{s}}^{-T}\}$  is a feasible solution to (12), then  $\tilde{\mathbf{a}} = (\tilde{\mathbf{a}}^+ - \tilde{\mathbf{a}}^-)$  is a feasible solution to (11). For any  $\tilde{\mathbf{x}}$  such that  $\tilde{\mathbf{a}}^{+T}\tilde{\mathbf{a}}^- = 0$ :

$$g(\tilde{\mathbf{a}}) = \sum_{i=1}^N |\tilde{a}(i)|^p = \sum_{i=1}^N |\tilde{a}^+(i)|^p + |\tilde{a}^-(i)|^p = h(\tilde{\mathbf{x}}). \quad (18)$$

If there exists a feasible solution  $\mathbf{a}_1$ , to (11) such that  $g(\mathbf{a}_1) < h(\mathbf{a}_0)$ , then from Statement 1, one could construct a feasible solution  $\mathbf{x}_1$  such that  $\mathbf{a}_1^{+T}\mathbf{a}_1^- = 0$ . But then from (18),  $h(\mathbf{x}_1) < h(\mathbf{x}_0)$ , which is not possible if  $\mathbf{x}_0$  is a global minimum of (12). It follows that if  $\mathbf{x}_0$  is a global minimum of (12) then  $\mathbf{a}_0$  is a global minimum of (11).  $\square$

**Theorem 4:** Let  $V$  be the set of all globally optimal solutions to (7) with  $\{f(\mathbf{a}) = r \text{ for all } \mathbf{a} \in V\}$ . Let  $U$  be the set of all globally optimal solutions to (11). Let  $S_x$  be the set of basic feasible solutions to (12) such that:

$$S_x = \{\mathbf{x}^T = [\mathbf{a}^{+T}, \mathbf{a}^{-T}, \mathbf{s}^{+T}, \mathbf{s}^{-T}] \in \mathbf{R}^{N'}\},$$

$$\text{such that } \mathbf{x} \text{ is basic and } \mathbf{a}^{+T}\mathbf{a}^- = 0\}.$$

Let:

$$\Gamma = \max_{\mathbf{x} \in S_x} \{a_i^+, a_i^-, i = 1, 2, \dots, N\}$$

and

$$\Psi = \min_{\mathbf{x} \in S_x} \{a_i^+ \neq 0, a_i^- \neq 0, i = 1, 2, \dots, N\}.$$

Then, if  $\Gamma < \infty$ ,  $U$  is a subset of  $V$  for all  $0 < p < p_1$ , where

$$p_1 = \frac{\log[(r+1)/r]}{\log[\Gamma/\Psi]}$$

**Proof of Theorem 4:** If  $\mathbf{a}$  is feasible for (7) then  $\mathbf{a}$  is feasible for (11) since the constraints are identical. Statement 1 establishes that for each  $\mathbf{a} \in U$  and  $\mathbf{a} \in V$  there exists an  $\mathbf{x} \in S_x$ . Consequently:

$$g(\mathbf{a}) \leq r\Gamma^p, \quad \text{for all } \mathbf{a} \in V. \quad (19)$$

Let  $\bar{V}$  denote the set of feasible solutions not in  $V$ . Then the minimum number of nonzero elements in all vectors  $\mathbf{a} \in \bar{V}$  is  $(r+1)$ . Consequently:

$$g(\mathbf{a}) \geq (r+1)\Psi^p, \quad \text{for all } \mathbf{a} \in \bar{V}. \quad (20)$$

Thus, to ensure that a global minimum of (11) lies in the set  $V$ , it is sufficient that:

$$g(\mathbf{a}_1) < g(\mathbf{a}_2), \quad \text{for all } \mathbf{a}_1 \in V \text{ and } \mathbf{a}_2 \in \bar{V}.$$

From (19) and (20), this condition is guaranteed provided  $p$  satisfies the inequality

