Total least-squares reconstruction with wavelets for optical tomography

Wenwu Zhu and Yao Wang

Department of Electrical Engineering, Polytechnic University, Brooklyn, New York 11201

Jun Zhang

Department of Electrical Engineering and Computer Science, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201

Received March 19, 1998; accepted May 11, 1998; revised manuscript received June 10, 1998

In a previous paper [Zhu *et al.*, J. Opt. Soc. Am. A **14**, 799 (1997)] an iterative algorithm for obtaining the total least-squares (TLS) solution of a linear system based on the Rayleigh quotient formulation was presented. Here we derive what to our knowledge are the first statistical properties of this solution. It is shown that the Rayleigh-quotient-form TLS (RQF-TLS) estimator is equivalent to the maximum-likelihood estimator when noise terms in both data and operator elements are independent and identically distributed Gaussian. A perturbation analysis of the RQF-TLS solution is derived, and from it the mean square error of the RQF-TLS solution is obtained in closed form, which is valid at small noise levels. We then present a wavelet-based multiresolution scheme for obtaining the TLS solution. This method was employed with a multigrid algorithm to solve the linear perturbation equation encountered in optical tomography. Results from numerical simulations show that this method requires substantially less computation than the previously reported one-grid TLS algorithm. The method also allows one to identify regions of interest quickly from a coarse-level reconstruction and restrict the reconstruction in the following fine resolutions to those regions. Finally, the method is less sensitive to noise than the one-grid TLS and multigrid least-squares algorithms.

[S0740-3232(98)00910-7]

OCIS codes: 100.0100, 100.2000, 100.3010, 100.3190, 100.6950, 100.7410.

1. INTRODUCTION

Recently there has been growing interest and rapid progress in medical optical imaging systems that emit near-infrared light into human tissues to determine their structures and functional status. The imaging problem is difficult, because in this frequency range, photons propagate through tissues in a highly diffused manner, and the relationship between the measured signal and the absorption properties of the medium is nonlinear. Over the past few years Wang's group has developed an iterative perturbation approach for continuous-wave, time-resolved, and frequency-domain data.[1–3] Independently, Singer *et al.*,[4] Arridge *et al.*,[5] and Arridge[6] described alternative iterative schemes derived from random-walk and diffusion theory. Mathematically, all these schemes require the solution of a linear perturbation equation at each iteration:

$$\mathbf{Hx} = \mathbf{y}, \tag{1}$$

where $\mathbf{x}$ is an $n \times 1$ vector of differences in the absorption properties between a reference and a test medium; $\mathbf{y}$ is an $m \times 1$ vector of detector-reading changes between the two media; and $\mathbf{H}$, a linear operator, is an $m \times n$ matrix of weights describing the influence of each volume element (voxel) on the detector readings, which are essentially the derivatives of the detector readings with respect to the absorption coefficients in the reference medium. In general, the perturbation equation is both underdetermined ($m < n$) and ill-posed. To solve it, several iterative algorithms have been developed, including projection onto convex sets (POCS),[2] conjugate gradient descent,[2] multiresolution regularized least-squares using wavelets,[7] (RLS), and total least squares (TLS).[8] These methods are developed for application to continuous-wave measurements. Layer-stripping[9] and regularized-layer-stripping[10] methods for TR data have also been proposed. The POCS method is known to have slow convergence. The conjugate gradient descent algorithm solves a least-squares (LS) problem. The regularization technique used in Refs. 10 and 7 is intended to reduce the sensitivity to noise embedded in the measurement data, whereas the TLS method in Ref. 8 deals with noise in both the data vector and the operator matrix.

In optical tomographic imaging systems the linear operator $\mathbf{H}$ is also subject to errors or noise, which may result from both the approximations used in deriving the linear model and numerical errors in computing the operator. In these cases the TLS solution is optimal and better than the LS in the sense that errors in both the operator and measurement data are minimized.[11] Regularized constrained TLS was proposed by Mesarović *et al.*[12] for image-restoration problems in which noise in $\mathbf{H}$ and $\mathbf{y}$ are linearly dependent (algebraically correlated). There $\mathbf{H}$, called the point-spread function, is a convolution operator. A TLS solution based on singular-value decomposition (SVD) was employed by Li *et al.*[13] in phase array

imaging where noise in **H** and **y** are assumed independent. However, the SVD-based TLS solution has computation difficulties for large-scale systems. This is because the conventional SVD computation needs $O(N^3)$ multiplication, where N is the length of vector **x**. The application of TLS to optical tomography was first considered by Zhu *et al.* in Ref. 8. Since computation is a major concern in optical tomography, a Rayleigh-quotient-form TLS (RQF-TLS) approach was developed, which requires $O(N^2)$ multiplication.[8] In this paper the equivalence of the RQF-TLS estimator to the maximum-likelihood (ML) estimator is established for the case in which the noise in both data and operator are independent and identically distributed (i.i.d) Gaussian. A perturbation analysis of the RQF-TLS is also performed to derive the mean square error (MSE) of the TLS estimate.

A challenge in solving the perturbation equation of Eq. (1) is that the computation is usually extremely expensive, owing to the extremely large dimension of the weight matrix. To alleviate this problem, a wavelet-based multiresolution reconstruction algorithm was recently developed.[7] This algorithm is based on the LS and RLS formulation, which was developed to overcome noise in the measurement data. Such strategies (RLS-based recovery algorithms) for overcoming measurement noise by means of wavelet transform have also been reported in other imaging and image processing applications, e.g., in Refs. 14–17. In Ref. 14 Wang *et al.* proposed a RLS-based multigrid algorithm for image restoration through wavelet expansion. In Ref. 15 Blanc-Feraud *et al.* compared one-dimensional (1D) and two-dimensional (2D) wavelet transforms for a RLS solution. In Ref. 16 Delaney *et al.* proposed a multiresolution tomographic reconstruction algorithm (filtered backprojection) for emission tomography. In Ref. 17 Banham *et al.* proposed a RLS-based multichannel subband approach for image restoration. However, as stated above, in optical tomography the imaging operator usually contains errors or noise and the TLS solution is needed. In this paper we propose a new wavelet-based multiresolution total least-squares algorithm. In this algorithm, the unknown, the data, and the weight matrix are expanded by means of wavelets, thus yielding a multiresolution representation of the Rayleigh quotient function. The transformed Rayleigh quotient function is minimized from coarse to fine resolutions. At each resolution, it uses the solution obtained from the previous resolution as the initial solution. Compared with that of the previous single-grid technique, which solves the Rayleigh quotient equation in the finest resolution directly,[8] the computational complexity of the transformed Rayleigh quotient function is reduced significantly.

The rest of this paper is organized as follows. In Section 2, after reviewing the RQF-TLS formulation, we show its equivalence to the ML estimator in the Gaussian setting and derive the MSE by a perturbation analysis. In Section 3 the wavelet-based RQF-TLS algorithm is presented and its equivalence to the original RQF-TLS is shown. Section 4 presents the multigrid conjugate gradient (CG) algorithm that is used to obtain the RQF-TLS in the wavelet domain. Section 5 presents experimental results. Finally, Section 6 provides some discussion and conclusions.

2. RAYLEIGH QUOTIENT FORM OF THE TOTAL LEAST-SQUARES SOLUTION AND ITS PROPERTIES

A. Total Least-Squares Solution of a Linear System

Consider an imaging system described by Eq. (1), where **x** represents the unknown values to be imaged, **y** represents the measurement data, and **H** represents the linear operator that models the relation between **x** and **y**. In the imaging problem, the operator is usually obtained by forward studies.[6] In the conventional LS formulation, it is assumed that the measurement **y** is noisy but the operator **H** is correct. The LS solution of **x** is determined by the following minimization:

$$\text{minimize } \|\Delta\mathbf{y}\|_2, \tag{2}$$

$$\text{subject to } \mathbf{y} + \Delta\mathbf{y} \in \text{range}(\mathbf{H}). \tag{3}$$

Here $\|\cdot\|_2$ denotes the Euclidean norm and $\Delta\mathbf{y}$ is the perturbation of **y**. When the smallest $\Delta\mathbf{y}$ satisfying Eq. (3) is found, any **x** satisfying

$$\mathbf{H}\mathbf{x} = \mathbf{y} + \Delta\mathbf{y} \tag{4}$$

is said to solve the LS problem.

It is well known that the solution of the LS problem amounts to solving the normal equation

$$\mathbf{H}^T\mathbf{H}\mathbf{x} = \mathbf{H}^T\mathbf{y}, \tag{5}$$

where T denotes transpose. The unique minimal two-norm solution to the LS problem is given by

$$\mathbf{x}_{\text{LS}} = \mathbf{H}^{\dagger}\mathbf{y}, \tag{6}$$

where $\mathbf{H}^{\dagger}$ denotes the Moore–Penrose pseudoinverse of **H**. In an overdetermined case, $\mathbf{H}^{\dagger} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$. For an underdetermined case, $\mathbf{H}^{\dagger} = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}$. In the above LS solution the assumption is that all errors or noise are confined to the observation and that the matrix **H** is correct. However, in reality there may be errors in the weight matrix **H** as well.

To alleviate this problem, the TLS was introduced by Golub and Van Loan.[11] The TLS approach attempts to minimize the errors in both **y** and **H**; i.e.,

$$\text{minimize } \|[\Delta\mathbf{H}|\Delta\mathbf{y}]\|_F, \tag{7}$$

$$\text{subject to } \mathbf{y} + \Delta\mathbf{y} \in \text{range}(\mathbf{H} + \Delta\mathbf{H}). \tag{8}$$

Here $\|\cdot\|_F$ denotes the Frobenius norm, which is defined, for an arbitrary $m \times n$ matrix B, as

$$\|B\|_F^2 = \text{tr}(B^TB) = \sum_{i=1}^{m}\sum_{j=1}^{n}(b_{ij})^2. \tag{9}$$

When the smallest perturbations in both **y** and **H** are found, then any **x** satisfying

$$(\mathbf{H} + \Delta\mathbf{H})\mathbf{x} = \mathbf{y} + \Delta\mathbf{y} \tag{10}$$

is said to solve the TLS problem.

