Nonlinear inverse scattering methods for thermal-wave slice tomography: a wavelet domain approach

Eric L. Miller

Department of Electrical and Computer Engineering, Northeastern University, 235 Forsyth Building, Boston, Massachusetts 02115

Lena Nicolaides and Andreas Mandelis

Photothermal and Optoelectronic Diagnostics Laboratory, Department of Mechanical Engineering, University of Toronto, 5 King's College Road, Toronto M5S 3G8, Ontario, Canada

Received July 30, 1997; revised manuscript received December 22, 1997; accepted January 12, 1998

A wavelet domain, nonlinear inverse scattering approach is presented for imaging subsurface defects in a material sample, given observations of scattered thermal waves. Unlike methods using the Born linearization, our inversion scheme is based on the full wave-field model describing the propagation of thermal waves. Multiresolution techniques are employed to regularize and to lower the computational burden of this ill-posed imaging problem. We use newly developed wavelet-based regularization methods to resolve better the edge structures of defects relative to reconstructions obtained with smoothness-type regularizers. A nonlinear approximation to the exact forward-scattering model is introduced to simplify the inversion with little loss in accuracy. We demonstrate this approach on cross-section imaging problems by using synthetically generated scattering data from transmission and backprojection geometries. © 1998 Optical Society of America [S0740-3232(98)00906-5]

OCIS codes: 100.7410, 290.0290, 170.6960.

1. INTRODUCTION

Thermal-wave slice tomography (TWST) has evolved in recent years as a useful tool for noninvasively imaging and detecting defects in the bulk properties of a material sample.[1–3] This nondestructive evaluation technique uses a modulated laser source illuminating an external surface of the material under test to induce internal thermal waves. Interaction of the thermal-wave field with material inhomogeneities gives rise to scattered fields that propagate and are ultimately measured at the surface of the material. The problem of interest in this paper is, given knowledge of the applied thermal-wave field and the observed scattered fields, to produce a reconstruction of the internal structure that reproduces as faithfully as possible features of interest such as defects.

The techniques that we use to solve this inverse problem are based on the results of Mandelis, who has shown that the thermal-wave field obeys a scalar Helmholtz equation with a complex-valued, space-varying propagation constant.[4,5] The spatial structure of this constant is related to the thermal diffusivity of the material. Because defects are reflected in changes in the thermal diffusivity, a reconstruction of the propagation constant, or a normalized form known as the object function,[1] yields quantitative information about the material's bulk structure.

Mathematically, the TWST inverse problem is equivalent to an inverse electrical conductivity problem that has been studied extensively in the geophysical and electromagnetics literature.[6–8] These ill-posed inverse problems possess a collection of interesting and well-established difficulties, not the least of which is that a complicated, nonlinear relationship exists between the observed scattered fields and the object function. The first efforts in employing the forward-scattering model of Refs. 4 and 5 for TWST inversion have been to consider a linearized form of the exact physics obtained under the first Born approximation.[1–3,9] This model takes into account diffractive effects and is most accurate when the thermal diffusivity perturbation is small in both size and amplitude relative to a known background.[10] The initial inversion results with the Born model have been quite encouraging in that images containing quantitatively useful information about the spatial location and magnitude of defects have been reconstructed with both simulated and experimental data.[1–3,9]

Here we extend the use of wave-field inversion methods for the TWST problem in a number of ways. First, a full, nonlinear inverse scattering approach[11,12] is used to generate the reconstruction. The resulting inversion method is iterative in nature and allows for the more accurate reconstruction of defects whose structure falls outside the bounds where the Born approximation is valid. We also make extensive use of wavelet methods in the formulation and the solution of the problem. Our motivation for employing these mathematical techniques is based on our previous work in wavelet-based nonlinear inverse scattering.[13] First, the wavelet transform is known to make sparse the matrix representations of many integral operators, including those arising in the TWST scattering

0740-3232/98/061545-12$15.00 © 1998 Optical Society of America

problem.[14] Thus a transform domain formulation can build on this sparsity to reduce the computational burden of generating a reconstruction. Additionally, the TWST inverse problem is highly ill posed in the sense discussed in Ref. 15. That is, small perturbations in the data, as would come from noise, can result in reconstructions with high-amplitude, oscillatory structure. Roughly speaking, these nonphysical artifacts arise because the forward-scattering process is highly smoothing, so that the data contain little useful information about the high-frequency structure of the object function. Attempting to extract such information without some stabilization (also known as regularization) results in distorted reconstructions.

Our previous work has demonstrated that wavelet methods can be employed in two respects for regularizing these inverse problems.[13,16] First, it was shown that traditional Tikhonov-type smoothness regularization schemes[17] could be implemented quite easily in a wavelet transform formulation. Second, by exploiting the structure of the wavelet regularizer, we developed an adaptive method for determining those elements in the wavelet transform of the object function for which the data did and did not provide useful information at each stage of the inversion. This work provided insight into the manner in which the physics of the problem, along with factors such as source–receiver configuration, noise level, and prior information, determined the finest scale as a function of space that could reasonably be obtained in the reconstruction. Additionally, we used this decomposition to reduce further the complexity of solving the inverse scattering problem.

Here we build on these results in a number of ways. First, we introduce and demonstrate the utility of a class of wavelet regularization schemes that are appropriate for the reconstruction of objects (defects) with sharply defined boundaries. Tikhonov-type regularizers are typically designed to produce smooth, low-pass reconstructions that blur important features such as edges in the image.[15] By making use of certain mathematical results that state that wavelets can be used as bases for a wide range of function spaces including spaces containing edgy objects,[18] we develop a new scale-space regularization method that produces significantly sharper reconstructions.

Additionally, we develop a wavelet-based, reduced-complexity approximation to the forward-scattering TWST problem. As discussed in Section 4, at each stage of the iterative solution method we must compute explicitly the inverse of a large, dense matrix, a computationally intensive task. This matrix is related to the solution of the thermal-wave forward-scattering problem for an object function equal to the reconstruction obtained at the previous iteration plus a small correction. By linearizing the expression for this matrix about the previous estimate, we obtain a recursive formula in which the inverse matrix at the current iteration is equal to the one from the last iteration plus an increment that is due to the change in the material properties. Construction of this increment requires only the multiplication of three sparse matrices rather than the inversion of a single, large, dense matrix.

The remainder of this paper is organized as follows. In Section 2 the mathematical model underlying the TWST problem is defined. Section 3 is devoted to an overview of the wavelet transform and its application to the thermal-wave tomographic inverse. The nonlinear inverse scattering algorithm is developed in Section 4, with examples of its application to synthetic problems provided in Section 5. Finally, in Section 6, conclusions and future work are discussed.

2. PHYSICAL MODEL FOR THERMAL-WAVE SLICE TOMOGRAPHY

As illustrated in Fig. 1, we consider TWST problems in which a modulated laser illuminates a point at the top of the material sample, inducing thermal waves in the bulk. Upon scattering from inhomogeneities, whose spatial structure is defined by the real-valued object function $g(\mathbf{r})$, the fields are measured along an array of points at either the top (backpropagation geometry) or the bottom (transmission mode). The inversion routines are based on K such scattering experiments. The data for the ith experiment form the vector of in-phase and quadrature components of the scattered thermal-wave field measurements that are obtained along one of the arrays and that are due to illumination by the laser at a given point on the top surface. One such data point for the ith experiment at location $\mathbf{r}_k$, denoted $y_i(\mathbf{r}_k)$, is

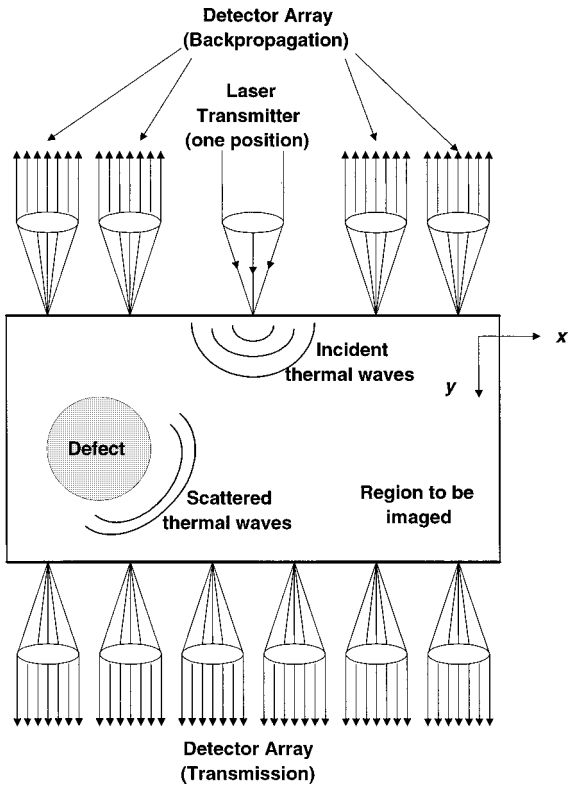


Fig. 1. Experimental setup for thermal-wave slice tomography. Incident thermal waves originating from a point on the top of the material sample interact with defects, giving rise to scattered fields, whose effects are measured by arrays located at the top and bottom of the sample. The objective of the inverse problem is to image the internal structure of the material on the basis of these measurements. The incidence point is generally scanned across the top.

