Radio Science, Volume 31, Number 1, Pages 51–65, January–February 1996

Wavelet-based methods for the nonlinear inverse scattering problem using the extended Born approximation

Eric L. Miller
Center for Electromagnetics Research, Department of Electrical and Computer Engineering
Northeastern University, Boston, Massachusetts

Alan S. Willsky
Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts

Abstract. In this paper, we present an approach to the nonlinear inverse scattering problem using the extended Born approximation (EBA) on the basis of methods from the fields of multiscale and statistical signal processing. By posing the problem directly in the wavelet transform domain, regularization is provided through the use of a multiscale prior statistical model. Using the maximum a posteriori (MAP) framework, we introduce the relative Cramér-Rao bound (RCRB) as a tool for analyzing the level of detail in a reconstruction supported by a data set as a function of the physics, the source-receiver geometry, and the nature of our prior information. The MAP estimate is determined using a novel implementation of the Levenberg-Marquardt algorithm in which the RCRB is used to achieve a substantial reduction in the effective dimensionality of the inversion problem with minimal degradation in performance. Additional reduction in complexity is achieved by taking advantage of the sparse structure of the matrices defining the EBA in scale space. An inverse electrical conductivity problem arising in geophysical prospecting applications provides the vehicle for demonstrating the analysis and algorithmic techniques developed in this paper.

1. Introduction

The desire to characterize the composition of a medium on the basis of observations of scattered radiation is a common problem in a variety of application areas [*Kak and Slaley*, 1987; *Bates et al.*, 1991; *Torres-Verdín and Habashy*, 1994]. Despite the ubiquity of such inverse scattering problems, generating a solution can be quite difficult because of the computational burden associated with the nonlinearity of the problem and the fact that these problems are highly ill posed. In this paper, we present a collection of methods for overcoming these difficulties based upon techniques drawn from the disciplines of multiscale and statistical signal processing. We employ estimation-theoretic analysis techniques to identify those degrees of freedom in a wavelet representation of the quantity to be reconstructed for which the data provide significant information. Such a formulation represents a natural framework for the analysis of issues such as the trade-off between reconstruction accuracy and resolution, as well as the development of bounds on our ability to localize spatial anomalies in the region of interest. Direct incorporation of this information into a nonlinear inversion algorithm coupled with a multiscale implementation of the forward scattering model results in substantial computational savings with little loss in reconstruction fidelity. We apply our method to an inverse electrical conductivity problem encountered in geophysical exploration applications.

The extended Born approximation (EBA) developed by *Habashy et al.* [1993] is used to lower the computational complexity of the forward modeling portion of our inverse scattering algorithm. The

Paper number 95RS03130.
0048-6604/96/95RS-03130$08.00

EBA provides a simple functional relationship between the conductivity perturbation and the scattered field, which makes this model ideally suited for the computationally efficient implementation of a complete, gradient-based nonlinear optimization algorithm for solving the inverse scattering problem [*Torres-Verdín and Habashy*, 1994]. We further improve the efficiency of our inversion algorithm by exploiting the fact that the matrices defining the extended Born approximation are of the class which are "nearly diagonalized" by the wavelet transform [*Beylkin et al.*, 1991].

The traditional technique for overcoming the ill posed nature of these inverse scattering problems is through the use of a regularization method [*Bertero et al.*, 1988; *Groetsch*, 1984; *Kress*, 1989]. From an estimation-theoretic perspective, the regularizer represents a prior statistical model for the quantity under investigation [*Miller and Willsky*, 1995b]. Here we demonstrate the utility of specifying such a model directly in the wavelet transform domain, where there exists a wide variety of scale-space modeling structures well suited for this task [*Miller and Willsky*, 1995a, 1995b]. These models are useful for representing many common, naturally occurring self-similar phenomena, are easily specified in scale-space, and for certain problems, lead to fast inversion algorithms. Moreover, in the linearized inverse scattering context, a wavelet representation of the conductivity provided a natural basis for defining the space-varying optimal level of detail in a reconstruction as a function of the resolution, quality, and spatial distribution of the data [*Miller and Willsky*, 1995b].

The solution to the nonlinear inverse scattering problem is obtained as the maximum a posteriori (MAP) estimate of the wavelet components of the conductivity field. Using the MAP formulation, we generalize the results from *Miller and Willsky* [1995b] by introducing the relative Cramér-Rao bound (RCRB) for quantitatively evaluating the information provided by the data with respect to the conductivity's wavelet transform. At each stage of the inversion, the RCRB is used to identify those wavelet coefficients for which the data do and do not provide significant information. Use of this partition leads to substantial computational savings with little loss in reconstruction fidelity.

In section 2, we present the physical space formulation of the inverse electrical conductivity problem. Section 3 is devoted to a review of the wavelet transform and a derivation of the scale-space MAP estimation problem. The relative Cramér-Rao bound is defined and its properties discussed in section 4. Our nonlinear inversion scheme is presented in section 5 with a collection of examples discussed in section 6. The conclusions to be drawn from this work are presented in section 7.

2. Physical Space Problem Formulation

2.1. The Forward Problem

We consider a two-dimensional inverse conductivity problem, illustrated in Figure 1, where there exists a set of electromagnetic line sources oriented perpendicularly to the page emitting time-harmonic waves into a lossy medium characterized by a constant background conductivity of g_0 Siemens per meter and the free space values for electric permittivity, ϵ_0, and magnetic permeability, μ_0. The objective of the problem is to reconstruct a conductivity perturbation, $g(\mathbf{r})$, in A given noisy, pointwise observations of that component of the scattered electric field oriented perpendicularly to the page. These measurements are obtained along receiver arrays positioned on the vertical edges of A from a collection of K scattering experiments. Each experiment produces a vector of measurements, $\mathbf{y}_i$, comprising the in-phase and quadrature components of the scattered field obtained over a single receiver array due to energy put into the medium from one of the sources operating at a particular frequency.

As discussed by *Torres-Verdín and Habashy* [1994], the relationship between $g(\mathbf{r})$ and the jth element of $\mathbf{y}_i$, that is, the measured scattered field at position $\mathbf{r}_j$, is

$$y_i(\mathbf{r}_j) = w\omega_i\mu_0 \int_A G_i(\mathbf{r}_j, \mathbf{r}')g(\mathbf{r}')E_i(\mathbf{r}')d\mathbf{r}' + n_i(\mathbf{r}_j), \tag{1}$$

where $n_i(\mathbf{r}_j)$ represents additive noise, $G_i(\mathbf{r}, \mathbf{r}')$ is the Green's function for the problem, and $E_i(\mathbf{r})$ is the component of the total electric field perpendicular to the page. From *Torres-Verdín and Habashy* [1994], $E_i(\mathbf{r})$ satisfies

$$E_i(\mathbf{r}) = \bar{E}_i(\mathbf{r}) + w\omega_i\mu_0 \int_A G_i(\mathbf{r}, \mathbf{r}')g(\mathbf{r}')E_i(\mathbf{r}')d\mathbf{r}' \tag{2}$$

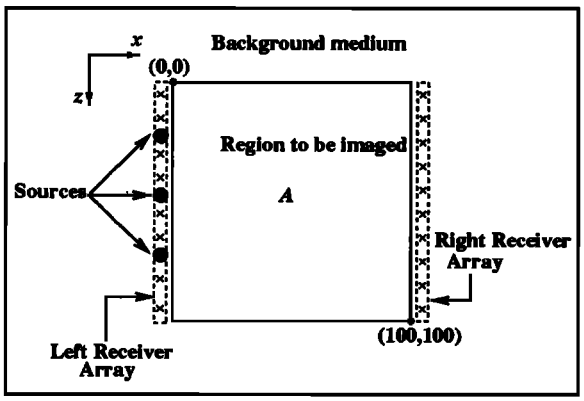

Figure 1. Inverse conductivity problem of interest in this paper.

with $\bar{E}_i(\mathbf{r}) = I_i \imath\omega_i\mu_0 G_i(\mathbf{r}, \mathbf{r}_i)$ the background field generated by a line source with current density I_i positioned at $\mathbf{r}_i$. Finally, $G_i(\mathbf{r}, \mathbf{r}')$ is

$$G_i(\mathbf{r}, \mathbf{r}') = \frac{\imath}{4} H_0^{(1)}(k_{i,0}|\mathbf{r} - \mathbf{r}'|) \tag{3}$$

with $k_{i,0}^2 = \omega_i^2\mu_0(\epsilon_0 + \imath g_0/\omega_i)$ and $H_0^{(1)}(z)$ the zeroth order Hankel function of the first kind.

A discrete representation of (1) is obtained using the method of moments (MoM) [*Harrington*, 1968], where $g(\mathbf{r})$ and $E_i(\mathbf{r})$ are expanded in pulse bases. Similarly, a Galerkin scheme is used to discretize (2) on the basis of the same pulse basis expansions (i.e., we use the pulse basis for both test and weighting functions). Thus, upon discretization, (1) and (2) reduce to

$$\mathbf{y}_i = \mathbf{G}_{i,s}\mathcal{D}(\mathbf{E}_i)\mathbf{g} + \mathbf{n}_i \tag{4a}$$

$$\mathbf{E}_i = \bar{\mathbf{E}}_i + \mathbf{G}_i\mathcal{D}(\mathbf{g})\mathbf{E}_i, \tag{4b}$$

where $\mathbf{E}_i$ (respectively $\bar{\mathbf{E}}_i$) is a vector of pulse basis expansion coefficients for $E_i(\mathbf{r})$ (respectively $\bar{E}_i(\mathbf{r})$), $\mathbf{G}_{i,s}$ (respectively $\mathbf{G}_i$) is a matrix representation of the integral kernel in (1) (respectively (2)), and $\mathcal{D}(\mathbf{x})$ is the diagonal matrix whose (i,i)th element is the ith component of the vector $\mathbf{x}$.

