Recovery of a Sparse Spike Time Series by L_1 Norm Deconvolution

Michael S. O'Brien, Anthony N. Sinclair, *Member, IEEE*, and Stuart M. Kramer

Abstract—An L_1 norm minimization scheme is applied to the determination of the impulse response vector $\mathbf{h}$ of flaws detected in practical examples of ultrasonic nondestructive evaluation in CANDU nuclear reactors. For each problem, parametric programming is applied to find the optimum value of the damping parameter that will yield the best estimate of $\mathbf{h}$ according to a quantified performance factor. This performance factor is based on a quantified analysis of the transitions in estimates of $\mathbf{h}$ as the damping parameter is varied over a wide range of possible values. It is shown that for the examined cases in which the true impulse response is a sparsely filled spike strain, the L_1 norm provides significantly better results than the more commonly used L_2 norm minimization schemes. These results are shown to be consistent with theoretical predictions.

I. NOMENCLATURE

$e(t)$	total error waveform
$\mathbf{e}$	discretized error vector of length R
e_i	member of the error vector $\mathbf{e}$, $1 \leq i \leq R$
$h(t)$	true impulse response
h_i	member of the true impulse response vector $\mathbf{h}$, $1 \leq i \leq Q$
$\hat{h}(t)$	estimated impulse response
$\mathbf{h}$	discretized true impulse response vector
$\hat{\mathbf{h}}$	discretized estimated impulse response vector
$\hat{h}_i$	member of the estimated response vector $\hat{\mathbf{h}}$, $1 \leq i \leq Q$
L_p	type p norm
p	integer designating type of norm, $p = 1$ or 2 in this study
P	number of terms in input vector $\mathbf{x}$
Q	number of terms in response vector $\mathbf{h}$
R	number of terms in output vector $\mathbf{y}$
SSD	figure of merit: sum of squared deviations between $\hat{\mathbf{h}}$ and $\mathbf{h}$
$T_D(\bar{\lambda}_1)$	transition parameter
t	time
V_{ij}	speed of bulk wave traveling in direction i with polarization j , where i, j designate R (radial), L (axial), or C (circumferential) directions.

Manuscript received September 20, 1993; revised February 21, 1994. This work was supported by the CANDU Owners Group (COG) and the Natural Sciences and Engineering Research Council (NSERC) of Canada. The associate editor coordinating the review of this paper and approving it for publication was Prof. Russell M. Mersereau.

M. S. O'Brien and A. N. Sinclair are with the Department of Mechanical Engineering, University of Toronto, Toronto, Canada M5S-1A4.

S. Kramer is with Metallurgical Research, Ontario Hydro, Toronto, Canada M8Z-5S4.

IEEE Log Number 9406024.

$x(t)$	input waveform (volts)
$\mathbf{x}$	discretized input vector (volts)
$\mathbf{X}$	matrix formed from the elements of $\mathbf{x}$
$y(t)$	output waveform (volts)
$\mathbf{y}$	discretized output vector (volts)
Φ_p	objective function; minimum value corresponds to good estimated solution $\hat{h}(t)$
λ_p	damping parameter corresponding to p -norm
$\bar{\lambda}_1$	normalized value of damping parameter λ_1
$\lambda_{p,\text{optimum}}$	value of λ_p corresponding to optimum estimate of response vector $\hat{\mathbf{h}}$ within a solution set, according to some given figure of merit
$\lambda_{1,\text{max}}$	maximum value of damping parameter that yields a nontrivial solution $\hat{\mathbf{h}}$
Γ	figure of merit: deviation between $\hat{\mathbf{h}}$ and $\mathbf{h}$ at location of spikes in $\mathbf{h}$
Ψ	figure of merit: measure of noise in calculated impulse response vector $\hat{\mathbf{h}}$

II. INTRODUCTION

MANY practical problems in acoustics can be modeled as a linear system where the measured output signal $y(t)$ is the convolution of an input waveform $x(t)$ with a sparse spike time series $h(t)$. The function $h(t)$ represents the system's impulse response, with the spikes corresponding to discrete reflecting surfaces. In theory, characteristics of the system can be revealed by calculating $h(t)$ as a straightforward deconvolution of the output signal by the input. However, factors such as the limited bandwidth of digitized waveforms, and particularly the presence of noise in the measured input and output signals, have been shown capable of grossly distorting the estimate of $h(t)$ obtained by a conventional deconvolution. Much effort to address this problem has concentrated on utilizing the *a priori* knowledge that $h(t)$ is "simple," i.e., a sparse spike time series. One way that this can be done is to minimize an objective function Φ_p , which is a weighted sum of the L_p norms of the estimated solution $\hat{h}(t)$ and the "error" $e(t)$ which corrupts the measured output $y(t)$:

$$\Phi_p = \|e(t)\|_p + \lambda_p \|\hat{h}(t)\|_p, \quad \text{where } e(t) = y(t) - \hat{h}(t) * x(t) \quad (1)$$

The parameter λ_p is a relative weighting or damping factor selected to balance the conflicting priorities of data accountability, i.e., minimizing $e(t)$, and addressing the *a priori* assumption that the true solution $h(t)$ is sparse.

The common choice of $p = 2$ in the application of (1) has the advantage of computational convenience [1], [2]. One of the earlier attempts at L_1 norm deconvolution was a time-domain approach presented by Claerbout and Muir [3]. Their proposed method involved solving (1) in the time domain via linear programming with $\lambda_1 = 0$, i.e., the undamped minimization of the residual error vector alone. Although pursued in some subsequent works [4], [5], it is clear that without explicit control on the estimated response vector $\mathbf{h}$, the method is ill-suited to the recovery of a sparse spike series.

The damped L_1 norm technique with $\lambda_1 \neq 0$ emerged in the work by Taylor *et al.* [6]. The data investigated were primarily synthetic with "pseudo-white" noise added (two examples featured signal-to-noise ratios of 4). The results presented were sharp and visually satisfactory. A significant step in Taylor's work was the introduction of parametric programming as an aid in the selection of λ_1 ; plots of $\|\mathbf{e}\|_1$ vs λ_1 , the number of nonzero response terms versus λ_1 , and $\|\mathbf{e}\|_1$ versus the number of nonzero response terms were all suggested as being "useful" for selection of an appropriate value for λ_1 . However, no systematic procedure for utilizing these plots to select λ_1 was presented.

A similar time-formulation to that of Taylor, *et al.* was given by Drachman [7], where the inexact convolution equations were presented as a set of inequality constraints. This defined an infinite strip in the space of possible solutions, within which Drachman defined the "best" solution to be the one for which the L_1 norm of the response series alone, $\|\mathbf{h}\|_1$, was minimized. The absence of an explicit damping parameter was compensated by the addition of a conditioning term; the problems of how to select this conditioning factor and the residual errors e_i were left unresolved.