$$y_i(\mathbf{r}_k) = \int_A G(\mathbf{r}_k, \mathbf{r}')T_i(\mathbf{r}')g(\mathbf{r}')\mathrm{d}\mathbf{r}' + n_i(\mathbf{r}_k), \quad (1)$$

where the integral on the right-hand side of Eq. (1) is the scattered field and $n_i(\mathbf{r}_k)$ is taken to be additive measurement noise. Region A is the area of space in which the defects might be found. The quantity $G(\mathbf{r}, \mathbf{r}')$ is the Green's function associated with the pseudowave Helmholtz equation for a homogeneous medium characterized by a complex propagation constant that is infinite in two dimensions and bounded by two planes in the third.[4,5] Finally, $T_i(\mathbf{r})$ is the thermal-wave field internal to A, which satisfies[4,5]

$$T_i(\mathbf{r}) = \bar{T}_i(\mathbf{r}) + \int_A G(\mathbf{r}, \mathbf{r}')T_i(\mathbf{r}')g(\mathbf{r}')\mathrm{d}\mathbf{r}', \quad (2)$$

where $\bar{T}_i(\mathbf{r})$ is the time-harmonic, incident thermal-wave field generated by the laser source at a given point on the top surface. The detailed expressions for $G(\mathbf{r}, \mathbf{r}')$ and $\bar{T}_i(\mathbf{r})$ for the planar geometry of interest here may be found elsewhere.[1–5]

Our algorithms are based on discrete representations of the integral equations (1) and (2) obtained by using the method of moments[19] with a pulse basis and Dirac testing functions. That is, region A is pixelated into an $N_y \times N_x$ array of rectangular pixels, and the fields and the object function are expanded in a series of flattop functions (i.e., zero-order splines), indicating that these quantities are constant over small pixels in region A. After these expansions are placed into Eqs. (1) and (2) and integration is performed with respect to $\mathbf{r}'$, the variable $\mathbf{r}$ is discretized in Eq. (2) by requiring equality of the left- and right-hand sides for all points $\mathbf{r}_j$ located at the center of each pixel. Upon performance of this discretization process, Eqs. (1) and (2) become

$$\mathbf{y}_i = \mathbf{L}_i \mathscr{D}(\mathbf{T}_i)\mathbf{g} + \mathbf{n}_i, \quad (3)$$

$$\mathbf{T}_i = \bar{\mathbf{T}}_i + \mathbf{G}\mathscr{D}(\mathbf{g})\mathbf{T}_i. \quad (4)$$

In Eq. (3) $\mathbf{y}_i$ is the vector of observations along the array for the ith experiment. If we regard $g(\mathbf{r})$ as an $N_y \times N_x$ pixelated image with the pixel value at row m and column n given by the flattop expansion coefficient $g_{m,n}$, then $[\mathbf{g}]_i$, the ith component of the vector $\mathbf{g}$, is related to the pixel values through the index mapping $i = N_x(n - 1) + m$ for $n = 1, 2, \ldots, N_x$ and $m = 1, 2, \ldots, N_y$. A similar construction holds for the internal field vector $\mathbf{T}_i$. In Eq. (4) $\bar{\mathbf{T}}_i$ is the background field vector obtained from point matching, $\mathscr{D}(\mathbf{x})$ is a diagonal matrix whose entries are the elements of the vector $\mathbf{x}$, and $\mathbf{L}_i$ and $\mathbf{G}$ are the matrices obtained by discretizing the integral kernels in Eqs. (1) and (2). $\mathbf{n}_i$ are taken to be mutually uncorrelated, zero-mean, white Gaussian noise vectors. By solving for $\mathbf{T}_i$ in Eq. (4) and substituting the result into Eq. (3), we relate the data to the object function through the nonlinear model

$$\mathbf{y}_i = \mathbf{h}_i(g) + \mathbf{n}_i, \quad (5)$$

$$\mathbf{h}_i(\mathbf{g}) = \mathbf{L}_i \mathscr{D}([\mathbf{I} - \mathbf{G}\mathscr{D}(\mathbf{g})]^{-1}\bar{\mathbf{T}}_i)\mathbf{g}. \quad (6)$$

The TWST problem may now be stated as follows: Given data from K scattering experiments defined by the physical model in Eq. (5), determine $\mathbf{g}$, the vector of expansion coefficients characterizing the object function.

There are two primary difficulties in recovering $\mathbf{g}$ from $\mathbf{y}_i$. The first challenge is caused by the physics of thermal-wave propagation. The strongly lossy nature of the Helmholtz pseudowave equation underlying Eq. (5) causes $\mathbf{h}_i$ to act essentially as a spatial low-pass filter when applied to $\mathbf{g}$. The data contain predominantly coarse-scale averages (i.e., low frequency) as information about $\mathbf{g}$. Finer-scale information is available primarily in the areas close to the source and receiver locations where the Green's functions are singular.[4,5,13] Thus attempts to reconstruct a uniform, fine-scale pixelated version of the object function are prone to instabilities, resulting in images that are typically characterized by nonphysical, oscillatory artifacts. As described more fully in Ref. 15, for linear inverse problems where $\mathbf{h}(\mathbf{g}) = \mathbf{H}\mathbf{g}$ for a matrix $\mathbf{H}$, such artifacts arise from noise-induced amplification associated with the inversion of small singular values of $\mathbf{H}$. The corresponding singular vectors typically possess an oscillatory structure. Such difficulties also arise in the nonlinear case such as the TWST problem considered here.

The second problem is computational. The nonlinearity of $\mathbf{h}_i$ implies that an iterative, "hill-climbing" approach must be used to generate the reconstruction. As described in Section 4, such an approach requires the explicit inversion of $\mathbf{I} - \mathbf{G}\mathscr{D}(\mathbf{g})$ at each iteration, with $\mathbf{g}$ equal to the current estimate of the object function. Because this matrix is both large and dense, this operation represents a substantial computational burden.

3. WAVELET DOMAIN MODEL

We pursue a wavelet-based solution to the TWST problem to address both the computational and stability problems. The expansion of $\mathbf{g}$ in a wavelet basis provides a natural mechanism for adapting the level of detail in the reconstruction to the information content in the data, thereby stabilizing the solution procedure.[13,16] For example, near the center of A, one may desire only a coarse-scale estimate of $\mathbf{g}$, with added detail near the edges, where the sharply peaked nature of the Green's functions near the source and receiver locations provides the additional information. Also, the wavelet domain representations of $\mathbf{L}_i$ and $\mathbf{G}$ are sparse, thereby lowering the computational costs associated with their manipulation. In the remainder of this section we provide an overview of the wavelet transform and describe the scale-space representation of the TWST problem. For simplicity, we limit most of the discussion to the one-dimensional (1D) case. Extensions to multiple dimensions are obtained through the use of separable transforms.[20]

Like the Fourier series, a wavelet series represents a (square integrable) function as the superposition of a set of orthonormal functions. Whereas the Fourier basis employs complex exponentials, the wavelet basis consists of all dyadic dilations and shifts of a single wavelet function $\psi(x)$, generally designed to be well localized in space. Thus the wavelet expansion for a 1D function $a(x)$ is

$$a(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \alpha_{j,k} \psi_{j,k}(x), \tag{7}$$

where $\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)$ is the wavelet function at scale j and shift k and the wavelet coefficient $\alpha_{j,k}$ is the inner product of a with $\psi_{j,k}$. Similar to the complex exponentials, the wavelet functions form an orthonormal basis of the space of square integrable functions. In Fig. 2 we plot a number of the Daubechies four-tap wavelet basis functions at different scales and positions. This figure illustrates that as j increases, the wavelet coefficients represent the inner product of $a(x)$ with increasingly compressed basis functions and therefore convey localized, fine-scale/high-frequency information about a near the point $2^{-j}k$.

For most problems we deal with truncated versions of the infinite sums. To truncate the scale index, we introduce the scaling function $\phi(x)$, which represents the coarse-scale information from $j = -\infty$ to some arbitrary, finite, coarse scale that we label $j = 0$. Because the $\phi_{0,k}$ are orthonormal to one another, as well as to the wavelets $\psi_{j,k}(x)$, Eq. (7) becomes

$$a(x) = \sum_{k=-\infty}^{\infty} a_{0,k}\phi_{0,k}(x) + \sum_{j=0}^{F_a} \sum_{k=-\infty}^{\infty} \alpha_{j,n}\psi_{j,n}(x), \tag{8}$$

where the scaling coefficients $a_{0,k}$ are the inner products of $\phi_{0,k}(x)$ with $a(x)$ and we have assumed that $a(x)$ is scale limited to $j = F_a$. Essentially, in this finite-scale case, the scaling functions capture the low-frequency behavior (including dc) of $a(x)$. The truncation of k required for considering functions defined on a bounded subset of the real line is achieved either through a process of periodization[21] or by using special edge wavelets built for multiscale expansion on an interval.[22] Following our previous work,[13,16] here we use the edge wavelets approach as such an approach that avoids any wrapping effects caused by linking one end of the interval to the other.

Like the Fourier case, there is an orthonormal discrete wavelet transform (DWT) that takes a vector of coefficients representing a function at some fine scale F_a into the wavelet coefficients at all scales $0 \leq j \leq F_a - 1$ along with the coarse scale scaling coefficients, $a_{0,k}$. The initial, fine-scale functions could be samples of the function, expansion coefficients $a_{F_a,k}$ in a fine scale, scaling function expansion of $a(x)$, etc. Although we will be concerned with both 1D and two-dimensional (2D) signals, we describe first the mechanics of the 1D DWT. We define $\mathbf{a}$, of dimension 2^{F_a}, to be the vector of fine-scale coefficients and will denote this vector by $\mathbf{a}(M_a)$, indicating that it is a representation of $\mathbf{a}$ at the finest scale M_a.

