GEOPHYSICS, VOL. 74, NO. 6 (NOVEMBER-DECEMBER 2009); P. WCC47–WCC58, 14 FIGS., 1 TABLE.
10.1190/1.3250203

Application of the finite-difference contrast-source inversion algorithm to seismic full-waveform data

Aria Abubakar[1], Wenyi Hu[1], Tarek M. Habashy[1], and Peter M. van den Berg[2]

ABSTRACT

We have applied the finite-difference contrast-source inversion (FDCSI) method to seismic full-waveform inversion problems. The FDCSI method is an iterative nonlinear inversion algorithm. However, unlike the nonlinear conjugate gradient method and the Gauss-Newton method, FDCSI does not solve any full forward problem explicitly in each iterative step of the inversion process. This feature makes the method very efficient in solving large-scale computational problems. It is shown that FDCSI, with a significant lower computation cost, can produce inversion results comparable in quality to those produced by the Gauss-Newton method and better than those produced by the nonlinear conjugate gradient method. Another attractive feature of the FDCSI method is that it is capable of employing an inhomogeneous background medium without any extra or special effort. This feature is useful when dealing with time-lapse inversion problems where the objective is to reconstruct changes between the baseline and the monitor model. By using the baseline model as the background medium in crosswell seismic monitoring problems, high quality time-lapse inversion results are obtained.

INTRODUCTION

Full-waveform seismic inversion is a powerful tool in reconstructing complex geological structures. This approach can be performed either in the time domain (Tarantola, 1986; Mora, 1987; Vigh and Starr, 2008) or in the frequency domain (Pratt and Worthington, 1990). Here we address only the frequency-domain case. Most inversion methods reported in the literature are based on gradient approaches, such as the nonlinear conjugate gradient methods (see Pratt and Worthington, 1990; Song et al., 1995; Liao and McMechan, 1996; Pratt, 1999; Shipp and Singh, 2002; Ravaut et al., 2004; Sirgue and Pratt, 2004; Ben-Hadj-Ali et al., 2008; Malinowski and Operto, 2008; Mulder and Plessix, 2008) and the Gauss-Newton methods (see Pratt et al., 1998; Hu et al., 2009). The Gauss-Newton method is preferable because of its faster convergence rate. On the other hand, nonlinear conjugate gradient methods do not require one to solve linear system of equations of the normal equations. Solving these normal equations (and constructing the Jacobian matrix) is the most expensive part of the Gauss-Newton method for problems with a large number of unknown parameters.

In each iteration of the nonlinear conjugate gradient or Gauss-Newton method, we have to solve at least several forward problems for calculating the data misfit, sensitivity matrix, and line search parameter. This computation cost can become quite high without the availability of an efficient forward solver. An alternative nonlinear inversion method is discussed in the literature, where the use of a full forward solver in each inversion iteration is not needed. This approach is called the contrast-source inversion (CSI) method, described in van den Berg and Kleinman (1997). The CSI method is a variant of the so-called source-type integral equation (STIE) approach introduced in Habashy et al. (1994). In each inversion iteration, the unknown contrast source (a quantity given in terms of fields and material parameters) and the unknown contrast function (based on material parameters) are updated by one conjugate gradient (CG) step to minimize the appropriate cost function. This CSI method is based on the integral equation (IE) formulation. In addition, it is applied in the seismic time-lapse inversion problem in Abubakar et al. (2003) and in the acoustic 3D problem (see van Dongen and Wright, 2007). For 2D electromagnetic problems, the time-domain counterpart of the CSI method is studied in Bloemenkamp and van den Berg (2000). Because the integral equation CSI (IECSI) method is based on the IE formulation, this method is very efficient when the Green's function is available in semiclosed form, such as in the case of a homogeneous or a layered background medium. In this IECSI approach, an arbitrary inhomogeneous background medium would require the Green's function to be constructed numerically. This extra cost of computing the Green's function numerically can be very expensive. Abubakar et al. (2008b) extend the CSI method for electromagnetic applications for reconstructing the unknown configuration

---

Manuscript received by the Editor 27 January 2009; revised manuscript received 30 June 2009; published online 3 December 2009; corrected version published online 17 December 2009.

[1] Schlumberger-Doll Research, Cambridge, Massachusetts, U.S.A. E-mail: aabubakar@slb.com; whu@slb.com; habashy@slb.com.
[2] Delft University of Technology, Delft, The Netherlands. E-mail: p.m.vandenberg@tudelft.nl.

© 2009 Society of Exploration Geophysicists. All rights reserved.

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

of inhomogeneous objects immersed in a known inhomogeneous background medium. This is achieved by using the finite-difference (FD) formulation. Similar to IECSI, the finite-difference contrast-source inversion (FDCSI) method alternately updates the unknown contrast source and contrast function to reconstruct the scatterers without requiring the explicit solution of the full forward problem at each iteration step in the inversion process.