A significant departure to the formulation of Taylor *et al.* was presented by Chapman and Barrodale [8]. Their "one-at-a-time spike extraction" scheme operated by making successive passes through the data, extracting one additional spike at each pass. The location and magnitude of each spike were selected to minimize the L_1 norm of the residuals ($\lambda_1 = 0$). Although Barrodale and others [8]–[10] reported significant computational advantages over Taylor's formulation, certain problems remain. The solution obtained at the end of each pass is not globally "best," as the locations of all spikes but one are fixed by the results of earlier passes through the data. The effects of noise on the results have received no detailed investigation.

In a time-frequency approach formulated by Levy and Fullagar [4], the number of equations is reduced by discarding data corresponding to frequencies where very little information is contained in the input vector $\mathbf{x}$; in this way, the high condition number of the problem is significantly lowered. The solution is then determined by L_1 norm minimization of the response series, subject to the remaining portion of the convolution equations. Good results have been reported [11], [12], despite unresolved uncertainties regarding the effects of noise and arbitrary selection of parameter values.

Selection of one such parameter, λ_2 in the case of L_2 deconvolution, has been addressed by a generalized cross-validation (GCV) criterion, e.g., [13]. In a practical application

of the theory to image restoration, it was shown that the computation time for the deconvolution can be reduced by assumptions on the nature of the image degradation.

Mendel has extensively investigated the use of maximum-likelihood deconvolution (MLD) methods for the extraction of a sparse spike train from noisy data, and offered some insight on its merits as compared to L_1 and minimum entropy methods [14]–[16]. Several techniques were explored for reducing the computational effort required for deconvolution, based on the use of a low-order autoregressive moving average (ARMA) model for the source wavelet.

In this work, the damped L_1 norm formulation of Taylor, *et al.* [6], as given by (1), is selected for more detailed study, and the following specific questions are to be addressed:

- To date, a practical procedure to select an appropriate value for λ_1 continues to be an elusive goal [6], [12], [13]; trial-and-error approaches are still frequently adopted to produce a calculated solution $\hat{h}(t)$ that has the best visual appeal. In this paper, a figure of merit will be sought to promote a more objective analysis of the proximity of $\hat{h}(t)$ to $h(t)$ as a function of λ_p ; it is emphasized, however, that such methods are not applicable to practical deconvolution problems where the true solution is unknown.
- Building on the results of objective (a) above, a rational scheme will be developed to select a weighting parameter λ_p for $p = 1$ when $h(t)$ is *not* known.
- The studies cited in parts (a) and (b) above are to be based on synthetic data series and added noise components. The proposed scheme will then be applied to two practical problems in the ultrasonic nondestructive evaluation (NDE) of CANDU nuclear reactor fuel channels. Results will be compared to those achieved with L_2 norm deconvolution, which continues to see extensive use in the NDE field.

III. IMPLEMENTATION OF AN L_1 NORM MINIMIZATION SCHEME

A. System Equations

The "inexact" convolution equations corresponding to an input vector $\mathbf{x}$ of length P and output vector $\mathbf{y}$ of length R are defined by:

$$y_k = \sum_{j=1}^{k \leq Q} h_j x_{k-j+1} + e_k, \quad k = 1, 2, 3, \dots, R; \quad (2)$$

$$Q = R - P + 1$$

or in matrix format $\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{e}$, where $\mathbf{X}$ is a $R \times Q$ matrix of bandwidth P . This yields an overdetermined system of R equations for the Q unknowns of the true response vector $\mathbf{h}$.

The elements of the error vector $\mathbf{e}$ are called fitting errors or residuals. Given that the e_k can be arbitrarily selected, there are an infinite number of calculated response vectors $\mathbf{h}$ that can satisfy (2). Under the L_p norm minimization scheme, the selection of the most appropriate $\mathbf{e}$ and corresponding $\hat{\mathbf{h}}$ is based on minimization of the function Φ_p of (1); in discrete

form, this is given by:

$$\begin{aligned} \text{Minimize: } \Phi_p(\hat{\mathbf{h}}, \mathbf{e}) &= \sum_{i=1}^R \|e_i\|_p + \lambda_p \sum_{j=1}^Q \|\hat{h}_j\|_p \\ &= \sum_{i=1}^R |e_i|^p + \lambda_p \sum_{j=1}^Q |\hat{h}_j|^p \quad (3) \\ &= \|\mathbf{e}\|_p + \lambda_p \|\hat{\mathbf{h}}\|_p, \\ \text{subject to: } \mathbf{X}\hat{\mathbf{h}} + \mathbf{e} &= \mathbf{y}. \end{aligned}$$

An explicit solution to (3) exists for the case $p = 2$. Although iterative techniques have seen extensive application, linear programming techniques such as the Simplex method [17], [18] can be used to find a solution corresponding to $p = 1$ in a finite number of steps. First, (3) must be reformatted. Based on the assumption that the $\hat{h}_j$ and e_i are uncorrelated, the minimum value for Φ_1 can be achieved by partial differentiation of (3) with respect to each term of the response vector, and setting the result to zero:

$$\begin{aligned} \frac{\partial \Phi_1(\hat{\mathbf{h}}, \mathbf{e})}{\partial \hat{h}_k} &= 0 = \frac{\partial \left(\sum_{i=1}^R |e_i| + \lambda_1 \sum_{j=1}^Q |\hat{h}_j| \right)}{\partial \hat{h}_k}, \\ k &= 1, 2, \dots, Q \\ 0 &= \sum_{i=1}^R \frac{\partial \sum_{j=1}^Q (y_i - X_{ij}\hat{h}_j)}{\partial \hat{h}_k} \text{sgn}(e_i) \\ &\quad + \lambda_1 \sum_{j=1}^Q \frac{\partial \hat{h}_j}{\partial \hat{h}_k} \text{sgn}(\hat{h}_j) \quad (4) \\ 0 &= - \sum_{i=1}^R X_{ik} \text{sgn}(e_i) + \lambda_1 \text{sgn}(\hat{h}_k) \end{aligned}$$

$$\text{subject to: } \mathbf{X}\hat{\mathbf{h}} + \mathbf{e} = \mathbf{y}$$

where the *signum* function of a real number w is defined by $\text{sgn}(w) = w/|w|$.

The implications of (4) are far-reaching. It suggests that each of the Q columns of the $\mathbf{X}$ matrix could be divided into two groups; the sum of the elements of the two groups would differ by λ_1 . However, when it is considered that the x_i and y_i variables originate from experimental measurement, it is clear that such conditions would not in general be met. To satisfy the Q conditions of (4), then, a total Q of the error and response terms must be zero; this would make Q of the *signum* functions indeterminate (one for each of the Q equations, see [3]).