To alleviate the computational burden associated with the use of (4b) and (4a) in an inversion routine [*Torres-Verdín and Habashy*, 1994], we employ a forward model based on the extended Born approximation. For the geometry of interest here, the EBA amounts to approximating $E_i(\mathbf{r})$ in A as [*Torres-Verdín and Habashy*, 1994]

$$E_i(\mathbf{r}) \approx X_i(\mathbf{r})\bar{E}_i(\mathbf{r}) \quad \mathbf{r} \in A \tag{5}$$

$$X_i(\mathbf{r}) = \left[1 - \imath\omega_i\mu_0 \int_A G_i(\mathbf{r}, \mathbf{r}')g(\mathbf{r}')d\mathbf{r}'\right]^{-1}. \tag{6}$$

Upon substituting (5) into (1) and discretizing again using the MoM approach described previously, we obtain the model [*Miller*, 1994]

$$\mathbf{y}_i = \underbrace{\mathbf{T}_i\mathcal{D}(\mathbf{X}_i)\mathbf{g}}_{\mathbf{h}_i(\mathbf{g})} + \mathbf{n}_i. \tag{7}$$

In (7), $\mathbf{T}_i$ is a matrix identical to that associated with the first-order Born approximation in the work by *Miller and Willsky* [1995b], and $\mathbf{X}_i$ is a vector whose nth element is

$$[\mathbf{X}_i]_n = \left(1 + [\mathbf{U}_i]_{(n,:)}\mathbf{g}\right)^{-1}, \tag{8}$$

where $[\mathbf{U}_i]_{(n,:)}$ is the nth row of the matrix $\mathbf{U}_i$ whose (n,m)th element is [*Miller*, 1994]

$$[\mathbf{U}_i]_{n,m} = -\frac{\omega_i\mu_0}{4(\text{area of } A_n)} \int_{A_n}\int_{A_m} H_0^{(1)}(k_{i,0}|\mathbf{r} - \mathbf{r}'|)d\mathbf{r}d\mathbf{r}'$$

with A_j the jth element in the pixelation of A.

Finally, we collect the data from the K scattering experiments into a single vector, $\mathbf{y}$, so that the overall observation model is given by

$$\mathbf{y} = \mathbf{h}(\mathbf{g}) + \mathbf{n}, \tag{9}$$

where $\mathbf{y}^T = [\mathbf{y}_1^T\ \mathbf{y}_2^T, \ldots, \mathbf{y}_K^T]$ and $\mathbf{h}(\mathbf{g})$ and $\mathbf{n}$ are defined accordingly.

2.2. The Inverse Problem and Its Probabilistic Interpretation

One method for recovering $\mathbf{g}$ from $\mathbf{y}$ is to define $\hat{\mathbf{g}}$, the estimate of the conductivity field, as the solution to the nonlinear least squares problem:

$$\hat{\mathbf{g}} = \arg\min_{\mathbf{g}} \|\mathbf{y} - \mathbf{h}(\mathbf{g})\|_{\mathcal{R}^{-1}}^2 + \|\mathbf{g}\|_{\mathbf{L}^T\mathbf{L}}^2, \tag{10}$$

where $\|\mathbf{x}\|_{\mathbf{A}} \equiv \mathbf{x}^T\mathbf{A}\mathbf{x}$. For completeness, the optimization problem should include a positivity constraint on the conductivity. We defer consideration of this question until section 5.2. As discussed in section 1, the matrix $\mathbf{L}$ in (10) is employed to regularize the problem and $\mathcal{R}$ is generally a diagonal matrix

whose entries reflect the noise levels present in the measurements, $\mathbf{y}$.

The Gauss-Newton (GN) algorithm represents one common method for performing the minimization in (10) [*Torres-Verdín and Habashy*, 1994; *Tarantola and Valette*, 1982]. Specifically, $\hat{\mathbf{g}}$ is generated iteratively starting from an initial guess $\hat{\mathbf{g}}^0$. At the $(k+1)$st stage of the algorithm, $\hat{\mathbf{g}}^{k+1}$ is

$$\hat{\mathbf{g}}^{k+1} = \hat{\mathbf{g}}^k + \mathbf{s}^k, \tag{11}$$

where $\mathbf{s}^k$ is given as the solution to the linear system

$$\left\{\mathbf{L}^T\mathbf{L} + \left[\nabla_g \mathbf{h}(\hat{\mathbf{g}}^k)\right]^T \mathcal{R}^{-1} \left[\nabla_g \mathbf{h}(\hat{\mathbf{g}}^k)\right]\right\} \mathbf{s}^k = \left[\nabla_g \mathbf{h}(\hat{\mathbf{g}}^k)\right]^T \mathcal{R}^{-1} \left[\mathbf{y} - \mathbf{h}(\hat{\mathbf{g}}^k)\right] - \mathbf{L}^T\mathbf{L}\hat{\mathbf{g}}^k \tag{12}$$

with $\nabla_g \mathbf{h}(\mathbf{g})$ as the matrix whose (i,j)th component is $\partial \mathbf{h}_i(\mathbf{g})/\partial \mathbf{g}_j$.

The nonlinear least squares formulation and the GN iteration admits a direct interpretation in the context of optimal statistical estimation. Equation (10) is equivalent to the MAP estimate [*Van Trees*, 1968] of $\mathbf{g}$ given $\mathbf{y}$ assuming

$$\mathbf{n}_i \sim \mathcal{N}(0, r_i\mathbf{I}) \tag{13a}$$

$$\mathbf{g} \sim \mathcal{N}(0, (\mathbf{L}^T\mathbf{L})^{-1}), \tag{13b}$$

and all the noises are uncorrelated so that $\mathcal{R} = diag(r_1\mathbf{I}, r_2\mathbf{I}, \ldots, r_K\mathbf{I})$. The notation $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{P})$ implies that $\mathbf{x}$ is a Gaussian random vector with mean $\mathbf{m}$ and covariance matrix $\mathbf{P}$.

An estimation-theoretic approach toward inversion is especially useful, because it provides measures of performance in terms of the second-order error statistics. For the nonlinear problem of interest here, explicit expressions for this information cannot generally be obtained; however, the Cramér-Rao bound (CRB) is a commonly used lower bound on the mean square error performance of the MAP estimator [*Van Trees*, 1968]. For the problem defined by (9), the CRB takes the form (A. S. Willsky, Class Notes, Spring 1990),

$$\mathcal{P}_{\mathrm{CRB}}(\mathbf{g}) = \left\{\mathbf{L}^T\mathbf{L} + \left[\nabla_g \mathbf{h}(\mathbf{g})\right]^T \mathcal{R}^{-1} \left[\nabla_g \mathbf{h}(\mathbf{g})\right]\right\}^{-1}. \tag{14}$$

Technically, the CRB is defined to be the inverse of the expected value of the matrix in braces on the right-hand side of (14), where the expectation is taken with respect to the distribution of $\mathbf{g}$. In this paper, it proves useful to consider the definition of (14) in which the expectation is replaced by evaluation at an arbitrary $\mathbf{g}$. Finally, comparing (14) to (12), we see that $\mathcal{P}_{\mathrm{CRB}}(\hat{\mathbf{g}})$ is the inverse of the matrix on the left-hand side of the linear system defining the Gauss-Newton iteration. In section 5, we make extensive use of this fact to reduce the computational complexity of our inversion algorithm.

3. Wavelet Domain Formulation

3.1. The Discrete Wavelet Transform

The basic idea behind the discrete wavelet transform is to decompose a signal, here represented as a vector, into a sequence of increasingly "coarser" representations while retaining the information lost in moving from a fine to a coarse scale. While we will be concerned both with one- and two-dimensional signals, we describe first the wavelet representation and notation for a one-dimensional signal, $\mathbf{a}$, of dimension 2^{M_a}. The elements of $\mathbf{a}$ are termed the finest scale scaling coefficients associated with $\mathbf{a}$, and the vector $\mathbf{a}$ is denoted by $\mathbf{a}(M_a)$, indicating that this is a representation of $\mathbf{a}$ at the finest scale, M_a.

Beginning with $\mathbf{a}(M_a)$, a coarser set of scaling coefficients, $\mathbf{a}(M_a - 1)$, is obtained by passing $\mathbf{a}(M_a)$ through a low-pass, finite impulse response filter, l, 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 filter 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 $m = M_a - 1, \ldots, L_a$ with L_a the coarsest scale at which $\mathbf{a}$ is represented.