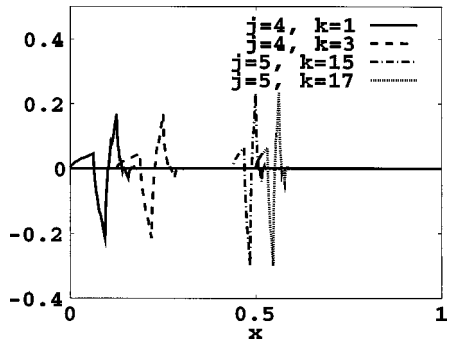


Fig. 2. Plots of Daubechies four-tap wavelet basis functions: $\psi_{4,1}(x)$ (solid curve), $\psi_{4,3}(x)$ (dashed curve), $\psi_{5,13}(x)$ (dotted–dashed curve), $\psi_{5,17}(x)$ (dotted curve).

Beginning with $\mathbf{a}(M_a)$, a coarser approximation $\mathbf{a}(M_a - 1)$ is obtained by passing $\mathbf{a}(M_a)$ through a low-pass, finite impulse response filter and decimating the filtered output by a factor of 2. Thus $\mathbf{a}(M_a - 1)$ is coarser than $\mathbf{a}(M_a)$ in that the filtering and downsampling procedure has removed the high-frequency structure from the original signal, and $\mathbf{a}(M_a - 1)$ is half as long as $\mathbf{a}(M_a)$. The detail lost in moving from $\mathbf{a}(M_a)$ to $\mathbf{a}(M_a - 1)$ [denoted $\boldsymbol{\alpha}(M_a - 1)$] is extracted by a high-pass filtering and decimation procedure. The filtering and decimation process is applied successively to the coarsened versions of $\mathbf{a}$, resulting in a sequence of scaling coefficient and detail vectors, $\mathbf{a}(m)$ and $\boldsymbol{\alpha}(m)$, respectively, each of dimension 2^m, for $m = M_a - 1, ..., 0$.

Low- and high-pass filters, which are closely related to the functions $\phi(x)$ and $\psi(x)$, can be constructed so that we may build a unitary matrix $\mathbf{W}_a$ relating the finest-scale scaling coefficients to the coarsest scaling coefficients and all detail coefficients.[23] We note that not all wavelet transforms result in unitary DWT matrices. In particular, so-called biorthogonal DWT's possess a variety of advantages over their orthonormal counterparts[24]; however, in this paper, we consider only the orthonormal DWT. We subscript the wavelet transform operator as $\mathbf{W}_a$ to make explicit that it is the transform for $\mathbf{a}$. We will use different wavelet transforms for the different variables.

The DWT of a 2D function is obtained by considering the image $a(m,n)$ as a matrix and applying one orthonormal wavelet transform, $\mathbf{W}_{a,x}$, to the columns and another, $\mathbf{W}_{a,y}$, to the rows. If we define $\check{\mathbf{a}}$ as the physical-space image—that is, $[\check{\mathbf{a}}]_{m,n} = a(m, n)$—and $\check{\boldsymbol{\alpha}}$ as the 2D DWT of $\check{\mathbf{a}}$, then the two are related according to

$$\check{\boldsymbol{\alpha}} = \mathbf{W}_{a,y}\, \check{\mathbf{a}} \mathbf{W}_{a,x}^{\mathrm{T}}. \tag{9}$$

Note that $\check{\boldsymbol{\alpha}}$ in Eq. (9) may be regarded as a multiscale image of $\check{\mathbf{a}}$. The columns and the rows of $\check{\boldsymbol{\alpha}}$ are independently indexed by scale/shift indices for the x and y variables, respectively. Thus the elements of $\check{\boldsymbol{\alpha}}$ represent wavelet coefficients for discrete, separable, 2D basis functions consisting of wavelets at scale/shift (j_x, k_x) in the horizontal direction and (j_y, k_y) in the vertical direction for all allowable combinations of these four indices. Additionally, there are components of $\check{\boldsymbol{\alpha}}$ corresponding to coarse-scale, x-oriented scaling functions with vertical wavelets, and vice versa. Finally, we note that Eq. (9) is a linear transformation of $\check{\mathbf{a}}$. Thus we define vectors $\mathbf{a}$ and $\boldsymbol{\alpha}$ obtained by stacking one column on top of another from $\check{\mathbf{a}}$ and $\check{\boldsymbol{\alpha}}$, respectively, and it is easily shown that

$$\boldsymbol{\alpha} = \mathbf{W}_a \mathbf{a}.$$

The matrix $\mathbf{W}_a$ is the linear operator mapping elements of $\mathbf{a}$ into elements of $\boldsymbol{\alpha}$ and represents the composition of $\mathbf{W}_{a,x}$ and $\mathbf{W}_{a,y}$. As in the 1D case, $\mathbf{W}_a$ as defined above is

unitary, so that $\mathbf{a}$ may be obtained from $\boldsymbol{\alpha}$ through multiplication by the transpose of $\mathbf{W}_a$: $\mathbf{a} = \mathbf{W}_a{}^{-1}\boldsymbol{\alpha} = \mathbf{W}_a^{\mathrm{T}}\boldsymbol{\alpha}$. Thus, when this column-stacking approach is followed, the mechanics and the symbolic manipulation of multidimensional wavelet transforms are identical to those in the 1D case.

We use the DWT to transform Eqs. (3) and (4) from physical space to scale space. Defining $\mathbf{W}_g$ as the 2D transform for $\mathbf{g}$ and $\mathbf{W}_i$ for $i = 1, 2, \ldots, K$ as the 1D transforms for $\mathbf{y}_i$ yields

$$\mathbf{W}_i\mathbf{y}_i = [\mathbf{W}_i\mathbf{L}_i\mathbf{W}_g^{\mathrm{T}}][\mathbf{W}_g\mathscr{D}(\mathbf{T}_i)\mathbf{W}_g^{\mathrm{T}}](\mathbf{W}_g\mathbf{g}) + \mathbf{W}_i\mathbf{n}_i,$$

$$\mathbf{W}_i\mathbf{T}_i = \mathbf{W}_i\bar{\mathbf{T}}_i + [\mathbf{W}_i\mathbf{G}\mathbf{W}_g^{\mathrm{T}}][\mathbf{W}_g\mathscr{D}(\mathbf{g})\mathbf{W}_i^{\mathrm{T}}](\mathbf{W}_i\mathbf{T}_i),$$

which, after we make the obvious definitions, become

$$\boldsymbol{\eta}_i = \boldsymbol{\Lambda}_i\boldsymbol{\Delta}(\boldsymbol{\theta}_i)\boldsymbol{\gamma} + \boldsymbol{\nu}_i, \tag{10}$$

$$\boldsymbol{\theta}_i = \bar{\boldsymbol{\theta}}_i + \boldsymbol{\Gamma}\boldsymbol{\Delta}(\boldsymbol{\gamma})\boldsymbol{\theta}_i, \tag{11}$$

where, for example, $\boldsymbol{\Delta}(\boldsymbol{\gamma})$ is the standard-form[14] wavelet transform of $\mathscr{D}(\mathbf{g})$ and is a function of $\boldsymbol{\gamma}$, since $\boldsymbol{\Delta}(\boldsymbol{\gamma}) = \mathbf{W}_g\mathscr{D}(\mathbf{W}_g^{\mathrm{T}}\boldsymbol{\gamma})\mathbf{W}_i^{\mathrm{T}}$. From Eqs. (10) and (11), the complete model relating the transform of the object to those of the data is

$$\eta_i = \underbrace{\boldsymbol{\Lambda}_i\boldsymbol{\Delta}([\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\boldsymbol{\gamma})]^{-1}\bar{\boldsymbol{\theta}}_i)\boldsymbol{\gamma}}_{\boldsymbol{\Upsilon}_i(\boldsymbol{\gamma})} + \boldsymbol{\nu}_i. \tag{12}$$

Finally, we aggregate the $\boldsymbol{\eta}_i$ into a single data vector and define the stacked system

$$\boldsymbol{\eta} = \boldsymbol{\Upsilon}(\boldsymbol{\gamma}) + \boldsymbol{\nu}, \tag{13}$$

with $\boldsymbol{\eta}^{\mathrm{T}} = [\boldsymbol{\eta}_1^{\mathrm{T}} \cdots \boldsymbol{\eta}_K^{\mathrm{T}}]$ and $\boldsymbol{\Upsilon}(\boldsymbol{\gamma})$ and $\boldsymbol{\nu}$ defined accordingly. With Eq. (13) the TWST inverse problem to be solved here is the recovery of $\boldsymbol{\gamma}$, the DWT of $\mathbf{g}$, from the measurements $\boldsymbol{\eta}$, a knowledge of $\boldsymbol{\Upsilon}$, and the statistics of $\boldsymbol{\nu}$.

4. INVERSION ALGORITHM

In this paper the reconstruction of $\boldsymbol{\gamma}$, denoted $\hat{\boldsymbol{\gamma}}$, given the data $\boldsymbol{\eta}$, is defined to be the solution to the following nonlinear least-squares type of optimization problem:

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\gamma} \mathscr{C}(\boldsymbol{\gamma}), \tag{14}$$

$$\mathscr{C}(\boldsymbol{\gamma}) = \frac{1}{2}\|\boldsymbol{\eta} - \boldsymbol{\Upsilon}(\boldsymbol{\gamma})\|_{\mathscr{R}^{-1}}{}^2 + \lambda^2\rho^{\mathrm{T}}(\boldsymbol{\gamma})\rho(\boldsymbol{\gamma}), \tag{15}$$

where $\|\mathbf{x}\|_{\mathbf{A}} \equiv \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, $\mathscr{R}$ is a diagonal weighting matrix whose entries reflect the noise levels in the data, and $\rho^{\mathrm{T}}(\boldsymbol{\gamma})\rho(\boldsymbol{\gamma})$ is used to regularize the problem. The process of minimizing the computational cost function $\mathscr{C}(\boldsymbol{\gamma})$ forces $\hat{\boldsymbol{\gamma}}$ to balance the effects of the two terms constituting $\mathscr{C}$. The first term enforces fidelity to the data. That is, our choice of $\hat{\boldsymbol{\gamma}}$ should be such that when put through the forward-scattering model $\boldsymbol{\Upsilon}$, it comes close to reproducing the data $\boldsymbol{\eta}$, where closeness is measured in the appropriate norm. Thus a smaller cost is associated with reconstructions that better explain the data that we have collected.