One additional note should be added regarding the structure of the objective function Φ_1 . For cases to be considered in this study for which the signal-to-noise ratio is relatively good, it makes only a moderate difference to the result whether one considers the L_1 or L_2 norm of the error vector in (3). The L_1 norm of $\mathbf{e}$ was selected to accommodate use of the SIMPLEX algorithm in the solution, even though this will force a somewhat sparse error vector. For cases where signals are very noisy, a combined $\|\mathbf{e}\|_2 + \lambda_1 \|\hat{\mathbf{h}}\|_1$ objective function would be preferable.

B. Geometrical Interpretation of L_1 Norm Minimization by Linear Programming

To minimize the objective function Φ_1 , there are $Q + R$ unknowns to be solved: $h_j, j = 1, 2, \dots, Q$ and $e_i, i = 1, 2, \dots, R$. According to (4), at least Q of these variable must be set to zero. Solution by the Simplex method requires that each of the possible negative response terms $\hat{h}_j$ and residual terms e_i be expressed as the difference between two strictly nonnegative variables. It can be shown [2] that minimization of the objective function $\hat{\Phi}_1$ defined below (instead of Φ_1 as defined by (3)), subject to certain transformed constraints, will yield the L_1 norm solution corresponding to any specified value of λ_i .

$$\text{Minimize: } \hat{\Phi}_1(\mathbf{e}, \hat{\mathbf{h}}) = \sum_{i=1}^R (e_i^+ + e_i^-) + \lambda_1 \sum_{j=1}^Q (h_j^+ + h_j^-) \quad (5)$$

$$\text{Subject to: } [\mathbf{X} - \mathbf{X}\mathbf{I} - \mathbf{I}][\mathbf{h}^+ \mathbf{h}^- \mathbf{e}^+ \mathbf{e}^-]^T = \mathbf{y} \quad (6)$$

$$\text{where } \hat{\mathbf{h}} = \mathbf{h}^+ - \mathbf{h}^-, \quad \mathbf{e} = \mathbf{e}^+ - \mathbf{e}^-. \quad (7)$$

The $N = 2(Q + R)$ unknowns of (6) define an N -dimensional variable space. Each of the R linear equations defines an infinite, $(N - 1)$ -dimensional plane within the space. The intersection of these planes then defines a convex $M = N - R$ dimensional subspace, within which any possible solution must lie. However, since all unknowns are constricted to nonnegative values, the valid solution space consists only of the positive portion of the M -dimensional space.

The linear objective function $\hat{\Phi}_1$ also represents an $(N - 1)$ -dimensional plane, called the objective plane. The objective plane intersects the solution space, thereby specifying the range of possible values for all the response variables and residuals for the corresponding value of $\hat{\Phi}_1$. Altering the value of λ_1 changes the "tilt" of the objective plane within the variable space. By contrast, altering the value of $\hat{\Phi}_1$ is equivalent to moving the objective plane normal to itself, towards positive sectors as $\hat{\Phi}_1$ is increased, and towards negative sectors as $\hat{\Phi}_1$ is decreased. By lowering the value of $\hat{\Phi}_1$, the objective plane eventually moves to the boundary of the solution space, until their intersection is in general only a single vertex. This vertex corresponds to the solution of (3) or (5)–(7), and uniquely defines all the response and error terms.

The vertices of the solution space merit further examination, as they are the potential candidates for the solution to the L_1 norm minimization problem. The vertices correspond to the intersections of the solution space with at least M zero-faces of the variable space, i.e., at least M of the N positive variables in the solution space are zero, or at most $N - M = R$ variables can be nonzero. It is this feature that gives the characteristic sparse response vector in the solution of a L_1 norm minimization problem. The larger the value of λ_1 , the greater the proportion of of the R nonzero values will be elements of the error vector, as opposed to the response vector.

An important feature of the solution to (5)–(7) is that for each pair of $+/-$ variables in the vectors $[\mathbf{h}^+, \mathbf{h}^-]$ and $[\mathbf{e}^+, \mathbf{e}^-]$, at least one element in each pair will be zero. This is illustrated by considering a single pair of $+/-$ variables,

say h_j^+ and h_j^- for a fixed value of $j = J$. The projection of the solution space onto the 2-D plane containing this pair of $+/-$ variables (while fixing the values of all other variables) is a line. The equation of this line is extracted from (6) as

$$X_{JJ}h_j^+ - X_{JJ}h_j^- = C_1 \quad \text{or} \quad h_j^+ = h_j^- + C_1/X_{JJ} \quad (8)$$

where C_1 incorporates the fixed values of all other variables contained in the J th line of the matrix equation (6). It is clear from (8) that the solution space must intersect at least one nonnegative axis of these two variables for a given set of values for the other $N-2$ variables. Therefore, an intersection of the solution space with at least one of these two nonnegative axes will be involved in all of the vertices of the solution space. This implies that at least one of the pair of variables must be zero at the solution $\hat{\mathbf{h}}$ (both will be zero if (8) corresponds to a line through the origin, i.e., $C_1 = 0$).

C. Solution Algorithm

Following the geometrical arguments of the previous section, a Simplex algorithm for the solution of (5) and (6) was implemented (Fig. 1). The algorithm determines estimated solution vectors $\hat{\mathbf{h}}$ corresponding to a number of damping parameter values, labeled $(\lambda_1)_n$, $n = 0, 1, 2, \dots$. The magnitudes of $(\lambda_1)_n$ range from just above $\lambda_{1,\max}$ for $n = 0$, down to zero, where $\lambda_{1,\max}$ is the highest damping parameter value for which a nontrivial solution is achievable. The value of $\lambda_{1,\max}$ is found by an iterative technique.

For each value of $(\lambda_1)_n$, the Simplex algorithm first requires a feasible solution, i.e., values for the unknowns $\mathbf{e}^+$, $\mathbf{e}^-$, $\mathbf{h}^+$, and $\mathbf{h}^-$ in order to start the iterative solution process. For the initial value of $(\lambda_1)_{n=0}$ just greater than $\lambda_{1,\max}$, the trivial response vector $\hat{\mathbf{h}} = 0$ is the solution to (5) and (6). As the damping parameter is then slowly decremented toward zero, very good initial estimates of the response and error vectors corresponding to $(\lambda_1)_n$ are obtained from the results at $(\lambda_1)_{n-1}$; this reduces the number of steps required to reach the solution at each value of the damping parameter. Because of this time-saving device, the cost of computing the L_1 norm solution set for 100 values of λ_1 was only 3 times that for a single calculation (Table I). It should be noted that the Simplex method does not guarantee that the solution at each value of λ_1 will be reached in the least possible number of steps; however, the saving in CPU time arises from the low cost of each step.