As described by *Daubechies* [1988], filters l and h can be constructed so that we may build an orthonormal matrix, $\mathcal{W}_a$ relating the finest scale scaling coefficients to the coarsest scaling coefficients and all detail coefficients. We subscript the wavelet transform operator as $\mathcal{W}_a$ to make explicit that this is the transform for $\mathbf{a}$. We will use different wavelet transforms for the different variables. Thus, we write

$$\boldsymbol{\alpha} = \mathcal{W}_a \mathbf{a}, \tag{15}$$

where $\boldsymbol{\alpha} = [\boldsymbol{\alpha}(M_a - 1)^T, \ldots, \boldsymbol{\alpha}(L_a)^T \, \mathbf{a}(L_a)^T]^T$ is the wavelet transform of $\mathbf{a}$. The nth element of $\boldsymbol{\alpha}(m)$ is denoted $\boldsymbol{\alpha}(m, n)$ and is referred to as the nth shift of $\boldsymbol{\alpha}$ at scale m. Similarly, $\mathbf{a}(m, n)$ is the nth element of the vector of scaling coefficients at scale m.

The wavelet decomposition of a two-dimensional function is obtained by considering $\mathbf{a}$ as a matrix and applying one orthonormal wavelet transform to the columns and another to the rows. We use $\mathcal{W}_a$ to represent the composition of these two operators. It is easily shown that the orthonormality of the row and column transforms ensures that $\mathcal{W}_a$ is also orthonormal. Finally, we denote a particular element of $\boldsymbol{\alpha}$ by $\boldsymbol{\alpha}(\mathbf{m}, \mathbf{n})$, where $\mathbf{m}$ and $\mathbf{n}$ are two vectors indexing the scales and shifts in the x and z directions, respectively.

3.2. Multiscale Prior Models

A key component in our formulation of the inverse problem is the use of a multiscale stochastic model for $\mathbf{g}$ to regularize the inversion and to capture prior information. To motivate the particular choice of prior model used here, consider (13b), the stochastic interpretation of the regularization term in the nonlinear least squares formulation of the inverse scattering problem. In the case where $\mathbf{g}$ is a function of one variable and $\mathbf{L}$ represents first-order differentiation, (13b) implies that $\mathbf{g}$ is a Brownian motion satisfying $\mathbf{Lg} = \mathbf{w}$ with $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Work by *Wornell* [1990] *Flandrin* [1992] and *Tewfick and Kim* [1992] has demonstrated that Brownian motions and other related fractal processes can be closely approximated via statistical models in which the wavelet and coarsest scale scaling coefficients of $\mathbf{g}$ are independent random variables distributed according to

$$\gamma(m, n) \sim \mathcal{N}(0, \kappa^2 2^{-\mu m}) \tag{16a}$$

$$\mathbf{g}(L_g, n) \sim \mathcal{N}(0, p_{L_g}). \tag{16b}$$

Here κ^2 controls the overall magnitude of the process and μ determines the fractal structure of sample paths. The scalar p_{L_g} is chosen to be a sufficiently large number so as to avoid any bias in the estimator of the low-frequency structure of $\mathbf{g}$. For these models, the resulting covariance matrix, $\mathbf{P}_0$, for $\boldsymbol{\gamma}$ is diagonal with nonzero entries corresponding to the variances of each component of $\boldsymbol{\gamma}$.

For the case where $\mathbf{g}$ is two-dimensional, we consider the separable representation with

$$\boldsymbol{\gamma}(\mathbf{m}, \mathbf{n}) \sim (0, \kappa_x^2 \kappa_z^2 2^{-(\mu_x m_x + \mu_z m_z)})$$

for $L_{g,x} \leq m_x \leq M_{g,x} - 1$ and $L_{g,z} \leq m_z \leq M_{g,z} - 1$. For $m_x = L_{g,x}$, we take $\boldsymbol{\gamma}(\mathbf{m}, \mathbf{n}) \sim (0, p_{L_{g,x}} \kappa_z^2 2^{-(\mu_z m_z)})$ with analogous results holding when $m_z = L_{g,z}$.

3.3. A Scale-Space Representation of the Extended Born Model

We use the wavelet transform to effect a change of basis with respect to the EBA model developed in section 2.1. To perform the scale-space transformation, we define one wavelet transform, $\mathcal{W}_g$, for the conductivity perturbation, $\mathbf{g}$, and for each scattering experiment a separate wavelet transform $\mathcal{W}_i$, $i = 1, 2, \ldots, K$, is specified. Taking advantage of the orthonormality of the wavelet transform, (7) is transformed as follows:

$$\begin{aligned} \mathcal{W}_i \mathbf{y}_i &\equiv \boldsymbol{\eta}_i \\ &= \underbrace{\left[\mathcal{W}_i \mathbf{T}_i \mathcal{W}_g^T\right]}_{\boldsymbol{\Theta}_i} \underbrace{\left[\mathcal{W}_g \mathcal{D}(\mathbf{X}_i) \mathcal{W}_g^T\right]}_{\boldsymbol{\Xi}_i(\boldsymbol{\gamma})} (\mathcal{W}_g \mathbf{g}) \\ &\quad + \mathcal{W}_i \mathbf{n}_i \\ &\equiv \boldsymbol{\Phi}_i(\boldsymbol{\gamma}) + \boldsymbol{\nu}_i, \end{aligned} \tag{17}$$

where $\boldsymbol{\Phi}_i(\boldsymbol{\gamma}) = \boldsymbol{\Theta}_i \boldsymbol{\Xi}_i(\boldsymbol{\gamma}) \boldsymbol{\gamma}$, and $\boldsymbol{\Xi}_i(\boldsymbol{\gamma})$ is related to $\boldsymbol{\gamma}$ by noting that (8) can be written as

$$\begin{aligned} [\mathbf{X}_i]_n &= \left(1 + [\mathbf{U}_i]_{(n,:)} \mathcal{W}_g^T \mathcal{W}_g \mathbf{g}\right)^{-1} \\ &\equiv \left(1 + [\boldsymbol{\Upsilon}_i]_{(n,:)} \boldsymbol{\gamma}\right)^{-1}. \end{aligned} \tag{18}$$

Gathering the transform domain observations, we obtain

$$\boldsymbol{\eta} = \boldsymbol{\Phi}(\boldsymbol{\gamma}) + \boldsymbol{\nu}, \tag{19}$$

where $\boldsymbol{\eta}$, $\boldsymbol{\Phi}(\boldsymbol{\gamma})$, and $\boldsymbol{\nu}$ are defined analogously to the quantities in (9).

In summary, the overall problem of interest in this paper is the determination of the MAP estimate of $\boldsymbol{\gamma}$ based on the observation model in (19) with $\boldsymbol{\gamma} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_0)$ and $\boldsymbol{\nu} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, where $\mathbf{P}_0$ is defined in section 3.2 and $\mathbf{R} = diag(r_1 \mathbf{I}, r_2 \mathbf{I}, \ldots, r_K \mathbf{I})$. Formally, this leads to a definition of $\hat{\boldsymbol{\gamma}}$, the estimate of the conductivity's wavelet transform, as

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} J(\boldsymbol{\gamma}), \tag{20}$$

where the cost function, $J(\gamma)$, is defined as

$$J(\gamma) = \|\eta - \Phi(\gamma)\|^2_{\mathbf{R}^{-1}} + \|\gamma\|^2_{\mathbf{P}_0^{-1}}. \quad (21)$$

While (20) is structurally identical to (10), a wavelet-domain formulation yields a variety of benefits explored in the following sections of this paper.

4. The Relative Cramér-Rao Bound

In the work by *Miller and Willsky* [1995b], the relative error covariance matrix (RECM) was introduced as a quantitative tool to analyze issues related to the manner in which the physics of the linearized inverse scattering problem impact the structure of the reconstruction. Without explicit knowledge of the error covariance matrix for the nonlinear inverse scattering problem, we use the Cramér-Rao bound as the basis for a generalized RECM. Specifically, we define the relative Cramér-Rao bound (RCRB) as

$$\mathbf{\Pi}_{\mathrm{CRB}}(\gamma) = \mathbf{P}_0^{-T/2}[\mathbf{P}_0 - \mathbf{P}_{\mathrm{CRB}}(\gamma)]\mathbf{P}_0^{-1/2} \quad (22)$$

which, for the model given by (19), is

$$\mathbf{P}_{\mathrm{CRB}}(\gamma) = [\boldsymbol{\mathcal{T}}^T(\gamma)\mathbf{R}^{-1}\boldsymbol{\mathcal{T}}(\gamma) + \mathbf{P}_0^{-1}]^{-1} \quad (23)$$

with

$$\boldsymbol{\mathcal{T}}(\gamma) = \left\{[\nabla_\gamma \Phi_1(\gamma)]^T, \ldots, [\nabla_\gamma \Phi_K(\gamma)]^T\right\}^T. \quad (24)$$

A detailed description of the straightforward yet tedious calculations required to compute $\boldsymbol{\mathcal{T}}(\gamma)$, as well as an expression for $\boldsymbol{\mathcal{T}}(\gamma)$, are presented by *Miller* [1994].