The TLS problem in Eqs. (7) and (8) can be restated as

$$\text{minimize } \|\Delta\mathbf{A}\|_F, \tag{11}$$

$$\text{subject to } (\mathbf{A} + \Delta\mathbf{A})\mathbf{q} = \mathbf{0}, \tag{12}$$

where

$$\mathbf{A} = [\mathbf{H}|\mathbf{y}], \qquad \Delta\mathbf{A} = [\Delta\mathbf{H}|\Delta\mathbf{y}], \tag{13}$$

$$\mathbf{q} = \begin{pmatrix} \mathbf{x} \\ -1 \end{pmatrix}.$$

The closed-form TLS solution is given by $\mathbf{x}_{\mathrm{TLS}} = (\mathbf{H}^T\mathbf{H} - \sigma_{n+1}^2\mathbf{I})^{-1}\mathbf{H}^T\mathbf{y}$, where σ_{n+1}^2 is the smallest nonzero singular value of the matrix $\mathbf{A}$.[11] Comparing the LS and TLS solutions, the only difference is the term σ_{n+1}^2, which helps to reduce the effect of the error in $\mathbf{H}$. Golub and Van Loan[11] showed that, when $\mathbf{A}$ has full column rank, the TLS solution is related to the right singular vector $\mathbf{v}_{n+1}$ of $\mathbf{A}$ associated with the smallest singular value σ_{n+1}, under the condition that the singular value $\sigma_n > \sigma_{n+1}$ and $v_{n+1,n+1} \neq 0$. Here σ_n represents the nth largest singular value and $v_{n+1,n+1}$ is the last component of the singular vector $\mathbf{v}_{n+1}$. Let $\mathbf{v}_{n+1}^T = [\tilde{\mathbf{v}}_{n+1}^T v_{n+1,n+1}]$, where T represents transpose. The TLS solution for a full-rank system is

$$\mathbf{x} = -\frac{1}{v_{n+1,n+1}}\tilde{\mathbf{v}}_{n+1}. \tag{14}$$

For a rank-deficient and hence underdetermined problem, the minimum norm solution is given by

$$\mathbf{x} = -\frac{\sum_{i=p+1}^{n+1} v_{i,n+1}\tilde{\mathbf{v}}_i}{\sum_{i=p+1}^{n+1} v_{i,n+1}^2}, \tag{15}$$

where $\sigma_p > \sigma_{p+1} = \ldots = \sigma_{n+1} = 0$ and $\mathbf{v}_i = [\tilde{\mathbf{v}}_i^T v_{i,n+1}]^T$ are the singular value and the singular vector of $\mathbf{A}$, respectively. Hence one way to obtain the TLS solution is through performing the SVD of $\mathbf{A}$.[11]

For large-scale systems the computation of SVD is difficult.[13,18] Therefore we have opted for iterative schemes that are more suitable for large-scale systems. It has been shown that the constrained minimization problem in Eqs. (11) and (12) is equivalent to the following minimization problem[19]:

$$\text{minimize}\left[F(\mathbf{q}) = \frac{\mathbf{q}^T\mathbf{A}^T\mathbf{A}\mathbf{q}}{\mathbf{q}^T\mathbf{q}} = \frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2}{\|\mathbf{x}\|^2 + 1}\right], \tag{16}$$

where $F(\mathbf{q})$ is called the Rayleigh quotient. The solution of Eq. (16) is equal to the eigenvector $\mathbf{q}$ associated with the smallest eigenvalue of $\mathbf{A}^T\mathbf{A}$.[19] We use the CG method[20,21] to iteratively solve the above optimization problem for $\mathbf{q}$. Recall that when $\mathbf{q}^T = [\mathbf{x}^T, -1]$, one can obtain the TLS solution by $\mathbf{x} = -(\mathbf{q}_n)/q_{n+1}$, where $\mathbf{q}_n$ is the vector containing the first n component of $\mathbf{q}$ and q_{n+1} is the $(n+1)$th component of $\mathbf{q}$. Note that when the problem is underdetermined, there is no guarantee that the minimization of the RQF by the CG method will yield the minimum norm solution. One way to guide the algorithm toward the minimum norm solution is to start with a zero initial solution in which all elements except the last are zero and the last element is set to 1. Then we use the convergence criterion in Ref. 8, which is essentially the difference between two successive eigenvalues.

B. Equivalence between the Rayleigh-Quotient-Form–Total Least-Squares Solution and the Maximum-Likelihood Estimate

It has been shown in the multivariate linear errors-in-variables (EIV) model that the TLS estimator is a ML estimator when the noise is additive and Gaussian.[22] In Ref. 23, by use of the regression model, it was also shown that the constrained TLS estimator is a ML estimator from a constrained state space. Here we show (also by use of the regression model) that the TLS solution that uses the formulation given in Eq. (16) (i.e., the RQF-TLS solution) is a ML estimator.

Let $\Delta\mathbf{H}$ and $\Delta\mathbf{y}$ represent the noise in the operator and data, respectively; then the imaging model is given by

$$\mathbf{y} = (\mathbf{H} + \Delta\mathbf{H})\mathbf{x} + \Delta\mathbf{y}. \tag{17}$$

Assume that the elements of $\Delta\mathbf{H}$ and $\Delta\mathbf{y}$ are i.i.d. Gaussian variables with zero mean and variance σ^2. Then the conditional mean of $\mathbf{y}$ given $\mathbf{x}$ is $\eta_{\mathbf{y}/\mathbf{x}} = \mathbf{H}\mathbf{x}$ and the covariance matrix is

$$\begin{aligned}\mathbf{C}_{\mathbf{y}/\mathbf{x}} &= E[(\mathbf{y} - \eta_{\mathbf{y}/\mathbf{x}})(\mathbf{y} - \eta_{\mathbf{y}/\mathbf{x}})^T] \\ &= E[(\Delta\mathbf{H}\mathbf{x} + \Delta\mathbf{y})(\Delta\mathbf{H}\mathbf{x} + \Delta\mathbf{y})^T] \\ &= E[\Delta\mathbf{H}\mathbf{x}\mathbf{x}^T\Delta\mathbf{H}^T] + \sigma^2\mathbf{I},\end{aligned} \tag{18}$$

where $E\{\cdot\}$ is the expectation operator and $\mathbf{I}$ is the $m \times m$ identity matrix. From the assumption that the elements of $\Delta\mathbf{H}$ are i.i.d. with zero mean and variance σ^2, we can obtain

$$E[\Delta\mathbf{H}\mathbf{x}\mathbf{x}^T\Delta\mathbf{H}^T] = \sigma^2\|\mathbf{x}\|^2\mathbf{I}. \tag{19}$$

Thus

$$\mathbf{C}_{\mathbf{y}/\mathbf{x}} = \sigma^2(\|\mathbf{x}\|^2 + 1)\mathbf{I}. \tag{20}$$

Because $\Delta\mathbf{H}$ and $\Delta\mathbf{y}$ are both Gaussian, $\mathbf{y}$ in Eq. (17) for a given $\mathbf{x}$ is also Gaussian. Therefore the conditional probability $P(\mathbf{y}/\mathbf{x})$ is

$$P(\mathbf{y}/\mathbf{x}) = [(2\pi)^m|\mathbf{C}_{\mathbf{y}/\mathbf{x}}|]^{-1/2}\exp[-\tfrac{1}{2}(\mathbf{y} - \mathbf{H}\mathbf{x})^T \times \mathbf{C}_{\mathbf{y}/\mathbf{x}}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x})], \tag{21}$$

where $|\mathbf{C}_{\mathbf{y}/\mathbf{x}}|$ is the determinant of $\mathbf{C}_{\mathbf{y}/\mathbf{x}}$. Taking the logarithm of Eq. (21) yields the following likelihood function:

$$\mathscr{L}(\mathbf{y}/\mathbf{x}) = -(\mathbf{y} - \mathbf{H}\mathbf{x})^T\mathbf{C}_{\mathbf{y}/\mathbf{x}}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) - \log[(2\pi)^m|\mathbf{C}_{\mathbf{y}/\mathbf{x}}|]. \tag{22}$$

The ML solution is obtained by maximizing the likelihood function $\mathscr{L}(\mathbf{y}/\mathbf{x})$ with respect to $\mathbf{x}$. It was tested and verified in Ref. 24 that the log term in Eq. (22) is negligible. Thus maximizing $\mathscr{L}(\mathbf{y}/\mathbf{x})$ can be accomplished by minimizing

$$\mathscr{L}(\mathbf{y}/\mathbf{x}) \approx (\mathbf{y} - \mathbf{H}\mathbf{x})^T\mathbf{C}_{\mathbf{y}/\mathbf{x}}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}). \tag{23}$$

When Eq. (20) is substituted into Eq. (23), the ML solution is obtained by minimizing

$$\frac{1}{\sigma^2(\|\mathbf{x}\|^2 + 1)}(\mathbf{y} - \mathbf{H}\mathbf{x})^T(\mathbf{y} - \mathbf{H}\mathbf{x}). \tag{24}$$

One can see immediately that the minimization of Eq. (24) is equivalent to the minimization of the Rayleigh quotient in Eq. (16). Thus we have shown that the RQF-

TLS solution is ML estimator. From this we know that the RQF-TLS solution has the desired statistical properties that have been shown for the SVD-based TLS solution, including asymptotic efficiency, among others.