The Simplex L_1 norm minimization technique is inherently more expensive than that of the explicit L_2 techniques. However, the calculation of $\hat{\mathbf{h}}$ for the L_2 norm is seen from Table I to require a computation time that varies linearly with the number of λ_2 values under consideration.

D. Figures of Merit

Many numerical studies of L_p norm minimization techniques make use of *a priori* knowledge about $\mathbf{h}$ to select a value of λ_p that yields a solution $\hat{\mathbf{h}}$ to (3) that is "similar" to $\mathbf{h}$. Often this "similarity" has been based on a subjective visual comparison of the calculated and true solutions. A degree of objectivity can be added to such comparisons by the introduction of three "figures of merit"; they have been

Fig. 1. Flow chart for obtaining impulse response vector $\hat{\mathbf{h}}$ as a function of λ_1 , using L_1 norm minimization.

selected solely on the basis of the subjective criteria suggested by several authors for evaluating the performance of L_1 norm deconvolution schemes, e.g., [10], [19], [20]:

$$\text{SSD} = \sum_i (\hat{h}_i - h_i)^2 \quad (9a)$$

$$\Gamma = \sum_{h_i \neq 0} (\hat{h}_i - h_i)^2 \quad (9b)$$

$$\Psi = \sum_{h_i=0} (\hat{h}_i - h_i)^2 = \sum_{h_i=0} \hat{h}_i^2 \quad (9c)$$

Here SSD is the sum of the squared deviations of each coefficient from its true counterpart, whereas Γ compares results only at locations of true spikes. The parameter Ψ compares the two series at locations where zeroes appear in the true response; it is proportional to the amount of energy that

TABLE I
COMPUTATION TIMES FOR L_1 AND L_2 NORM DECONVOLUTION OF
SYNTHETIC DATA ON AN ALR 33/386 POWER VEISA PERSONAL COMPUTER

Operation	Computation Time
Step-wise L_1 norm deconvolution at 100 values of $\bar{\lambda}_1$ between 0 and 1	35.6 s
Single L_1 norm deconvolution at $\bar{\lambda}_1 = 0$	11.9 s
Single L_1 norm deconvolution at $\bar{\lambda}_1 = 0.3$	8.9 s
Single L_2 norm deconvolution at any value of λ_2	0.102 s
L_2 norm deconvolution at 101 values of λ_2	10.2 s

has "leaked" from true spikes into phantom locations. Note that for each of the three parameters, low values are indicative of good solutions, i.e., small deviation between $\hat{h}_i$ and h_i .

Clearly these figures of merit are of no use in practical problems where the true solution $\mathbf{h}$ is unknown. However, they have been included in brief preliminary studies of noisy synthetic data (where $\mathbf{h}$ is known) to gain insight into the L_p norm deconvolution process. The sparse spike train response vector $\mathbf{h}$ used for this preliminary work is shown in Fig. 2(a). Its profile includes several spikes of various magnitudes, including closely spaced pairs of spikes with either matched or opposite polarities. The input vector $\mathbf{x}$ is shown in Fig. 2(b), and the noise-free output $\mathbf{y}$ of Fig. 2(c) was obtained by convolution of $\mathbf{x}$ with $\mathbf{h}$.

A series of tests was performed in which Gaussian, Laplacian, or uniform noise was added to the output vector of Fig. 2(c) to assess the relative performance of L_1 and L_2 norm minimization schemes at recovering $\mathbf{h}$. The influence of both the statistical probability distribution and magnitude of noise were topics of investigation, where noise magnitude was specified by the percentage of the maximum ideal output that its standard deviation represented:

$$\%RMS = \frac{\sigma_{\text{noise}}}{\max(y_i)} \times 100\%. \quad (10)$$

For 50 realizations of each of the three noise types, and noise magnitudes ranging from 0 to 20% RMS, a L_p norm solution set of possible solution vectors $\hat{\mathbf{h}}$ was determined as a function of λ_p for $p = 1, 2$. From each such set, an "optimum" solution and corresponding damping parameter $\lambda_{p,\text{optimum}}$ were then determined that would minimize each of the three figures of

Fig. 2. Test data: (a) True impulse response series $\mathbf{h}$; (b) noise-free input series $\mathbf{x}$; (c) noise-free output series $\mathbf{y}$.

merit of (9). The following conclusions were reached regarding the dependence of the solution set on the noise profile and value of $\lambda_p, p = 1, 2$, plus implications for the optimum solution:

- Far superior values of the figures of merit for the optimum solutions were achieved with L_1 as opposed to L_2 norm minimization schemes for a wide variety of noise magnitudes. This is illustrated by Fig. 3. It is noted that the superior performance is most apparent for cases where the %RMS noise is less than 10%, but the effect persists to significantly higher noise values. It was also found that very similar trends were observed for SSD, Γ , and Ψ , suggesting that the precise form of the figure of merit plays only a minor role. Further detailed studies could therefore be confined to a single one of these three figures of merit.
- As indicated by Fig. 3, the statistical nature of the noise had little impact on the optimum solution or value of $\lambda_{1,\text{optimum}}$ for the range of test cases considered. Instead, it was the statistical distribution of the response series itself that resulted in the favorable performance

Fig. 3. Relative performance of L_1 as opposed to L_2 norm minimization scheme at achieving a good (small) figure of merit Ψ . Tests were based on the synthetic data of Fig. 2 with added noise.

of the L_1 norm as opposed to L_2 . This indicated that future L_1 norm studies could reasonably be confined to a single form of noise.

- c) Values of SSD, Γ , and Ψ were all found to be smoothly dependent on λ_p for the case $p = 2$. In addition, this dependence was found to be relatively weak, e.g., Fig. 4(a). By contrast, the dependence of the three figures of merit on λ_p was seen to be far more strong and "choppy" for the case $p = 1$ (Fig. 4(b)). A geometrical interpretation for this difference was given in [2], in which it was shown that the choppiness in the L_1 norm minimization originates from the multi-faceted profile of its solution space boundary. It is therefore important that the selected value of $\lambda_{1,\text{optimum}}$ be not widely divergent from the ideal, or there will be significant degradation in the quality of the solution.

IV. SELECTION OF OPTIMUM DAMPING PARAMETER VALUE

In a search for the value of $\lambda_{1,\text{optimum}}$ to perform a deconvolution operation, the true response $\mathbf{h}$ will of course not be known; the figures of merit defined by (9) are therefore not applicable. The limited *a priori* knowledge of the general features of $\mathbf{h}$ was therefore used in the development of algorithms to identify $\lambda_{1,\text{optimum}}$ and the associated "best" or optimum solution $\hat{\mathbf{h}}$. In this section, one possible algorithm to determine $\lambda_{1,\text{optimum}}$ is presented.