As with the RECM defined for the linear case, $\mathbf{\Pi}_{\mathrm{CRB}}$ possesses a variety of useful properties. From (22), $\mathbf{\Pi}_{\mathrm{CRB}}$ is a symmetric matrix. Also, like the RECM, the diagonal elements of $\mathbf{\Pi}_{\mathrm{CRB}}$ are bounded between zero and one and provide a direct measure as to the information content of the data with respect to each component of γ. This property of the RCRB provides a mechanism for defining the space-varying appropriate level of detail to include in a reconstruction of $\mathbf{g}$ [*Miller and Willsky*, 1995b]. In particular, for each location $\mathbf{j}$ in the finest scale representation of $\mathbf{g}$, we say that the data support a reconstruction of $g(\mathbf{M}_g, \mathbf{j})$ at scale $\mathbf{m}$ if there exists a component in γ at scale $\mathbf{m}$ for which the data provide a sufficiently large quantity of information; that is, the diagonal element of $\mathbf{\Pi}_{\mathrm{CRB}}$ associated with that coefficient of γ is greater than a threshold, $\tau \in [0, 1)$. The finest scale detail to include in the reconstruction at that point $\mathbf{j}$ is just the finest scale for which a wavelet coefficient may be found that satisfies the above criterion. In addition to obtaining these detail maps, the ability of the RCRB to provide a measure of the information content in the data leads to a substantial reduction in model complexity for the inversion algorithm developed in the following section.

5. A Wavelet-Based Inversion Algorithm

5.1. The Basic Algorithm

On the basis of (11) and (12), the $(k+1)$st step of the scale-space form of the Gauss-Newton algorithm is

$$\hat{\gamma}^{k+1} = \hat{\gamma}^k + \varsigma^k \quad (25)$$

and

$$\left[\boldsymbol{\mathcal{T}}^T(\hat{\gamma}^k)\mathbf{R}^{-1}\boldsymbol{\mathcal{T}}(\hat{\gamma}^k) + \mathbf{P}_0^{-1}\right]\varsigma^k = \boldsymbol{\mathcal{T}}^T(\hat{\gamma}^k)\mathbf{R}^{-1}\left[\eta - \Phi(\hat{\gamma}^k)\right] - \mathbf{P}_0^{-1}\hat{\gamma}^k, \quad (26)$$

which we write in a more compact form as

$$\left[\boldsymbol{\mathcal{F}}^k + \mathbf{P}_0^{-1}\right]\varsigma^k = \mathbf{v}^k, \quad (27)$$

where $\boldsymbol{\mathcal{F}}^k = \boldsymbol{\mathcal{T}}^T(\hat{\gamma}^k)\mathbf{R}^{-1}\boldsymbol{\mathcal{T}}(\hat{\gamma}^k)$, and we recall from the discussion in section 2.2 that the matrix on the left-hand side of (26) is $\mathbf{P}^{-1}_{\mathrm{CRB}}(\hat{\gamma}^k)$.

At the kth step of the Gauss-Newton algorithm, we would like to use the diagonal components of $\mathbf{P}_{\mathrm{CRB}}(\hat{\gamma}^k)$ and $\mathbf{P}_0$ to construct the diagonal elements of the RCRB according to (22) so that we can determine and then compute only those elements of ς for which "significant" information exists. The difficulty here is that we only have access to the inverse of $\mathbf{P}_{\mathrm{CRB}}(\hat{\gamma}^k)$, and our desire is to avoid explicitly inverting this matrix. Thus we make two assumptions. First, at step k we assume that we know the diagonal elements of $\mathbf{P}_{\mathrm{CRB}}(\hat{\gamma}^{k-1})$, that is, the diagonals from the Cramér-Rao bound matrix of the previous iteration of the algorithm, from which we are able to

construct the diagonal elements of $\mathbf{\Pi}_{\mathrm{CRB}}(\hat{\gamma}^{k-1})$ using (22) (since $\mathbf{P}_0$ is diagonal). For the first iteration of the algorithm, we begin by explicitly inverting the matrix on the right-hand side of (27). Second, we assume that $\mathbf{\Pi}_{\mathrm{CRB}}(\hat{\gamma}^{k-1})$ is close to $\mathbf{\Pi}_{\mathrm{CRB}}(\hat{\gamma}^{k})$; that is, the bounds at successive iterations do not change dramatically.

Formally, we let $\mathbf{\Pi}^{\mathbf{m}}_{\mathrm{CRB},\mathbf{n}}(\hat{\gamma}^{k-1})$ be the component on the diagonal of $\mathbf{\Pi}_{\mathrm{CRB}}(\hat{\gamma}^{k-1})$ corresponding to the wavelet coefficient in γ at scale $\mathbf{m}$ and shift $\mathbf{n}$. Using $\mathbf{\Pi}^{\mathbf{m}}_{\mathrm{CRB},\mathbf{n}}(\hat{\gamma}^{k-1})$, we partition ς^k into ς_1^k and ς_2^k, where the component of ς^k at scale $\mathbf{m}$ and shift $\mathbf{n}$ is included in ς_1^k if $\mathbf{\Pi}^{\mathbf{m}}_{\mathrm{CRB},\mathbf{n}}(\hat{\gamma}^{k-1})$ is greater than some threshold $\tau \in [0,1)$. If this condition is not met, then that element of ς^k is placed into ς_2^k. Thus ς_1^k contains those components of ς^k for which significant information is available relative to that of the prior model where the level of significance is defined by the threshold τ.

On the basis of this decomposition of ς^k, the rows and columns of (27) are appropriately permuted so that the linear system at step k of the Gauss-Newton algorithm takes the block-partitioned form

$$\begin{bmatrix} \mathcal{F}_{1,1}^k + \mathbf{P}_{0,1}^{-1} & \mathcal{F}_{1,2}^k \\ \mathcal{F}_{2,1}^k & \mathcal{F}_{2,2}^k + \mathbf{P}_{0,2}^{-1} \end{bmatrix} \begin{bmatrix} \varsigma_1^k \\ \varsigma_2^k \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^k \\ \mathbf{v}_2^k \end{bmatrix}, \tag{28}$$

which we invert directly using the block matrix inversion formula [*Beyer*, 1987] to obtain

$$\begin{bmatrix} \varsigma_1^k \\ \varsigma_2^k \end{bmatrix} = \begin{bmatrix} \mathcal{Q}^k + \mathcal{Q}^k \mathcal{F}_{1,2}^k \mathcal{S}^{-k} \mathcal{F}_{2,1}^k \mathcal{Q}^k & -\mathcal{Q}^k \mathcal{F}_{1,2}^k \mathcal{S}^{-k} \\ -\mathcal{S}^{-k} \mathcal{F}_{2,1}^k \mathcal{Q}^k & \mathcal{S}^{-k} \end{bmatrix} \bullet \begin{bmatrix} \mathbf{v}_1^k \\ \mathbf{v}_2^k \end{bmatrix}. \tag{29}$$

In (29), $\mathcal{Q}^k$ and the Schur complement, $\mathcal{S}^k$, are defined as

$$\mathcal{Q}^k = (\mathcal{F}_{1,1}^k + \mathbf{P}_{0,1}^{-1})^{-1} \tag{30}$$

$$\mathcal{S}^k = \mathbf{P}_{0,2}^{-1} + \mathcal{F}_{2,2}^k - \mathcal{F}_{2,1}^k \mathcal{Q}^k \mathcal{F}_{1,2}^k \tag{31}$$

and $\mathcal{S}^{-k} = (\mathcal{S}^k)^{-1}$.

The utility of (29) through (31) arises from two observations. First, we anticipate from our work on linear inverse problems [*Miller and Willsky*, 1995b] that the dimension of ς_1^k should be much smaller than the dimension of the full vector ς^k. Hence the cost of computing the inverse in (30) will be small. Second, the Schur complement matrix is well approximated by only its diagonal components, so that evaluation of $\mathcal{S}^{-k}$ requires little computation. While a detailed verification of this approximation is provided by *Miller* [1994], here we demonstrate its validity through the accurate results obtained in our examples using the diagonal approximation to $\mathcal{S}^k$. The diagonal approximation to $\mathcal{S}^k$ means that ς_1^k may be obtained from $\mathbf{v}^k$ using the first block row of (29) with far fewer computations than direct inversion of $\mathcal{F}^k$. Moreover, rather than setting ς_2^k to zero, the presumed diagonal structure of $\mathcal{S}^k$ and the small size of $\mathcal{Q}^k$ imply that the second block row of (29) can be applied to $\mathbf{v}^k$ with little computational overhead. Finally, because the matrix on the right-hand side of (29) is a row- and column-permuted form of $\mathbf{P}_{\mathrm{CRB}}(\hat{\gamma}^k)$, (29) provides an efficient method for computing the diagonal elements of the RCRB to be used in the next iteration of the Gauss-Newton method. Specifically, the diagonals of $\mathbf{P}_{\mathrm{CRB}}(\hat{\gamma}^k)$ are obtained by inverting the diagonal elements of $\mathcal{S}^k$ and computing the diagonal elements of the upper left block of the matrix in (29), which is small in dimension.