C. Perturbation Analysis of the Rayleigh-Quotient-Form–Total Least-Squares Estimator

In the past, statistical properties of the TLS solution have been analyzed from the EIV model,[22] or from the formulation described in Refs. 12 and 23 for the constrained TLS from the regression model. Specifically, with the EIV model it has been shown that the TLS estimator is unbiased and that statistical properties of the TLS estimate have been derived. In Refs. 12 and 23 it was shown that the constrained TLS estimator is also unbiased and that other statistical properties are obtained from the constrained TLS formulation. In this paper we analyze the statistical properties of the RQF-TLS estimator. Because the RQF-TLS solution is nonlinear, error analysis is difficult. As in Refs. 12 and 23, we will perform a perturbation analysis to derive an analytic closed-form formula for the covariance matrix of the RQF-TLS solution for a small noise level.

Let $\mathbf{H}_o$ and $\mathbf{y}_o$ represent the noise-free system matrix (operator) and data vector, and let $\mathbf{x}_o$ represent the correct solution; then these must satisfy

$$\mathbf{A}_o\mathbf{q}_o = \mathbf{0}, \tag{25}$$

where

$$\mathbf{A}_o = [\mathbf{H}_o|\mathbf{y}_o], \qquad \mathbf{q}_o = \begin{pmatrix} \mathbf{x}_o \\ -1 \end{pmatrix}. \tag{26}$$

Now we perturb $\mathbf{A}_o$ by adding a small noise term $\Delta\mathbf{A} = [\Delta\mathbf{H}|\Delta\mathbf{y}]$; this will lead to a perturbation of the solution by $\Delta\mathbf{q}$. The necessary condition for minimizing Eq. (16) is that the gradient of $F(\mathbf{q})$ with respect to $\mathbf{q}$ be zero. That is,

$$\frac{\partial F(\mathbf{q})}{\partial \mathbf{q}} = \frac{2}{(\mathbf{q}^T\mathbf{q})^2}[\mathbf{q}^T\mathbf{q}\mathbf{A}^T\mathbf{A} - \mathbf{q}^T\mathbf{A}^T\mathbf{A}\mathbf{q}\mathbf{I}]\mathbf{q} = \mathbf{0}. \tag{27}$$

Substituting $\mathbf{A} = \mathbf{A}_o + \Delta\mathbf{A}$ and $\mathbf{q} = \mathbf{q}_o + \Delta\mathbf{q}$ into Eq. (27) and ignoring high-order terms of $\Delta\mathbf{q}$ and $\Delta\mathbf{A}$, we obtain

$$\mathbf{q}_o^T\mathbf{q}_o(\mathbf{A}_o^T\mathbf{A}_o + \mathbf{A}_o^T\Delta\mathbf{A} + \Delta\mathbf{A}^T\mathbf{A}_o)\mathbf{q}_o + \mathbf{q}_o^T\mathbf{q}_o\mathbf{A}_o^T\mathbf{A}_o\Delta\mathbf{q} + 2\mathbf{q}_o^T\Delta\mathbf{q}\mathbf{A}_o^T\mathbf{A}_o\mathbf{q}_o = \|\mathbf{A}_o\mathbf{q}_o\|^2(\mathbf{q}_o + \Delta\mathbf{q}). \tag{28}$$

Using the condition in Eq. (25) and after some algebraic manipulations, we obtain

$$\Delta\mathbf{q} = -(\mathbf{A}_o^T\mathbf{A}_o)^\dagger\mathbf{A}_o^T\Delta\mathbf{A}\mathbf{q}_o, \tag{29}$$

which is the closed-form perturbation solution. If we assume $\Delta\mathbf{A}$ has zero mean, then $E[\Delta\mathbf{q}] = \mathbf{0}$; i.e., the RQF-TLS estimator is unbiased.

Next, we derive the MSE of the RQF-TLS estimate. By definition,

$$\mathrm{MSE}_{\mathrm{TLS}} = E[\Delta\mathbf{q}^T\Delta\mathbf{q}] = \mathrm{trace}(\mathbf{R}), \tag{30}$$

where $\mathbf{R} = E[\Delta\mathbf{q}\Delta\mathbf{q}^T]$ is the correlation matrix of $\Delta\mathbf{q}$, which is equivalent to the covariance matrix of $\Delta\mathbf{q}$, since $\Delta\mathbf{q}$ has zero mean. Using the solution in Eq. (29), we have

$$\mathbf{R} = (\mathbf{A}_o^T\mathbf{A}_o)^\dagger\mathbf{A}_o^T E[\Delta\mathbf{A}\mathbf{q}_o\mathbf{q}_o^T\Delta\mathbf{A}^T]\mathbf{A}_o(\mathbf{A}_o^T\mathbf{A}_o)^{\dagger T}. \tag{31}$$

As above, if we assume that the elements of $\Delta\mathbf{A}$ are i.i.d. with zero mean and variance σ^2, then

$$E[\Delta\mathbf{A}\mathbf{q}_o\mathbf{q}_o^T\Delta\mathbf{A}^T] = \sigma^2\|\mathbf{q}_o\|^2\mathbf{I}. \tag{32}$$

Thus

$$\mathbf{R} = \sigma^2\|\mathbf{q}_o\|^2(\mathbf{A}_o^T\mathbf{A}_o)^\dagger. \tag{33}$$

Substituting Eq. (33) into Eq. (30) yields

$$\mathrm{MSE}_{\mathrm{TLS}} = \mathrm{trace}(\mathbf{R}) = \sigma^2\|\mathbf{q}_o\|^2\mathrm{trace}[(\mathbf{A}_o^T\mathbf{A}_o)^\dagger]. \tag{34}$$

Assume that the rank of $(\mathbf{A}_o^T\mathbf{A}_o)$ is r, and let $\lambda_1, \lambda_2, \ldots, \lambda_r$ represent the r positive eigenvalues of $(\mathbf{A}_o^T\mathbf{A}_o)$. Then Eq. (34) can be rewritten as

$$\mathrm{MSE}_{\mathrm{TLS}} = \sigma^2\|\mathbf{q}_o\|^2\sum_{i=1}^{r}\lambda_i^{-1} = \sigma^2(1 + \|\mathbf{x}_o\|^2)\sum_{i=1}^{r}\lambda_i^{-1}. \tag{35}$$

Equation (35) is useful for predicting the accuracy of the RQF-TLS estimator at small noise levels.

3. WAVELET-BASED TOTAL LEAST-SQUARES SOLUTION

A. Two-Dimensional Multiresolution Representation by a Wavelet Transform

A wavelet is a function whose dilations and shifts form an orthonormal basis for the space of all energy-finite signals.[25] For a given (energy-finite) signal, the wavelet basis provides a multiresolution representation of the signal. For a 1D discrete signal this involves taking the discrete wavelet transform (DWT) of the signal. For example, for a 1D signal that has length N and is denoted by $\mathbf{x}$, a two-level DWT is an N-dimensional vector $\tilde{\mathbf{x}}$,

$$\tilde{\mathbf{x}}_N = \mathbf{W}\mathbf{x} = [A_{-1}\mathbf{x}_{(N/2)}\, D_{-1}\mathbf{x}_{(N/2)}]^T, \tag{36}$$

where $A_{-1}\mathbf{x}$ and $D_{-1}\mathbf{x}$ are the lowpass and bandpass components, respectively, of $\mathbf{x}$, both of dimension $N/2$. In the above equation the subscript denotes the dimension of the underlying vector.

Since we are dealing with images, which are 2D, in this paper we use the 2D DWT. In general, a 2D image can be arranged into a vector by lexicographic ordering. The 1D wavelet transform described above can then be applied to the vector. One can then reorder the vector as a 2D image. However, a more efficient way of implementing a wavelet transform of a 2D image is by first applying a 1D transform to each row of the 2D image and then applying a 1D wavelet transform to each column of the resulting image from the previous step. This is known as separable wavelet transform.

Let $\mathbf{F}$ represent a 2D image with size $M \times N$, and let $\mathbf{f}$ be the vector consisting of elements of $\mathbf{F}$ arranged in lexicographic order, with size $K = M \times N$. Let $\mathbf{W}_M$ and $\mathbf{W}_N$ represent the 1D wavelet transform matrix of size $M \times M$ and $N \times N$, respectively; the separable wavelet transform of $\mathbf{F}$ can be described by

$$\tilde{\mathbf{F}} = \mathbf{W}_M\mathbf{F}\mathbf{W}_N^T. \tag{37}$$

This is equivalent, in the 1D notation, to

$$\tilde{\mathbf{f}} = (\mathbf{W}_M \otimes \mathbf{W}_N)\mathbf{f}, \tag{38}$$

where $\otimes$ denotes the Kronecker product and $\tilde{\mathbf{f}}$ is the vector consisting of elements of $\tilde{\mathbf{F}}$ arranged in lexicographic order. Therefore performing the 2D wavelet transform on a $M \times N$ 2D image by means of a separable transform is equivalent to performing a 1D wavelet transform on the corresponding 1D vector by means of $\mathbf{W}_K = \mathbf{W}_M \otimes \mathbf{W}_N$. The transformed image can be grouped into four subimages consisting of an approximation image $A_{-1}\mathbf{F}$ and three detail subimages $D^1_{-1}\mathbf{F}$, $D^2_{-1}\mathbf{F}$, and $D^3_{-1}\mathbf{F}$; i.e.,

$$\tilde{\mathbf{F}}_{(M\times N)} = \begin{bmatrix} A_{-1}\mathbf{F}_{(M/2\times N/2)} & D^1_{-1}\mathbf{F}_{(M/2\times N/2)} \\ D^2_{-1}\mathbf{F}_{(M/2\times N/2)} & D^3_{-1}\mathbf{F}_{(M/2\times N/2)} \end{bmatrix}.$$

This is a two-level representation of $\mathbf{F}$ and can be implemented by a set of QMF's as illustrated in Fig. 1 (for two levels). Usually we refer to these QMF's as LL, LH, HL, and HH subimages, from top to bottom, respectively. One can decompose the subimage $A_{-1}\mathbf{F}$ further to the next scale (coarser resolution), which leads to a three-level representation:

$$\tilde{\mathbf{F}}_{(M\times N)} = \begin{bmatrix} A_{-2}\mathbf{F}_{(M/4\times N/4)} & D^1_{-2}\mathbf{F}_{(M/4\times N/4)} & D^1_{-1}\mathbf{F}_{(M/2\times N/2)} \\ D^2_{-2}\mathbf{F}_{(M/4\times N/4)} & D^3_{-2}\mathbf{F}_{(M/4\times N/4)} & \\ D^2_{-1}\mathbf{F}_{(M/2\times N/2)} & & D^3_{-1}\mathbf{F}_{(M/2\times N/2)} \end{bmatrix}.$$

The above procedure can be applied to the approximation subimage repeatedly to obtain an $L \leq \min\{\log_2 M, \log_2 N\}$-level representation.