From (3), it is evident that for a damping parameter greater than a critical value $\lambda_{1,\text{max}}$, the response vector $\hat{\mathbf{h}}$ will be completely damped out, yielding the trivial solution $\hat{h}_i = 0$. This leads to the definition for a normalized damping parameter that would be an aid to the display and analysis of results

$$\bar{\lambda}_1 = \frac{\lambda_1}{\lambda_{1,\text{max}}} \quad (11)$$

Tests on synthetic data indicated that for RMS noise magnitudes up to 20%, the value of $\lambda_{1,\text{max}}$ is only weakly dependent on the noise level. The dependence of the L_1 norm of the calculated solution vector $\hat{\mathbf{h}}$ and error vector $\mathbf{e}$ on the

Fig. 4. Sum of squared deviations (SSD) as a function of damping parameter for various levels of Laplacian noise. For each noise level, the "best" value to use for the damping parameter according to the SSD figure of merit is the value that yields the lowest SSD value: (a) L_2 norm minimization routine; (b) L_1 norm minimization routine.

normalized damping parameter is displayed in Fig. 5(a)–(b), respectively. A key feature of Fig. 5(a) is the choppy dependence of the L_1 norm solution set on $\bar{\lambda}_1$. This choppy dependence yields a number of *solution regimes*. Inside each regime, the solution $\hat{\mathbf{h}}$ to (5) and (6) and its L_1 norm $\|\hat{\mathbf{h}}\|_1$ are relatively stable across a range of damping parameter values. A similar trend is noted for the L_1 norm of the error vector $\|\mathbf{e}\|_1$ in Fig. 5(b). In both figures, the solution regimes are linked together by sharp discontinuities, labeled "transitions." It can be seen that the transitions become somewhat less distinct with increasing noise levels; however, the major solution regimes are still clearly visible for the cases of up to 20% RMS noise considered here. The prominence of these solution transitions is also evident in the profiles of the response series $\hat{h}_i$, illustrated by Fig. 6(a)–(b) for the cases of zero and 5% RMS added Laplacian noise.

The calculated response series of Fig. 6(a) for noise-free data indicate the presence of a damping parameter value $\lambda_{1,\text{min}}$ below which the true impulse response is achieved by solving (3). Calculations performed on these same data with various amounts of noise added to $\mathbf{y}$ show that the value of $\lambda_{1,\text{min}}$ (or its normalized equivalent $\bar{\lambda}_{1,\text{min}}$) is only weakly dependent on

Fig. 5. L_1 norm parametric plots for synthetic data with added Laplacian noise: (a) Norm of calculated solution $\hat{\mathbf{h}}$; (b) norm of corresponding calculated error vector $\mathbf{e}$.

noise level. Furthermore, the optimum value for λ_1 lies within a narrow range of values that are just below $\lambda_{1,\min}$, as gauged by all three figures of merit. This point is illustrated in Fig. 7 for the Ψ figure of merit; the figure indicates that an accurate estimate of $\lambda_{1,\min}$ would be a valuable aid to the determination of a suitable $\lambda_{1,\text{optimum}}$. The error bars indicate the standard deviation of results, based on a number of separate realizations of noise superimposed on the vector $\mathbf{y}$ of Fig. 2(c).

Consideration of Figs. 5 and 7 indicates that a major transition is seen in $\|\hat{\mathbf{h}}\|_1$ at damping parameter values in the vicinity of $\lambda_{1,\min}$. Determining the location of this transition could then be a useful indicator of the numerical value of $\lambda_{1,\min}$. This might be accomplished by a visual examination of the solution set, but this introduces an undesirable degree of subjectivity into the process. A more objective method for estimating $\lambda_{1,\min}$ is to define a *transition derivative* parameter $T_D(\lambda_1)$ (or $T_D(\bar{\lambda}_1)$) whose value is highly sensitive to the “jumps” noted in Fig. 5. In normalized format,

$$T_D(\bar{\lambda}_1) = \frac{\partial \|\hat{\mathbf{h}}\|_1}{\partial \bar{\lambda}_1} \times \frac{\partial \|\mathbf{e}\|_1}{\partial \bar{\lambda}_1} \quad (12)$$

where the partial derivatives are estimated by a forward differencing operation on the calculated solution set and corresponding error vectors.

Fig. 8 shows T_D plotted as a function of normalized damping parameter for the case of synthetic data superimposed with 5% Laplacian noise. For values of $\bar{\lambda}_1$ less than $\bar{\lambda}_{1,\min}$, the solution should be highly stable and therefore T_D should

Fig. 6. Calculated impulse response series $\hat{\mathbf{h}}$ as a function of normalized damping parameter using L_1 norm deconvolution for synthetic data: (a) Noise-free data; (b) 5% RMS Laplacian noise added to data.

Fig. 7. Optimum value of L_1 norm damping parameter corresponding to Ψ figure of merit, calculated as a function of Laplacian RMS noise level superimposed on synthetic data. Data points indicate average values and standard deviations corresponding to multiple realizations of noise.

be small. This corresponds to the region $0.1 < \bar{\lambda}_1 < 0.3$. (For values of $\bar{\lambda}_1$ less than 0.1, large values are seen in T_D as the response series $\hat{h}_i$ is forced to fit the noisy data; this noise-controlled regime need not be included in the search for $\lambda_{1,\text{optimum}}$). The upper bound of this region is marked by a clearly-identifiable large spike in T_D , at $\bar{\lambda}_1 \approx 0.33$. This value is then taken as the estimate for $\bar{\lambda}_{1,\min}$, and also the estimate

Fig. 8. Transition parameter T_D shown as a function of damping parameter for synthetic data superimposed with 5% Laplacian noise.

for $\bar{\lambda}_{1,\text{optimum}}$. The associated response vector is then taken as “best” available estimate for $\mathbf{h}$.

It is noted that there are limitations to this algorithm, in that this estimate realized for $\lambda_{1,\text{optimum}}$ (or its normalized counterpart $\bar{\lambda}_{1,\text{optimum}}$) might more accurately be regarded as its upper bound. However, the technique does represent a distinct improvement over a subjective visual examination of $\hat{\mathbf{h}}$ versus λ_1 in search of the optimum solution.

Several variations in the definition of the transition factor were explored, but all were based on locating the points where $\|\hat{\mathbf{h}}\|_1$ and $\|\mathbf{e}\|_1$ undergo significant discontinuities. For the range of data examined, all transition factors explored yielded similar results.

V. APPLICATION OF L_1 NORM DECONVOLUTION TO TEST CASES

A. Synthetic Test Data With Added Noise

In this section, the transition parameter T_D will be employed in estimates of $\lambda_{1,\text{optimum}}$ for the synthetic data of Fig. 2 with added Laplacian noise. Results will be assessed according to the determined values of optimum damping parameter value and $\hat{h}_i$, compared to those calculated by minimization of the figures of merit of (9).