$$r\Gamma^p < (r+1)\Psi^p$$

or equivalently

$$p < \frac{\log[(r+1)/r]}{\log[\Gamma/\Psi]}. \quad \square$$

#### REFERENCES

- [1] D. King, R. Packard, and R. Thomas, "Unequally-spaced, broad-band antenna arrays," *IRE Trans. Antennas Propagat.*, vol. AP-8, pp. 380-384, 1960.
- [2] A. Maffett, "Array factors with nonuniform spacing arrays," *IRE Trans. Antennas Propagat.*, vol. AP-10, pp. 131-136, 1962.
- [3] N. Balakrishnan, P. Murthy, and S. Ramakrishna, "Synthesis of antenna arrays with spatial and excitation constraints," *IEEE Trans. Antennas Propagat.*, vol. AP-29, pp. 690-696, 1979.
- [4] H. Unz, "Linear arrays with arbitrarily distributed elements," *IRE Trans. Antennas Propagat.*, vol. AP-8, pp. 222-223, 1960.
- [5] A. Ishimaru and Y-S Chen, "Thinning and broadbanding antenna arrays by unequal spacings," *IEEE Trans. Antenna Propagat.*, vol. AP-13, pp. 34-42, 1965.
- [6] Y. Lo and S. Lee, "A study of space-tapered arrays," *IEEE Trans. Antennas Propagat.*, vol. AP-14, pp. 22-30, 1966.
- [7] Y. Lo, "A mathematical theory of antenna arrays with randomly spaced elements," *IEEE Trans. Antennas Propagat.*, vol. AP-12, pp. 257-268, 1964.
- [8] Y. Kim and D. Jaggard, "The fractal random array," *Proc. IEEE*, vol. 74, pp. 1278-1280, 1986.
- [9] M. Skolnik, G. Nemhauser, and J. Sherman, "Dynamic programming applied to unequally spaced elements," *IEEE Trans. Antennas Propagat.*, vol. AP-12, pp. 35-43, 1964.
- [10] F. Bratkovic, "Computer determination of spaces for broad-band nonuniform arrays," *IEEE Trans. Antennas Propagat.*, vol. AP-21, pp. 211-213, 1973.
- [11] R. Redlich, "Iterative least squares synthesis on nonuniformly spaced linear arrays," *IEEE Trans. Antennas Propagat.*, vol. AP-21, pp. 106-108, 1973.
- [12] N. Mitrou, "Results on nonrecursive digital filters with nonequidistant taps," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1621-1624, 1985.
- [13] G. Kishi, K. Sakanjwa, and T. Uyematsu, "A design method for transversal filters with nonuniform tap spacings based on the mean square error criterion," in *Proc. 5th Int. Symp. Network Theory*, Sept. 1984, pp. 363-368.
- [14] P. Jarski, T. Saramaki, S. Mitra, and U. Neuvo, "On properties and design of nonuniformly spaced linear arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 372-380, 1988.
- [15] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoust., Speech, Signal Processing, Magazine*, Apr. 1988.
- [16] B. D. Steinbert, "The peak sidelobe of the phased array having randomly located elements," *IEEE Trans. Antennas Propagat.*, vol. AP-20, pp. 127-136, 1972.
- [17] T. W. Parks and C. S. Burrus, *Digital Filter Design*. New York: Wiley-Interscience, 1987.
- [18] L. Scales, *Introduction to Non-Linear Optimization*. New York: Springer-Verlag, 1985.
- [19] D. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Reading, MA: Addison-Wesley, 1984.
- [20] A. Pietsch, "Approximation spaces," *J. Approx. Theory*, vol. 32, pp. 115-134, 1981.
- [21] W. Gray, "Variable norm deconvolution," Ph. D. dissertation, Stanford Univ., Stanford, CA, 1979.
- [22] I. Barrodale and F. Roberts, "Application of mathematical programming to  $l_p$  approximation," in *Nonlinear Programming*, J. Rosen et al., Eds. New York: Academic, 1970.
- [23] T. Pham and R. deFigueiredo, "Maximum likelihood estimation of a class of non-Gaussian densities with application to  $l_p$  deconvolution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 73-82, 1989.
- [24] J. Falk and K. Hoffman, "Concave minimization via collapsing polytopes," *Oper. Res.*, vol. 34, pp. 919-929, 1986.
- [25] P. Zwart, "Global maximization of a convex function with linear inequality constraints," *Oper. Res.*, vol. 22, pp. 602-609, 1974.
- [26] N. Thoai and H. Tuy, "Convergent algorithms for minimizing a concave function," *Math. Oper. Res.*, vol. 5, pp. 556-566, 1980.
- [27] B. Jeffs, "Maximally sparse constrained optimization for signal pro-

cessing applications," Ph. D. dissertation, Univ. Southern California, Los Angeles, CA, Jan. 1989.

[28] A. Schrijver, *Theory of Linear and Integer Programming*. New York, Wiley, 1986.

[29] H. Kruse, "Degeneracy graphs and the neighborhood problem," lecture notes in Econ. Math. Systems, vol. 260, Springer-Verlag, 1986.

[30] C. Dolph, "A current distribution for broadside arrays which optimizes the relationship between beamwidth and sidelobe level," *Proc. IRE*, vol. 34, pp. 335-348, 1946.

[31] R. Leahy, B. Jeffs, and Z. Wu, "A DSP algorithm for minimum order solutions," in *Proc. 21st Asilomar Conf. Signals, Syst. Comp.*, Nov. 1987.

[32] J. Mendel, *Optimal Seismic Deconvolution*. New York: Academic, 1983.

signal processing, and other applications of optimization theory and statistics in signal and image processing.



**Richard M. Leahy** (M'85) was born in Surrey, England, in 1960. He received the B.Sc. and Ph.D. degrees in electrical engineering from the University of Newcastle upon Tyne, England, in 1981 and 1985, respectively.

In 1985 he joined the University of Southern California, Los Angeles, as an Assistant Professor in the Signal and Image Processing Institute of the Department of Electrical Engineering-Systems. He currently also holds a joint appointment with the Department of Radiology at USC. His research

interests include medical image reconstruction and analysis, biomedical



**Brian D. Jeffs** (M'90) was born at Fort Meade, MD, on January 25, 1954. He received the B.S. and M.S. degrees from Brigham Young University, Provo, UT, in 1978 and 1982, respectively, and the Ph.D. degree from the University of Southern California, Los Angeles, in 1989, all in electrical engineering.

From 1978 to 1982 he worked as a digital system engineer for Eyring Research Institute, Provo, UT, and from 1982 to 1989 as a signal processing system engineer for Hughes Aircraft

Company Ground Systems Group, Fullerton, CA. His areas of activity at Hughes included digital signal processing design and analysis for active and passive SONAR systems. Since 1990 he has been an Assistant Professor in the Department of Electrical and Computer Engineering at Brigham Young University. His current research interests include array signal processing, mathematical image reconstruction from indirect observation, biomedical image processing, and acoustic imaging.

Backscatter ..... E. J. ROWE ..... 123

Theoretical Design and Predictions of Volute Antenna Performance ..... F. Manshadi ..... 124

Wide Band Transient Electromagnetic Scattering Laboratory ..... ..

Wide Band Transient Electromagnetic Scattering Laboratory ..... ..