GEOPHYSICS, VOL. 64, NO. 3 (MAY-JUNE 1999); P. 888–901, 11 FIGS., 1 TABLE.

Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model

R. Gerhard Pratt*

ABSTRACT

Seismic waveforms contain much information that is ignored under standard processing schemes; seismic waveform inversion seeks to use the full information content of the recorded wavefield. In this paper I present, apply, and evaluate a frequency-space domain approach to waveform inversion. The method is a local descent algorithm that proceeds from a starting model to refine the model in order to reduce the waveform misfit between observed and model data. The model data are computed using a full-wave equation, viscoacoustic, frequency-domain, finite-difference method. Ray asymptotics are avoided, and higher-order effects such as diffractions and multiple scattering are accounted for automatically.

The theory of frequency-domain waveform/wavefield inversion can be expressed compactly using a matrix formalism that uses finite-difference/finite-element frequency-domain modeling equations. Expressions for fast, local descent inversion using back-propagation techniques then follow naturally. Implementation of these methods depends on efficient frequency-domain forward-modeling solutions; these are provided by recent developments in numerical forward modeling. The inversion approach resembles prestack, reverse-time migration but differs in that the problem is formulated in terms of velocity (not reflectivity), and the method is fully iterative.

I illustrate the practical application of the frequency-domain waveform inversion approach using tomographic seismic data from a physical scale model. This allows a full evaluation and verification of the method; results with field data are presented in an accompanying paper. Several critical processes contribute to the success of the method: the estimation of a source signature, the matching of amplitudes between real and synthetic data, the selection of a time window, and the selection of suitable sequence of frequencies in the inversion.

An initial model for the inversion of the scale model data is provided using standard traveltime tomographic methods, which provide a robust but low-resolution image. Twenty-five iterations of wavefield inversion are applied, using five discrete frequencies at each iteration, moving from low to high frequencies. The final results exhibit the features of the true model at subwavelength scale and account for many of the details of the observed arrivals in the data.

INTRODUCTION

During the last decades the seismic waveform inverse problem has been tackled by many applied mathematicians and geophysicists with an increasing degree of success. The aim of such research, simply stated, is to develop processing methods that take data from a seismic survey and (automatically, if possible) compute the best earth model consistent with the data (and consistent with any existing a priori information). An important step was taken by Lailly (1983) and Tarantola (1984), who recognized that models could be improved iteratively by back-propagating the data residuals and correlating the result with forward-propagated wavefields in a manner very similar to many prestack migration algorithms. Numerical results using this approach are given by Kolb et al. (1986), Gauthier et al. (1986), and Mora (1987). These authors all formulate their methods in the time domain. Shin (1988), Pratt and Worthington (1990), Pratt (1990b), Geller and Hara (1993), and Song et al. (1995) apply the same idea to inverse problems in the frequency domain and use an implicit frequency-domain numerical algorithm to provide the forward model.

Several researchers have achieved considerable improvements in tomographic (transmission) imaging from cross-borehole data by utilizing gradient methods for waveform inversion

Published on Geophysics Online January 11, 1999. Manuscript received by the Editor November 6, 1997; revised manuscript received September 10, 1998. Published on line January 11, 1999.
*Formerly Department of Geology, Imperial College, London SW7 2BP, United Kingdom; currently Queen's University, Department of Geological Sciences, Kingston, Ontario K7L 3N6, Canada. E-mail: pratt@geol.queensu.ca.

in conjunction with finite-difference modeling. Results with real crosshole data have been demonstrated by Zhou et al. (1995), who use a time-domain approach, and by Song et al. (1995) and Pratt et al. (1995), who use a frequency-domain approach. The motivation in the use of waveform inversion for crosshole data (or other transmission data) is to improve the resolution when compared with traveltime inversion methods (Williamson, 1991; Williamson and Worthington, 1993). Although crosshole seismic methods are a specialized technique, in many respects the successful inversion of crosshole data, with the difficult phases often observed in the data [see, for example, Van Schaack et al. (1995)], paves the way for the application of such techniques to more conventional reflection seismic data.

In this paper, and in the accompanying paper (Pratt and Shipp, 1999), I return to the frequency-domain approach for seismic waveform inversion. Here I provide a new, compact derivation of the theory and use the methodology on scale model data; in the accompanying paper the method is applied to a real crosshole data set. Time-domain and frequency-domain methods are analytically equivalent; to some extent the nonspecialist (or end user) need not be concerned in which domain the results are obtained. However, the numerical advantages of a frequency-domain implementation are formidable. There are also other good reasons for preferring the frequency domain. Most of the advantages are linked to the forward modeling algorithm. Marfurt (1984) was the first to point out that the frequency domain is the method of choice for finite-difference/finite-element modeling if a significant number of source locations are involved. I pointed out later (Pratt, 1990a) that large aperture seismic surveys could be inverted effectively using only a limited number of frequency components, thus reducing the number of frequency-domain forward solutions in a manner not possible with time-domain methods. To these important advantages, we may add the ease with which viscous attenuation and dispersion are incorporated into frequency-domain methods and the ease with which inverse methods can be implemented to use the lowest data frequencies first, thus mitigating the notorious nonlinearities in the seismic inverse problem.

Several recent developments in waveform inversion make this paper timely and relevant:

1) Advances in seismic modeling by finite differences in the frequency domain (Jo et al., 1996; Štekl and Pratt, 1998) that make the method significantly more efficient than the original second-order method of Pratt (1990a);
2) Advances in computer hardware—particularly in the wider availability of cheap random access memory—that make the routine application of the frequency domain method possible on standard workstations;
3) A new description of the theory based on matrix algebra, rather than functional analysis, that makes the method more accessible (Pratt et al., 1998); and
4) A practical method for minimizing artifacts that makes use of the capability of incorporating viscous attenuation into the forward model.

The method of frequency-domain waveform inversion is verified in this paper by applying the method to tomographic seismic data from a physical scale model. The use of scale model data allows a full evaluation and verification of the method. As in synthetic studies the final result can be compared directly to the true model. Unlike computer-generated data, however, these data come from a real, physical experiment and contain both random and signal-generated noise—we are not simply testing the computer software in forward and inverse modes. These same scale model data have been used before to demonstrate a variety of imaging and inversion methods (Pratt and Goulty, 1991; Pratt et al., 1991; Leggett et al., 1993; Rowbotham and Goulty, 1995; Reiter and Rodi, 1996); an objective comparison can thus be made between those results and the results I present here.

The paper is divided into two main sections. In the first section I provide a review of the theory for seismic waveform inversion, specifically cast into the frequency-space domain. By specifying the frequency-domain method from the outset, one can introduce a discrete matrix formalism for the seismic waveform inverse problem directly in the frequency-space domain. The matrix formalism replaces the more difficult notions of functional analysis of, for example, Tarantola (1987). This section of the paper concludes with a matrix algebra demonstration of Lailly's (1983) fast back-propagation method for computing the gradient of the misfit function for the waveform inversion problem.

In the second section of the paper I give a complete account of the inversion of the physical scale model data. The model geometry simulates a seismic crosshole experiment. This section deals with the generation (and validation) of an appropriate starting model, the time windowing of the data, the estimation of a source signature, the matching of amplitudes between model and observed data, the selection of a suitable sequence of frequencies in the inversion, and the verification of the final results. The final results exhibit the features of the true model at subwavelength scales. Moreover, modeling in the final result provides model data that match many of the waveform details of the observed data.

THEORY

The mathematics of descent-type inversion methods have been given in the past in terms of functional analysis, using infinite dimensional spaces to represent the data and the model [see, for example, Tarantola (1987)]. These developments are often difficult to follow; ultimately, the results must be expressed in the form of computer programs that deal with finite-dimensional representations of the data and the models. In this section I present an alternative approach (which can be followed in more detail in Pratt et al., 1998), in which finite-dimensional (i.e., discrete valued) data and model representations are assumed from the outset. The integral operations of functional analysis are thereby replaced with matrix operations.

Full-wave equation forward modeling in the space-frequency domain

I begin by assuming that the forward problem can be adequately simulated using a finite-element or a finite-difference method. These numerical methods for wave modeling avoid the use of ray theory, which is an asymptotic theory that ignores crucial low-frequency effects in the data. I use the full-wave equation, and not a one-way wave equation, which ignores backscattering and wide-angle effects. Usually in seismic wave modeling, explicit time marching schemes are used. Here

I assume instead that an implicit, time-independent scheme has been obtained by temporal Fourier transformation of the full-wave equation. The resultant numerical system can be represented compactly by

$$\mathbf{S}(\omega)\,\mathbf{u}(\omega) = \mathbf{f}(\omega) \quad \text{or} \quad \mathbf{u}(\omega) = \mathbf{S}^{-1}(\omega)\,\mathbf{f}(\omega) \qquad (1)$$

[See, for example, Marfurt (1984) or Pratt (1990a)], where $\mathbf{u}$ is the Fourier-transformed, complex-valued, discretized wavefield (i.e., the pressure or the displacement), sampled at l node points and arranged as an $l \times 1$ column vector, $\mathbf{S}$ is an $l \times l$ complex-valued impedance matrix, and $\mathbf{f}$ are the source terms, also arranged as an $l \times 1$ column vector. Although I will use the second mathematical form in equation (1), which makes use of the inverse impedance matrix, $\mathbf{S}^{-1}(\omega)$, this is purely symbolic; the impedance matrix is very large (sparse) and its inverse is difficult and expensive to compute. Interestingly, the inverse impedance matrix is identical in form to Berkhout's response matrix (Berkhout, 1982), although Berkhout's approach to imaging and inversion is not used in this paper.