Fig. 7 shows the values of $\bar{\lambda}_{1,\text{optimum}}$ as a function of RMS noise level; these data were determined by locating the minimum in Ψ , using *a priori* knowledge of the true solution $\mathbf{h}$. In comparison, Fig. 9 also shows the dependence of values of $\bar{\lambda}_{1,\text{optimum}}$ on noise level; however, these damping parameter values were determined without use of the true solution $\mathbf{h}$, but instead using T_D (Fig. 9(a)) and visual examination (Fig. 9(b)) of the solution set. The following conclusions are reached from analysis of Figs. 7 and 9:

- The optimum value of $\bar{\lambda}_1$ for this set of data generally lies in the range of 0.20-0.30 for Laplacian RMS noise levels of up to 20%. This estimate is based on the figure of merit Ψ used to generate Fig. 7. It is noted that this range of values is distinctly below the value of $\bar{\lambda}_{1,\text{min}} = 0.33$ for the noise-free data.
- In the search for $\bar{\lambda}_{1,\text{optimum}}$ where $\mathbf{h}$ was unknown, estimates based both on visual examination of the so-

(a)

(b)

Fig. 9. Optimum value of L_1 norm damping parameter, calculated as a function of RMS noise level superimposed on synthetic data. Data points indicate average values and standard deviations corresponding to multiple realizations of noise. Unlike Fig. 4, these estimates of $\bar{\lambda}_{1,\text{optimum}}$ were made without explicit knowledge of the true solution $\mathbf{h}$, using the following schemes: (a) Transition parameter T_D ; (b) visual examination for the most “visually satisfactory” response series $\hat{\mathbf{h}}$.

lution set (Fig. 9(b)) and identification of the first major spike in T_D (Fig. 9(a)) yielded values similar to those observed in Fig. 7 that were largely independent of noise level at up to 20% RMS noise. This independence is due to the increase in the error norm term with increasing noise levels; this leads to a corresponding increase in $\lambda_{1,\text{max}}$, which is then used to normalize the damping parameter.

B. Idealized Experimental Data

In preparation for application of L_1 norm deconvolution techniques to practical problems, a test was performed in which the data were acquired from an idealized experimental setup: The ultrasonic signal $\mathbf{y}$ reflected by a pyrex plate in a normal-beam, immersion, pulse-echo setup was deconvolved with the incident waveform $\mathbf{x}$. The unfocused transducer had a nominal central frequency of 30 MHz. The pyrex plate was only 205 microns thick, such that the echo received back at the transducer contained multiple pulses from internal reverberations between the faces of the plate.

Fig. 10. Discretized signals corresponding to idealized immersion ultrasonic test, where a normal beam signal is reflected from a pyrex plate: (a) Input signal; (b) output signal.

The input x and output y are shown in Fig. 10; both signals are seen to have low noise levels. Fig. 11 shows the results of calculating $\hat{h}$ by various deconvolution techniques. Fig. 11(a) is the theoretical impulse response function of the plate for a normally incident plane compression wave; there would be no mode conversion to a shear component in this idealized model. Fig. 11(b) illustrates the relatively poor performance of the L_2 norm minimization at handling this type of problem. (In L_2 norm minimization, there are no features in the solution set that can be used as unbiased guides to selecting the optimum value of damping parameter. Therefore, for all L_2 norm results given in this study, a value of $\bar{\lambda}_{2,\text{optimum}}$ was selected to yield the most visually satisfactory response series from the solution set.)

In contrast to the L_2 norm, good results were achieved using L_1 norm minimization when either T_D or visual examination were employed to obtain $\bar{\lambda}_{1,\text{optimum}}$ (Fig. 11(c)–(d)). The corresponding values obtained for $\bar{\lambda}_{1,\text{optimum}}$ of 0.77 and 0.51, respectively, were sufficiently close together that very similar profiles for the optimum solution vector were obtained. It is noted that $\hat{h}$ in these last two figures shows a sequence of satellite pulses that are absent in the idealized output of Fig. 11(a); these extra pulses correspond to the slow-moving shear wave components inside the pyrex. The shear waves originate from components of the incident wave that were not perpendicular to the pyrex plate, and therefore underwent mode conversion at the plate-water interface. Such mode conversion is inevitable in any real experiment, as a

transducer of finite diameter cannot generate the idealized plane compression wave assumed in the numerical generation of Fig. 11(a).

VI. APPLICATION OF L_1 NORM DECONVOLUTION TO ENGINEERING PROBLEMS

As a final test, L_1 norm deconvolution was applied to two ultrasonic NDE problems encountered in the maintenance of CANDU nuclear reactors. In both cases, it was desired to improve the temporal resolution of spikes in the impulse response of the inspected system beyond that which could be achieved by conventional signal processing techniques. It is noted that noise in such signals is inevitable due to the high radiation fields and correspondingly long signal cables, plus the electrically noisy environment of a power generation station.

A. Water Gap Measurement

An important objective in the nondestructive examination of CANDU reactors is the measurement of the water-filled gap between two concentric cylinders called the calandria tube and pressure tube [21]. A mock-up of the system, used to calibrate equipment that will go into the nuclear reactor, is shown in Fig. 12. A broadband ultrasonic signal is directed at normal incidence to the cylinder assembly in a pulse-echo configuration. Provided the water gap is large, reflections from the pressure tube are well-separated from those originating from ringing inside the calandria tube wall. Accurate time-of-flight measurements of the spacing between the two cylinders can be made visually in such cases by the difference in travel time of rays A and B in Fig. 12. However, if the water-gap is small, then there is overlap of rays A and B with internal reflections from inside the calandria tube; some temporal resolution enhancement using deconvolution is then required.

The system geometry of Fig. 12 was simulated using a pyrex plate with the same through-thickness acoustic travel time as a calandria tube. The use of the flat plate eliminated the complication of beam diffraction that would be caused by the calandria tube wall; alternatively, the curvature effects could be addressed by using a contact transducer with an appropriately curved surface. Unlike visually-analyzed time-of-flight measurements where distortion of pulses is of minor consequence, proper consideration of distortional effects on the pulse shape is essential where deconvolution techniques are to be applied.

Tests were performed using a 6.35-mm diameter contact probe with 5 MHz nominal central frequency. (Noise problems associated with the nuclear reactor environment hamper attempts to use very high frequency probes for which temporal resolution would be expected to be better.) The output signal y was acquired using a set-up corresponding to Fig. 12 with a water-gap of 800 μm . The input signal x used in the deconvolution procedure was obtained from the frontwall reflection of a smooth flat plate. The two signals were acquired using the same effective probe-to-reflector distance. Both x and y are shown in Fig. 13.