We apply FDCSI to full waveform seismic inversion problems. We use the FD frequency-domain (FDFD) method of Hu et al. (2009) incorporated with a perfectly matching layer (PML) absorbing boundary condition. An attractive feature of the CSI algorithm is that the stiffness matrix is dependent only on the background medium, which is invariant throughout the inversion process. Therefore, by introducing the FD operator with the CSI method and at least for 2D configurations where the size of the stiffness matrix is manageable, this stiffness matrix can be inverted using a direct solver such as an LU decomposition method (see Davis and Duff, 1997). Hence, this FD operator needs to be inverted only once and the results can be reused for multiple source positions and in successive iterations of the inversion.

By using the well-known Marmousi model, we show that the FDCSI method produces inversion results that are comparable with those obtained from the Gauss Newton approach (Hu et al., 2009) with less computational effort. We also show that FDCSI inversion results are better than those obtained from the unpreconditioned, nonlinear conjugate gradient method in Hu et al. (2009). Furthermore, because the FDCSI method is readily capable of using an inhomogeneous background medium, it is very attractive for time-lapse data inversion applications. The idea is to use the baseline model (which is inhomogeneous) as the background medium for the FDCSI algorithm. By doing so, the inversion algorithm will always honor the baseline model always and will reconstruct incremental changes only as a function of time. We demonstrate the advantages of such an approach by applying it to a crosswell seismic monitoring.

FORWARD MODELING

Next we present the 2D FDFD forward-modeling algorithm (Dablain, 1986; Pratt, 1990; Harari and Turkel, 1995; Hustedt et al., 2004) used as the forward engine in our inversion method. This simulator models acoustic-wave propagation in an isotropic, lossless, but inhomogeneous medium. The Cartesian coordinates are denoted by x, y, and z. Here we assume that the computational domain is in the xz-plane and there is no variation along the y-direction. We further assume that the mass density ρ (kg/m^3) is a constant value. The governing equation describing acoustic wave propagation in the frequency-domain is given by (Fokkema and van den Berg, 1993)

$$[\nabla^2 + k^2(\mathbf{r})]u(\mathbf{r}) = -Q(\mathbf{r},\mathbf{r}^{\mathrm{S}}), \quad (1)$$

where ∇ denotes differentiation with respect to the spatial position $\mathbf{r} = (x,z)$, $u(\mathbf{r})$ is the pressure field (Pa), and the wavenumber $k(\mathbf{r})$ is given by

$$k(\mathbf{r}) = \omega\sqrt{\rho\kappa(\mathbf{r})}, \quad (2)$$

in which ω is the angular frequency and $\kappa(\mathbf{r})$ is the compressibility (Pa^{-1}). Source term $Q(\mathbf{r},\mathbf{r}^{\mathrm{S}})$ is given by

$$Q(\mathbf{r},\mathbf{r}^{\mathrm{S}}) = -\mathrm{i}\omega\rho q(\mathbf{r},\mathbf{r}^{\mathrm{S}}), \quad (3)$$

where $\mathrm{i}^2 = -1$ and $q(\mathbf{r},\mathbf{r}^{\mathrm{S}})$ is the volume density of the injected volume (s^{-1}), in which $\mathbf{r}^{\mathrm{S}}$ is the source position. As absorbing boundary conditions, we employed the complex stretched coordinate PML (Bérenger, 1994; Chew and Weedon, 1994).

The computation domain is divided into $(P+2P_{\mathrm{PML}})\times(Q+2Q_{\mathrm{PML}})$ uniform grids, where P and Q are the number of grids in the x- and z-directions in the regular domain, whereas P_{PML} and Q_{PML} are the number of grids in the PML domain in the x- and z-directions, respectively. After discretization, equation 1 can be written in the matrix form as

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad (4)$$

where $\mathbf{A}$ is the stiffness matrix, $\mathbf{x}$ is the state variable vector, and $\mathbf{b}$ is the source term. Stiffness matrix $\mathbf{A}$ is asymmetric but very sparse. To solve equation 4, we use a direct solver based on an LU decomposition (see Davis and Duff, 1997) rather than an iterative solver. The main advantage of using a direct solver is that we can obtain all solutions for all source excitations simultaneously by factoring the stiffness matrix only once (Pratt and Worthington, 1990). However, because stiffness matrix $\mathbf{A}$ is frequency dependent, a new matrix needs to be inverted for every new frequency.

INVERSION ALGORITHM

Formulation

We denote the inversion domain as the object domain D and the data domain S as the domain where the sources and the receivers are located. Together, D and S are located within the total domain T, which is also the finite-difference computational domain. We assume we have N^{S} sources and the index of the source is denoted by subscript j.