Were the first term in Eq. (15) the only portion of the cost, the ill-posed nature of the TWST problem would result in a reconstruction with large-amplitude, high-frequency components. To counter this effect, the regularization term $\rho^{\mathrm{T}}\rho$ is included. Generally, $\rho^{\mathrm{T}}\rho$ is used to constrain the reconstruction to have properties such as small energy (i.e., small L_2 norm), minimal gradient norm, or other smoothness-type characteristics.[25] Hence vectors corresponding to such oscillatory reconstructions are less desirable and therefore incur a higher cost. Finally, λ^2, the regularization parameter, is specified to balance the relative impact of the two terms on the reconstruction procedure. To summarize, the overall problem is to find a vector $\hat{\boldsymbol{\gamma}}$ that minimizes a cost function. This function imposed higher cost (i.e., higher computational penalty) for $\boldsymbol{\gamma}$, which fails to replicate the measured data and which possesses unfavorable structure, such as oscillatory behavior.

We employ a form of the Levenberg–Marquardt algorithm (LMA)[26] for finding $\hat{\boldsymbol{\gamma}}$. This iterative technique defines a sequence of reconstructions $\hat{\boldsymbol{\gamma}}_n$ whose costs as measured by Eq. (15) are steadily decreasing. As with all gradient-descent-type optimization approaches, the LMA converges to a local minimum of the cost function. Despite this restriction, in Section 5 we demonstrate that the reconstructions obtained for realistic TWST problems are quite accurate.

Starting from an initial guess $\hat{\boldsymbol{\gamma}}_0$, the form of the LMA used here is

$$\hat{\boldsymbol{\gamma}}_{n+1} = \hat{\boldsymbol{\gamma}}_n + \mathbf{s}_n, \tag{16}$$

$$\mathbf{s}_n = \arg\min_{\sigma^2} \mathscr{C}(\hat{\boldsymbol{\gamma}}_n + \mathbf{s}),$$

$$\mathbf{s} = [\mathscr{T}^{\mathrm{T}}(\hat{\boldsymbol{\gamma}}_n)\mathscr{R}^{-1}\mathscr{T}^{\mathrm{T}}(\hat{\boldsymbol{\gamma}}_n) + \sigma^2\mathscr{L}^{\mathrm{T}}(\hat{\boldsymbol{\gamma}}_n)\mathscr{L}(\hat{\boldsymbol{\gamma}}_n)]^{-1} \times \mathscr{T}^{\mathrm{T}}(\hat{\boldsymbol{\gamma}}_n)\mathscr{R}^{-1}[\boldsymbol{\eta} - \boldsymbol{\Upsilon}(\hat{\boldsymbol{\gamma}}_n)] - \mathscr{L}^{\mathrm{T}}(\hat{\boldsymbol{\gamma}}_n)\mathscr{L}(\hat{\boldsymbol{\gamma}}_n). \tag{17}$$

In Eq. (17) $\mathscr{T}(\hat{\boldsymbol{\gamma}}_n)$ is the Jacobian matrix of $\boldsymbol{\Upsilon}$ evaluated at the vector $\hat{\boldsymbol{\gamma}}_n$. To build the (j, k)th element of $\mathscr{T}$, the jth component of $\boldsymbol{\Upsilon}$ is differentiated with respect to the kth element of $\boldsymbol{\gamma}$. This scalar function will depend, in general, on all components of $\boldsymbol{\gamma}$. The Jacobian is obtained by evaluating each such derivative at the point $\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}$. Similarly, $\mathscr{L}(\hat{\boldsymbol{\gamma}}_n)$ is the Jacobian of ρ evaluated at $\hat{\boldsymbol{\gamma}}_n$. Finally, σ is a regularization parameter whose value is determined adaptively at each iteration of the algorithm.

In the remainder of this section, we discuss the choice of ρ, examine an approximation to the forward-scattering model designed to reduce the complexity of the LMA including the construction of $\mathscr{T}$, and overview the computational burden of this inversion approach.

A. Edge-Preserving Wavelet Regularization

In our previous work,[13,16] we have concentrated on the use of wavelet domain regularizers with $\rho(\boldsymbol{\gamma}) = \mathbf{D}\boldsymbol{\gamma}$, where the matrix $\mathbf{D}$ was diagonal, with

$$[\mathbf{D}]_{ii} \equiv d_i = 2^{-(\alpha_x j_{x,i} + \alpha_y j_{y,i})}. \tag{18}$$

In Eq. (18) $j_{x,i}$ and $j_{y,i}$ are the horizontal and vertical scale indices for the ith wavelet coefficient, and α_x and α_y

are constants. For this regularization approach, $\rho^{\mathrm{T}}(\hat{\boldsymbol{\gamma}}_n)\rho(\hat{\boldsymbol{\gamma}}_n)$ is a weighted two-norm of $\hat{\boldsymbol{\gamma}}_n$. This choice of ρ enforces smoothness in the reconstruction and the functions in much the same way as traditional Tikhonov-type regularizers.[13,16] One consequence of this choice is that edges and other sharp discontinuities that may be of interest in localizing defects or quantitatively characterizing their structure tend to be blurred in the final reconstruction.

Recently, there has been considerable work performed in the area of edge-preserving regularizers.[27,28] The idea is to construct a physical-space regularization scheme that results in reconstructions whose discontinuities are better preserved as compared with those from a Tikhonov approach. One way of implementing this regularization technique is to choose an expression for $\rho^{\mathrm{T}}\rho$ that, instead of being a weighted two-norm of the object, is a norm in a function space containing edgy objects. Adding this as the second term in Eq. (15) produces an object that lies in such a space and therefore retains the desired edgelike structure.

It is the case that, in addition to spanning the space of square integrable functions, orthonormal wavelets are also bases for these more exotic function spaces.[18] Here we use the fact that the norm in such a space may be computed in terms of the wavelet coefficients through $\Sigma_i d_i |\gamma_i|^p$, with $1 \leqslant p \leqslant 2$ and d_i exactly the same as in Eq. (18).

The use of p in this range to describe edge characteristics is best examined with a 1D example. In Fig. 3 we plot a 1D signal with two edges, along with the wavelet coefficients of this signal over a collection of scales. The original function is the topmost trace, and each box on the subsequent traces represents a single wavelet coefficient at a given scale. Finest-scale coefficients are closest to the top of the picture. These sequences are basically zero except for coefficients describing the behavior of the function near the edges of the step, where there are spikes. Thus the wavelet sequences provide localized information about the discontinuity structure (i.e., local smoothness) of the underlying function. If we were attempting to reconstruct such a signal, the cost of the spikes as measured by the regularization term $\Sigma_i d_i |\gamma_i|^p$ decreases as p drops from 2 to 1. In other words, a $p = 2$ regularization scheme seeks smooth reconstructions and thus penalizes the presence of such spikes, thereby resulting in blurry reconstruction. If we take $p < 2$, the cost of the spikes is smaller and the reconstruction should more faithfully reproduce the underlying edges.

With this motivation we make the following choice for $\rho(\boldsymbol{\gamma})$:

$$\rho^{\mathrm{T}}(\boldsymbol{\gamma}) = [d_1^{1/2}|\boldsymbol{\gamma}_1|^{p/2} \quad d_2^{1/2}|\boldsymbol{\gamma}_2|^{p/2} \quad \cdots \quad d_N^{1/2}|\boldsymbol{\gamma}_N|^{p/2}], \tag{19}$$

which results in a diagonal $\mathscr{L}$ with

$$[\mathscr{L}(\hat{\boldsymbol{\gamma}}_n)]_{i,i} = d_i^{1/2}\frac{p}{2}\, l([\hat{\boldsymbol{\gamma}}_n]_i), \tag{20}$$

$$l(x) = \mathrm{sign}(x)|x|^{p/2-1}, \tag{21}$$

where $[\hat{\boldsymbol{\gamma}}_n]_i$ is the ith element of the vector $\hat{\boldsymbol{\gamma}}_n$.

As seen from the plot of $l(x)$ in Fig. 4, the presence of a singularity at $x = 0$ will lead to numerical difficulties in the implementation of this regularization scheme. For the examples in Section 5, we replace $l(x)$ with the function $\bar{l}(x)$, defined as

$$\bar{l}(x) = \frac{x}{\epsilon + |x|^{2-p/2}}, \tag{22}$$

where ϵ is a small positive number governing the structure of $\bar{l}(x)$ for $x \approx 0$. For example, Fig. 4 shows that for $p = 1.2$ and $\epsilon = 10^{-3}$, $\bar{l}(x)$ is better behaved near the origin while still retaining the essential shape of $l(x)$. Also, $\bar{l}(x)$ possesses the same asymptotic structure as that of $l(x)$ for $|x| \gg 0$. Taken together, these facts indicate that $\bar{l}(x)$ should perform comparably with $l(x)$ in the LMA.