Fig. 11. Response series $\hat{h}$ corresponding to various deconvolution schemes for idealized experimental data of Fig. 10: (a) Theoretical response series; (b) typical L_2 response ($\lambda_2 = 5$); (c) L_1 norm response indicated by T_D , shown for $\bar{\lambda}_{1,\text{optimum}} = 0.77$; (d) L_1 norm response indicated by visual inspection, shown for $\bar{\lambda}_{1,\text{optimum}} = 0.51$.

Fig. 12. Experimental geometry used to simulate an ultrasonic water gap measurement in a CANDU reactor.

The calculated impulse response $\hat{h}$ is shown in Fig. 14, using three different deconvolution techniques. Fig. 14(a)–(b) show the results of L_1 norm deconvolution, using visual inspection and T_D , respectively, to obtain $\lambda_{1,\text{optimum}}$. Both techniques led to clear visualization of spikes originating from ringing within the pyrex plate, as well as the two spikes of interest corresponding to rays A and B in Fig. 12.

Fig. 14(c) shows the results of L_2 norm deconvolution for the same input data. The temporal resolution is far worse than

for the L_1 norm, and the background noise in the result could even inhibit correct identification of the various spikes in $\hat{h}$.

B. Zr-2.5% Nb Shear Wave Birefringence Measurement

The uptake of hydrogen by CANDU reactor pressure tubes can lead to potential shortening of tube life [22]. A nondestructive measurement of the hydrogen concentration of these tubes would therefore be a great asset to an evaluation of their condition. As a minimum, it is desired to measure the percentage of the Zr-2.5% Nb material that is sufficiently saturated with hydrogen that it has led to a perturbed form of its crystal structure called zirconium hydride. The challenge to such measurements stems from the fact that the hydrided metal has very similar acoustic properties to the hydride-free material. Although straight-forward time-of-flight techniques may have sufficient sensitivity to detect the small perturbation in acoustic velocity associated with hydride formation, there is sufficient uncertainty in the tube wall thickness to render such techniques impractical.

One possible solution to this problem is based on shear wave velocity birefringence measurements. Radially propagating shear waves polarized axially along the tube ($V_{RL} = 2290 \pm 20$ m/s) have a slightly smaller velocity than those polarized circumferentially ($V_{RC} = 2390 \pm 20$ m/s) [23]. This difference is due to the anisotropic crystal structure

Fig. 13. Discretized signals for an ultrasonic water gap measurement: (a) Input signal; (b) output signal.

originating from the extrusion and cold drawing operations during the tubes' manufacture. Hydriding the metal causes significantly different perturbations in V_{RC} and V_{RL} , such that the difference between these two wave velocities would be a possible indicator of hydrided volume fraction. Even though both V_{RC} and V_{RL} would be obtained by time-of-flight measurements, the *difference* in estimates of the two wave velocities would not be strongly sensitive to the small uncertainty in wall thickness.

As a first step toward adoption of this concept, it is useful to determine whether time-of-flight data can be acquired corresponding to both shear wave polarizations for a piece of nonhydrided Zr-2.5% Nb pressure tube material of standard thickness (approximately 4.5 mm). A simple normal beam pulse-echo immersion geometry was used for these tests. The broadband transducer had a nominal central frequency of 10 MHz, and a spherical focusing distance in water of 25.4 mm. A focused transducer was used to enhance the extent of mode conversion to shear waves at the tube-water interface. Two complications in the test procedure needed to be addressed:

- a) Similar to the case of water gap measurement, it is possible that the curvature of the pressure tube wall would cause sufficient diffraction and distortion of the beam to seriously corrupt the results. (Radius of curvature of the pressure tube wall is approximately 5.5 cm.) Two sets of experiments were therefore performed, one set using a piece of curved pressure tube with the transducer located on the inside, and the other set using a flat plate of the same material and thickness as the tube section. The two sets of results were similar, such that only the set

Fig. 14. Response series $\hat{h}$ corresponding to various deconvolution schemes for ultrasonic water gap measurement data of Fig. 11: (a) L_1 norm response with $\lambda_{1,\text{optimum}} = 0.59$ based on visual inspection; (b) L_1 norm response with $\lambda_{1,\text{optimum}} = 0.57$ based on T_D ; (c) L_2 norm response with $\lambda_2 = 5$.

corresponding to the curved pressure tube wall will be presented here.

- b) In the evaluation of the pressure tube wall's impulse response, it is necessary to identify the spikes of interest corresponding to the two shear wave polarizations. Profiles for $\hat{h}$ determined by several deconvolution techniques are shown in Fig. 15. In each of the graphs of this figure, the first two positive spikes correspond to the birefringent shear waves of interest. The next major spike is a multiple longitudinal backwall reflection. The train of spikes on the right side of each graph is due to another pair of birefringent shear waves plus another multiply-reflected pulse. For each set of results, this se-

Fig. 15. Response series $\hat{h}$ corresponding to various deconvolution schemes for ultrasonic birefringence data: (a) L_1 norm response with $\lambda_{1,\text{optimum}} = 0.66$ based on visual inspection; (b) L_1 norm response with $\lambda_{1,\text{optimum}} = 0.60$ based on T_D ; (c) L_2 norm response with $\lambda_2 = 5$.

quence of pulses was used to identify the first two spikes upon which the birefringent velocity measurements were based.

Comparison of the profiles in the three graphs of Fig. 15 provides further confirmation of the trends noted earlier. Selection of $\lambda_{1,\text{optimum}}$ by visual inspection and the transition parameter T_D yielded very similar results, as shown in Fig. 15(a)–(b). Results using L_2 norm deconvolution, shown in Fig. 15(c), are sufficiently degraded that resolution of the two shear wave pulses is severely inhibited.

From the profile of $\hat{h}$ in Fig. 15(b), which was derived using the transition parameter T_D , estimates of the two shear wave velocities were obtained corresponding to the nonhybridized

metal: $V_{RL} = 2290 \pm 30$ m/s, and $V_{RC} = 2390 \pm 30$ m/s. These are virtually identical to values of 2290 ± 20 m/s and 2390 ± 20 m/s, respectively, obtained from independent research programs [23].