The impedance matrix is computed by numerical approximation of the underlying (spatial) partial differential equations, given the discretized physical parameters (for example, the seismic velocities sampled at each of the node points). The boundary conditions are included implicitly; these are incorporated directly into the various matrix coefficients. Equation (1) can be used to represent acoustic, elastic, or electromagnetic wave propagation as well as potential field problems or diffusion problems.

Viscous damping (attenuation) and dispersion may be included when generating the impedance matrix simply by allowing the material properties to be complex valued and frequency dependent. As we shall see, it is convenient (but not required) that source-receiver reciprocity holds in equation (1). If the numerical problem is exactly reciprocal with respect to an interchange of source and receiver elements, then both $\mathbf{S}$ and $\mathbf{S}^{-1}$ are symmetric matrices.

Matrix solution methods.—The discrete frequency-domain modeling problem represented by equation (1) is an implicit specification of the wavefield; the second, explicit form shown in equation (1) is only representational, as it is not generally possible (or desirable) to actually invert the very large impedance matrix $\mathbf{S}$. For multiple source problems, equation (1) is best solved using direct matrix factorization methods, such as LU decomposition (Press et al., 1992; Pratt, 1990b). It is critical to use ordering schemes that allow maximum advantage to be taken of the sparsity of both $\mathbf{S}$ and its LU factorization; nested dissection (Liu and George, 1981; Marfurt and Shin, 1989) is such a method. If LU decomposition is used to solve equation (1), the matrix factors can be reused to rapidly solve the forward problem for any new source vector, $\mathbf{f}$. This is especially important in the iterative solution of the inverse problem, in which many forward solutions for the true sources and for the virtual sources are required at each iteration.

The inverse problem in the space-frequency domain

Mathematically, the inverse problem is to estimate a set of model parameters from the recorded data. The model can be thought of as being specified at each of the l node points, as above. However, it is more sensible to define the model in terms of a more general parameter set, $\mathbf{p}$, an $m \times 1$ column vector, where $m \neq l$ in general (and usually $m < l$). In terms of generating forward modeled data, it is only necessary to be able to compute the $l \times l$ impedance matrix $\mathbf{S}$ from the parameters $\mathbf{p}$, for example, by interpolation. As we shall see below, we will also require expressions for the partial derivatives $\partial\mathbf{S}/\partial p_i$ for all m model parameters.

Suppose we have n experimental observations, $\mathbf{d}$ (after Fourier transformation at a particular frequency, ω), recorded at a subset of nodal points corresponding to the receiver locations. The explicit dependence of each term in equation (1) on the angular frequency, ω, will be dropped from now on. We proceed in the inversion by reducing the misfit between these observed data and the modelled data by iteratively updating the model. As always with such descent methods, I assume the existence of a suitable initial model, $\mathbf{p}^{(0)}$, that is close enough to the global solution to allow successively more accurate data fitting to be achieved using local information about the misfit function. This is appropriate for seismic tomography, since good starting models can be generated by the more linear but more approximate methods of traveltime analysis. Given the initial model, we can calculate the impedance matrix, $\mathbf{S}$, and the initial wavefield estimate by forward modeling.

The residual error at the n receiver node points, $\delta\mathbf{d}$ is defined as the difference between the initial model response and the observed data at the receiver locations. Thus

$$\delta d_i = u_i - d_i, \qquad i = (1, 2, \ldots, n), \qquad (2)$$

where the subscripted quantities are the individual components of $\delta\mathbf{d}$, $\mathbf{u}$, and $\mathbf{d}$ and the subscript i represents the receiver number. (It is convenient to assume the node points are ordered in such a way that the first $n \leq l$ node points are receiver locations, but the results I shall describe are not specific to such an ordering scheme.) As is common in many inverse problems, we seek to minimize the sum of squared residual errors

$$E(\mathbf{p}) = \tfrac{1}{2}\delta\mathbf{d}^t\delta\mathbf{d}^* \qquad (3)$$

(the l_2 norm misfit function), where (here and later) the superscript t represents the ordinary matrix transpose and the superscript * represents the complex conjugate.

Gradient method of inversion.—To proceed, we use the direction in parameter space defined by the gradient of the misfit function, $\nabla_p E$. We produce a new parameter set by (iteratively) updating the parameter vector according to

$$\mathbf{p}^{(k+1)} = \mathbf{p}^{(k)} - \alpha^{(k)}\nabla_p E^{(k)}, \qquad (4)$$

where k is an iteration number and α is a step length (a positive scalar) chosen by simple line search methods to minimize the l_2 norm in the direction given by the gradient of $E(\mathbf{p})$. The iteration in equation (4) is performed until some suitable stopping criteria is reached. The convergence rate of the gradient method is generally quite slow; convergence can be improved by adopting a conjugate gradient approach [see, for example, Mora (1987)] that does not require any significant additional computations.

The key quantity we require for such descent-type inverse methods is the gradient direction, $\nabla_p E$. We may evaluate the

gradient direction by taking partial derivatives of equation (3) with respect to each of the m model parameters, $\mathbf{p}$, yielding

$$\nabla_p E = \frac{\partial E}{\partial \mathbf{p}} = \mathcal{R}eal\{\mathbf{J}^t \delta \mathbf{d}^*\}. \quad (5)$$

The matrix $\mathbf{J}^t$ is the transpose of the $n \times m$ Frechét derivative matrix, $\mathbf{J}$, the (complex valued) elements of which are given by

$$J_{ij} = \frac{\partial u_i}{\partial p_j} \qquad i = (1, 2, \ldots, n); \quad j = (1, 2, \ldots, m). \quad (6)$$

Computation of the Frechét derivatives.—We now wish to explicitly link the computation of the gradient vector to the forward frequency-domain modeling problem given in equation (1). To do this we first augment the $m \times n$ matrix $\mathbf{J}$ with the additional rows required to define partial derivatives at all node points—not just at the receiver locations—to obtain a new $m \times l$ matrix $\hat{\mathbf{J}}$. We may then write a new equation equivalent to equation (5):

$$\nabla_p E = \mathcal{R}eal\{\hat{\mathbf{J}}^t \delta \hat{\mathbf{d}}^*\}, \quad (7)$$

where $\delta \hat{\mathbf{d}}^*$ is the data residual vector of length n, augmented with $(l - n)$ zero values to produce a new vector of length l.

An expression for each column of the augmented Frechét derivative matrix, $\hat{\mathbf{J}}$, can be obtained by taking the partial derivative of both sides of equation (1) with respect to the ith parameter p_i:

$$\mathbf{S}\frac{\partial \mathbf{u}}{\partial p_i} = -\frac{\partial \mathbf{S}}{\partial p_i}\mathbf{u} \quad \text{or} \quad \frac{\partial \mathbf{u}}{\partial p_i} = \mathbf{S}^{-1}\,\mathbf{f}^{(i)}. \quad (8)$$

From a comparison with equation (1), we see that the benefit of writing this solution using the matrix inverse is that we may obtain the required partial derivatives by solving a different forward problem in which we have introduced the ith virtual source term

$$\mathbf{f}^{(i)} = -\frac{\partial \mathbf{S}}{\partial p_i}\mathbf{u}, \quad (9)$$

itself an $l \times 1$ vector [see Rodi (1976) and Oristaglio and Worthington (1980), who originally suggested this approach for the electromagnetic problem]. The computation of the partial derivatives of the impedance matrix, $\partial \mathbf{S}/\partial p_i$, depends on the specific details of the finite approximation method used. However, in all cases these are relatively trivial to compute.

Since we could generate an equation similar to equation (8) for any choice of i, we can represent all of the partial derivatives simultaneously by the matrix equation

$$\hat{\mathbf{J}} = \left[\frac{\partial \mathbf{u}}{\partial p_1}\frac{\partial \mathbf{u}}{\partial p_2}\cdots\frac{\partial \mathbf{u}}{\partial p_m}\right] = \mathbf{S}^{-1}\left[\mathbf{f}^{(1)}\,\mathbf{f}^{(2)}\cdots\mathbf{f}^{(m)}\right]$$
$$\text{or} \quad \hat{\mathbf{J}} = \mathbf{S}^{-1}\mathbf{F}, \quad (10)$$

where $\mathbf{F}$ is an $l \times m$ matrix, the columns of which are the virtual source terms for each of the m physical parameters. Computation of the elements of $\hat{\mathbf{J}}$ using equation (10) would require m forward propagation problems to be solved (one for each column in $\mathbf{F}$), in addition to the one required to compute the virtual sources using equation (9). This is unfortunate since m can be a large number and forward modeling is expensive. Computing the partial derivatives explicitly is, however, not the most efficient way to proceed.

Efficient calculation of the gradient direction.—We can significantly reduce the cost of computing the gradient vector by some simple matrix manipulations. Substituting equation (10) into equation (7), we obtain

$$\nabla_p E = \mathcal{R}eal\{\hat{J}^t \delta \hat{\mathbf{d}}^*\} = \mathcal{R}eal\{\mathbf{F}^t[\mathbf{S}^{-1}]^t \delta \hat{\mathbf{d}}^*\} \quad (11)$$

or

$$\nabla_p E = \mathcal{R}eal\{\mathbf{F}^t \mathbf{v}\}, \quad (12)$$

where

$$\mathbf{v} = [\mathbf{S}^{-1}]^t \delta \hat{\mathbf{d}}^* \quad (13)$$

is the back-propagated wavefield. If the inverse impedance matrix $\mathbf{S}^{-1}$ is symmetric, as we expect for reciprocal problems, then $[\mathbf{S}^{-1}]^t = \mathbf{S}^{-1}$ and

$$\mathbf{v} = \mathbf{S}^{-1}\delta \hat{\mathbf{d}}^*. \quad (14)$$

In either equation (13) or (14), computing the new back-propagated wavefield $\mathbf{v}$ requires only one additional forward problem to be solved.