B. Total Least-Squares Solution with One-Dimensional Wavelet Representation

Considering the perturbation equation in Eq. (1), to represent the perturbation equation in a wavelet domain let the transform matrices for $\mathbf{x}$ and $\mathbf{y}$ be represented by $\mathbf{W}_\mathbf{x}$ and $\mathbf{W}_\mathbf{y}$, respectively. We consider only orthonormal wavelets that satisfy $\mathbf{W}_\mathbf{x}^T\mathbf{W}_\mathbf{x} = \mathbf{W}_\mathbf{x}\mathbf{W}_\mathbf{x}^T = \mathbf{I}$ and $\mathbf{W}_\mathbf{y}^T\mathbf{W}_\mathbf{y} = \mathbf{W}_\mathbf{y}\mathbf{W}_\mathbf{y}^T = \mathbf{I}$. Multiplying Eq. (1) from the left by $\mathbf{W}_\mathbf{y}$ and inserting $\mathbf{W}_\mathbf{x}^T\mathbf{W}_\mathbf{x} = \mathbf{I}$ in between $\mathbf{H}$ and $\mathbf{x}$, we obtain

$$\tilde{\mathbf{H}}\tilde{\mathbf{x}} = \tilde{\mathbf{y}}, \tag{39}$$

where $\tilde{\mathbf{H}} = \mathbf{W}_\mathbf{y}\mathbf{H}\mathbf{W}_\mathbf{x}^T$, $\tilde{\mathbf{y}} = \mathbf{W}_\mathbf{y}\mathbf{y}$ and $\tilde{\mathbf{x}} = \mathbf{W}_\mathbf{x}\mathbf{x}$. Equation (39) is the perturbation equation in the wavelet domain. From this equation one can formulate the TLS in the WT domain. Let

$$\tilde{\mathbf{A}} = [\tilde{\mathbf{H}}|\tilde{\mathbf{y}}],$$

$$\tilde{\mathbf{q}} = \begin{pmatrix} \tilde{\mathbf{x}} \\ -1 \end{pmatrix}. \tag{40}$$

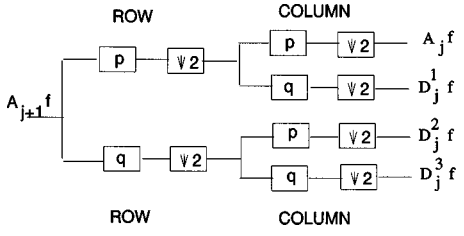

Fig. 1. Wavelet decomposition of 2D image from level $j+1$ to j.

From the orthonormality of the wavelet transform it can easily be shown that minimizing the Rayleigh quotient function defined in Eq. (16) is equivalent to the following problem:

$$\text{minimize}\left[F(\tilde{\mathbf{q}}) = \frac{\tilde{\mathbf{q}}^T\tilde{\mathbf{A}}^T\tilde{\mathbf{A}}\tilde{\mathbf{q}}}{\tilde{\mathbf{q}}^T\tilde{\mathbf{q}}}\right]. \tag{41}$$

That is, the TLS solution to the original Eq. (1) is equivalent to the TLS solution to the transform equation (39). Once we get the solution $\tilde{\mathbf{q}}$, and consequently $\tilde{\mathbf{x}}$, we can take the inverse wavelet transform to obtain the desired solution $\mathbf{x} = \mathbf{W}_\mathbf{x}^T\tilde{\mathbf{x}}$.

C. Transform of the Perturbation Equation by Use of Separable Operation

Suppose that $\mathbf{x}$ and $\mathbf{y}$ in Eq. (1) are vectors obtained from 2D $N \times N$ image $\mathbf{X}$ and $M \times M$ image $\mathbf{Y}$, respectively. Here, for notational simplicity, we consider only the case in which both $\mathbf{X}$ and $\mathbf{Y}$ are square images. Let $\mathbf{W}_\mathbf{X}$ and $\mathbf{W}_\mathbf{Y}$ represent 1D transform matrices for $\mathbf{X}$ and $\mathbf{Y}$, respectively. Then $\mathbf{W}_\mathbf{y} = \mathbf{W}_\mathbf{Y} \otimes \mathbf{W}_\mathbf{Y}$ and $\mathbf{W}_\mathbf{x} = \mathbf{W}_\mathbf{X} \otimes \mathbf{W}_\mathbf{X}$. The transforms of $\mathbf{X}$ and $\mathbf{Y}$ can be performed easily with separable processing. The question is how to calculate the wavelet transform of the operator $\tilde{\mathbf{H}} = \mathbf{W}_\mathbf{y}\mathbf{H}\mathbf{W}_\mathbf{x}^T$. Our method, slightly different from that in Ref. 15, is to do the following. First, we reorder each row $\mathbf{h}_{r_i}$, $i = 1, 2, \ldots, M^2$, of $\mathbf{H}$ as a 2D image of size $N \times N$, $\mathbf{H}_{r_i}$, and then use $\mathbf{W}_X$ to perform a separable transform. The result can be represented as $\tilde{\mathbf{H}}_{r_i} = \mathbf{W}_\mathbf{X}\mathbf{H}_{r_i}\mathbf{W}_\mathbf{X}^T$. This transformed image is then reordered back into a row vector $\tilde{\mathbf{h}}_{r_i}$. Let the resulting image be $\tilde{\mathbf{H}}_r = \mathbf{H}\mathbf{W}_\mathbf{X}^T$. Then we first order each column of $\tilde{\mathbf{H}}_r$, $\tilde{\mathbf{h}}_{c_j}$, $j = 1, 2, \ldots, N^2$, as a 2D image $\tilde{\mathbf{H}}_{c_j}$ and then perform a separable transform with $\mathbf{W}_\mathbf{Y}$; i.e., $\tilde{\tilde{\mathbf{H}}}_{c_j} = \mathbf{W}_\mathbf{Y}\tilde{\mathbf{H}}_{c_j}\mathbf{W}_\mathbf{Y}$, $j = 1, 2, \ldots, N^2$. This transformed image is then ordered back into a column vector, $\tilde{\tilde{\mathbf{h}}}_{c_j}$, which forms the jth column of the transformed operator $\tilde{\mathbf{H}}$.

4. MULTIGRID ALGORITHM

As shown in Subsection 3.B, the TLS solution can be obtained by directly minimizing the Rayleigh quotient in Eq. (16) or Eq. (41). When $\mathbf{x}$, $\mathbf{y}$, or $\mathbf{H}$ is slowly varying, its wavelet transform coefficients at high frequencies are zero or negligible. Then the minimization can be performed in the transform domain more efficiently by taking advantage of the sparsity of $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{A}}$. Even when the operator $\mathbf{H}$ is not smooth in the original domain, we can exploit the multiresolution property of the wavelet transform to reduce computation time.

In this paper we propose a multigrid scheme that uses a modified V-cycle algorithm. The principle of this algorithm is illustrated in Fig. 2. It differs from the classical V-cycle method[26] in two aspects. First, instead of downward (from fine to coarse) restriction, in which error is calculated and restricted, we use wavelet decomposition to reach the lower vertex (coarsest grid) of the V cycle. For the upward reconstruction (from coarse to fine), we use multigrid reconstruction instead of error prolongation

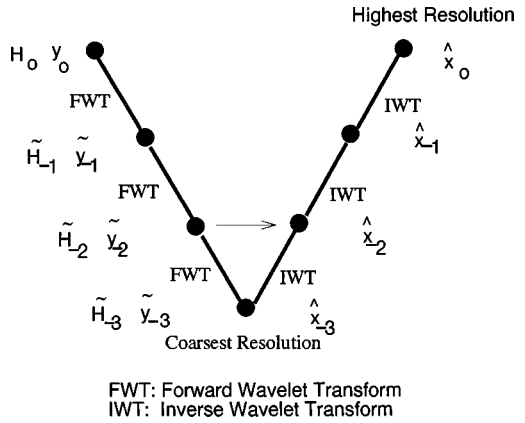

Fig. 2. Modified V-cycle multigrid algorithm.