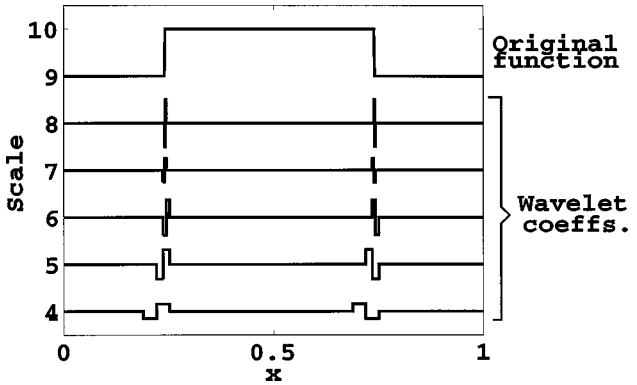


Fig. 3. Wavelet transform of a pulse function. The original function is given by the top trace, and the wavelet coefficients at a variety of scales are shown in the lower traces. Finer-scale information is conveyed in traces closer to the top. The wavelet coefficients characterize the local discontinuity structure of the function.

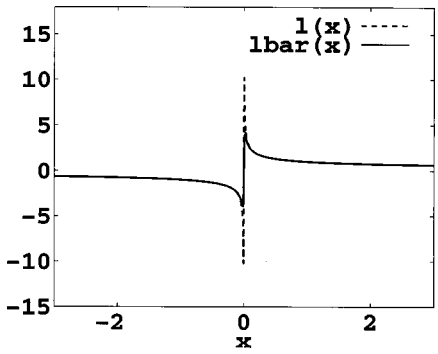


Fig. 4. Plots of $l(x)$ (dashed curve) and $\bar{l}(x)$ (solid curve) used to implement wavelet domain edge-preserving regularization. Both functions are identical for $x \gg 0$, but $\bar{l}(x)$ is better behaved near the origin, thereby aiding in numerical implementation.

B. Approximate Physical Model

Two factors related to the forward-scattering model dominate the computational complexity of the LMA. The first component is the evaluation of $\mathbf{Y}(\hat{\boldsymbol{\gamma}}_n)$ as required by Eq. (17). According to Eq. (12), this problem requires the inversion of $\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\hat{\boldsymbol{\gamma}}_n)$. The quantity $\boldsymbol{\Xi}_n \equiv [\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\hat{\boldsymbol{\gamma}}_n)]^{-1}$ has the interpretation as the wavelet transform of the discretized resolvent associated with the

second-kind integral equation (2), with $g(\mathbf{r})$ equal to the estimate of the object function at the nth iteration of the LMA.

The second computationally intensive task is the construction of $\mathscr{J}(\hat{\boldsymbol{\gamma}}_n)$. To obtain an expression for $\mathscr{J}(\hat{\boldsymbol{\gamma}}_n)$, we note first that this matrix is built by stacking $\mathscr{J}_i(\hat{\boldsymbol{\gamma}}_n)$, the Jacobian matrices for the K scattering experiments. For an arbitrary $\boldsymbol{\gamma}$, $\mathscr{J}_i(\boldsymbol{\gamma})$ is[29]

$$\mathscr{J}_i(\boldsymbol{\gamma}) = \boldsymbol{\Lambda}_i\boldsymbol{\Delta}(\boldsymbol{\theta}_i) + \boldsymbol{\Lambda}_i[\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\boldsymbol{\gamma})]^{-1}\boldsymbol{\Gamma}\boldsymbol{\Delta}(\boldsymbol{\theta}_i), \quad (23)$$

with $\boldsymbol{\theta}_i = [\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\boldsymbol{\gamma})]^{-1}\bar{\boldsymbol{\theta}}_i$. The construction of $\mathscr{J}_i$ clearly requires a number of matrix multiplications and, like the evaluation of $\mathbf{Y}_i$, the inversion of $\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\boldsymbol{\gamma})$. Although the matrix–matrix multiplications can be computationally intensive, the relevant matrices are sparse, thereby reducing the time overhead of this task (see Subsection 4.C). Thus our primary concern here is to develop a method that avoids the need to invert $\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\boldsymbol{\gamma})$ explicitly, as such an approach aids both in evaluating $\mathbf{Y}$ and in building the Jacobian.

The method that we propose is based on the observation that at the beginning of stage $n + 1$ of the LMA, we require that

$$\boldsymbol{\Xi}_{n+1} \equiv [\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\hat{\boldsymbol{\gamma}}_{n+1})]^{-1} = [\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\hat{\boldsymbol{\gamma}}_n + \mathbf{s}_n)]^{-1},$$

where the second equality follows from Eq. (16) and where $\mathbf{s}_n$ is typically a small correction to $\hat{\boldsymbol{\gamma}}_n$. Using the fact that $\boldsymbol{\Delta}$ is a linear operator, we obtain

$$\begin{aligned}\boldsymbol{\Xi}_{n+1} &= [\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\hat{\boldsymbol{\gamma}}_n + \mathbf{s}_n)]^{-1} \\ &= [\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\hat{\boldsymbol{\gamma}}_n) - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\mathbf{s}_n)]^{-1}, \quad (24)\\ &= \boldsymbol{\Xi}_n[\mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Delta}(\mathbf{s}_n)\boldsymbol{\Xi}_n]^{-1}, \quad (25)\\ &\approx \boldsymbol{\Xi}_n + \boldsymbol{\Xi}_n\boldsymbol{\Gamma}\boldsymbol{\Delta}(\mathbf{s}_n)\boldsymbol{\Xi}_n, \quad (26)\end{aligned}$$

where relation (26) follows from Eq. (24) under the assumption that $\boldsymbol{\Gamma}\boldsymbol{\Delta}(\mathbf{s}_n)\boldsymbol{\Xi}_n$ is small relative to $\mathbf{I}$. This assumption of smallness is quite valid for the TWST examples considered in Section 5. Indeed, the matrix two-norm of $\boldsymbol{\Gamma}\boldsymbol{\Delta}(\mathbf{s}_n)\boldsymbol{\Xi}_n$ was almost always less than 0.03 and was usually substantially smaller. Thus, relative to $\mathbf{I}$, whose norm is 1, $\boldsymbol{\Gamma}\boldsymbol{\Delta}(\mathbf{s}_n)\boldsymbol{\Xi}_n$ can really be viewed as a perturbation, so that the linearization is well justified.

As an approximation to $\boldsymbol{\Xi}_{n+1}$, relation (26) possesses some interesting and useful properties. First, we note that, unlike the Born approximation, this is not a linearization of the physics about the current estimate of the object function. Indeed, although relation (26) *is* linear in the increment $\mathbf{s}_n$, the presence of the two $\boldsymbol{\Xi}_n$ in the second term makes it nonlinear in $\hat{\boldsymbol{\gamma}}_n$. Second, because the left-hand side of Eq. (24) is, by definition, $\boldsymbol{\Xi}_{n+1}$, our approximation provides a recursive method for updating the resolvent from one iteration of the LMA to the next. Specifically, we have

$$\boldsymbol{\Xi}_{n+1} = \boldsymbol{\Xi}_n + \boldsymbol{\Xi}_n\boldsymbol{\Gamma}\boldsymbol{\Delta}(\mathbf{s}_n)\boldsymbol{\Xi}_n, \quad (27)$$

which states that the new resolvent is equal to the old resolvent plus an increment that is a function of both $\boldsymbol{\Xi}_n$ and $\mathbf{s}_n$, the update to the reconstruction at iteration n.

Finally, we note that despite the use of this linearization in computing the Jacobian, we are still employing a *nonlinear* inverse scattering approach to the recovery of the contrast function. Because we are using a full wave scattering model (i.e., no Born or Rytov approximation) along with the edge-preserving regularization, Eqs. (14) and (15) represent a nonlinear least-squares optimization problem. Any solution to such a problem, including the variant of the LMA discussed in this section, must be a nonlinear function of the data. Indeed, it is clearly the case that $\hat{\boldsymbol{\gamma}}_n$ is *some* function of the data, and the fact that $\boldsymbol{\Xi}_n$ is nonlinear in the elements of $\hat{\boldsymbol{\gamma}}_n$ already ensures that, even with the approximation to the Jacobian, the overall reconstruction must be a nonlinear function of $\boldsymbol{\eta}$. Hence the reconstruction approach falls in the class of nonlinear inverse scattering algorithms.

C. Computational Considerations

Each iteration of the LMA defined by Eqs. (16) and (17) requires first the construction of $\mathbf{Y}(\hat{\boldsymbol{\gamma}}_n)$, $\mathscr{J}(\hat{\boldsymbol{\gamma}}_n)$, and $\mathscr{L}(\hat{\boldsymbol{\gamma}}_n)$ and then the solution of a sequence of least-squares-type problems to determine $\mathbf{s}_n$. According to Eq. (20), determining $\mathscr{L}(\hat{\boldsymbol{\gamma}})$ requires roughly $2N$ floating-point operations. Although the need to solve the least-squares problems can be intensive, we have previously developed techniques based on our wavelet representation for $\mathbf{g}$ for reducing the computational overhead of this task.[13] As discussed in Subsection 4.B, the computational costs of the remaining tasks, evaluating $\mathbf{Y}$ and building $\mathscr{J}$, are dominated by the need to compute $\boldsymbol{\Xi}_n$.

By using Eq. (27), we replace the matrix inversion by three matrix–matrix products. The primary reduction in computational complexity comes from the fact that these matrices are sparse by construction [e.g., $\boldsymbol{\Delta}(\hat{\boldsymbol{\gamma}}_n)$][13] or can be made so with little loss (e.g., $\boldsymbol{\Gamma}$ and $\boldsymbol{\Xi}_n$) by truncating small elements. Although it is not necessarily the case that products of sparse matrices are themselves sparse, we ensure sparsity and therefore low complexity by implementing Eq. (27) as

$$\boldsymbol{\Xi}_{n+1} = \mathrm{trunc}_\delta(\boldsymbol{\Xi}_n + \mathrm{trunc}_\delta(\boldsymbol{\Xi}_n\boldsymbol{\Gamma})\mathrm{trunc}_\delta[\boldsymbol{\Delta}(\hat{\boldsymbol{\gamma}}_n)\boldsymbol{\Xi}_n]), \quad (28)$$

where, for a matrix $\mathbf{M}$, $\mathrm{trunc}_\delta(\mathbf{M})$ is the sparse approximation to $\mathbf{M}$ by means of the method in Ref. 30 with a threshold δ. Under this truncation scheme, we set to zero all elements of the $m \times n$ matrix $\mathbf{M}$ whose absolute values are less than $(\delta/n)\|\mathbf{M}\|_\infty$.