Table 1. Data Set Definitions for Observation Processes of Interest in the Paper

Experiment Number	Source Position	Frequency of Source kHz	Receiver Array
1–6	1–6	$f_{\mathrm{LO}} = 0.1$	Right
7–12	1–6	$f_{\mathrm{MID}} = 1.0$	Left
13–18	1–6	$f_{\mathrm{HI}} = 10.0$	Left

Table 2. Common Parameters Defining Numerical Experiments

Parameter	Value
g_0	0.1 S/m
z Wavelet	Daubechies 2-tap
x Wavelet	Daubechies 2-tap
$N_{g,z} = N_{g,x}$	16
$M_{g,z} = M_{g,x}$	4
$L_{g,z} = L_{g,x}$	2
$\mu_z = \mu_x$	1
$p_{L_g,z} = p_{L_g,x}$	16
$\kappa_z^2 = \kappa_x^2$	0.1

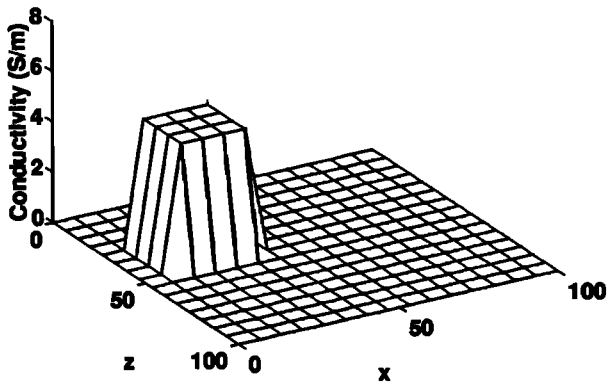

Figure 2. True conductivity profile for the first example.

5.2. Implementation Details

To improve the convergence of the nonlinear optimization procedure, we implement a form of the Gauss-Newton algorithm known as the Levenberg-Marquardt (LM) method [*Gill et al.*, 1981]. Essentially, this approach replaces (25) and (27) by

$$\hat{\gamma}^{k+1} = \hat{\gamma}^k + \varsigma^k \tag{32a}$$

$$\left[\mathcal{F}^k + (\kappa^{-2} + \alpha^k)\tilde{\mathbf{P}}_0^{-1}\right]\varsigma^k = v^k. \tag{32b}$$

where $\tilde{\mathbf{P}}_0 = \mathbf{P}_0/(\kappa_x^2\kappa_z^2)$ with $\kappa_x^2\kappa_z^2$ defined in section 3.2. As discussed by *Gill et al.* [1981], proper selection of α^k guarantees that $J(\hat{\gamma}^{k+1}) < J(\hat{\gamma}^k)$, so that the LM procedure does in fact converge to a minimum of the cost function. Here we choose α^k so as to minimize the cost associated with the resulting $\hat{\gamma}^{k+1}$, that is,

$$\alpha^k = \arg\min_{\alpha} J\left\{\hat{\gamma}^k + \left[\mathcal{F}^k + (\kappa^{-2} + \alpha^k)\tilde{\mathbf{P}}_0^{-1}\right]^{-1}\mathbf{v}^k\right\}. \tag{33}$$

We use a line minimization procedure to solve (33), resulting in around 20 calls to the EBA forward solver per iteration in the examples presented in section 6. Clearly, a more efficient implementation of the LM algorithm would avoid much of this burden.

Physical principles dictate that the overall conductivity, $g_0 + g$, in region A be positive. In order to enforce this constraint on the inversion algorithm, we assume that the ith element of the vector $\mathbf{g}$ is given

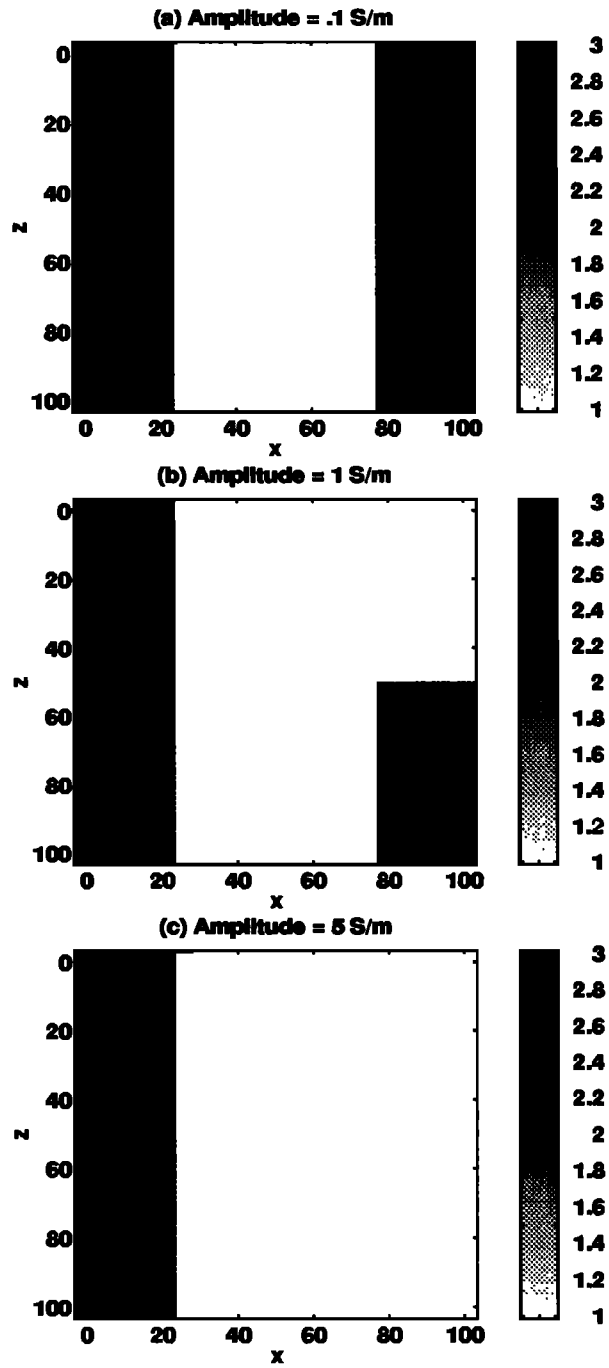

Figure 3. Bounds on the space-varying optimal level of x-oriented detail in a reconstruction when the structure in Figure 2 has an amplitude of (a) 0.1 S/m, (b) 1 S/m, and (c) 5 S/m. All figures are computed using a threshold of 0.5 in the relative Cramér-Rao bound analysis. The corresponding z-oriented detail maps are constant over the region A.

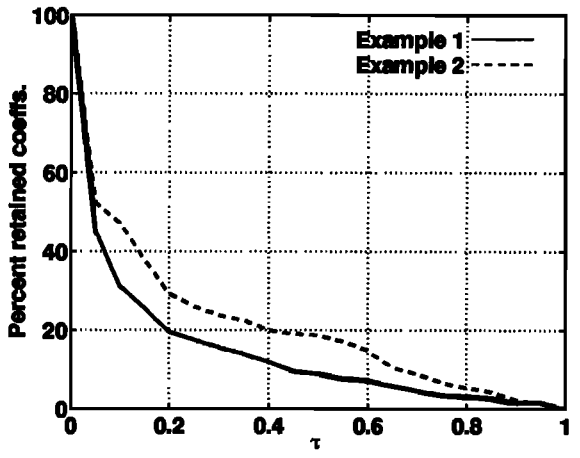

Figure 4. Percentage of elements expected to be considered significant as a function of truncation parameter. Solid line: results for first example; dashed line: results for second example.

as

$$\mathbf{g}_i = e^{\check{\mathbf{g}}_i} - g_0, \tag{34}$$

where we estimate the vector $\check{\boldsymbol{\gamma}} = \mathcal{W}_g \check{\mathbf{g}}$ rather than $\mathbf{g}$. Here $\check{\mathbf{g}}$ is obtained by lexicographically ordering the $\check{\mathbf{g}}_i$. As in the work by *Wang et al.* [1994], this change of variables only results in a slight modification to the structure of $\Phi(\boldsymbol{\gamma})$ and its gradient matrix. The general structure of the LM algorithm, including the computational benefits associated with a wavelet domain implementation, are not affected by this substitution.

Finally, the matrices defining the extended Born approximation are either sparse by construction (Ξ_i in (17) which is the wavelet transform of a diagonal matrix) or are of the variety which are made sparse by the wavelet transform (Θ_i in (17) and Υ_i in (19)). In the examples considered in section 6, we explore the effects of truncating the small elements of Θ_i and Υ_i in the performance of the inversion algorithm. For this, we follow the strategy of *Alpert*

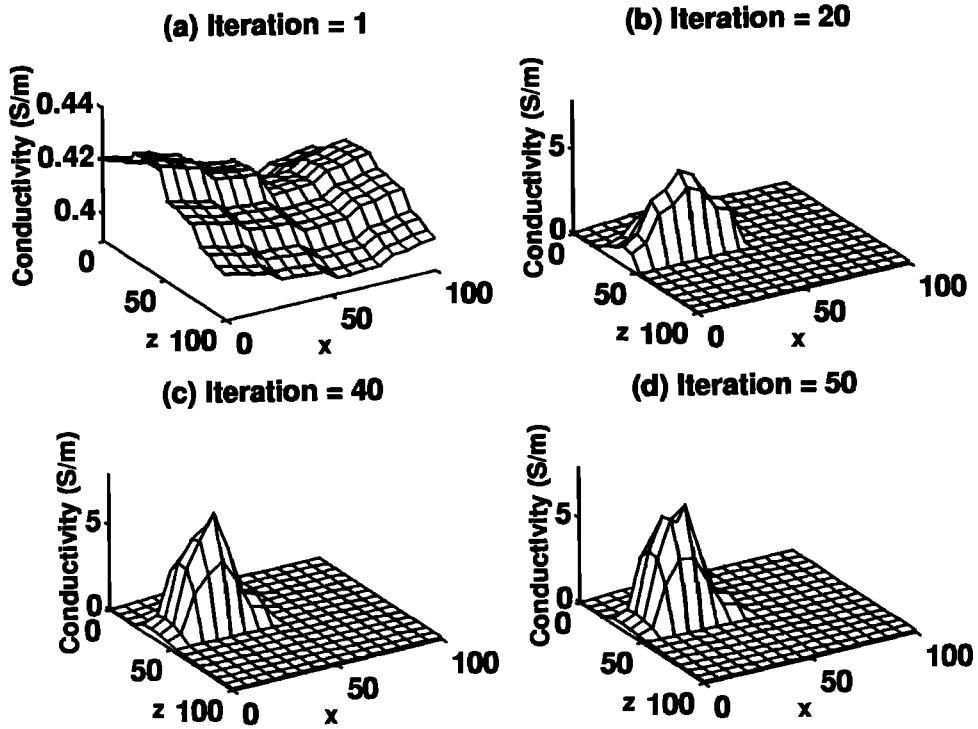

Figure 5. Estimated conductivity structure after 1, 20, 40, and 50 iterations of the nonlinear, multiscale inverse scattering algorithm. The true conductivity is shown in Figure 2. Note that the z-axis scaling in (a) is different from that in (b) to (d).