(error compensation). Specifically, this modified V-cycle algorithm consists of the following steps:

1. Perform the wavelet transform of $\mathbf{y}$ and $\mathbf{H}$ to obtain $\tilde{\mathbf{H}}_l$ and $\tilde{\mathbf{y}}_l$, $l = -1, -2, \ldots, -L$. Set $l = -L$, the coarsest level.
2. Solve for the TLS solution $\tilde{\mathbf{q}}_l$ and hence $\tilde{\mathbf{x}}_l$ at the lth level, by minimizing $F(\mathbf{q}_l) = (\tilde{\mathbf{q}}_l^T \tilde{\mathbf{A}}_l^T \tilde{\mathbf{A}}_l \tilde{\mathbf{q}}_l)/(\tilde{\mathbf{q}}_l^T \tilde{\mathbf{q}}_l)$ using the CG algorithm.
3. Prolongate from $\tilde{\mathbf{x}}_l$ to $\tilde{\mathbf{x}}_{l+1}$ by padding zeros; i.e., $\tilde{\mathbf{x}}_{l+1} = [\tilde{\mathbf{x}}_l^T, \mathbf{0}]^T$.
4. Let $l = l + 1$. If $l = 0$, stop; otherwise, go back to Step 2, using the solution obtained in the previous step as the initial solution.

The multigrid algorithm described above differs slightly from the one proposed by Wang *et al.*[14] In the latter method, at each resolution level the LL, LH, HL, and HH components are reconstructed alternatingly until the solution converges. In our method, at each level only the LL component is solved, assuming that the other components are zeros.

In real clinic applications most of the time one is interested only in certain regions of the image with abnormal features. The above multigrid algorithm can not only save computational time compared with the single-grid algorithm but also allows one to zoom in on special regions of interest (ROI). Suppose we are interested in a region that appears questionable in a coarser-level reconstruction. Then only the corresponding region in the next-finer resolution needs to be further refined by the CG algorithm, while other regions in the finer resolution can be kept fixed. This will reduce the computation time for this grid significantly. This process can be repeated so that only a small number of unknowns are solved in each grid level.

In our study, we implemented a two-grid algorithm ($L = 1$) by which we first reconstruct the LL component of the wavelet transform $\tilde{\mathbf{x}}$ of $\mathbf{x}$. This yields the coarse-level reconstruction $\tilde{\mathbf{x}}_{-1}$. Starting from this solution, we then solve all the components of $\tilde{\mathbf{x}}$.

5. EXPERIMENTAL RESULTS

In our experiment the length-4 Daubechies' wavelet[27] was used because it achieves a good trade-off between computational precision and computational complexity. To demonstrate the effectiveness of the wavelet-based multigrid TLS algorithm compared with a single-grid algorithm, three experiments were performed. We used the MSE and correlation coefficient (CC) between the original image $\mathbf{x}$ and reconstructed image $\hat{\mathbf{x}}$ to evaluate the reconstruction accuracy quantitatively. The MSE and the CC are defined as

$$\mathrm{MSE} = \frac{1}{n}\|\mathbf{x} - \hat{\mathbf{x}}\|^2, \tag{42}$$

$$\mathrm{CC} = \frac{\langle \mathbf{x}, \hat{\mathbf{x}} \rangle}{\|\mathbf{x}\|\|\hat{\mathbf{x}}\|}, \tag{43}$$

respectively.

In the first experiment a 1-cm-diameter rod was placed in an infinite homogeneous medium. The rod was homogeneous, with absorption and scattering coefficients $\mu_a = 0.05\ \mathrm{cm}^{-1}$ and $\mu_s = 10\ \mathrm{cm}^{-1}$, respectively. The optical properties of the background medium were $\mu_a^b = 0.02\ \mathrm{cm}^{-1}$ and $\mu_s^b = 10\ \mathrm{cm}^{-1}$, respectively. The imaged region occupied an area of 10 cm $\times$ 10 cm, which was discretized to 32×32 pixels. A total of 16 sources and 16 detectors were evenly spread around the rod in a ring geometry with a diameter of 8 cm (see Fig. 3). Therefore, in our simulation, $M = 16$ and $N = 32$, or $m = 16^2$, $n = 32^2$. The solution to the forward problem (to get $\mathbf{H}$ and $\mathbf{y}$) was obtained by analytically solving the diffusion equation with the normal-mode series method described in Ref. 28. The data were corrupted by white Gaussian noise with a 30-dB signal-to-noise ratio (SNR). The weights were also corrupted by white Gaussian noise with the same noise variance as that in the detector reading. We performed this experiment five times for different noise realizations. The LS algorithm we used here was described in Ref. 2. The average MSE and CC obtained with different methods are given in Table 1. Figure 4 shows reconstructed images that use one-grid and two-grid algorithms, respectively, for noise realization. The coarse-level reconstruction solves only a quarter of the pixels in the fine resolution. In our computer implementation, one iteration of the coarse level roughly re-

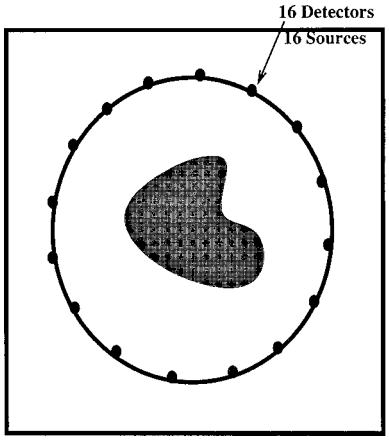

Fig. 3. The source–detector configurations of the cylindrical rod.

Table 1. MSE and CC for Different Methods for a Test Medium Containing a Rod Absorber at the Center

	One-Grid TLS with 85 Iterations	Two-Grid TLS with 1200 Iterations at the Coarse Grid	One-Grid TLS with 885 Iterations	Two-Grid TLS with 800 Iterations at the Fine Grid	Two-Grid TLS with ROI
Average CC	$2.27E \times 10^{-1}$	$4.95E \times 10^{-1}$	$4.95E \times 10^{-1}$	$5.22E \times 10^{-1}$	$5.29E \times 10^{-1}$
Average MSE	$1.80E \times 10^{-4}$	$1.51E \times 10^{-4}$	$1.50E \times 10^{-4}$	$1.50E \times 10^{-4}$	$1.46E \times 10^{-4}$

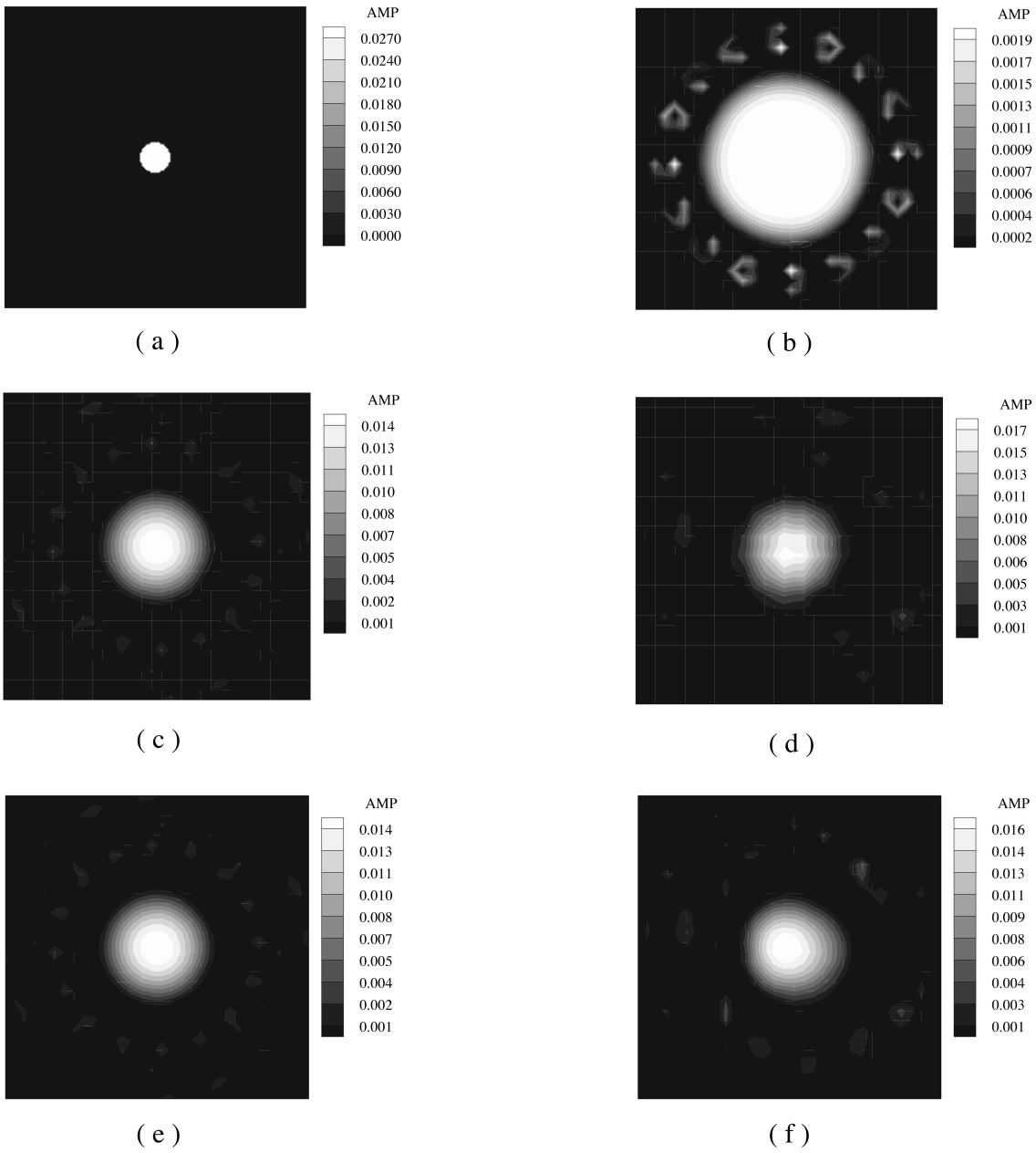

Fig. 4. Reconstruction results for a medium with a centered rod. The absorption coefficient distribution is a constant in the rod. The SNR in the data is 30 dB. The weights are corrupted by the same noise as that in the data. (a) Original image, (b) reconstructed image with one-grid TLS with 85 iterations, (c) reconstruction result with one grid with 885 iterations, (d) reconstruction image with two-grid TLS algorithm with 1200 iterations at the coarse grid only, (e) reconstruction image with two-grid algorithm with additional 800 iterations at the fine grid, (f) reconstructed image with 3200 iterations in a localized region at the fine grid with (d) as the initial solution. The total computation time for (b) and (d) are roughly the same, and the total computation time for (c), (e), and (f) are also similar. The time for (d) is approximately 1/10 of (c). Note that although (f) looks the same as (e), it is quantitatively more accurate.