Since the computation of the step length also requires the solution of at least one more forward problem, this brings the total number of required forward solutions to three. The forward and back-propagation problems are solved in the same model; hence, the stored LU factors of $\mathbf{S}$ can still be used to compute these solutions rapidly. This back-propagation method parallels the back-propagation method derived by Lailly (1983), who used the (more difficult) methods of functional analysis. Tarantola (1986), Kolb et al. (1986), Mora (1987), Pratt (1990b), Pratt and Worthington (1990), and Chavent and Jacewitz (1995) all exploit the back-propagation technique in computing the gradient of the misfit function for nonlinear waveform inversion. Further details of the matrix methods presented above are found in Pratt et al. (1998).

In many of the references cited in the previous paragraph, comparisons are drawn between the inverse method and prestack migration. There is indeed very little difference between the first iteration of the inverse method given above and prestack, reverse time migration. The inverse method is distinct because the problem is formulated in terms of fundamental model parameters (e.g., velocity, not reflectivity) and because the method is fully iterative (i.e., the model is updated and the process is reiterated to convergence).

APPLICATION TO PHYSICAL SCALE MODEL DATA

Scale model data

The physical scale model experiment was conducted in an ultrasonic modeling tank at the University of Durham (Sharp et al., 1985). Figure 1 is a photograph of the epoxy resin scale model used in this study; data from this model have also been used by others to demonstrate a variety of imaging and inversion schemes (Pratt and Goulty, 1991; Pratt et al., 1991; Leggett et al., 1993; Rowbotham and Goulty, 1995; Reiter and Rodi, 1996). As shown in Figure 2a, the model contains a number of horizontal layers of different velocities and thicknesses, a single channel feature, and a single dipping layer with a small fault. The model is sufficiently wide in the out-of-plane direction to allow us to assume the geometry is two dimensional.

The source and receiver piezoelectric transducers were operated at frequencies between approximately 200 and 800 kHz. Geometric and static corrections to account for the finite dimensions of the source and receiver transducers were applied to the data before any further processing. The full survey consisted of a total of 51 source positions and 51 receiver positions, each at 2.5-mm intervals and generating 2601 records. The target region was 55 mm wide by 125 mm long (i.e., 125 mm deep).

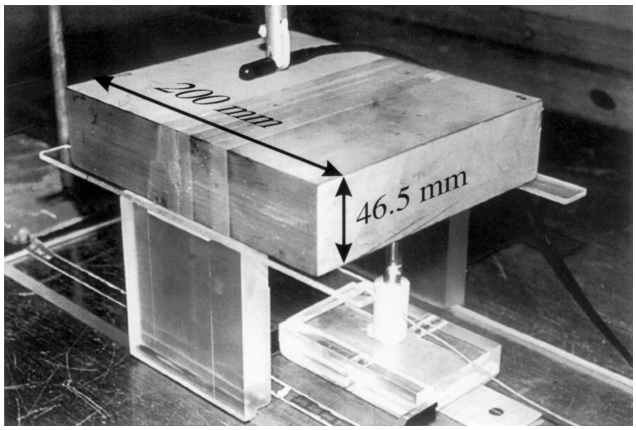

FIG. 1. A photograph of the epoxy resin model and the source transducer before immersion in water. The model has been tipped on one side in preparation for the experiment; the geometry of the epoxy resin layers is faintly visible. The top of the survey is at the right end of the photograph. See Figure 2a for a diagram of the layer geometry.

Scale dimensions.—All distances, times, and frequencies may be scaled by a factor of 1000 to produce cross-borehole dimensions that are realistic, although somewhat smaller than some reported real data examples. The scaled survey is 55 m across and 120 m deep, and scaled seismic frequencies lie between 200 and 800 Hz (used hereafter instead of the true lengths and frequencies). Time, measured on the model in microseconds, is given in scaled dimensions of milliseconds.

Data quality.—A representative common source gather from the survey is depicted in Figure 2b. This gather was synthesized by regathering the original common receiver gathers. The original picked arrival times are also shown superimposed on the data. The data are of good quality, with clear first arrival time breaks and only a low level of random noise. The source wavelet reverberates for several cycles at the dominant frequency, making the identification of individual phases difficult; but there is a clear time advance opposite the high-velocity layer in which the source is situated, with some evidence of secondary, reflected phases at slightly later times. Amplitudes fall off strongly with distance (and angle) at the top and at the bottom of the gather. A weak shear-wave phase, generated by mode conversions at the two edges of the model, is visible at late times near the top and the bottom of the gather.

Traveltime tomography

Since the waveform inversion method is a local, descent-type method, the first stage is to generate an appropriate starting model. The starting model must be sufficiently accurate to

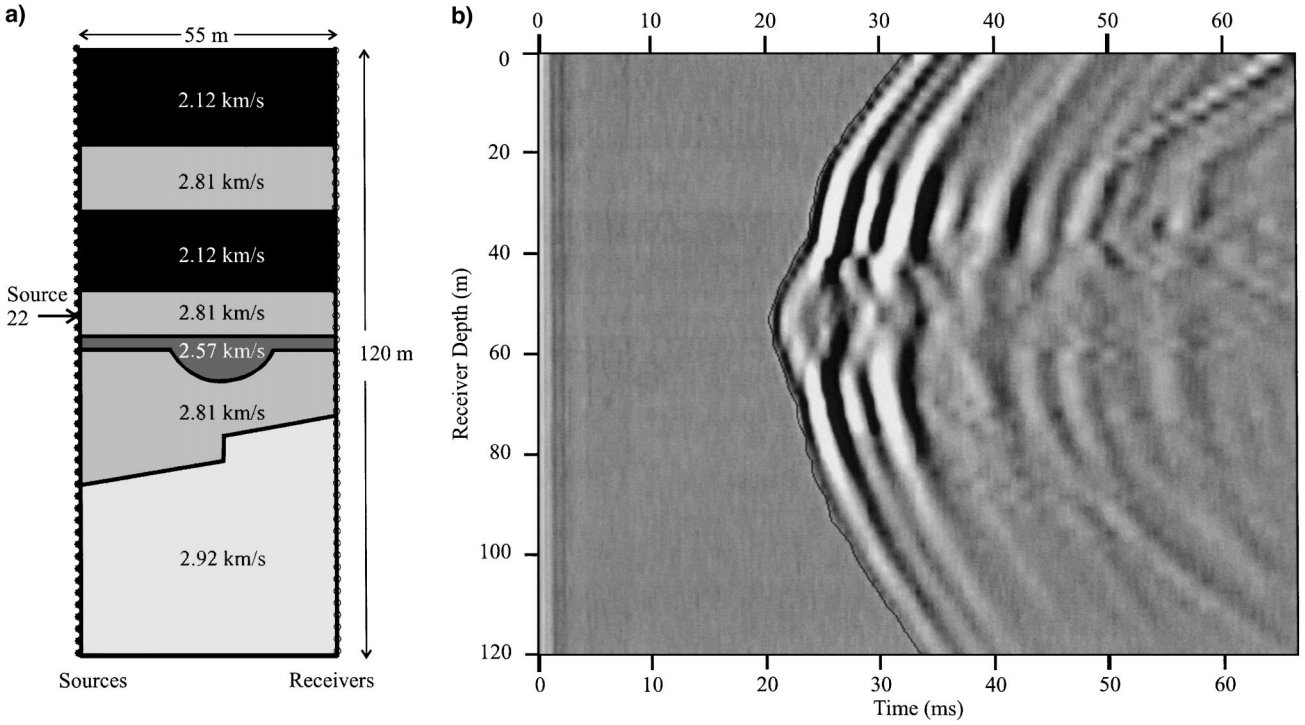

FIG. 2. The physical scale model and some representative data. (a) Schematic section through the epoxy resin scale model, showing the layer geometry, the nominal velocities, the source/receiver locations, and the location of source 22. All dimensions are given in scaled units, which are 1000 times the true dimensions (see text). (b) Common shot gather 22 from the epoxy resin model prior to data processing. Times are given in scaled units (1000 times the true times). The curve on the figure represents the picked arrival times.

allow us to descend to the global minimum of the misfit function. In practice, with seismic waveform inversion this implies that much of the waveform energy must be predicted to within a half-wavelength of the observed waveforms; if not, a minimum misfit model will be obtained when the predicted waveforms match the wrong cycle of the observed waveforms. Fortunately, in the crosshole configuration we can generate an appropriate model reasonably rapidly using traveltime methods.

Traveltimes for all raw data were picked (as in Figure 2b) and used in a tomographic inversion for the velocity field. I used the curved-ray traveltime tomographic methods described by Pratt et al. (1993), which include regularization methods for minimizing the first and second spatial derivatives. (I did not include any of the anisotropy parameters suggested in that paper.) The result is shown in Figure 3, together with the geometry and nominal acoustic velocities of the epoxy resin model. The tomogram exhibits the expected velocity variations, but at a very poor resolution. The large, strong, horizontal low-velocity layers are correctly imaged, but the central, semicircular channel feature is only very faintly present on the image. The dipping interface is recovered, but there is no trace of the small fault on this image.

Despite the apparent poor resolution of the tomogram in Figure 3b, this velocity structure predicts the traveltimes to within an rms level of 0.2 ms (compare this with the sample interval of 0.25 ms, in scaled time). Any attempt to reduce the level of residuals by reducing the level of constraints only resulted in the generation of noise in the images and not in the recovery of additional model features. The result appears to be consistent with Williamson's (1991) prediction of a resolving length for traveltime tomography of the first Fresnel zone width, approximately given by $\sqrt{\lambda L}$. Here, the dominant wavelength, $\lambda \approx 5$ m, and the propagation distance, $L \approx 50$ m, yield a resolving length of approximately 15 m.