For the examples in Section 5, the use of Eq. (28) results in sparse approximations to the Jacobian, which translates into substantial computational savings. To obtain an accurate initial estimate of the Jacobian, the first iteration of the LMA is executed with no approximation. That is, we construct $\mathscr{J}_i$ directly according to Eq. (23). Thus the operation counts and execution times for this iteration may be compared with those of subsequent iterations, where Eq. (28) is used to build $\mathscr{J}_i$. For the examples in Section 5, on average, when computed according to Eq. (28), $\boldsymbol{\Xi}_n$ is between 8% and 12% full across all iterations of the algorithm. This level of sparsity, coupled with the savings afforded by the linearization process, translates into a factor-of-3–4 savings relative to direct use of Eq. (23), in terms of both the number of floating point-operations and the time to run each iteration of the algorithm.

5. EXAMPLES OF THERMAL-WAVE TOMOGRAPHIC IMAGING

To illustrate the inversion algorithm developed in the previous sections, we consider the imaging of single and multiple defects in a 3-mm × 3-mm block of aluminum (thermal diffusivity of 0.82 $cm^2\,s^{-1}$). Depending on the example, the reconstructions are based on the individual or joint processing of backpropagation and transmission data collected for one or more locations of the modulated heating laser equally spaced across the top of the material sample. A given transmission or backpropagation data set consists of measurements of noisy scattered fields obtained at 32 equally spaced points along either the top or the bottom of the material. For all cases the top of the sample is taken to be the line $x = 0$, and the bottom is at $x = 3$ mm. The angular modulation frequency of the laser for all experiments is taken to be 8 Hz. Finally, the signal-to-noise ratio (SNR) in decibels (dB) for the ith experiment is defined as

$$\mathrm{SNR}_i = 10 \log_{10} \frac{\|\mathbf{h}(\mathbf{g})\|_2^2}{N_i q_i^2},$$

where q_i^2 is the variance of the ith noise process, N_i is the number of data points in the ith data vector, and $\mathbf{h}^{\mathrm{T}}(\mathbf{g}) = [\mathbf{h}_1^{\mathrm{T}}(\mathbf{g}) \quad \mathbf{h}_2^{\mathrm{T}}(\mathbf{g}) \quad \cdots \quad \mathbf{h}_K^{\mathrm{T}}(\mathbf{g})]$, with $\mathbf{h}_i(\mathbf{g})$ given by Eq. (6).

For purposes of inversion, we seek a reconstruction of g on a 16 × 16 grid. A Haar wavelet[24] is used to transform both the data vectors and $\mathbf{g}$. The parameter λ^2 in Eq. (15) is 10, and $\alpha_x = \alpha_y = 1$ in the regularization method. The value of δ used to truncate the operators is 0.01. These quantities were chosen by trial and error. The quality of the reconstructions was relatively insensitive to the choices for α_x, α_y, and δ. Choosing the correct value for the regularization parameter λ^2 can be a delicate procedure. For the TWST examples presented here, values between 1 and 100 tended to produce similar results. We leave for future efforts the development of a more automated procedure for selecting this parameter.

To verify the utility of the edge-preserving regularization method, we examine reconstructions obtained with both $p = 2$ and $p = 1.2$ in Eq. (22). The $p = 2$ case corresponds to the wavelet domain smoothness regularization previously employed,[13,16] and the $p = 1.2$ selection will be shown to result in more accurately recovered edges. For the $p = 2$ reconstructions, the LMA is initialized with $\hat{\boldsymbol{\gamma}}_0 = 0$. For the $p = 1.2$ case, we first run the LMA with $p = 2$ for three iterations and then switch to $p = 1.2$ for the remainder of the reconstruction process. This strategy avoids instabilities associated with the sharp transition of the $p = 1.2$ regularization function that can arise when the LMA is seeded with the zero vector. Finally, by starting the LMA at zero, the first iteration of the algorithm is mathematically equivalent to reconstruction under the Born approximation with the $p = 2$ wavelet domain regularizer. Thus we can compare the reconstructions for the nonlinear algorithm with those obtained when a Born model with wavelet regularization is used to process the data.

The first example that we consider is the reconstruction of a square air hole centered near the top of the sample. The object function for this case is shown in Fig. 5(a) and consists of a square hole of amplitude[4,5] $(\alpha_{\mathrm{Al}}/\alpha_{\mathrm{air}}) - 1 \approx 3$. The SNR for all experiments is 50 dB. The reconstruction obtained under the Born approximation is displayed in Fig. 5(b). For this case the Born approximation yields coarse-scale localization of the defect. That is, the reconstruction is nonzero over a region of space that includes the area of the true defect. However, the amplitude of the reconstruction is at best a third of the true amplitude, and the shape of the reconstructed defect is in fact larger than that of the true structure. The execution time for the Born algorithm on a Sparc 20 was approximately 11 min, with much of this time devoted to the search for the optimal regularization parameter σ in Eq. (17). To summarize, the Born reconstruction provides relatively fast access to an image that provides a rough idea of the location of the defect with little information regarding the amplitude.

The final results of the $p = 2$ and the $p = 1.2$ LMA are shown in Figs. 5(c) and 5(d) respectively. In Fig. 5(e) the value of $\|\hat{\boldsymbol{\gamma}}_n - \boldsymbol{\gamma}\|_2$ is plotted for both regularization schemes as a function of iteration. It is evident that both regularization schemes produce reconstructions that are better localized with more accurate amplitude information than those from the Born inversion. The $p = 1.2$ case is slightly better than the $p = 2$ reconstruction from both a quantitative and a qualitative perspective. Quan-

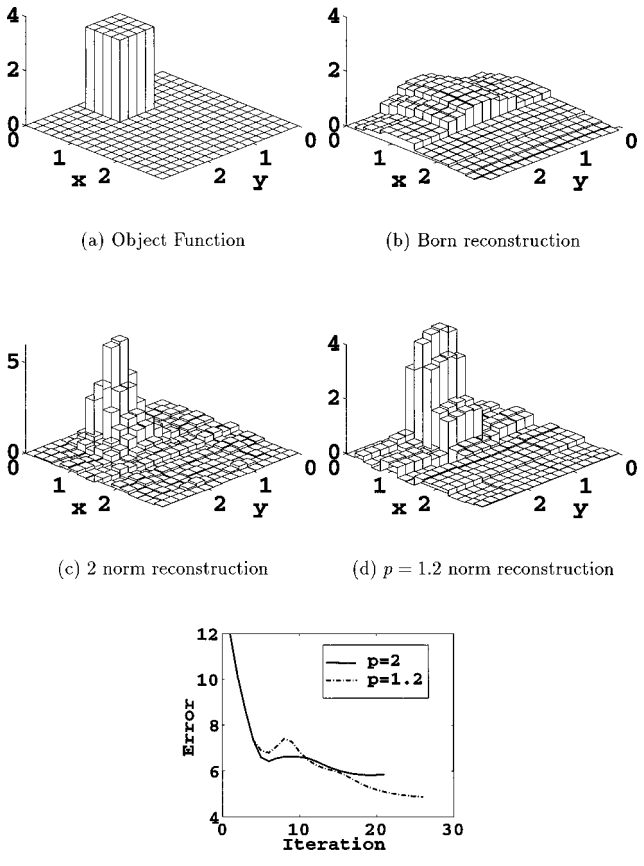
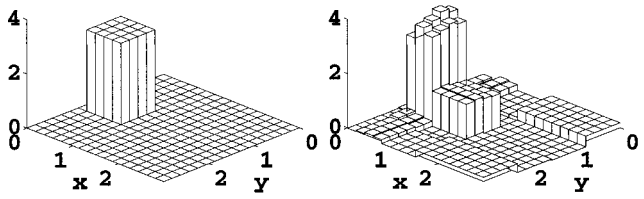
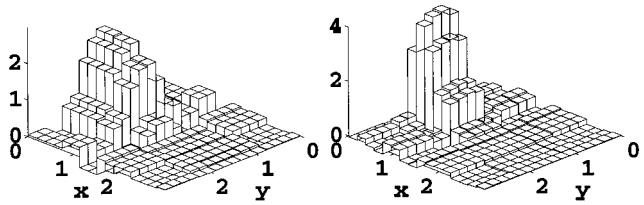


Fig. 5. True object function, inversion results, and error versus iteration for a single defect located near the top of the material. The laser source position is along the line $x = 0$.