VII. CONCLUSION

Methods for applying L_1 norm minimization techniques to the deconvolution of a sparse spike time series have been explored. Implementation of the technique was by the Simplex linear programming method. Where applicable, results have been compared to those achievable by L_2 norm deconvolution—a technique which continues to enjoy widespread use. Major findings are as follows:

- A transition parameter $T_D(\bar{\lambda}_1)$ was introduced as an aid to selecting the best damping parameter value $\lambda_{1,\text{optimum}}$ for L_1 norm deconvolution. For the range of synthetic and experimental data considered, it was found that $T_D(\lambda_1)$ was effective in obtaining a highly resolved response vector $\hat{h}$ believed to closely approximate the true solution h . In all cases, it was found that estimates of $\lambda_{1,\text{optimum}}$ achieved through T_D were very similar to those achieved by subjective visual examination of the solution set.
- The L_1 norm deconvolution technique was successfully applied to two practical problems encountered in the ultrasonic NDE of CANDU reactor pressure tube components. The low noise level in the calculated response aided both in the identification of each spike's origin, and in the temporal resolution of adjacent spikes.
- Consistent with theoretical predictions, L_2 norm minimization techniques were shown to be markedly less suitable than L_1 at recovering a sparse spike time series. Although the computation time for a single application of L_1 norm minimization was found to be a factor of 100 larger than for L_2 , the need for parametric programming in the search for $\lambda_{p,\text{optimum}}$ reduced this difference to a factor of only 3 to 4 for a solution set of 100 vectors.
- It is noted that only RMS noise levels of up to 20% were considered in this study; this was consistent with the relatively low noise levels in the ultrasonic NDE signals under investigation. Further work would be required to evaluate the use of the transition parameter at higher noise levels. Under such conditions, the form of the objective function Φ should also be re-examined for replacement of the L_1 norm of the error vector by the L_2 norm.

REFERENCES

- J. B. Bednar, R. Yarlagadda, and T. Watt, " L_1 deconvolution and its application to seismic signal processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 6, pp. 1655–1658, Dec. 1986.
- M. S. O'Brien, A. N. Sinclair and S. Kramer, "High resolution deconvolution using least absolute values minimization" in *Proc. IEEE 1990 Ultrasonics Symp.*, 1991, pp. 1151–1156, vol. 2.
- J. Claerbout and F. Muir, "Robust modeling with erratic data," *Geophys.*, vol. 38, no. 5, pp. 826–844, 1973.
- R. Yarlagadda, J. B. Bednar, and T. Watt, "Fast algorithms for L_p deconvolution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 1, pp. 174–182, 1985.

- [5] J. B. Bednar, R. Yarlagadda, and T. Watt, " L_1 deconvolution and its application to seismic signal processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 6, pp. 1655-1658, 1986.
- [6] H. Taylor, S. Banks, and F. McCoy, "Deconvolution with the L_1 norm," *Geophys.*, vol. 44, no. 1, pp. 39-52, 1979.
- [7] B. Drachman, "Two methods to deconvolve: L_1 -method using Simplex algorithm and L_2 -method using least squares and a parameter," *IEEE Trans. Antenn., Propagat.*, vol. TAP-32, no. 3, pp. 219-225, 1984.
- [8] N. Chapman and I. Barrodale, "Deconvolution of marine seismic data using the L_1 norm," *Geophys. J. R. Astronom. Soc.*, vol. 72, pp. 93-100, 1983.
- [9] I. Barrodale, C. Zala, and N. Chapman, "Comparison of the L_1 and L_2 norms applied to one-at-a-time spike extraction from seismic traces," *Geophys.*, vol. 49, no. 11, pp. 2048-2052, 1984.
- [10] C. A. Zala, "Enhancement of NDE ultrasonic images by high resolution deconvolution," *Defense Res. Establishment Pacific*, Tech. Rpt. W7708-6-7245, Department of National Defense, Canada (1987).
- [11] C. Chen and S. Sin, "On effective spectrum-based ultrasonic deconvolution techniques for hidden flaw characterization," *J. Amer. Stat. Assn.*, vol. 87, no. 3, pp. 976-987, 1990.
- [12] C. A. Zala, I. Barrodale, and J. Kennedy, "High-resolution signal and noise field estimation using the L_1 (least absolute values) norm," *IEEE J. Oceanic Eng.*, vol. 12, no. 1, pp. 253-264, 1987.
- [13] J. R. Reeves and R. M. Mersereau, "Optimal estimation of the regularization parameter and stabilizing functional for regularized image restoration," *Optical Eng.*, vol. 29, no. 5, pp. 446-454, 1990.
- [14] J. Mendel, *Maximum-Likelihood Deconvolution: A Journey into Model-Based Signal Processing*. New York: Springer-Verlag, 1990.
- [15] J. J. Kormylo and J. M. Mendel, "Maximum-likelihood seismic deconvolution," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-21, no. 1, pp. 72-82, 1983.
- [16] A. Swami and J. M. Mendel, "ARMA parameter estimation using only output cumulants," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, pp. 1257-1265, 1990.
- [17] F. A. Ficken, *The Simplex Method of Linear Programming*. New York: Holt, Reinhart, and Winston, 1961.
- [18] S. I. Gass, *Linear Programming: Methods and Applications*. New York: McGraw-Hill, 5th ed., 1985.
- [19] S. Levy and P. Fullagar, "Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution," *Geophys.*, vol. 46, no. 9, pp. 1235-1243, 1981.
- [20] D. Oldenburg, S. Levy, and K. Stinson, "Inversion of band-limited reflection seismograms: Theory and practice," *Proc. IEEE*, vol. 74, no. 3, pp. 487-497, 1986.
- [21] M. P. Dolbey and G. N. Jarvis, "Inspection and monitoring of CANDU reactor fuel channels," tech. supplement to *CNS Bulletin*, Canadian Nuclear Society, Spring 1990.
- [22] D. O. Northwood and U. Kosasih, "Hydrides and delayed hydrogen cracking in zirconium and its alloys," *Int. Metals Rev.*, vol. 28, no. 2, pp. 92-121, 1983.
- [23] H. D. Mair and Earl, "Zr-2.5% Nb pressure tube ultrasonic data," CANDU Owner's Group Rpt. COG 90-247-P, 1990.

Michael S. O'Brien obtained the B.S. degree in engineering science in 1988 and the M.A.Sc. in mechanical engineering in 1991, specializing in digital signal processing for nondestructive evaluation. His research interests are in the areas of geophysical inverse theory in geomagnetism, and construction of geomagnetic field models which satisfy the frozen flux assumption for fluid flow in the earth's outer core. He is currently pursuing the doctoral degree at the Department of Scripps Institution of Oceanography at the University of California at San Diego.

Anthony N. Sinclair (M'89) obtained the B.S. degree in engineering science at the University of Toronto in 1976, and the M.S. and Ph.D. degrees in nuclear engineering at the University of Michigan in 1978 and 1980, respectively.

He is currently an Associate Professor in the Department of Mechanical Engineering at the University of Toronto. He has published extensively in the areas of acoustic wave propagation and ultrasonic nondestructive evaluation, particularly as applied to nuclear power plant safety.

Stuart M. Kramer received the B.Sc. degree in engineering physics and the M.Sc. degree in physics, both from Queen's University, Kingston, Canada.

Since 1983 he has been active in the research and development of ultrasonic digital signal processing techniques for nondestructive evaluation, first at the Royal Military College Department of Physics and later at Ontario Hydro Technologies. He is currently a safety specialist at Ontario Hydro, Whitby, Canada.