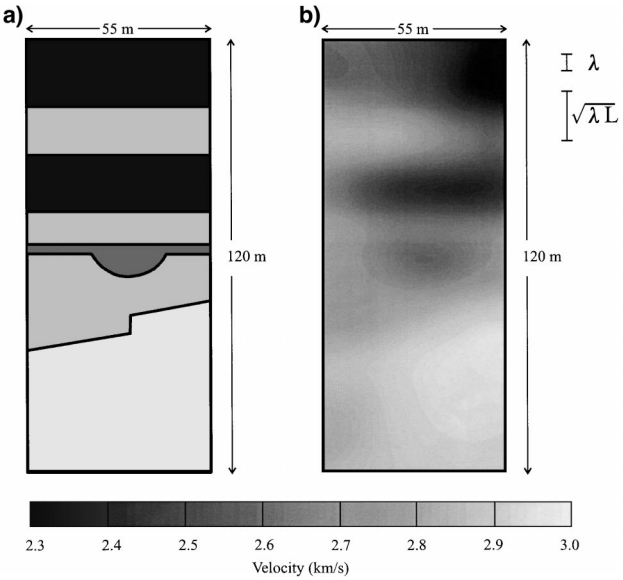

FIG. 3. (a) The true velocity model. Velocity values represent the nominal velocities of the epoxy resins used in the model. (b) Curved-ray tomography result from the picked traveltimes. The bars next to the figure depict the dominant wavelength and the approximate width of the first Fresnel zone.

Initial forward modeling

The suitability of the tomographic velocity image in Figure 3b as a starting model for waveform inversion can be verified by forward modeling and comparing the resultant waveforms with the original data. I carried out the modeling (and the subsequent inversions) using an acoustic-wave equation finite-difference method. We are primarily interested in modeling and inverting the dominant phases in the data—the P-wave first arrivals plus the reflections or diffractions created within a few milliseconds of these arrivals.

The waveform modeling was carried out using a 2-D, visco-acoustic, frequency-domain, finite-difference method. The second-order rotated operator method of Jo et al. (1996) was used. Grid-point intervals of $\Delta x = \Delta z = 0.8333$ m were used for the finite-difference modeling, corresponding approximately to five grid points per wavelength at 600 Hz. The total grid, including a five-grid-point buffer zone outside the edges of the survey, consisted of $75 \times 160 = 12\,000$ node points. The node velocities were taken directly from the traveltime tomography result of Figure 3b. A constant value of 2.5 kg/m^3 for the density was used in all the modeling and inversion results to follow. As a first guess for appropriate attenuation values, I used a constant Q model with a (homogeneous) quality factor of $Q = 100$; this guess is refined below. The source signature was estimated directly from the deconvolved data; this step is also described in more detail below. Time-domain records 64 ms long were synthesized by inverse Fourier transformation of the results; frequency-domain modeling was carried out every $\Delta f = 1/0.064 = 15.625$ Hz. A total of 77 frequency components were computed, up to 1200 Hz, beyond which the source spectrum contains virtually no energy.

The result of forward modeling in the traveltime tomography velocity field is shown in Figure 4a; the first arrival times are clearly matched to within a half-cycle, as required, but there are major differences in the appearance of these modeled data and the real data, shown again in Figure 4b (after deconvolution). In particular, the synthetic data show only weak evidence of the reflections and diffractions associated with the direct arrivals on the real data. There is also evidence on the real data of geometric amplitude focusing effects (at the top and bottom of the high-velocity layer); these amplitude effects are not predicted by the waveform modeling. The challenge for waveform inversion is to produce an improved velocity model that can be used to predict waveform data that look much more like the real, observed data and to demonstrate that the new velocity model is a better representation of the true model than that obtained from traveltime tomography.

Waveform inversion preprocessing

Having chosen to use an acoustic modeling/inversion approach, we must ensure that data phases which do not conform to the physics of the forward model (in this case, the shear waves) are eliminated, or at least suppressed.

Deconvolution.—Deconvolution is not strictly necessary in seismic waveform inversion. The source wavelet must be known, or estimated, but the detailed shape of this wavelet should not play any role in the success (or otherwise) of the method. Nevertheless, I applied deconvolution to these data

before proceeding. This was carried out because the source wavelet can potentially change with angle as a result of the source and receiver characteristics and because late time arrivals will be muted at a later stage (see below). Before windowing the data, deconvolution is used to reposition some of the reverberating signal to earlier times. I followed a simple, deterministic deconvolution process [described earlier by Pratt and Goulty (1991) and later used in a modified form by Rector et al. (1995)]. The process is adapted from standard vertical seismic profile (VSP) processing: a Wiener filter is computed offset by offset (or, equivalently, angle by angle) that converts the estimated direct wavefield into a minimum-phase equivalent wavelet. The result of this deconvolution process is depicted in Figure 4b: the compression of the wavelet is evident, although there is still some strong remnant reverberation left in the signal. In comparison with the original data (Figure 2b), it is now much easier to visually separate and follow individual phases. Because of the local nature of the deconvolution operator, any angle-dependent variation in the source signature is at least partially deconvolved.

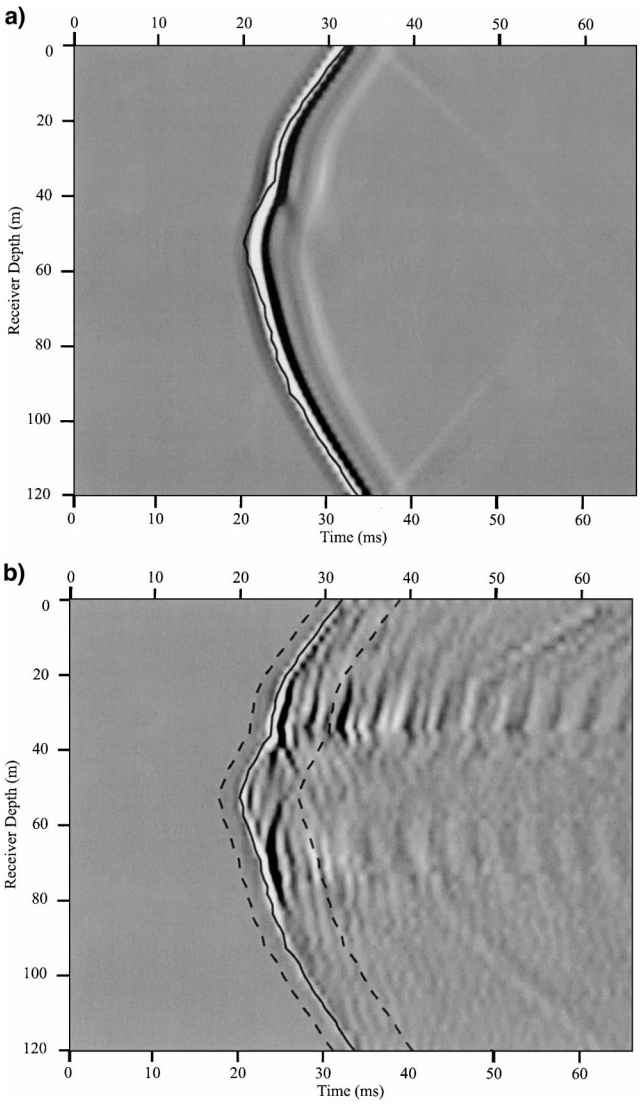

FIG. 4. (a) Synthetic common shot gather 22, obtained by full wave equation, viscoacoustic, finite-difference forward modeling in the traveltime tomography result (see Figure 3b). A quality factor of $Q = 100$ was used to generate these data. (b) Common shot gather 22 from the model, after the application of deterministic deconvolution. This figure may be compared with Figure 2b, which shows the same data before deconvolution. Dashed lines show the extent of the data window used in the subsequent waveform inversion.

Data windowing.—Following the deconvolution process, the data were then muted to remove all energy except that falling within a window starting 2.5 ms before the first break and ending 7.5 ms following the first break. Cosine tapers 1 ms long were used at each end of the time window. The extent of the window is indicated in Figure 4b. The window length was selected to be as short as possible (to enhance the S/N ratio and eliminate *S*-waves) but still to include the visible diffractions and reflections associated with the direct arrival.

The time windowing of the direct arrivals is a crucial step in the inversion process. The objective is to mute any energy arriving later than a few cycles following the direct arrivals [as, for example, in Pratt and Worthington (1988)]. This technique is somewhat ad hoc; it is primarily required to precondition the data in order to force the inversion to fit the direct arrivals, which contain the critical information on the low and intermediate wavenumbers in the model. At a later stage, the window can always be increased in length to include more of the data; this was not considered beneficial or necessary in this example. A further discussion of the advantages and disadvantages of the windowing operation is given at the end of the accompanying paper.

Source signature estimation.—In frequency-domain forward modeling we can account for the source signature explicitly by assuming the source terms in the modeling are multiplied by an unknown, complex-value scalar, s. This leads to the following modification to equation (1), the original forward-modeling equation:

$$\mathbf{Su} = s\mathbf{f}. \tag{15}$$

The source signature estimation problem is to find s for each frequency in the survey (and, if necessary, for each distinct source in the survey).

Because equation (15) is linear with respect to changes in the source signature, it is trivial to obtain an exact solution for the source signature by minimizing the misfit function (Song and Williamson, 1995). Using the same definition of the misfit function,

$$E = \tfrac{1}{2}\delta\mathbf{d}^t\delta\mathbf{d}^* \tag{16}$$

[equation (3) again], it is not difficult to show that the minimum misfit is found when

$$s = \frac{\mathbf{u}^t\mathbf{d}^*}{\mathbf{u}^t\mathbf{u}^*}. \tag{17}$$

This is a complex-valued expression that yields both the amplitude and phase of the desired source signature at the given frequency. To evaluate this expression, we need only compute one forward model (per source and per frequency). In contrast to the algorithm suggested by Zhou et al. (1997), this algorithm

is independent of the starting model for the source wavelet, and it converges in one iteration.