The total field $u_j(\mathbf{r})$ satisfies the equation

$$\mathcal{H}[u_j] = [\nabla^2 + k^2(\mathbf{r})]u_j(\mathrm{r}) = -Q(\mathbf{r},\mathbf{r}_j^{\mathrm{S}}), \quad \mathbf{r}\in T. \quad (5)$$

We split the total field into its incident and scattered parts, $u_j = u_j^{\mathrm{inc}} + u_j^{\mathrm{sct}}$. The incident field satisfies the equation

$$\mathcal{H}_{\mathrm{b}}[u_j^{\mathrm{inc}}] = [\nabla^2 + k_{\mathrm{b}}^2(\mathbf{r})]u_j^{\mathrm{inc}}(\mathbf{r}) = -Q(\mathbf{r},\mathbf{r}_j^{\mathrm{S}}), \quad \mathbf{r}\in T, \quad (6)$$

where $k_{\mathrm{b}}(\mathbf{r})$ is the spatially varying wavenumber of the background medium. By subtracting equation 6 from equation 5 and using the definition of the scattered field, the scattered field $u_j^{\mathrm{sct}}(\mathbf{r})$ can be shown to satisfy

$$[\nabla^2 + k_{\mathrm{b}}^2(\mathbf{r})]u_j^{\mathrm{sct}}(\mathbf{r}) = -k_{\mathrm{b}}^2(\mathbf{r})w_j(\mathbf{r}), \quad \mathbf{r}\in T, \quad (7)$$

where $w_j(\mathbf{r})$ are the contrast sources, defined as

$$w_j(\mathbf{r}) = \chi(\mathbf{r})u_j(\mathbf{r}), \quad (8)$$

in which the contrast $\chi(\mathbf{r})$ is given by

$$\chi(\mathbf{r}) = \left[\frac{k(\mathbf{r})}{k_{\mathrm{b}}(\mathbf{r})}\right]^2 - 1 = \frac{c^{-2}(\mathbf{r}) - c_{\mathrm{b}}^{-2}(\mathbf{r})}{c_{\mathrm{b}}^{-2}(\mathbf{r})}, \quad \mathbf{r}\in D. \quad (9)$$

Here, $c^{-2}(\mathbf{r}) = \rho\kappa(\mathbf{r})$ is the velocity of the scattering object, and $c_{\mathrm{b}}^{-2}(\mathbf{r}) = \rho\kappa_{\mathrm{b}}(\mathbf{r})$ is the velocity of the background medium.

Equation 7 can be written using operator notation as

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

$$\mathcal{H}_{\mathrm{b}}[u_j^{\mathrm{sct}}(\mathbf{r})] = -k_{\mathrm{b}}^2(\mathbf{r})w_j(\mathbf{r}), \quad \mathbf{r}\in T. \tag{10}$$

The solution of equation 10 can be formally written as

$$u_j^{\mathrm{sct}}(\mathbf{r}) = \mathcal{H}_{\mathrm{b}}^{-1}[-k_{\mathrm{b}}^2(\mathbf{r})w_j(\mathbf{r})] = \mathcal{L}_{\mathrm{b}}[w_j(\mathbf{r})], \quad \mathbf{r}\in T, \tag{11}$$

where operator $\mathcal{L}_{\mathrm{b}}$ is defined as

$$\mathcal{L}_{\mathrm{b}}[\cdot] = \mathcal{H}_{\mathrm{b}}^{-1}[-k_{\mathrm{b}}^2(\mathbf{r})(\cdot)]. \tag{12}$$

Introducing an operator $\mathcal{M}^S$ that selects the field points on the measurement domain S among the field points in the total domain T, the data equation can be written as

$$u_j^{\mathrm{sct}}(\mathbf{r}) = \mathcal{M}^S\{\mathcal{L}_{\mathrm{b}}[w_j(\mathbf{r}')]\}, \quad \mathbf{r}\in S, \mathbf{r}'\in T. \tag{13}$$

By introducing an operator $\mathcal{M}^D$ that selects the fields inside the inversion domain D, we can write the object (or domain) equation as

$$u_j(\mathbf{r}) = u_j^{\mathrm{inc}}(\mathbf{r}) + \mathcal{M}^D\{\mathcal{L}_{\mathrm{b}}[w_j(\mathbf{r}')]\}, \quad \mathbf{r}\in D, \mathbf{r}'\in T. \tag{14}$$

By multiplying both sides of equation 14 by χ and rearranging terms, equation 14 becomes

$$w_j(\mathbf{r}) = \chi(\mathbf{r})u_j^{\mathrm{inc}}(\mathbf{r}) + \chi(\mathbf{r})\mathcal{M}^D\{\mathcal{L}_{\mathrm{b}}[w_j(\mathbf{r}')]\}, \quad \mathbf{r}\in D, \mathbf{r}' \in T. \tag{15}$$

For simplicity, we drop the explicit dependence of the field quantities on $\mathbf{r}$ and $\mathbf{r}'$ in the remainder of this manuscript.