quired 1/10 of the computational time taken by one iteration in the fine resolution. From Fig. 4 and Table 1 we see that the proposed multigrid algorithm provided significantly better reconstruction than the one-grid method under similar computation times [compare Figs. 4(b) and 4(d)] or required significantly shorter time to reach the

same resolution quality [compare Figs. 4(c) and 4(d)]. Figure 4(f) was obtained by restricting reconstruction in the fine grid within a ROI that was selected on the basis of the coarse-grid resolution. The ROI was a square block (a quarter of the size of the original image domain) in the center. By restricting the unknowns within this

Table 2. MSE and CC for Different Methods for a Test Medium Containing an Off-Centered Sinelike Absorber

	One-Grid TLS with 85 Iterations	Two-Grid TLS with 1200 Iterations at the Coarse Grid	One-Grid TLS with 885 Iterations	Two-Grid TLS with 600 Iterations at the Fine Grid	Two-Grid LS
Average CC	$1.85E \times 10^{-1}$	$4.43E \times 10^{-1}$	$4.65E \times 10^{-1}$	$4.63E \times 10^{-1}$	$4.60E \times 10^{-1}$
Average MSE	$1.77E \times 10^{-4}$	$1.66E \times 10^{-4}$	$1.65E \times 10^{-4}$	$1.65E \times 10^{-4}$	$1.80E \times 10^{-4}$

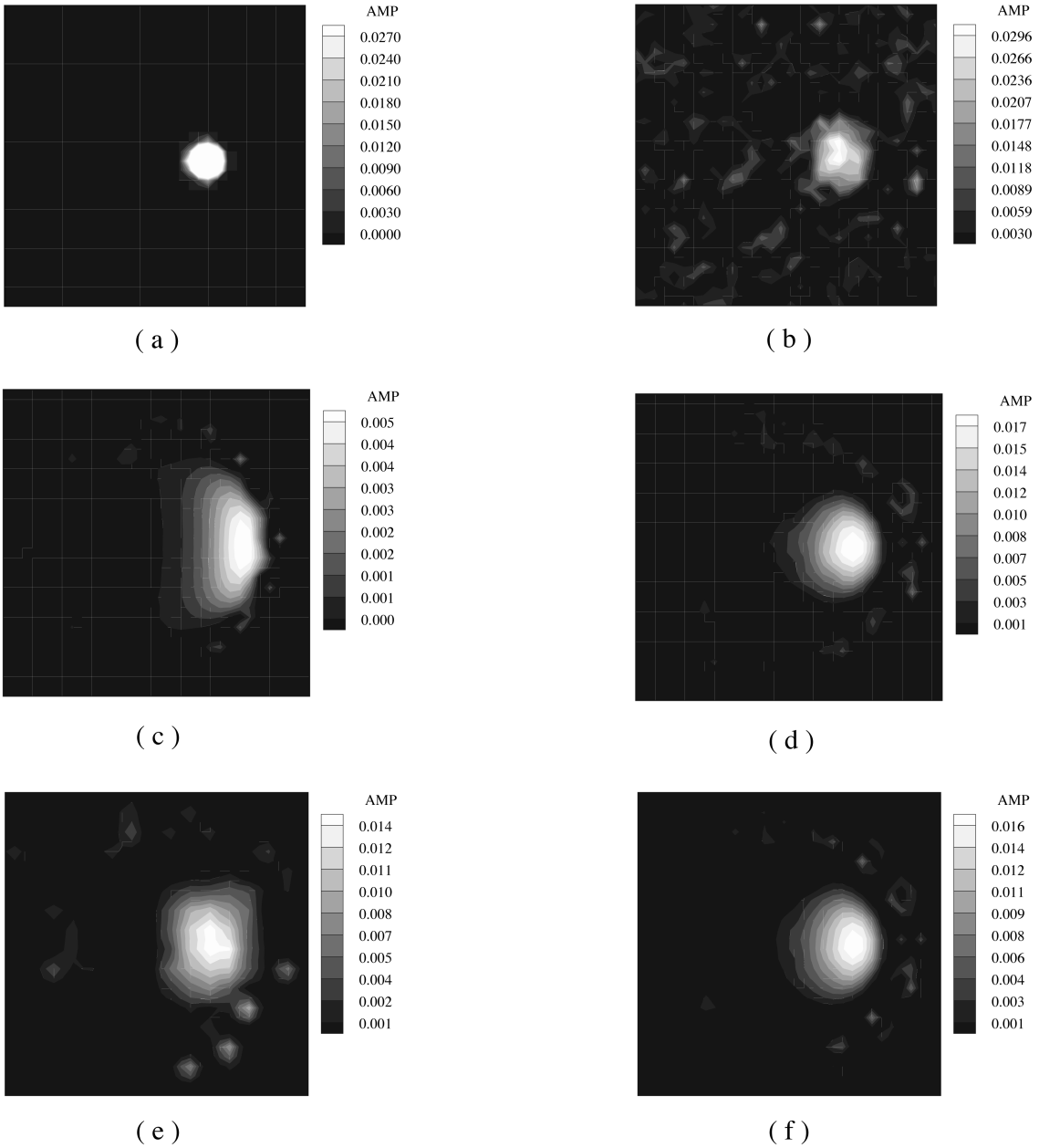

Fig. 5. Reconstruction results of a medium with an off-center sinelike absorber. The SNR in the data is 30 dB. The weights are corrupted by the same noise as that in the data. (a) Original image, (b) reconstruction result with multigrid LS with 1200 iterations at the coarse grid and 800 at the fine grid, (c) reconstructed image with one-grid TLS with 85 iterations, (d) reconstructed image with one-grid TLS with 885 iterations, (e) reconstructed image with two-grid TLS algorithm with 1200 iterations in the coarse grid only, (f) reconstructed image with the two-grid TLS algorithm with an additional 600 iterations at the fine grid with (e) as initial solution. The computation time for (c) and (e) are roughly same. The time for (f) is approximately 3/4 of (d) and that for (c) or (e) is approximately 1/10 of (d).

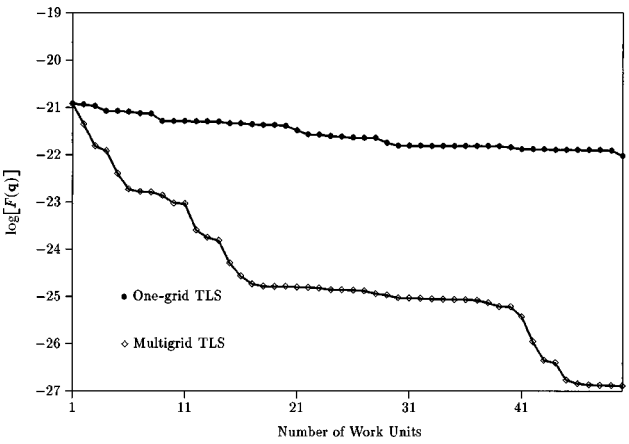

Fig. 6. Comparison of computation time required by multigrid TLS versus one-grid TLS in terms of work units for the test medium considered in Fig. 5.

ROI, we convert the original underdetermined problem into a determined one and therefore can achieve a more accurate result than that shown in Fig. 4(e).

In the second experiment a 1.5-cm-diameter rod was placed at an off-center position with respect to the source and the detector ring. The rod had a nonhomogeneous absorption distribution, following a sinusoidal pattern (one positive cycle only). We will refer to an absorber with such an absorption profile as a sinelike absorber. The forward solution in this experiment was obtained with a multigrid finite-difference solver described in Ref. 3. In this experiment white Gaussian noise was added to the weights and the detector readings in the same way as in the first experiment. We performed this experiment five times for different noise realizations; the means of the MSE's and CC's are given in Table 2. Note that $\|\mathbf{x}\|^2 = 1.26 \times 10^{-1}$ in the calculation. Figure 5 presents the reconstruction results for one noise realization. From Fig. 5 and Table 2 we can see that accurate results can be obtained from the coarse-level reconstruction alone [see Fig. 5(e)] in approximately one tenth of computation time for the fine-grid reconstruction, because the energy of the original image is contained mostly in the LL band. This energy-compaction property of the wavelet transform makes the reconstruction in the wavelet domain more efficient for images that are smooth. We can also see that, as expected, the two-grid TLS solution provided more accurate results than the two-grid LS solution [compare Figs. 5(b) and 5(f)]. To evaluate more quantitatively the savings in computation time by the multigrid over the single-grid algorithm, Fig. 6 compares the value of the the Rayleigh quotient $F(\mathbf{q})$ (in log scale) after successive numbers of work units. A work unit here is defined as the amount of computation for performing one iteration in the finest grid.[29] It can be seen that the multigrid method achieves better results and is significantly faster.