In deriving the solution above, I implicitly assumed that the source signature was the only unknown, i.e., I assumed that the velocity model was correct. In practice, for the crosshole problem a very good representation of the source signature can be obtained when the velocity model is only approximate. To verify this, I show in Figure 5 the minimum-phase desired output wavelet used in the deconvolution of the data (see previous section), together with the source signature estimated from the deconvolved data, using the traveltime velocity model of Figure 3b. The source signature was estimated by a straightforward application of equation (17) to the deconvolved data, frequency by frequency, followed by inverse Fourier transformation of the results. It was assumed that all 51 sources in the survey could be modeled using the same source signature.

Some differences between the two wavelets in Figure 5 are observed: most obviously, the estimated wavelet is more advanced in phase with respect to the desired output. This, to some degree, is expected because of the 2-D nature of the simulation, which generates a phase delay or wake (Morse and Feschbach, 1953). Thus, to obtain a desired seismic arrival pulse, a phase advance must be applied to the source signature. Note that the desired output is not a perfect representation of the output of the result of the deconvolution process. Because the original, undeconvolved wavelet was clearly not minimum phase, the optimum Wiener filter was not perfectly causal; the result of the deconvolution also exhibits a slight phase advance. This signature can be updated during the waveform inversion process as the velocity model improves in resolution—I used this strategy for the following inversion results.

Modeling amplitude variations.—The comparison of the deconvolved data and the current synthetic data (from the traveltime velocity model) in Figure 4 reveals significant amplitude discrepancies, in spite of the fact that the source amplitudes for the synthetic data were obtained directly from the real data. The real data exhibit geometrical focusing effects and a strong loss in amplitude with increasing source-receiver offset; neither of these effects is observed on the synthetic data. The geometric effects are absent because the velocity image does not (yet) contain sufficient variation to predict these additional effects. There are several good reasons for the remaining discrepancy in the amplitude behavior.

1) The real amplitudes are affected by transmission from the water to the epoxy resin and back into the water at the source and receiver edges of the model. These likely lead to amplitudes that fall off with increasing angle (and, as a result, offset). These interfaces are not included in the synthetic model.
2) The real amplitudes are affected by mode conversion to *S*-waves; the synthetic amplitudes are generated within the acoustic approximation.
3) The epoxy resins are likely to be highly attenuating (Leggett et al., 1993). The synthetic data in Figure 4b were generated using a constant Q value of 100. This may be too high.
4) The real amplitudes experience 3-D geometrical spreading. The synthetic amplitudes only spread in 2-D and thus have a smaller amount of geometrical spreading.

The last of these discrepancies can be formally accounted for by using 2.5-D forward modeling (Song and Williamson, 1995), but this is expensive and the remaining discrepancies may be equally serious. An approximate amplitude correction (from three dimensions to two dimensions) can be applied by the use of a $\sqrt{t}$ filter, but this fails in areas of strong velocity gradients or when phases travel with different velocities (i.e., when both *P*- and *S*-waves are present).

The amplitude discrepancy between the real and synthetic data is further illustrated in Figure 6, in which the rms amplitude values for a set of constant offset (constant angle) bins have been computed and are depicted as a function of the offset. Clearly, synthetic data generated with a value of $Q = 100$ fail to match the observed decay in amplitude with distance offset. We can force the synthetic data to match the observed

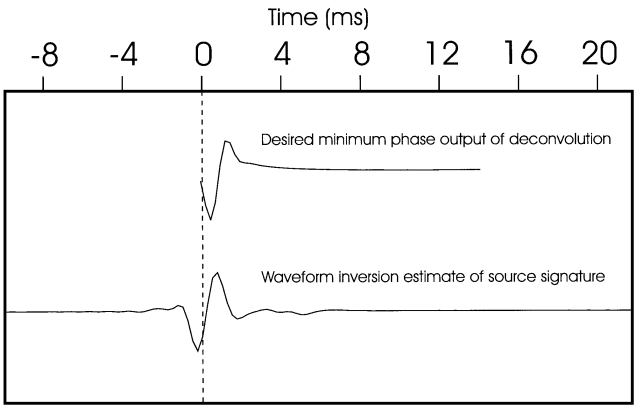

FIG. 5. (Top) Desired minimum-phase output wavelet used in the deconvolution of the data. (Bottom) Source signature extracted from the deconvolved data by waveform inversion, using the traveltime inversion result from Figure 3b as a velocity model.

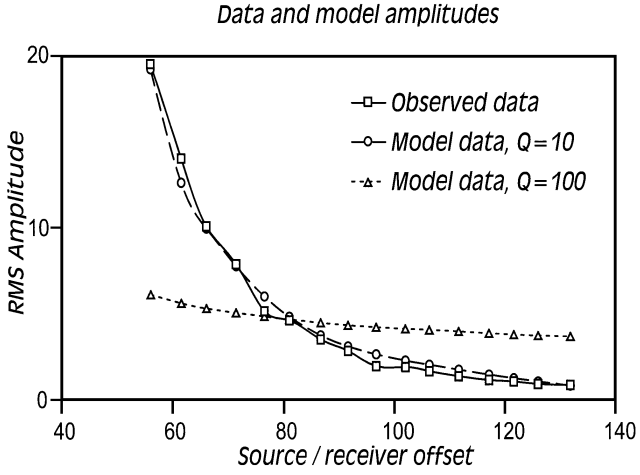

FIG. 6. The amplitude versus source-receiver offset relationship of the direct arrival from the observed data (solid curve), compared with the same relationship for the synthetic data, using a value of $Q = 100$ (dotted curve) and $Q = 10$ (dashed curve). Offsets represent the total (oblique) distance between source and receiver in the model.

amplitude decay by reducing the quality factor, Q (i.e., increasing the intrinsic attenuation). The amplitude results of using a value of $Q = 10$ in the forward modeling are also depicted in Figure 6; this leads to a much improved fit in the amplitude behavior in the data. By matching the amplitudes in this fashion, we ensure that no offset range dominates the data residuals during the inversion, and we avoid the need for amplitude normalization (advocated by Zhou et al., 1995). However, as a result of the additional amplitude effects that have not been included, we cannot claim that $Q = 10$ is better than an approximate representation of the true attenuation properties of the scale model [note, however, that Leggett et al. (1993) found similar Q values from the same data].

Waveform inversion

We now have (1) a starting model (from the traveltime tomography, with a quality factor, $Q = 10$, chosen in the manner described above), (2) a preprocessed data set (deconvolved and windowed), and (3) a reasonable estimate of the source signature for the survey. These form the input data required to apply the iterative waveform inversion formula, equation (4), to the physical scale model data. The following strategy was used to invert the data.

First, data frequencies were selected and Fourier transformed. The frequencies to be used in the inversion were grouped into five overlapping groups of increasing frequency content. Each group contained five frequencies, according to the schedule given in Table 1. The frequency interval 31.25 Hz was chosen to oversample the data within the 10-ms time window used. The total frequency bandwidth, 125–750 Hz, was chosen to match the observed bandwidth of the data. Complex-valued spectral components at each of these frequency values were obtained by direct application of the Discrete Fourier transform to the deconvolved and windowed observed data.

For each frequency group, the gradient was computed using all five frequencies simultaneously. A set of five iterations of the inverse method [equation (4)] was carried out for each group, proceeding through all groups. By the end of the inversion, 25 iterations had been carried out. The strategy of moving up through the frequency spectrum helps to mitigate the nonlinearity of the problem: for lower frequencies, the method is more tolerant of velocity errors because these are less likely to lead to errors of more than a half-cycle in the waveforms. As the inversion proceeds, the velocity model improves and we move to progressively higher frequencies.

Computation times.—For the $75 \times 160 = 12\,000$ nodes used in this example, the forward modeling required less than 8 s (CPU time) per frequency, for all 51 sources, on a DEC Alpha 500 5/500 (915 Specfp rating). The 77 frequencies required to synthesize the time-domain wavefield (see Figure 4) were computed in approximately 10 minutes (again, for all 51 sources). Storage requirements are also important for frequency-domain methods: the finite-difference matrix and its LU decomposition required 11 Mbytes of RAM for storage (using the nested dissection ordering scheme of Liu and George (1981). Larger models would require significantly more RAM (see the accompanying paper).

Table 1. Schedule of frequencies used in the inversion of the data to produce the results shown in Figure 7.

Group	Frequencies (Hz)					Iterations
1	125.00	156.25	187.50	218.75	250.00	5
2	250.00	281.25	312.50	343.75	375.00	5
3	375.00	406.25	437.50	468.75	500.00	5
4	500.00	531.25	562.50	593.75	625.00	5
5	625.00	656.25	687.50	718.75	750.00	5

For inversions, the computer time is dominated by the forward-modeling runs, with some additional overhead costs associated with updating the models. As described above, three modeling runs are required per iteration. Thus, the total time per iteration can be estimated by multiplying the time for single-frequency forward modeling first by the number of frequencies used in each iteration (five) and then by three. Since each forward model requires 8 s, this yields a figure of 2 minutes per iteration. In actuality, each iteration took slightly longer (140 s). For this rather small model (in terms of wavelengths) the total time required to carry out frequency-domain waveform inversion was comparable to the time required to complete the traveltime inversion. Further comments on the computation time for waveform inversion are made in the accompanying paper with reference to a real data set, with a conclusion that frequency-domain waveform inversion is one to two orders of magnitude faster than an equivalent time-domain waveform inversion scheme.