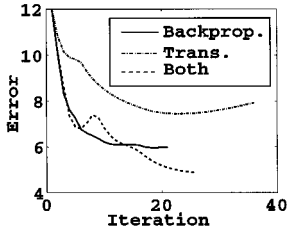
(a) Object Function

(b) Backpropagation reconstruction



(c) Transmission reconstruction

(d) Combined reconstruction



(e) $\|\hat{\gamma}_n - \gamma\|_2$

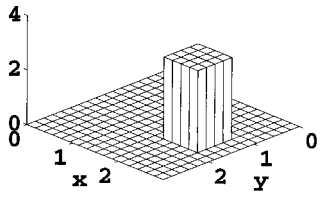
Fig. 6. True object function, inversion results, and error versus iteration for a single defect located near the top of the material. Here we compare reconstructions obtained by using (b) only backpropagation data, (c) only transmission data, and (d) a combination of both in the nonlinear inversion algorithm. The laser source position is along the line $x = 0$.

titatively, the error in the $p = 1.2$ estimate is somewhat lower than that of the $p = 2$ case after ten iterations. This improvement is due to the fact that the $p = 2$ case overestimates the amplitudes in a couple of pixels, whereas the amplitude of the $p = 1.2$ version is quite accurate. Visually, as we expect, the edges on the $p = 1.2$ reconstruction are sharper, and the overall reconstruction looks much more like a box than a smoothed blob, as is the case in Fig. 5(c). The price paid for the increase in the accuracy of both reconstructions relative to the Born case is primarily computational. The processing time for both LMA reconstructions of this 16×16 grid of pixels is of the order of 2 h, again with the use of a Sparc 20.

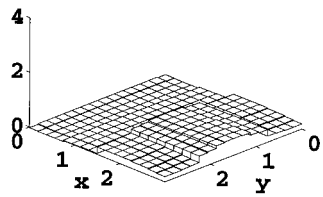
In Fig. 6 the utility of jointly processing both transmission and backpropagation data is illustrated. In all cases we use the same object as before and the $p = 1.2$ regularization. The final reconstruction obtained with the use of only backpropagation data is displayed in Fig. 6(b). Here we see that the object is relatively well localized and that the edges are nicely captured; however, the reconstruction of the rear portion of the defect is clearly inferior to that obtained when both transmission and backpropagation data are employed [Fig. 6(d)]. Indeed, this result is not unexpected, given that the rear of the defect is effectively shadowed for the backpropagation case.

The transmission-only reconstruction is shown in Fig. 6(c). This image is significantly more blurred than the two others, despite the use of edge-preserving regularization. The longer propagation distances from sources to receivers basically remove from the data nearly all of the fine-scale information about the object function, resulting in a severely degraded reconstruction. The results of using both transmission and backpropagation data are seen in Fig. 6(d). This reconstruction combines the best features of the previous two. The object and its edges are well localized as in the backpropagation reconstruction, and the use of the transmission data aids in improving the resolution of the back side of the defect. These advantages are seen as well in the error curves of Fig. 6(e). The transmission-only inversion is clearly inferior to both of the others. Although the backpropagation produces lower-error reconstructions in the opening iterations, the joint inversion is ultimately superior.

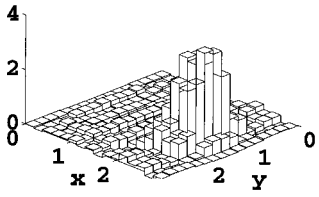
The second example that we consider is the reconstruction of a square hole located near the bottom of the material sample at a SNR of 50 dB. As seen in Fig. 7, the structure of this defect is identical to that of the previous example, except for the location. In Figs. 7(b)–7(e) we again plot the Born reconstructions, the final estimates obtained with the $p = 2$ and $p = 1.2$ regularizations, and



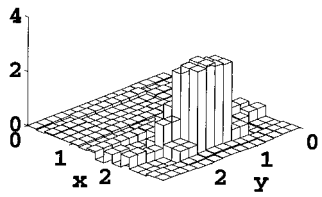
(a) Object Function



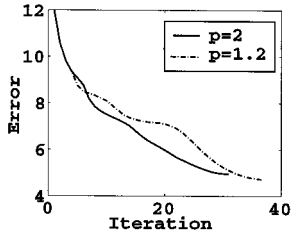
(b) Born reconstruction



(c) 2 norm reconstruction

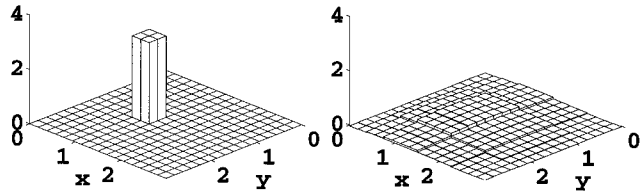


(d) $p = 1.2$ norm reconstruction



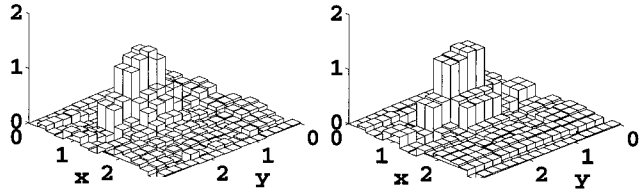
(e) $\|\hat{\gamma}_n - \gamma\|_2$

Fig. 7. True object function, inversion results, and error versus iteration for a single defect located near the bottom of the material. The laser source position is along the line $x = 0$.



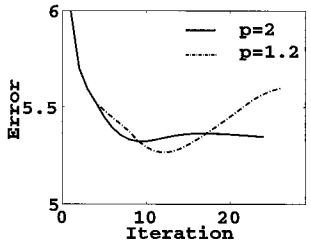
(a) Object Function

(b) Born reconstruction



(c) 2 norm reconstruction

(d) $p = 1.2$ norm reconstruction



(e) $\|\hat{\gamma}_n - \gamma\|_2$

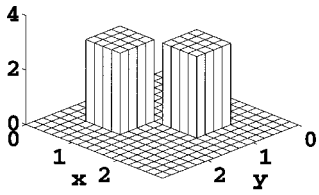
Fig. 8. True object function, inversion results, and error versus iteration for a single, small-defect problem. The laser source position is along the line $x = 0$.

the error as a function of iteration. As in the previous case, the Born estimate provides a rough localization of the underlying defect with limited quantitative accuracy. For both regularization schemes, the final reconstructions improve significantly on the Born inversion. The final results for the $p = 1.2$ case are, however, better than the $p = 2$ estimate in capturing the sharp edges in the defect and in providing a flattop to the estimated object function.

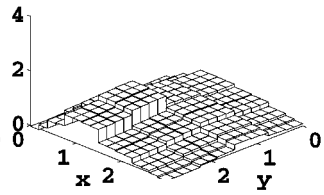
In Fig. 8 we display the inversion results for the case of a smaller defect located in the middle of the region. Unlike the previous two examples, where the hole was 4×4 pixels in size, for this problem we consider a 2×2 case. The Born reconstruction shows only a small-amplitude, poorly localized disturbance in the top half of the medium. The final estimates for both the $p = 2.0$ and $p = 1.2$ cases clearly indicate the presence of the hole in approximately the correct location; however, the accuracy of the localization is degraded here relative to that seen in the previous examples. Although the extent of the object in the x direction is well captured, Figs. 8(c) and 8(d) both show an object spread in y over approximately 4, rather than 2, pixels. The amplitudes of these reconstructions are approximately 1.2, as opposed to 3.0, which is the amplitude of the true defect. Although the error of the final $p = 1.2$ estimate is larger than that of the $p = 2.0$ case, we see that, as in the previous examples, the $p = 1.2$ estimate is somewhat better in preserving the blocky nature of the true structure.

The results in Fig. 8 demonstrate the well-known difficulty associated with the imaging of small defects near the middle of the material sample. In Ref. 13 we provided quantitative results describing the loss in accuracy suffered by attempts to recover fine-scale information in regions far from the sources and the receivers. The work here does point to some interesting and perhaps useful methods for improving the resolution in this area. From an experimental perspective, it would clearly be helpful to obtain data from more angles than just transmission and backscatter. Additionally, the incorporation of data from higher-frequency experiments might aid in the resolution of these smaller structures. Just as the current algorithm can handle data from both transmission and backscatter experiments, it is an easy matter to process data jointly from multiple frequencies.

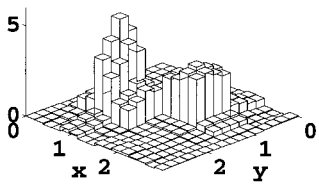
In terms of processing, the results in Figs. 8(c) and 8(d), although not as accurate as those in the previous examples, still provide much information about the location of the defect. These images could be used to generate an accurate initial guess for a more object-based processing approach such as that of Ref. 31. In that work, rather than looking for a pixelated representation of the region, we seek to extract information about the location, the shape, and the amplitude of rectangular defects in the



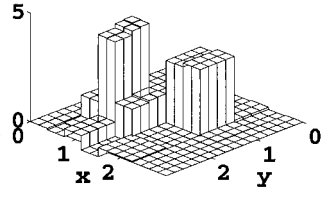
(a) Object Function



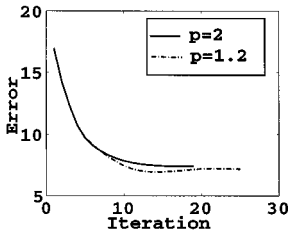
(b) Born reconstruction



(c) 2 norm reconstruction



(d) $p = 1.2$ norm reconstruction



(e) $\|\hat{\gamma}_n - \gamma\|_2$

Fig. 9. True object function, inversion results, and error versus iteration for a two-defect problem. The laser source position is along the line $x = 0$.

medium. By using the data to determine this relatively parsimonious set of parameters describing the defect, we can achieve high accuracy even for objects in the middle of the region.

As a final test, we demonstrate the performance of our approach in resolving adjacent subsurface defects on an example with two square defects located near the center of the material [shown in Fig. 9(a)]. As seen in Fig. 9(b), the Born inversion basically indicates that there is some gross disturbance located near the center of the region. Although the top anomaly (i.e., the structure nearer to the line $x = 0$) is slightly better captured than the bottom one, in neither case is the amplitude well represented.