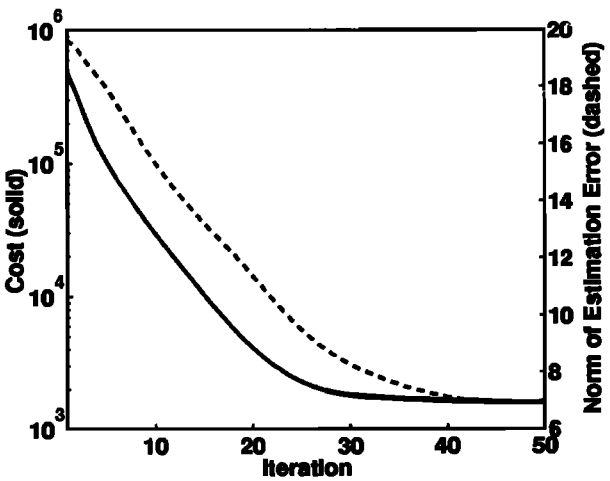

Figure 6. Performance curves as a function of iteration numbers for the first example. Solid line: value of cost function, $J(\gamma)$; dashed line: norm of estimation error, $(\|\mathbf{g} - \hat{\mathbf{g}}^k\|)$.

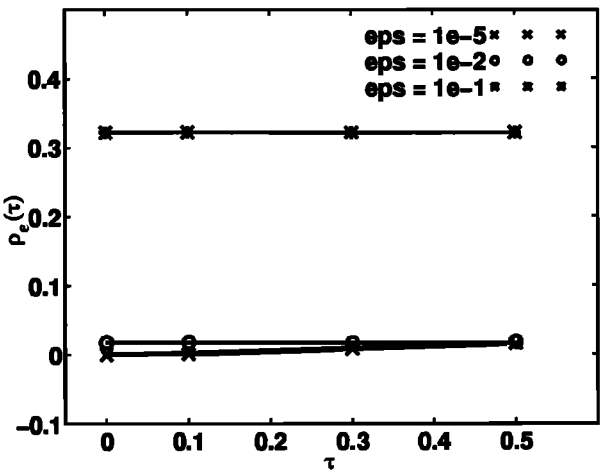

Figure 7. Plots of $\rho_\epsilon(\tau)$ for the first example.

et al. [1993] in which the level of truncation is governed by a parameter $\epsilon > 0$ with zero corresponding to no truncation.

6. Examples

We now consider the problem of recovering the electrical conductivity in region A of Figure 1 on the basis of observations of scattered radiation from the 18 scattering experiments described in Table 1. The source frequencies are 100 Hz (used in a cross-well configuration), as well as 1 and 10 kHz (used to obtain information near the left edge of A). Region A is 100 m x 100 m with the top left corner at $(0, 0)$. For the method of moments discretization, A is decomposed into a 16 x 16 array of square pixels so that the overall dimensionality of the problem is 256. Both receiver arrays consist of 32 equally spaced elements extending from $z = -0.05$ m to $z = 100.05$ m. The left array is located at $x = -0.05$ m and the right at $x = 100.05$ m. The six sources are located along the line $x = -0.05$ m equally spaced from $z = 0.05$ m to $z = 99.95$ m. All other parameters describing the examples presented in this section are listed in Table 2. We define the signal-to-noise ratio (SNR) for the model $\boldsymbol{\eta}_i = \Phi_i(\gamma) + \boldsymbol{\nu}_i$ with $\boldsymbol{\eta}_i \in \mathbb{R}^{N_i}$ and $\boldsymbol{\nu}_i \sim \mathcal{N}(\mathbf{0}, r_i^2\mathbf{I})$ as

$$SNR_i^2 = \frac{\Phi_i^T(\gamma)\Phi_i(\gamma)}{N_i r_i^2}. \tag{35}$$

While the inverse algorithm is based on the extended Born approximation, all data are generated using the exact physical model in (4a). Finally, for all examples, $\hat{\gamma}^0 = \mathbf{0}$.

As a first example, we examine the inverse problem for which the geometric structure of the underlying conductivity field is shown in Figure 2. Unlike the linear inverse problem in which performance is independent of the true conductivity field [*Miller and Willsky*, 1995b], in the nonlinear case, expected performance of the imaging algorithm does depend on the configuration of the underlying conductivity profile. As an example of this relationship, we use the RCRB-based analysis methods, and we display in Figures 3a–3c the finest scales of x-oriented detail supported in a reconstruction when the amplitude of the rectangular structure in Figure 2 is 0.1 S/m,

Table 3. Mean Length of ς_1 for the First Example, Expressed as a Percent of the Maximum Length (256 in This Case)[a]

	ϵ		
τ	10^{-5}	0.01	0.1
0.0	100	100	100
0.1	24	25	30
0.3	14	14	18
0.5	11	11	14

[a]Each entry in this table is obtained as an average over the 50 iterations of the corresponding inversion.

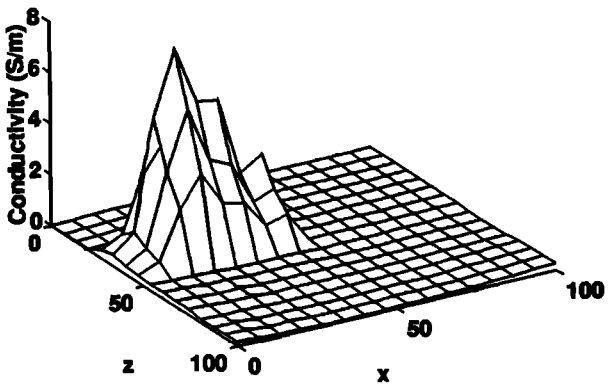

Figure 8. Final conductivity reconstruction for the first example with $\tau = 0.5$ and $\epsilon = 0.1$.

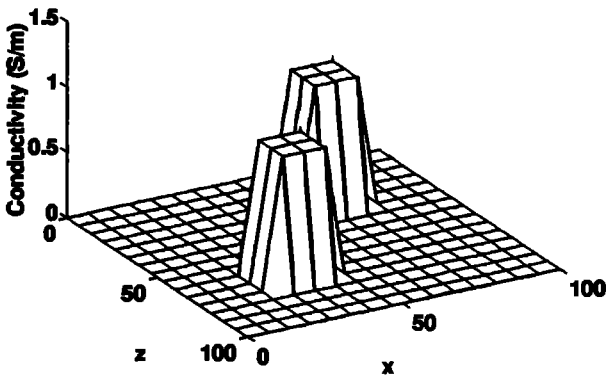

Figure 9. True conductivity profile for the second example.

1 S/m, and 5 S/m, corresponding to 1:1, 10:1, and 50:1 contrasts, respectively. The finest z-oriented detail for this problem is basically constant over A because of the dense sampling pattern of the receivers along each vertical edge of the region. All figures were obtained using an RCRB threshold parameter $\tau = 0.50$. These illustrations demonstrate that as the amplitude of the conductivity block increases, the level of resolvable detail decreases. Comparing Figure 3a to 3b, we see a drop in the scale of detail for the region directly behind the conductivity structure. This loss in detail is explained by the fact that the dissipation of electrical energy in a medium is directly related to the level of the conductivity with larger values corresponding to greater loss.

The spatial structure of Figure 3 is related closely to the three different frequencies used to probe the medium. The ability to recover finer scale information near the left vertical edge of A extends to ranges on the order of a skindepth (16 m and 50 m for the 10 kHz and 1 kHz sources, respectively) for the two higher frequencies. Although the skin depth for the 100-Hz source is 160 m while the size of A is only 100 m, the frequency is sufficiently low such that only the coarsest scale of information regarding $\mathbf{g}$ may be recovered further into the medium. Finally, note that the high detail which can be recovered near the two vertical edges also can be partially attributed to the singularities of the Green's functions for this problem which exist at the source and receiver locations. A more detailed treatment of these issues in the context of a linearized inverse conductivity problem similar to the one considered here may be found in the work by *Miller and Willsky* [1995b].