Results.—Two wavefield inversion results (velocity images) are depicted in Figure 7. For reference, the true geometry of the epoxy resin model is given in Figure 7a. Figure 7b depicts the result of inverting the data, using the original attenuation value of $Q = 100$. Figure 7c depicts the result obtained when an attenuation value of $Q = 10$ was used in the inversion. Both images show a reasonable representation of the true model geometry and velocities, and both images represent a significant gain in resolution when compared with the starting model for the waveform inversion, the traveltime tomogram (see Figure 3b). On both images each of the individual layers can be clearly discerned and followed from one edge of the model to the other, including the thin, low-velocity layer containing the simulated channel feature. The channel feature itself is unambiguous on the images. These waveform inversion images are consistent with the expectation that the resolution should be comparable with the dominant wavelength (Wu and Töksoz, 1987). This is in contrast to the traveltime inversions shown earlier, for which the width of the first Fresnel zone controls the resolution of the result (Williamson, 1991). Both the wavelength and the width of the first Fresnel zones are shown on the figures.

The first of the two images, Figure 7b, is marred by the presence of a strong, X-shaped artifact that crosses the image and obscures, in particular, the reconstruction of the small, vertical fault on the lowest, dipping interface. Such artifacts are often obtained in cross-borehole tomographic inversions, especially when waveform inversion is attempted. In this case the artifact appears to be associated with the large level of data residuals obtained for the largest offsets (see Figure 6). As the iteration proceeds, these residuals, which have large noise levels because of the poor S/N ratios at large offsets, apparently exert a disproportionate influence on the inversion. This conjecture seems to be confirmed by the spatial association of the artifact with the largest offset data.

In an attempt to remove the X-shaped artifact observed on Figure 7b, I re-computed the inversion using exactly the same strategy but with a value of $Q = 10$ in the model. By doing this we match the amplitude versus offset behavior of the observed data (see Figure 6), yielding the second, improved image in Figure 7c. The artifact appears to have been successfully removed by this strategy, although there is a still a vestige of the problem at the bottom right of the image. Most importantly, the location and geometry of the small, subtle fault on the dipping interface is now correctly imaged.

In addition to the large, X-shaped artifact, the original image in Figure 7b showed some distortion near each source and receiver location. Source/receiver location artifacts are also common on tomographic reconstructions. These arise because of singularities in the Green's functions associated with each source and receiver location, which affect the gradient calculation disproportionately, and because of the poor coverage of the target near the source and receiver locations. These artifacts have also been limited through the use of matched amplitudes in the inversion.

Verification of final results

An important verification that should be provided along with any inversion result is a physical comparison of the observed and predicted data, preferably together with a subtraction to show the final residual level. I show such a result in Figure 8 but for only one frequency (187.50 Hz). By using only a single frequency, I am able to depict the data (actually, the real parts of the data) for all 2601 source-receiver pairs. The figure caption describes the manner in which these displays are organized; there is a strong analogy between this manner of displaying the data and the interference data recorded in optical holography.

The three rows in Figure 8 represent, respectively, results for the starting model (Figure 3b), the initial ($Q = 100$) inversion result (Figure 7a), and the final inversion result (Figure 7b). The observed data are common to each row, and the two columns in the figure depict the predicted data and the data residuals for each result. At first sight the three rows appear very similar. On closer examination we can observe subtle phase variations on the observed data that are not present on the predicted data from the starting model (in particular, the red region at the center of the panels). These variations are recreated in the synthetic data during the waveform inversion procedure. Note that these subtle features on the frequency domain data are obtained from the direct-arrival waveforms; any muting of the direct arrival [as advocated, for example, by Zhou et al. (1995)] would have changed the character of these displays significantly. Finally, note that the high offset data regions, at the upper right and lower left of these panels, are clearly best matched by the final ($Q = 10$) inversion results.

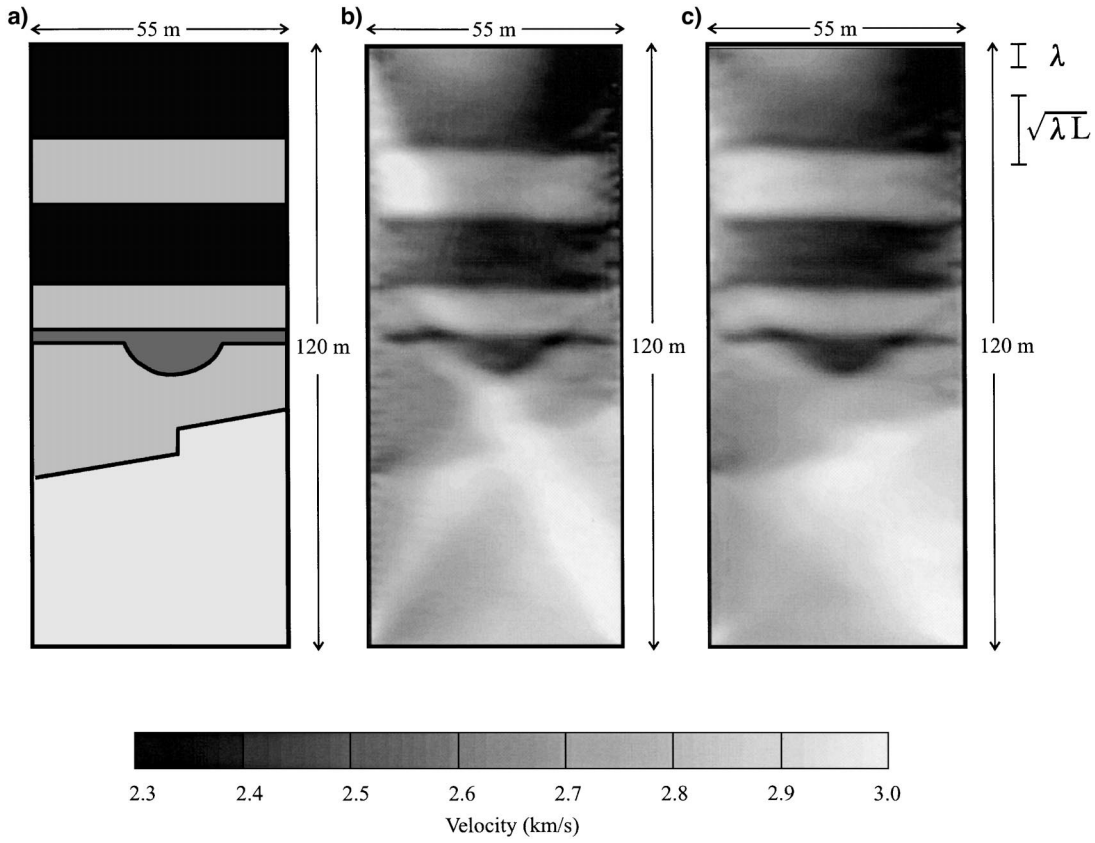

FIG. 7. (a) Geometry and nominal velocities of the original epoxy resin model. (b) Initial wavefield inversion result, obtained using a value of $Q = 100$ throughout the processing and inversion of the data. (c) Final wavefield inversion result, obtained by using a value of $Q = 10$. As in Figure 3, the bars next to the figure depict the dominant wavelength and the approximate width of the first Fresnel zone.

In Figure 9a I again depict the observed data, the predicted data for the starting (traveltime) model, and the predicted data for the final result. In this case, however, the real valued parts of the data are shown as curves for only one common source gather. The differences between the curves are very subtle; there is only a slight phase mismatch observable on the curves. Nevertheless, the more accurate fit of the final model is clear. On Figure 9b I show the velocity variation with model depth from a single, constant lateral position in the model for the true velocities (in this case recovered from the straight across traveltimes), the traveltime result, and the final waveform inversion result. The traveltime result fails to delineate layer boundaries accurately or to respond at all to the thin, low-velocity layer at 62 m depth. In contrast, the waveform inversion result is effective in representing all interfaces in the model and clearly indicates the presence and thickness of the thin layer.

As a final test of the validity of the waveform inversion model, I computed synthetic, time-domain, finite-difference seismograms using, not the acoustic modeling software used in the inversion, but a full viscoelastic code (Štekl and Pratt, 1997). This result is displayed along with the corresponding common source gather from the observed data in Figure 10. The real data and the synthetic data match each other in some detail, especially within the data window used in the inversion. The synthetic data are cleaner than the observed data. This is

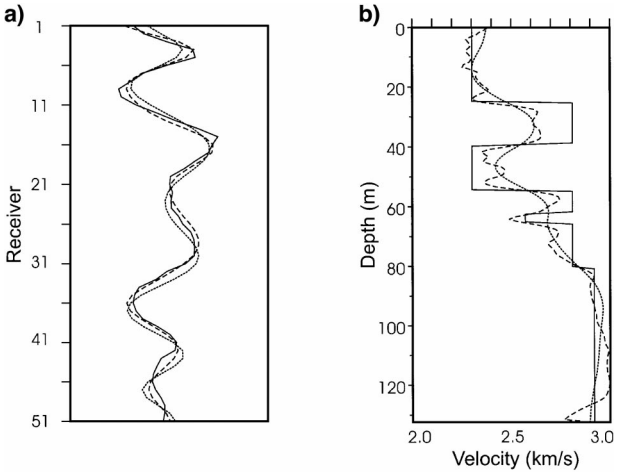

FIG. 9. (a) Frequency-domain data (real values at 187.5 Hz) from shot gather 22, showing the real data (solid line), the modeled data in the tomographic velocity model (dotted line), and the modeled data in the final waveform inversion velocity model (dashed line). (b) Velocities from the center of the model, showing the nominal velocities of the epoxy resins (solid line), the results from curved-ray tomography (dotted line), and the results from the final wavefield inversion model (dashed line).