In the third experiment a 1.5-cm-diameter rod was placed at the center position with respect to the source and the detector ring. As with the second experiment, the rod had a nonhomogeneous absorption distribution, following a sinusoidal pattern (one positive cycle only). The forward solution in this experiment was also obtained by the multigrid finite-difference solver described in Ref. 3. In this experiment, white Gaussian noise with 20-dB SNR was added to the detector readings. The same amount of additive white Gaussian noise (the same noise variance as in the detector readings) was also added to the weights. We performed this experiment five times for different noise realizations; the mean of the MSE's and CC's are given in Table 3. Note that $\|\mathbf{x}\|^2 = 5.96 \times 10^{-1}$ in the calculation. Figure 7 shows the reconstruction results for a noise realization. From Fig. 7 and Table 3 we see again that the two-grid TLS solution provided more accurate results than the two-grid LS solution, both qualitatively and quantitatively, with the same or less computation time. We can also see that the coarse-level reconstruction alone [see Fig. 7(e)] produced fairly good results compared with the final solution in Fig. 7(f) but with only one tenth of the computation time. Notice that the multigrid method not only reduced computation time but also suppressed noise artifacts [compare Figs. 7(d) and 7(f)]. This is because the coarse solution is quite robust to noise and, when used as the initial solution in the finer resolution, makes the entire algorithm more stable.

Figure 8 illustrates the wavelet decomposition of the original image and the final reconstructed image for experiments 2 and 3. From this figure we can see that the LL bands (lower left corners in the displays) of the reconstructed images are quite close to the LL bands of the original images [compare the lower left of Figs. 8(a) and 8(b); also see lower left of Figs. 8(c) and 8(d)].

6. DISCUSSION AND CONCLUSIONS

In this paper the RQF-TLS method is used for solving the perturbation equation in optical tomography. The statistical properties of the RQF-TLS estimate are derived by means of a perturbation analysis. The equivalency between the RQF-TLS estimator and ML estimator is shown when the noise is Gaussian and i.i.d. To reduce computation complexity, a wavelet-based multiresolution TLS reconstruction scheme is proposed for solving the perturbation equation. Experimental results indicate that the proposed multigrid algorithm provides significantly better reconstruction than the single-grid method with comparable computation time.

In general, the inverse problem in optical tomography is ill-posed. In the LS formulation, which minimizes the perturbation in the measurement vector, a penalty term that forces the solution vector $\mathbf{x}$ to be smooth or to have finite energy can be imposed (Tikhonov regularization) to overcome the ill-posed nature of the matrix $\mathbf{H}$. With the RQF-TLS, the finite energy constraint is embedded in the Rayleigh quotient function. The first- or second-order derivative type of penalty terms as well as others might also be incorporated in the minimization of the Rayleigh quotient. The incorporation of such constraints to further improve the stability of the TLS solution is under investigation.

The proposed wavelet-based multigrid algorithm allows one to focus on special ROI's when going from a coarse grid to a fine grid. Further study is needed for automatic detection of ROI's from coarse solutions.

ACKNOWLEDGMENTS

This work was supported in part by the National Institutes of Health under grants RO1-CA59955 and RO1-CA66184. The authors thank Nikolas P. Galatsanos at the Illinois Institute of Technology and Randall L. Barbour at the State University of New York Health Science Center of Brooklyn for many valuable discussions and

Table 3. MSE and CC for Different Methods for a Test Medium Containing a Centered Sinelike Absorber

	One-Grid TLS with 85 Iterations	Two-Grid TLS with 1200 Iterations at the Coarse Grid	One-Grid TLS with 885 Iterations	Two-Grid TLS with 600 Iterations at the Fine Grid	Two-Grid LS
Average CC	$3.63E \times 10^{-1}$	$4.71E \times 10^{-1}$	$5.07E \times 10^{-1}$	$5.31E \times 10^{-1}$	$2.23E \times 10^{-1}$
Average MSE	$1.74E \times 10^{-4}$	$1.63E \times 10^{-4}$	$1.60E \times 10^{-4}$	$1.60E \times 10^{-4}$	$1.77E \times 10^{-4}$

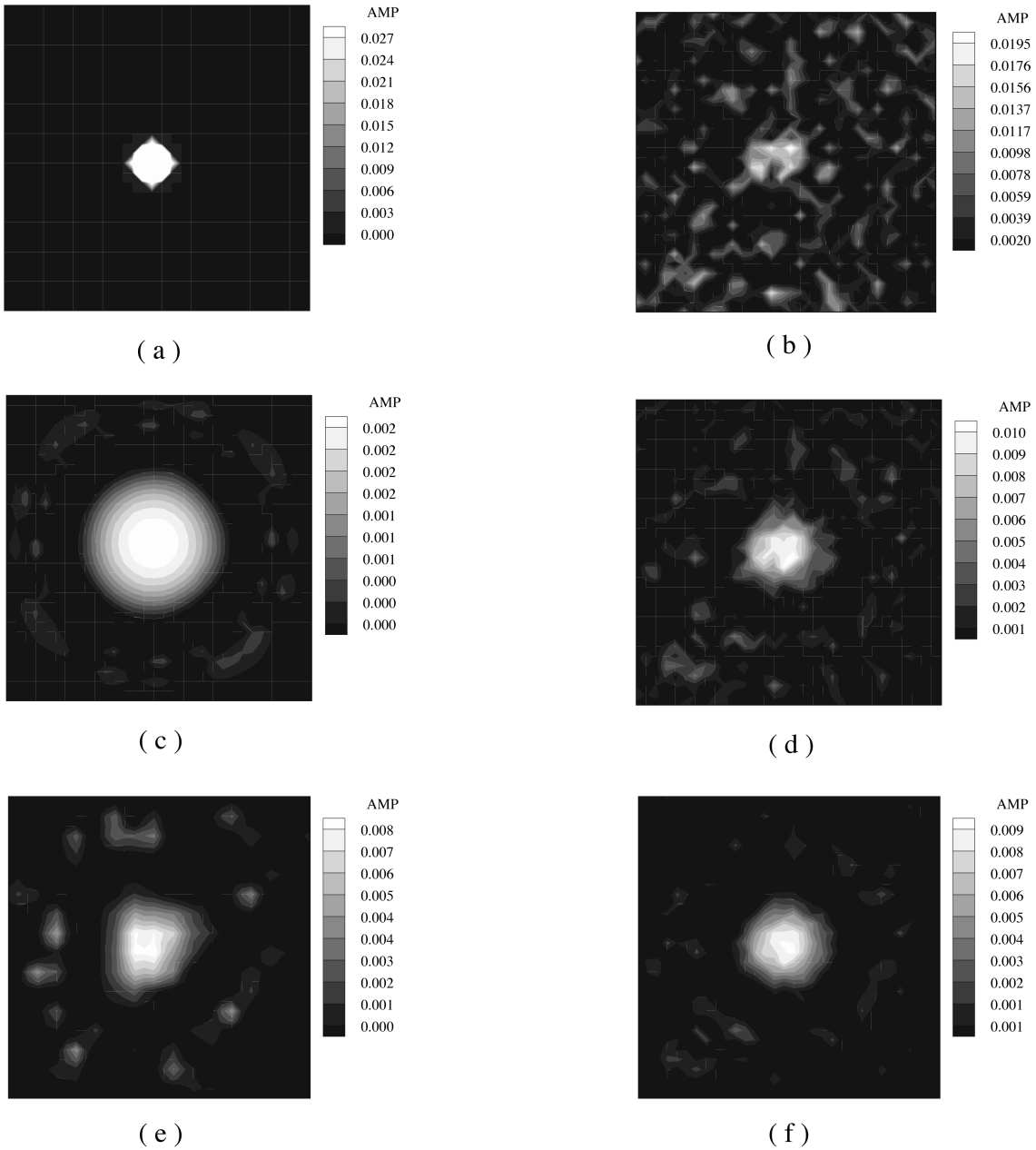

Fig. 7. Reconstruction results of a medium containing a centered sinelike absorber. The SNR in the data is 20 dB. The weights are corrupted by the same noise as that in the data. (a) Original image, (b) reconstruction result with multigrid LS with 1200 iterations at the coarse grid and 800 at the fine grid, (c) reconstructed image with one-grid TLS with 85 iterations, (d) reconstructed image with one-grid TLS with 885 iterations, (e) reconstructed image with two-grid algorithm with 1200 iterations at the coarse grid only, (f) reconstructed image with the two-grid algorithm with an additional 600 iterations at the fine grid with (e) as the initial solution. The computation time for (c) and (e) are roughly same. The time for (f) is approximately 3/4 of (d) and that for (c) or (e) is approximately 1/10 of (d).