To gauge the expected savings associated with the use of the RCRB at each iteration of our inversion algorithm, in Figure 4 we plot (solid line), as a function of τ, the percent of coefficients for which the data are deemed to supply a reasonable amount of information (i.e., the percentage of elements in γ for which the value of $\mathbf{\Pi}^{\mathbf{m}}_{\mathbf{CRB},\mathbf{n}}$ exceeds the threshold τ). This illustration is obtained for the case in which the amplitude of the structure in Figure 2 is set to 5 S/m and provides a benchmark for the expected performance of the inversion routine explored in subsequent paragraphs. Specifically, Figure 4 is generated using the CRB matrix evaluated at the true conduc-

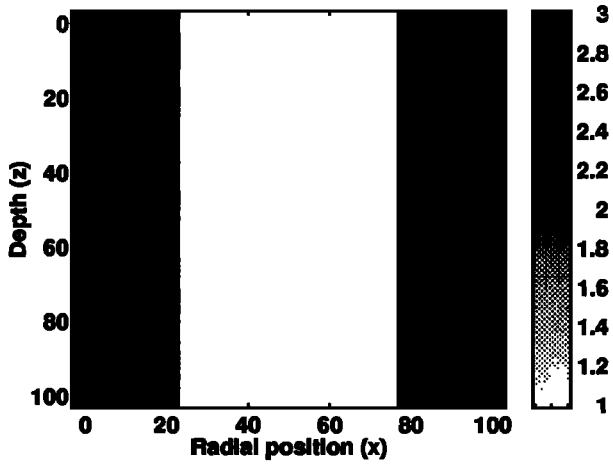

Figure 10. Bounds on space-varying optimal level of x-oriented detail for $\tau = 0.5$ in a reconstruction when the structures in Figure 9 both have amplitudes of 1 S/m. The corresponding z-oriented detail maps are constant over the region A.

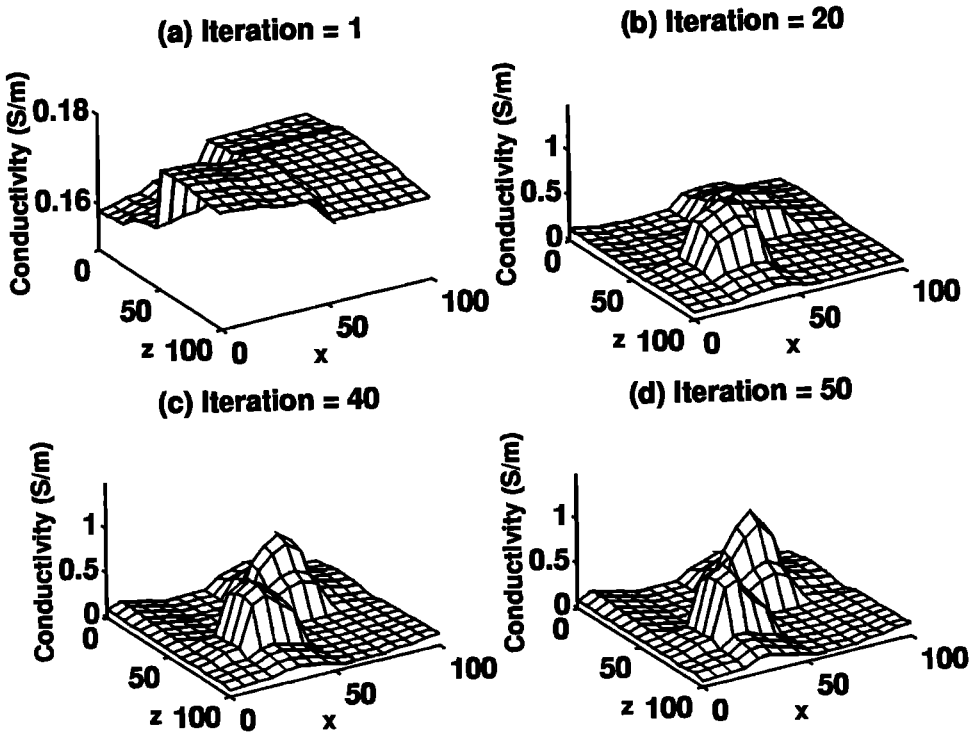

Figure 11. Estimated conductivity structure after 1, 20, 40, and 50 iterations of the nonlinear, multiscale inverse scattering algorithm. The true conductivity is shown in Figure 9. Note that the z-axis scaling in (a) is different from that in (b) to (d).

tivity profile, while the CRB used at the kth step in the inversion algorithm is evaluated at $\hat{\gamma}^k$. Thus there will be some discrepancy between the number of coefficients retained during the execution of the algorithm and the number indicated in Figure 4; however, as we shall see shortly, this difference is small.

The estimates of the conductivity field after 1, 20, 40, and 50 iterations of the LM algorithm are shown in Figure 5. These results are obtained with $\epsilon = 10^{-5}$ (corresponding to EBA matrices all greater than 99% full) and $\tau = 0$ so that no approximation was made in the least squares problems. In Figure 6, we plot the values of the cost function $J(\gamma)$ in (20) and the norm of the estimation error $\|\mathbf{g} - \hat{\mathbf{g}}^k\|$, as a function of iteration number. We observe that after 50 iterations, the estimate of the conductivity has converged to at least a local minimum of the cost function. Moreover, this reconstruction captures the basic features of the true conductivity field in that we have isolated both the location of the perturbation as well as its amplitude.

In Table 3, the average length of the vector ς_1 over the 50 iterations of the LM procedure is shown as a percent of the size of the problem, 256, as we vary the parameters ϵ and τ. As ϵ is raised to 0.01 (respectively 0.10) the sparsity of the EBA matrices drops to around 30% (respectively less than 10%). Examining Table 3, we see that for $\tau = 0.5$ and $\epsilon = 0.01$, on average only 11% of the elements of $\hat{\gamma}^k$ (corresponding to 28 rather than 256 degrees of freedom) are deemed important. Finally, comparing Table 3 and the solid line in Figure 4 shows excellent agreement between the average number of coefficients retained at each iteration of the inversion and the theoretical bound obtained using the RCRB analysis based on the true conductivity profile.

In Figure 7, we observe that this truncation procedure results in little degradation in inversion performance. Define $\hat{\gamma}^{50}_{\tau,\epsilon}$ as the estimate of γ obtained after 50 iterations of the LM algorithm with truncation parameters τ and ϵ. In Figure 7, we plot as a function of τ and for $\epsilon \in \{10^{-5}, 10^{-2}, 10^{-1}\}$, the

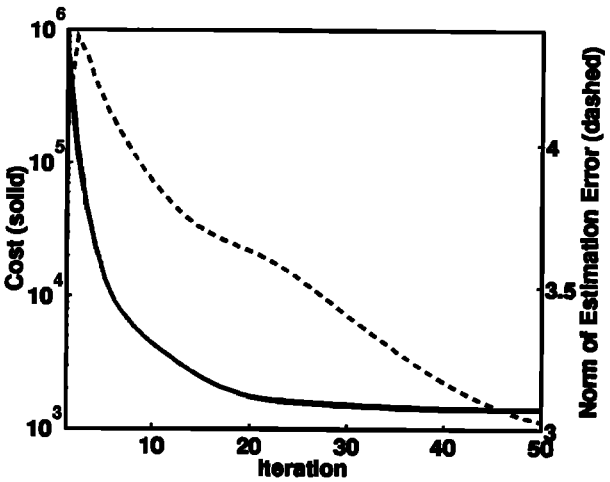

Figure 12. Performance curves as a function of iteration numbers for the second example. Solid line: value of cost function, $J(\gamma)$; dashed line: norm of estimation error, $(\|\mathbf{g} - \hat{\mathbf{g}}^k\|)$.

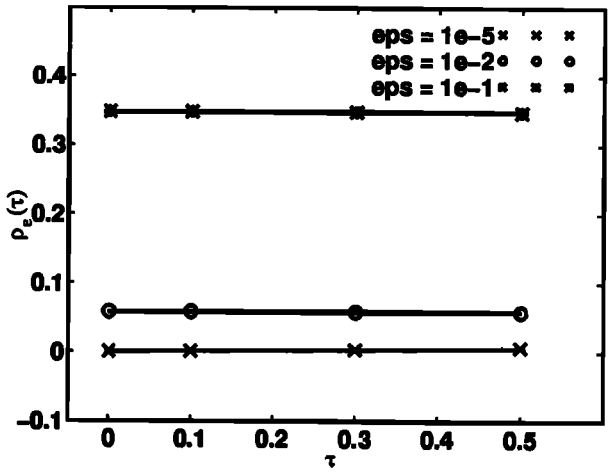

Figure 13. Plots of $\rho_\epsilon(\tau)$ for the second example.

quantity

$$\rho_\epsilon(\tau) = \frac{\|\hat{\gamma}^{50}_{0,10^{-5}} - \hat{\gamma}^{50}_{\tau,\epsilon}\|_2}{\|\hat{\gamma}^{50}_{0,10^{-5}}\|_2}, \tag{36}$$

which provides a measure of the effects of operator and model truncation on the reconstruction. For $\epsilon = 10^{-5}$ and $\epsilon = 10^{-2}$, we see negligible performance degradation; however, with $\epsilon = 0.1$ there is some loss in fidelity. Specifically, in Figure 8, we display $\hat{\mathbf{g}}^{50}_{0.5,10^{-5}}$, the inverse wavelet transform of $\hat{\gamma}^{50}_{0.5,10^{-5}}$. Comparing this image to that obtained with no truncation (Figure 5d), shows little difference in the two reconstructed conductivity profiles. Thus, even with this severe level of truncation, the conductivity structure remains well localized, and the final amplitude is still close to that of the true profile.