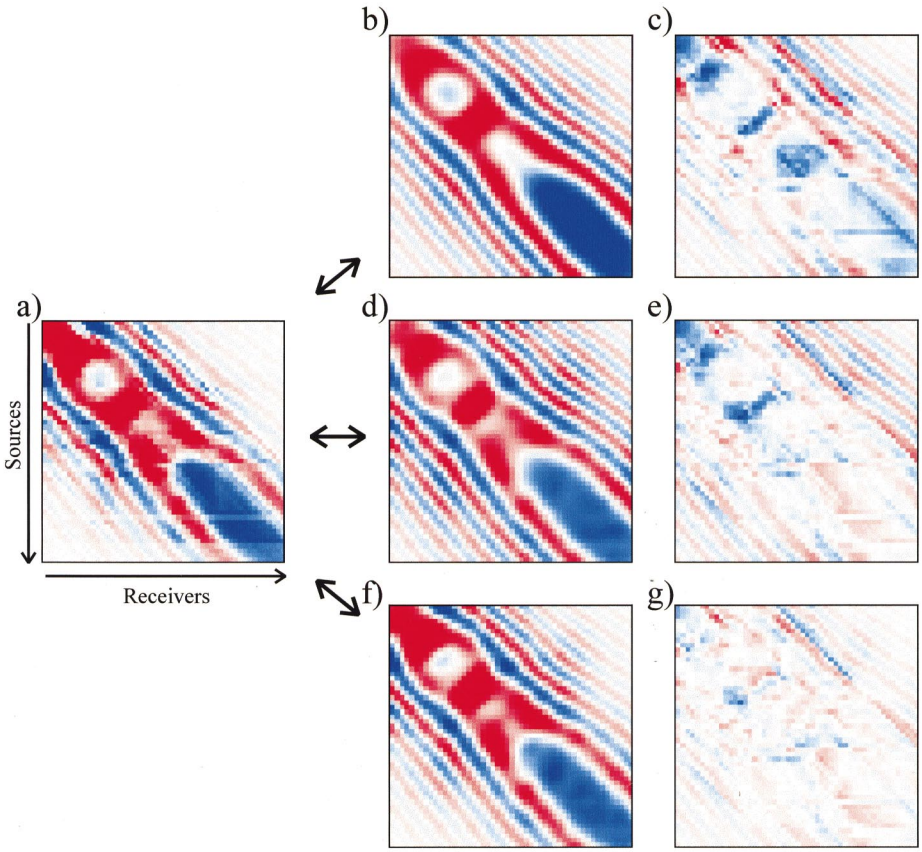

FIG. 8. Frequency-domain data. Each panel represents the real part of the Fourier components of the data at 187.5 Hz, with source number increasing from top to bottom and receiver number increasing from left to right. (a) Observed data. (b), (c) Modeled data and residuals for the starting model (i.e., the tomogram in Figure 3b). (d), (e) Modeled data and residuals for the initial wavefield inversion result (Figure 7a). (f), (g) Modeled data and residuals for the final wavefield inversion result (Figure 7b).

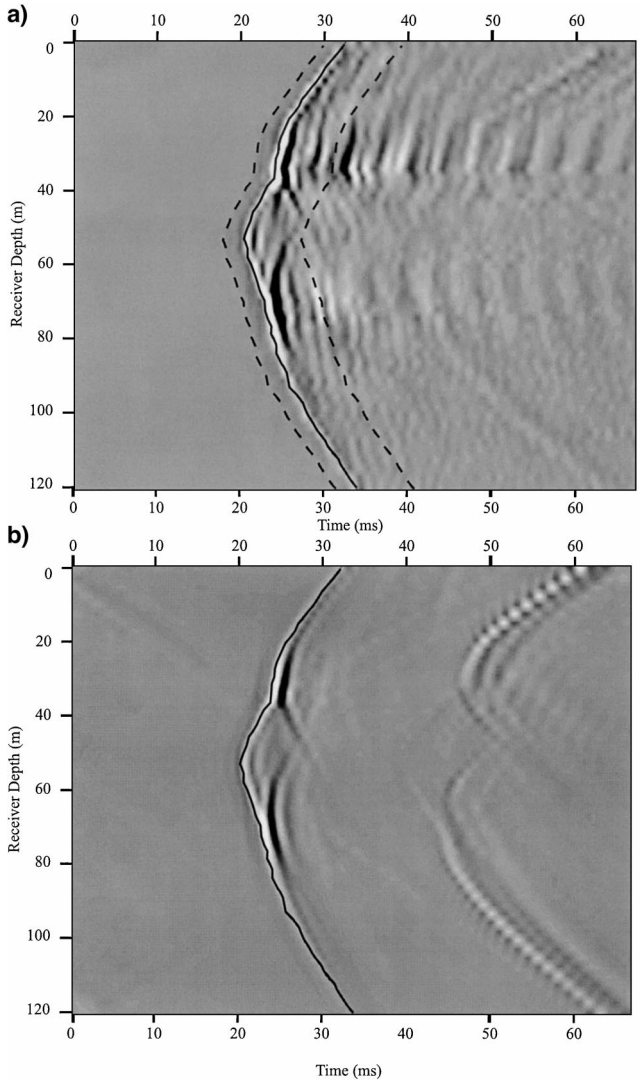

FIG. 10. Time-domain data, showing (a) the deconvolved real data (Figure 4b) and (b) synthetic, viscoelastic data, modeled in the final wavefield inversion result. In (a), the dashed lines show the extent of the data window used for the wavefield inversion.

mainly because of the imperfect deconvolution of the real data, which leaves some reverberatory energy present—especially in areas of high amplitude. The reflected and diffracted modes in the real data within the data window have been predicted correctly in the waveform model, as have the geometric amplitude effects above and below the high-velocity layer. The weak, low-velocity phase at late time on the real data correlates well in time and amplitude with the direct shear arrival on the synthetic viscoelastic data.

Finally, in Figure 11 I show three velocity images as shaded-relief, color-contoured maps. The three images correspond to the true velocity model, the traveltime tomography result, and the final wavefield inversion result. This kind of display is well suited for detecting and interpreting discontinuities. The traveltime tomogram allows only a general interpretation of the target region to be inferred, whereas the final waveform inversion result has clearly resolved all geometric features in the true model, including the small fault on the dipping layer in the bottom half of the model. The shaded-relief display, however, slightly overexaggerates the importance of the artifacts associated with each source and receiver location (see above).

CONCLUSIONS

I have described the theory and application of a frequency-space seismic waveform inversion process. In spite of the attraction of an intuitive, time-domain method for waveform inversion, the frequency-space domain appears to be better suited for tomographic inversion purposes. There are significant computational advantages, both in the forward modeling and in the inverse problem: modeling the seismic response in the frequency domain is fast for multiple source problems, and inverting the seismic data in the frequency domain allows us to limit the number of forward models to a reduced number of frequency components. Additional benefits include being able to progress through the data frequencies from the lowest to the highest (an effective strategy for mitigating nonlinearities) and being able to easily incorporate any desired viscous attenuation and dispersion laws (an effective strategy for matching the amplitudes of the observed data).

This paper began with a restatement of the basic theory of seismic waveform inversion, cast into the frequency-space domain and making use of the implicit matrix formulation for discrete boundary value problems. This allowed me to restate some of the classic advances of Lailly (1983), Tarantola (1984), and others using a simple matrix formalism. I then applied the frequency-domain wavefield inversion method to tomographic data from a physical scale model. The method yielded an image on which the individual layers and structures in the model could be unambiguously interpreted. The results are a clear improvement when compared with the starting model (i.e., the traveltime tomography result). The results were further verified by a detailed comparison of the data with the modeled data in the final inversion result, in both the time and frequency domains. Much of the high-resolution information about the model is contained in subtle and difficult waveform effects in the data, arising from the superposition of many different coincident phases.

To achieve the successful inversion results we showed in Figures 7 and 11, we made a number of preprocessing decisions. The most critical of these was to window the original data in the time domain before Fourier transformation and subsequent inversion. This effectively forces the inversion to fit the direct arrivals, which contain the crucial information on the low and medium wavenumbers in the model. It was also found that the use of a nonzero attenuation factor in the model to simulate the amplitude behavior of the observed data effectively suppressed some of the severe imaging artifacts often encountered in waveform inversion of tomographic data.

REFERENCES

Berkhout, A. J., 1982, Seismic migration: Imaging of acoustic energy by wavefield extrapolation A: Theoretical aspects, 2nd Ed.: Elsevier Science Publ. Co., Inc.

Chavent, G., and Jacewitz, C. A., 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490.

Gauthier, O., Virieux, J., and Tarantola, A., 1986, Two-dimensional non-linear inversion of seismic waveforms: Numerical results: Geophysics, **51**, 1387–1403.