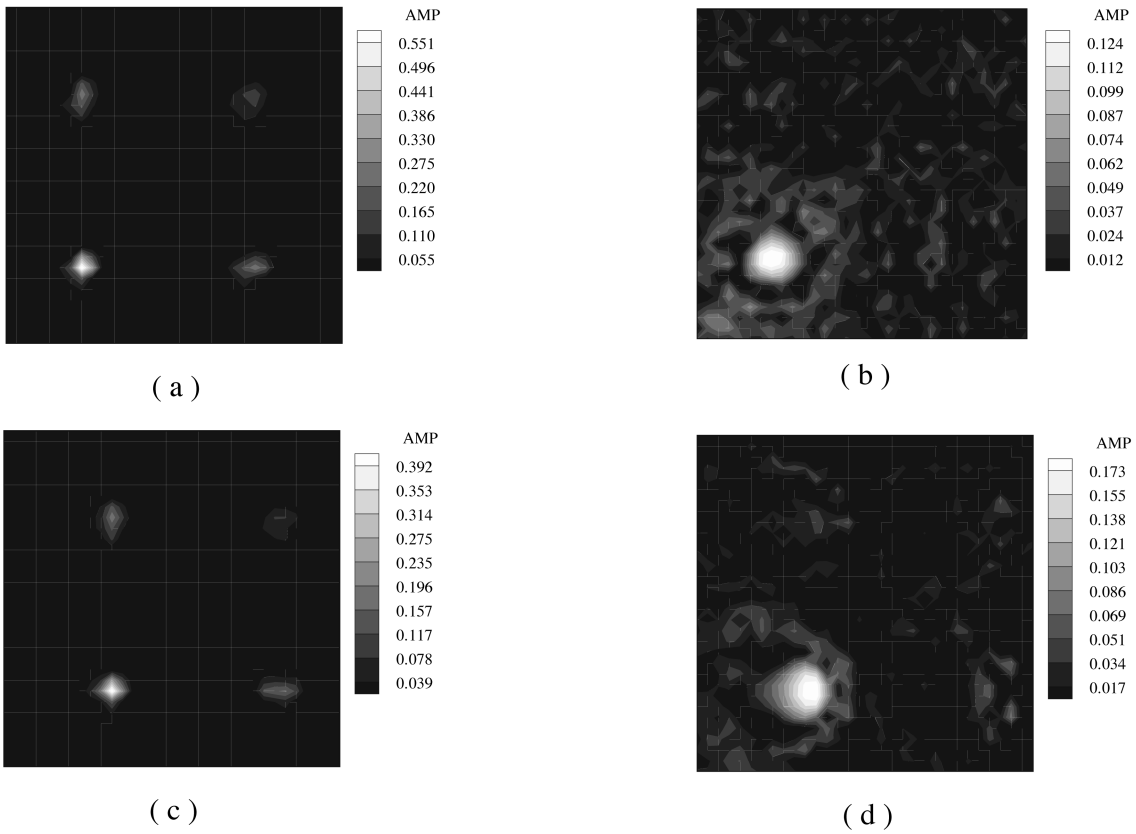

Fig. 8. Wavelet-transform analysis. (a) Wavelet transform of Fig. 7(a), (b) wavelet transform of Fig. 7(f), (c) wavelet transform of Fig. 6(a), (d) wavelet transform of Fig. 6(f). To reveal the high-frequency subsignals, we applied a square-root mapping to the signal magnitude when plotting.

suggestions. We also thank the reviewers for their valuable suggestions and comments for improving the presentation of this paper.

W. Zhu is now with Bell Labs, Lucent Technologies, Holmdel, New Jersey.

REFERENCES

1. R. L. Barbour, H. L. Graber, Y. Wang, J. Chang, and R. Aronson, "A perturbation approach for optical diffusion tomography using continuous-wave and time-resolved data," in *Medical Optical Tomography: Functional Imaging and Monitoring*, G. Mueller, B. Chance, R. Alfano, S. Arridge, J. Beuthan, E. Gratton, M. Kaschke, B. Masters, S. Svanberg, and P. van der Zee, eds., Vol. IS11 of Institute Series (Society of Photo-optical Instrumentation Engineers, Bellingham, Washington, 1993), pp. 87–120.
2. Y. Wang, J. Chang, R. Aronson, R. L. Barbour, H. L. Graber, and J. Lubowsky, "Imaging scattering media by diffusion tomography: an iterative perturbation approach," in *Physiological Monitoring and Early Detection Diagnostic Methods*, T. S. Mang, ed., Proc. SPIE **1641**, 58–71 (1992).
3. Y. Q. Yao, Y. Wang, Y. L. Pei, W. Zhu, J. H. Hu, and R. L. Barbour, "Frequency domain optical tomography in human tissue," in *Experimental and Numerical Methods for Solving Ill-Posed Inverse Problems: Medical and Nonmedical Application*, R. L. Barbour, M. J. Carvlin, and M. A. Fiddy eds., Proc. SPIE **2570**, 254–266 (1995).
4. J. R. Singer, F. A. Grünbaum, P. Kohn, and J. P. Zubelli, "Image reconstruction of the interior of bodies that diffuse radiation," Science **248**, 990–993 (1990).
5. S. R. Arridge, P. van der Zee, M. Cope, and D. Delpy, "New results for the development of infrared-red absorption imaging," in *Biomedical Image Processing*, A. C. Bovik and W. E. Higgins, eds., Proc. SPIE **1245**, 92–103 (1990).
6. S. R. Arridge, "The forward and inverse problems in time resolved infra-red imaging," in *Medical Optical Tomography: Functional Imaging and Monitoring*, G. Mueller, B. Chance, R. Alfano, S. Arridge, J. Beuthan, E. Gratton, M. Kaschke, B. Masters, S. Svanberg, and P. van der Zee, eds., Vol. IS11 of Institute Series (Society of Photo-optical Instrumentation Engineers, Bellingham, Washington, 1993), pp. 35–64.
7. W. Zhu, Y. Wang, Y. Deng, Y. Yao, and R. Barbour, "A wavelet-based multiresolution regularized least squares reconstruction approach for optical tomography," IEEE Trans. Med. Imag. **16**, 210–217 (1997).
8. W. Zhu, Y. Wang, Y. Yao, J. Chang, H. L. Graber, and R. Barbour, "Iterative total least-squares image reconstruction algorithm for optical tomography by the conjugate gradient method," J. Opt. Soc. Am. A **14**, 799–807 (1997).
9. J. Chang, Y. Wang, R. Aronson, H. L. Graber, and R. L. Barbour, "A layer-stripping approach for recovery of scattering media from time-resolved data," in *Inverse Problems in Scattering and Imaging*, M. A. Fiddy, ed., Proc. SPIE **1767**, 384–395 (1992).
10. J. Chang, W. Zhu, Y. Wang, H. L. Graber, and R. L. Barbour, "Regularized progressive expansion algorithm for recovery of scattering media from time-resolved data," J. Opt. Soc. Am. A **14**, 306–312 (1997).
11. G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," SIAM (Soc. Ind. Appl. Math.) J. Math. Anal. **17**, 883–893 (1980).
12. V. Z. Measarvić, N. P. Galatsanos, and A. Katsaggelos, "Regularized constrained total least squares image restoration," IEEE Trans. Image Process. **4**, 1096–1108 (1995).
13. P. Li, S. W. Flax, E. S. Ebbini, and M. O'Donnell, "Blocked

element compensation in phased array imaging," IEEE Trans. Ultrason. Frequencies Freq. Control **40**, 282–292 (1993).
14. G. Wang, J. Zhang, and G. W. Pan, "Solution of inverse problems in image processing by wavelet expansion," IEEE Trans. Image Process. **4**, 579–593 (1995).
15. L. Blanc-Feraud, P. Charbonnier, P. Lobel, and M. Barlaud, "A fast tomographic reconstruction algorithm in the 2D wavelet transform domain," in *Proceedings of the 1994 IEEE International Conference on Acoustics, Speech, and Signal Processing* (Institute of Electrical and Electronics Engineers, Piscataway, N.J., 1994), pp. 305–308.
16. A. H. Delaney and Y. Bresler, "Multiresolution tomographic reconstruction using wavelets," IEEE Trans. Image Process. **4**, 799–813 (1995).
17. M. R. Banham, Nikolas P. Galatsanos, H. L. Gonzalez, and A. K. Katsaggelos, "Multichannel restoration of single channel images using a wavelet-based subband decomposition," IEEE Trans. Image Process. **3**, 821–833 (1994).
18. M. T. Silvia and E. C. Tacker, "Regularization of Marchenko's integral equation by total least squares," J. Acoust. Soc. Am. **72**, 1202–1207 (1982).
19. G. H. Golub, "Some modified matrix eigenvalue problems," SIAM (Soc. Ind. Appl. Math.) Rev. **15**, 318–334 (1973).
20. H. Chen, T. K. Sarkar, S. A. Dianat, and J. D. Brule, "Adaptive spectral estimation by the conjugate gradient method," IEEE Trans. Acoust. Speech Signal Process. **34**, 272–284 (1986).
21. X. Yang, T. K. Sarkar, and E. Arvas, "A survey of conjugate gradient algorithms for solution of extreme eigen-problem of a symmetric matrix," IEEE Trans. Acoust. Speech Signal Process. **37**, 1550–1556 (1989).
22. S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis* (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1991).
23. T. J. Abatzoglou, J. M. Mendel, and G. A. Harada, "The constrained total least squares technique and its applications to harmonic superresolution," IEEE Trans. Signal Process. **39**, 1070–1087 (1991).
24. V. Z. Measarvić and N. P. Galatsanos, "MAP and regularized constrained total least squares image restoration," in *Proceedings of the 1994 IEEE International Conference on Image Processing* (Institute of Electrical and Electronics Engineers, Piscatway, N.J., 1994), Vol. 3, pp. 177–181.
25. S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Trans. Pattern. Anal. Mach. Intell. **11**, 674–693 (1989).
26. W. Hackbusch, *Multigrid Methods and Applications* (Springer-Verlag, Berlin, 1985).
27. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge U. Press, Cambridge, UK, 1992).
28. Y. Q. Yao, Y. Wang, R. L. Barbour, H. L. Graber, and J. W. Chang, "Scattering characteristics of photon density waves from an object in a spherically two-layer medium," in *Optical Tomography, Photon Migration, and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation*, B. Chance and R. R. Alfano, eds., Proc. SPIE **2389**, 291–303 (1995).
29. W. L. Briggs, *A Multigrid Tutorial* (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1987).