Table 4. Mean Length of ς_1 for the Second Example, Expressed as a Percent of the Maximum Length (256 in This Case)[a]

	ϵ		
τ	10^{-5}	0.01	0.1
0.0	100	100	100
0.1	38	41	54
0.3	21	22	34
0.5	18	18	25

[a]Each entry in this table is obtained as an average over the 50 iterations of the corresponding inversion.

We next examine the performance of our imaging algorithm for a more difficult problem in which the underlying conductivity distribution is illustrated in Figure 9. Here there are two structures to be resolved, each of which is smaller than that considered in the previous example and lies further from the two vertical edges where the sources and receivers are located. In Figure 10, we display the optimal x-oriented detail for this problem for $\tau = 0.5$. It should be noted that both conductivity structures are in the region for which it is anticipated that the least amount of detail can be recovered. As seen by the dashed line in Figure 4, the greater complexity of this problem is captured directly by an increase in the expected number of significant wavelet coefficients for all values of τ relative to the previous problem.

With $\epsilon = 10^{-5}$ and $\tau = 0$, the reconstructions obtained for this problem after 1, 20, 40, and 50 iterations of our inverse scattering algorithm are displayed in Figure 11, and the corresponding performance curves are shown in Figure 12. We observe that the algorithm is quite successful both in distinguishing the two structures from one another and in determining their amplitudes. The two reconstructed blocks are larger in area than the true structures; however, this is consistent with the lack of fine scale detail expected to be recoverable from the data in these regions of A. Moreover, we observe from Figure 11b that valuable geometric information is embedded in the reconstructed profile after only 20 iterations of the algorithm. In particular, the number (two) and

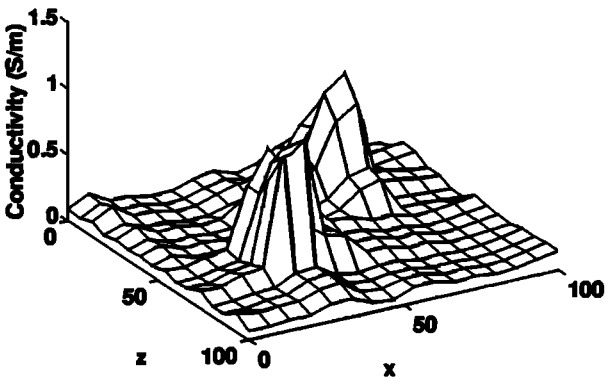

Figure 14. Final conductivity reconstruction for the second example with $\tau = 0.5$ and $\epsilon = 0.1$.

locations of both structures are clearly evident. In many applications, such information could be as important as the absolute amplitudes of the structures or the final image itself. Finally, from Figure 12, we see that after 50 iterations the algorithm has reached at least a local minimum of the cost function; however, the still-decreasing curve of the norm of the estimation error indicates that further iterations may be useful, particularly in refining the amplitudes of the two structures.

In Table 4 and Figure 13, the performance characteristics of the inversion method are displayed as a function of the truncation parameters τ and ϵ. The results here are similar to those seen in the first example. The average length of ς_1 is consistent with that predicted by the RCRB analysis in Figure 4. The effects of truncation on the quality of the reconstruction as measured by $\rho_\epsilon(\tau)$ defined in (36) are again small with the reconstructed profile obtained under the most severe truncation ($\tau = 0.5$ and $\epsilon = 0.1$) shown in Figure 14. Although slightly less smooth than Figure 11d, this estimate of the conductivity clearly retains all essential features of the untruncated version.

7. Conclusions

We have presented a multiscale, statistical approach to the nonlinear inverse scattering problem based on the extended Born approximation. We considered an inverse electrical conductivity problem arising in the field of geophysical prospecting. The application of statistical signal processing methods to the recovery of the conductivity's wavelet transform leads to the quantitative analysis of issues, such as reconstruction accuracy versus resolution and the development of bounds on our ability to localize conductivity anomalies in the region of interest. Finally, our approach allowed for the use of physically realistic, computationally efficient regularization methods which are described most naturally in scale-space.

We introduced the relative Cramér-Rao bound matrix as a generalization of the RECM employed by *Miller and Willsky* [1995a, 1995b]. The RCRB provided useful information regarding the level of detail expected to be recoverable from a given data set and played a central role in reducing the complexity of our inversion algorithm.

The reconstruction itself was specified as the solution to a maximum a posteriori estimation problem which was obtained using the the Levenberg-Marquardt method. At each iteration of the algorithm, the RCRB was used to identify those degrees of freedom in scale space for which the current linear problem provided substantial information relative to the prior model. The resulting block-partitioned form of the normal-type equations were directly and efficiently inverted, yielding not only the solution to the linear system but also the RCRB information required to compute the partition at the next step of the algorithm. The computational difficulties of solving the forward problem at each LM step were reduced significantly through the use of a scale-space form of the extended Born approximation. Specifically, the matrices defining the structure of the EBA model relating the conductivity to the observations could be made up to 90% sparse with little impact on inversion performance.

Acknowledgments. This work was supported in part by the Air Force Office of Scientific Research under grant AFOSR-F49620-95-1-0083, the Advanced Research Project Agency under Air Force grant F49620-93-1-0604, and the Office of Naval Research under grant N00014-91-J-1004. The work of the first author was also supported by a U.S. Air Force Laboratory Graduate Fellowship.

References

Alpert, B., G. Beylkin, R. Coifman, and V. Rokhlin, Wavelets for the fast solution of second-kind integral equations, *SIAM J. Sci. Comput.*, *14*(1), 159–184, 1993.

Bates, R., V. Smith, and R. Murch, Manageable multidimensional inverse scattering theory, *Phys. Rep.*, *201*(4), 185–277, 1991.

Bertero, M., C. D. Mol, and E. R. Pike, Linear inverse problems with discrete data, II, Stability and regularisation, *Inverse Probl.*, *4*, 573–594, 1988.

Beyer, W. H. (Ed.), *CRC Standard Math Tables*, 28th ed., CRC Press, Boca Raton, Fl., 1987.

Beylkin, G., R. Coifman, and V. Rokhlin, Fast wavelet transforms and numerical algorithms, I, *Commun. Pure Appl. Math.*, *44*, 141–183, 1991.

Daubechies, I., Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math.*, *41*, 909–996, 1988.

Flandrin, P., Wavelet analysis and synthesis of fractional Brownian motion, *IEEE Trans. Inf. Theory*, *38*, 910–917, 1992.

Gill, P. E., W. Murry, and M. H. Wright, *Practical Optimization*, Academic, San Diego, Calif., 1981.

Groetsch, C. W., *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, London, 1984.

Habashy, T. M., R. W. Groom, and B. R. Spies, Beyond the Born and Rytov approximations: A nonlinear approach to electromagnetic scattering, *J. Geophys. Res.*, *98*(B2), 1759–1775, 1993.

Harrington, R. F., *Field Computations by Moment Methods*, Macmillan, New York, 1968.

Kak, A. C., and M. Slaley, *Principles of Computerized Tomographic Imaging*, IEEE Press, Piscataway, N.J., 1987.

Kress, R., *Linear Integral Equations*, Springer-Verlag, New York, 1989.

Miller, E. L., The application of multiscale and statistical techniques to the solution of inverse problems, *Tech. Rep. LIDS-TH-2258*, MIT Lab. for Inf. and Decision Syst., Cambridge, Mass., 1994.

Miller, E. L., and A. S. Willsky, A multiscale approach to sensor fusion and the solution of linear inverse problems, *Appl. Comput. Harmonic Anal.*, *2*, 127–147, 1995a.

Miller, E. L., and A. S. Willsky, Multiscale, statistically-based inversion scheme for the linearized inverse scattering problem, *IEEE Trans. Geosc. Remote Sens.*, in press, 1995b.

Tarantola, A., and B. Valette, Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.*, *20*(2), 219–232, 1982.

Tewfick, A. H., and M. Kim, Correlation structure of the discrete wavelet coefficients of fractional Brownian motion, *IEEE Trans. Inf. Theory*, *38*, 904–909, 1992.

Torres-Verdín, C., and T. M. Habashy, Rapid 2.5-D forward modeling and inversion via a new nonlinear scattering approximation, *Radio Sci.*, *29*(4), 1051–1079, 1994.

Van Trees, H. L., *Detection, Estimation and Modulation Theory: Part I*, John Wiley, New York, 1968.

Wang, T., M. Oristaglio, A. Tripp, and G. Hohmann, Inversion of diffusive transient electromagnetic data by a conjugate-gradient method, *Radio Sci.*, *29*(4), 1143–1156, 1994.

Wornell, G. W., A Karhuenen-Loeve-like expansion for $1/f$ processes via wavelets, *IEEE Trans. Inf. Theory*, *36*, 859–861, 1990.

E. L. Miller, Center for Electromagnetics Research, Department of Electrical and Computer Engineering, 235 Forsyth Building, Northeastern University, 360 Huntington Ave., Boston, MA 002215. (e-mail: elmiller@cdsp.neu.edu)

A. S. Willsky, Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139. (e-mail: willsy@lids.mit.edu)

(Received March 30, 1995; revised August 10, 1995; accepted September 11, 1995.)