Figures 9(c) and 9(d) show that the nonlinear inversion routine is quite capable of resolving both structures and in providing quantitatively accurate contrast information. As was seen in the previous two examples, the $p = 2$ reconstruction is smoother than that of the $p = 1.2$ case and is slightly less accurate in terms of the amplitudes of the defects. For the $p = 1.2$ reconstruction, the back defect is quite well localized in space and has a close-to-flat contrast. The front structure is well defined in its y variation, with less accuracy in x. The amplitude of this reconstruction is again quite flat, with sharp, well-defined edges. The defects are well resolved in both the $p = 2$ and $p = 1.2$ reconstructions. Quantitatively, the error in the $p = 1.2$ estimate is lower than that of the $p = 2$ case after seven iterations.

6. CONCLUSIONS AND FUTURE WORK

We have presented a new approach to image formation from scattered thermal waves based on the use of nonlinear inverse scattering methods and wavelet domain techniques. We build our inversion routine on the full wavefield physics developed by Mandelis in Refs. 4 and 5, resulting in a highly nonlinear relationship between the data on which a reconstruction is to be based and the desired image of the object function. The reconstruction problem was formulated as a solution to a nonlinear least-squares-type optimization problem in the wavelet transform domain. We chose to work in a multiscale setting for a number of reasons. First, the matrices constituting the physical model are sparse in this domain, thereby lowering the computational cost of generating a reconstruction. Second, we were able to use a new class of edge-preserving regularization methods that are easily specified and implemented in the wavelet transform domain. Finally, the computational burden was further reduced by employing the methods of Ref. 13 for rapidly solving nonlinear inverse scattering problems in a multiscale domain.

The Levenberg–Marquardt algorithm (LMA) formed the basis for the inversion procedure. This approach provided a natural mechanism for the joint processing of data from an arbitrary number of scattering experiments. Rather than forming separate images for each source–receiver array combination and then averaging the results, we were able to produce a single image that was optimal across all data sets. Although the LMA is guaranteed to converge only to a local minimum of the cost function, our numerical experiments demonstrate that the reconstructions are still quite accurate for realistic thermal-wave slice tomography (TWST) experimental configurations. Still, one area of future work might be to explore the development of efficient methods designed to produce a reconstructed contrast function the cost of which is globally minimum.

A key component of this inversion technique was shown to be the explicit inversion of the resolvent for the forward-scattering model at each iteration of the algorithm. This matrix was required for the calculation of the Jacobian and played a central role in evaluating the data misfit at every stage of the reconstruction. We introduced a new approximation to this resolvent that replaced the need to explicitly invert a generally large matrix with only the requirement of multiplying a sequence of three matrices. The complexity of even this operation was further reduced by means of our wavelet approach, where the sparse-matrix structure could be exploited. A more rigorous analysis of the conditions under which the linearization is valid represents an interesting area of further investigation.

The algorithm was demonstrated for imaging defects in the bulk structure of an aluminum sample by using TWST. Inversions were considered in which synthetic transmission and backpropagation scattering data were both input to the algorithm. Our results indicated that the use of the nonlinear approach produces reconstructions that are significantly more accurate than those from a Born-based inversion in terms of both localizing the defects and obtaining quantitative contrast information. Additionally, relative to traditional Tikhonov-type regularizers, the edge-based regularization scheme produced reconstructions that better represented the blocklike nature of the defects. Although it was still difficult to reconstruct accurately small structures located far from the top and bottom edges, we feel that future efforts in the areas of multifrequency inversion and the development of more object-based inversion procedures should greatly improve this situation.

ACKNOWLEDGMENTS

The work of E. L. Miller was supported in part by U.S. Department of Energy contract DE-FC07-95ID13395, National Science Foundation grant MIP-9623721, and subcontract GC123920NDG from Boston University under the U.S. Air Force Office of Scientific Research Multidisciplinary University Research Initiative Program on Reduced Signature Target Recognition. The work of L. Nicolaides and A. Mandelis was supported by the Natural Sciences and Engineering Research Council of Canada through a research grant.

The authors can be reached as follows. E. L. Miller: telephone, 617-373-8386; fax, 617-373-8627; e-mail, elmiller@ece.neu.edu. A. Mandelis: telephone/fax, 416-978-5106; e-mail, mandelis@mie.utoronto.ca.

REFERENCES

1. L. Nicolaides and A. Mandelis, "Image-enhanced thermal-wave slice diffraction tomography with numerically simu-

lated reconstructions," Inverse Probl. **13**, 1339–1412 (1997).
2. L. Nicolaides, M. Munidasa, and A. Mandelis, "Image-enhanced thermal-wave slice diffraction tomography with backpropagation and transmission reconstructions: experimental," Inverse Probl. **13**, 1413–1425 (1997).
3. O. Pade and A. Mandelis, "Computational thermal-wave slice tomography with backpropagation and transmission reconstructions," Rev. Sci. Instrum. **64**, 3548–3562 (1993).
4. A. Mandelis, "Theory of photothermal-wave diffraction and interference in condensed media," J. Opt. Soc. Am. A **6**, 298–308 (1989).
5. A. Mandelis, "Green's functions in thermal-wave physics: Cartesian coordinate representations," J. Appl. Phys. **78**, 647–655 (1995).
6. C. Torres-Verdín and T. M. Habashy, "Rapid 2.5-D forward modeling and inversion via a new nonlinear scattering approximation," Radio Sci. **29**, 1051–1079 (1994).
7. T. M. Habashy, W. C. Chew, and E. Y. Chow, "Simultaneous reconstruction of permittivity and conductivity profiles in a radially inhomogeneous slab," Radio Sci. **21**, 635–645 (1986).
8. T. Wang, M. Oristaglio, A. Tripp, and G. Hohmann, "Inversion of diffusive transient electromagnetic data by a conjugate-gradient method," Radio Sci. **29**, 1143–1156 (1994).
9. O. Pade and A. Mandelis, "Thermal-wave slice tomography using wave-field reconstruction," Inverse Probl. **10**, 185–197 (1994).
10. T. M. Habashy, R. W. Groom, and B. R. Spies, "Beyond the Born and Rytov approximations: a nonlinear approach to electromagnetic scattering," J. Geophys. Res. **98**, 1759–1775 (1993).
11. D. Lesselier and B. Duchene, "Wave-field inversion of objects in stratified environments: from backpropagation schemes to full solutions," in *Review of Radio Science*, W. R. Stone, ed. (Oxford U. Press, Oxford, UK, 1996).
12. T. Isernia, V. Pascazio, and R. Pierri, "A nonlinear estimation method in tomographic imaging," IEEE Trans. Geosci. Remote Sens. **35**, 910–923 (1997).
13. E. L. Miller and A. S. Willsky, "Wavelet-based methods for the nonlinear inverse scattering problem using the extended Born approximation," Radio Sci. **31**, 51–67 (1996).
14. G. Beylkin, R. Coifman, and V. Rokhlin, "Fast wavelet transforms and numerical algorithms. I," Commun. Pure Appl. Math. **44**, 141–183 (1991).
15. M. Bertero, "Linear inverse and Ill-posed problems," in *Advances in Electronics and Electron Physics*, P. Hawkes, ed. (Academic, Boston, 1989), Vol. 75, pp. 1–120.
16. E. L. Miller and A. S. Willsky, "Multiscale, statistically-based inversion scheme for the linearized inverse scattering problem," IEEE Trans. Geosci. Remote Sens. **34**, 346–357 (1996).
17. M. Bertero, C. D. Mol, and E. R. Pike, "Linear inverse problems with discrete data. II. Stability and regularisation," Inverse Probl. **4**, 573–594 (1988).
18. Y. Meyer, *Wavelets and Operators* (Cambridge U. Press, Cambridge, 1995).
19. R. F. Harrington, *Field Computations by Moment Methods* (Macmillan, New York, 1968).
20. B. Wang, J. C. Moulder, and J. P. Basart, "Wavelets in the solution of the volume integral equation: application to eddy current modeling," J. Appl. Phys. **81**, 6397–6406 (1997).
21. M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software* (A. K. Peters, Wellesley, Mass., 1994).
22. A. Cohen, I. Daubechies, B. Jawerth, and P. Vial, "Multiresolution analysis, wavelets and fast algorithms on an interval," Appl. Comput. Harmon. Anal. **1**, 54–81 (1993).
23. I. Daubechies, "Orthonormal bases of compactly supported wavelets," Commun. Pure Appl. Math. **41**, 909–996 (1988).
24. G. Strang and T. Nguyen, *Wavelets and Filter Banks* (Wellesley-Cambridge Press, Wellesley, Mass., 1996).
25. M. Bertero, C. D. Mol, and E. R. Pike, "Linear inverse problems with discrete data. I: General formulation and singular system analysis," Inverse Probl. **1**, 301–330 (1985).
26. P. E. Gill, W. Murry, and M. H. Wright, *Practical Optimization* (Academic, San Diego, 1981).
27. A. Gersztenkorn, J. B. Bednar, and L. R. Lines, "Robust iterative inversion for the one-dimensional acoustic wave equation," Geophysics **51**, 357–368 (1986).
28. P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlund, "Deterministic edge-preserving regularization in computed imaging," IEEE Trans. Image Process. **6**, 298–311 (1997).
29. E. L. Miller, "The application of multiscale and statistical techniques to the solution of inverse problems," Tech. Rep. LIDS-TH-2258 (MIT Laboratory for Information and Decision Systems, Cambridge, Mass., 1994).
30. B. Alpert, G. Beylkin, R. Coifman, and V. Rokhlin, "Wavelets for the fast solution of second-kind integral equations," SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **14**, 159–184 (1993).
31. E. L. Miller and A. S. Willsky, "Multiscale, statistical anomaly detection analysis and algorithms for linearized inverse scattering problems," Multidimens. Syst. Signal Process. **8**, 151–184 (1997).