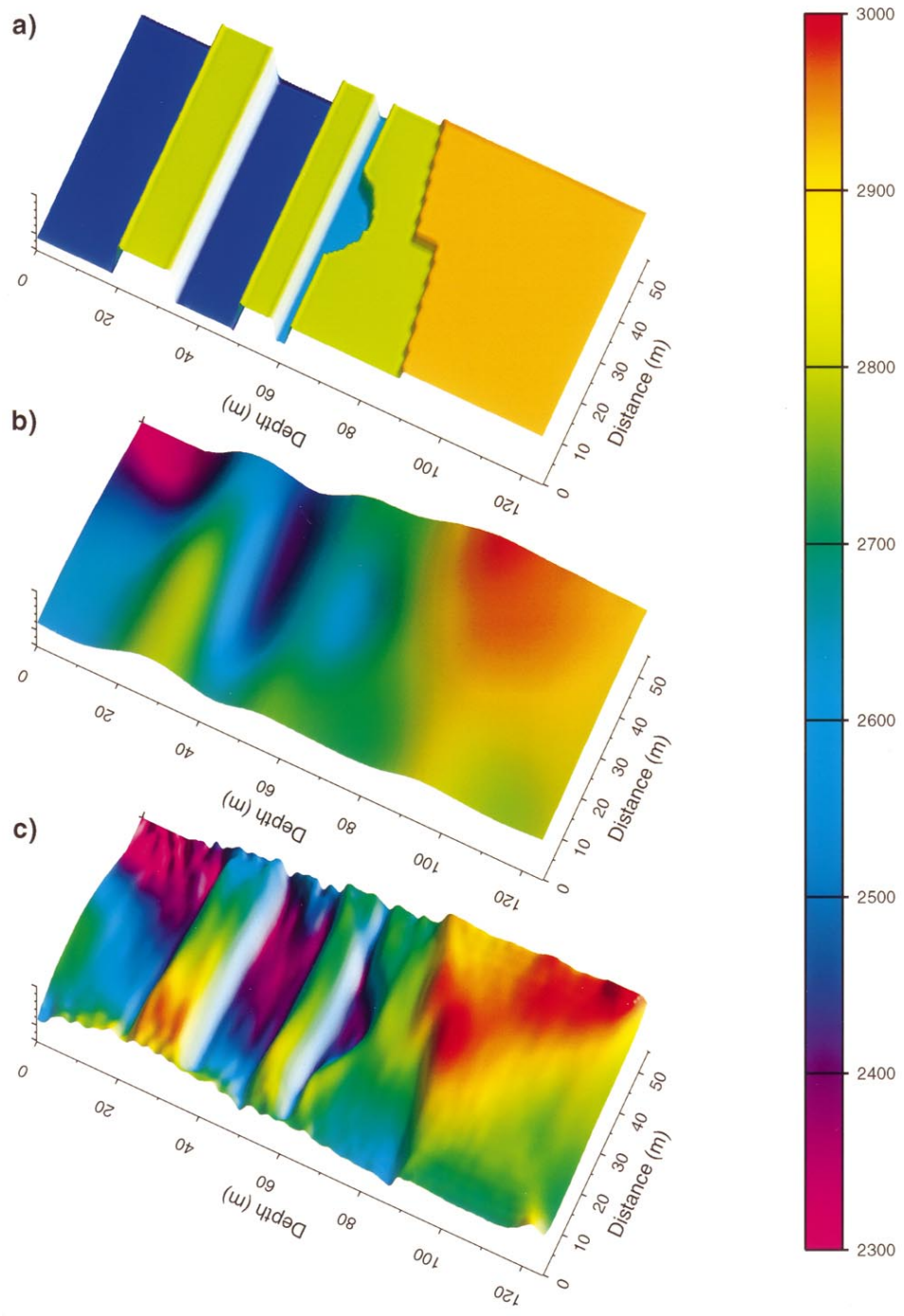

FIG. 11. Shaded-relief displays showing (a) the true velocity model, (b) the traveltime tomography result, and (c) the final wavefield inversion result.

Geller, R. J., and Hara, T., 1993, Two efficient algorithms for iterative linearized inversion of seismic waveform data: Geophys. J. Internat., **115**, 699–710.

Jo, C. H., Shin, C. S., and Suh, J. H., 1996, Design of an optimal 9 point finite difference frequency-space acoustic wave equation scheme for inversion and modeling: Geophysics, **61**, 529–537.

Kolb, P., Collino, F., and Lailly, P., 1986, Pre-stack inversion of a 1-D medium: Proc. IEEE, **74**, 498–508.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations, *in* Bednar, J. B., Redner, R., Robinson, E., and Weglein, A., Eds., Conference on Inverse Scattering: Theory and Application, Soc. Industr. Appl. Math.,

Leggett, M., Goulty, N. R., and Kragh, J. E., 1993, Study of traveltime and amplitude time-lapse tomography using physical model data: Geophys. Prosp., **41**, no. 5, 599–620.

Liu, J. W., and George, A., 1981, Computer solution of large sparse positive definite systems: Prentice-Hall, Inc.

Marfurt, K. J., 1984, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave-equations: Geophysics, **49**, 533–549.

Marfurt, K. J., and Shin, C. S., 1989, The future of iterative modeling of geophysical exploration, *in* Eisner, E., Ed., Supercomputers in seismic exploration: Pergamon Press, 203–228.
Mora, P. R., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228.
Morse, P. M., and Feschbach, H., 1953, Methods of theoretical physics: McGraw-Hill.
Oristaglio, M. L., and Worthington, M. H., 1980, Inversion of surface and borehole electromagnetic data for two-dimensional electrical conductivity models: Geophys. Prosp., **28**, no. 4, 633–657.
Pratt, R. G., 1990a, Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging (short note): Geophysics, **55**, 626–632.
———1990b, Inverse theory applied to multi-source cross-hole tomography **II**: Elastic wave-equation method: Geophys. Prosp., **38**, 311–330.
Pratt, R. G., and Shipp, R. M., 1999, Seismic waveform inversion in the frequency domain II—Fault delineation in sediments using crosshole data: Geophysics, **64**, 901–913.
Pratt, R. G., and Goulty, N. R., 1991, Combining wave-equation imaging with traveltime tomography to form high-resolution images from crosshole data: Geophysics, **56**, 208–224.
Pratt, R. G., Li Quan, Dyer, B. C., Goulty, N. R., and Worthington, M. H., 1991, Algorithms for EOR imaging: An experiment with scale model data: Geoexploration, **28**, 193–220.
Pratt, R. G., McGaughey, W. G., and Chapman, C. H., 1993, Anisotropic velocity tomography: A case study in a near-surface rock mass: Geophysics, **58**, 1748–1763.
Pratt, R. G., Shin, C. S., and Hicks, G. J., 1998, Gauss–Newton and full Newton methods in frequency-space seismic waveform inversion: Geophys. J. Internat., **133**, 341–362.
Pratt, R. G., Shipp, R. M., Song, Z. M., and Williamson, P. R., 1995, Fault delineation by wavefield inversion of cross-borehole seismic data: 57th Mtg., Eur. Assoc. Expl. Geophys., Extended Abstracts 95.
Pratt, R. G., and Worthington, M. H., 1988, The application of diffraction tomography to cross-hole seismic data: Geophysics, **53**, 1284–1294.
———1990, Inverse theory applied to multi-source cross-hole tomography I: Acoustic wave-equation method: Geophys. Prosp., **38**, 287–310.
Press, W. H., Teukolsky, S. A., Vettering, W. T., and Flannery, B. P., 1992, Numerical recipes in Fortran: The art of scientific computing, 2nd Ed.: Cambridge Univ. Press.
Rector, J. W., Harris, J. M., Lazaratos, S. K., and Van Schaack, M., 1995, High-resolution crosswell imaging of a West Texas carbonate reservoir 3—Wavefield separation of reflections: Geophysics, **60**, 692–701.
Reiter, D. L., and Rodi, W., 1996, Nonlinear waveform tomography applied to cross-hole seismic data: Geophysics, **61**, 902–913.
Rodi, W. L., 1976, A technique for improving the accuracy of finite element solutions for magnetotelluric data: Geophys. J. Roy. Astr. Soc., **44**, 483–506.
Rowbotham, P. S., and Goulty, N. R., 1995, Quantitative evaluation of crosshole seismic reflection images using physical model data: Geophys. Prosp., **43**, no. 4, 529–540.
Sharp, R. J., Peacock, J. H., and Goulty, N. R., 1985, Ultrasonic modelling system: Presented at the 47th Mtg., Eur. Assoc. Expl. Geophys.,
Shin, C. S., 1988, Nonlinear elastic wave inversion by blocky parameterization: Ph.D. thesis, Univ. of Tulsa.
Song, Z.-M., and Williamson, P. R., 1995, Frequency-domain acoustic-wave modeling and inversion of crosshole data I—2.5-D modeling method: Geophysics, **60**, 784–795.
Song, Z.-M., Williamson, P. R., and Pratt, R. G., 1995, Frequency-domain acoustic-wave modeling and inversion of crosshole data II—Inversion method, synthetic experiments and real-data results: Geophysics, **60**, 796–809.
Štekl, I., and Pratt, R. G., 1998, Accurate visco-elastic modeling by frequency-domain finite differences using rotated operators: Geophysics, **63**, 1779–1794.
Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.
———1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903.
———1987, Inverse problem theory: Methods for data fitting and parameter estimation: Elsevier Science Publ. Co., Inc.
Van Schaack, M., Harris, J. M., Rector, J. W., and Lazaratos, S. K., 1995, High-resolution crosswell imaging of a West Texas carbonate reservoir 2—Wavefield modeling and analysis: Geophysics, **60**, 682–691.
Williamson, P. R., 1991, A guide to the limits of resolution imposed by scattering in ray tomography: Geophysics, **56**, 202–207.
Williamson, P. R., and Worthington, M. H., 1993, Resolution limits in ray tomography due to wave behavior: Numerical experiments: Geophysics, **58**, 727–735.
Wu, R. S., and Töksoz, M. N., 1987, Diffraction tomography and multisource holography applied to seismic imaging: Geophysics, **52**, 11–25.
Zhou, C., Cai, W., Luo, Y., Schuster, G. T., and Hassanzadeh, S., 1995, Acoustic wave-equation traveltime and waveform inversion of crosshole seismic data: Geophysics, **60**, 765–773.
Zhou, C., Schuster, G. T., Hassanzadeh, S., and Harris, J. M., 1997, Elastic wave equation traveltime and waveform inversion of crosswell data: Geophysics, **62**, 853–869.