The Gauss-Newton and nonlinear conjugate gradient methods

The most popular approaches to solve this nonlinear inverse problem are the Gauss-Newton and nonlinear conjugate gradient methods (see Gauthier et al., 1986; Mora, 1987; Pratt and Worthington, 1990; Song et al., 1995; Liao and McMechan, 1996; Pratt et al., 1998; Pratt, 1999; Shipp and Singh, 2002; Sirgue and Pratt, 2004; Ben-Hadj-Ali et al., 2008; Hu et al., 2009). In these approaches, the contrast χ is reconstructed iteratively by minimizing the cost function

$$F(\chi) = \frac{\sum_j \|f_j - u_j^{\mathrm{sct}}(\chi)\|_S^2}{\sum_j \|f_j\|_S^2} + \lambda F^{\mathrm{R}}(\chi), \tag{16}$$

where f_j is the scattered measured data and the simulated data u_j^{sct} is given formally by

$$u_j^{\mathrm{sct}}(\chi) = \mathcal{M}^S\{\mathcal{L}_{\mathrm{b}}[(1-\chi\mathcal{M}^D\mathcal{L}_{\mathrm{b}})^{-1}(\chi u_j^{\mathrm{inc}})]\}. \tag{17}$$

Equation 17 is obtained by solving equation 15 for w_j and substituting the result in equation 13. Symbol λ denotes the regularization parameter and $F^{\mathrm{R}}(\chi)$ is the regularization part of the cost function. The L_2-norm $\|f_j\|_S^2$ on the data domain S is defined as

$$\|f_j\|_S^2 = \langle f_j, f_j\rangle_S = \int_S \gamma_j(\mathbf{r}^{\mathrm{R}})f_j(\mathbf{r}^{\mathrm{R}})\overline{f_j(\mathbf{r}^{\mathrm{R}})}\mathrm{d}\mathbf{r}^{\mathrm{R}}, \tag{18}$$

where the overbar denotes the complex conjugate of a quantity and $\gamma_j(\mathbf{r}^{\mathrm{R}})$ is a data weighting, which is a function of source j and receiver r^{R}. This data weighting can include a frequency weighting to deal with multifrequency data and also can include extra weighting based on data-error estimates.

These nonlinear inversion approaches require a solution of the full forward problem $(1-\chi\mathcal{M}^D\mathcal{L}_{\mathrm{b}})^{-1}(\chi u_j^{\mathrm{inc}})$ several times in each inversion iteration. Without a very efficient forward solver, the computation cost of these approaches can be quite high. Note that equation 17 also can be formulated directly in terms of the total field u_j by using equation 1. When employing a direct solver for the forward simulator, the computation complexity of the Gauss-Newton method and the nonlinear conjugate gradient method are given approximately by

$$\begin{aligned}\mathrm{Computations}_{\mathrm{GN}} \approx N^{\mathrm{iter}}\{&O(N^{1.5}) + O[(2N^{\mathrm{S}} + N^{\mathrm{R}})N\log N] \\ &+ O(N^{\mathrm{S}}N^{\mathrm{R}}N) + N_{\mathrm{CGLS}}^{\mathrm{iter}}O(N^2)\}\end{aligned} \tag{19}$$

and

$$\mathrm{Computations}_{\mathrm{CG}} \approx N^{\mathrm{iter}}\{O(N^{1.5}) + O[(2N^{\mathrm{S}} + N^{\mathrm{R}})N\log N]\}, \tag{20}$$

where N^{iter}, N, and N^{R} are the number of inversion iterations, the number of spatial discretization grids, and the number of receivers, respectively. The presence of N^{R} in equations 19 and 20 accounts for the sensitivity-matrix calculation using the adjoint approach. Term $N\log N$ accounts for cost of the back-substitution. Term $O(N^{\mathrm{S}}N^{\mathrm{R}}N)$ in equation 19 accounts for the cost of constructing the Jacobian matrix, whereas the term $N_{\mathrm{CGLS}}^{\mathrm{iter}}O(N^2)$ accounts for the cost of calculating the Gauss-Newton step using the full Hessian matrix inversion, in which $N_{\mathrm{CGLS}}^{\mathrm{iter}}$ denotes the number of iterations of the CG iterative solver. In these estimates, we assumed that both methods use at least one line-search step.

The finite-difference contrast-source inversion method

In this work, we proceed differently so we do not need to solve a full forward problem explicitly in each inversion iteration. Contrast-source inversion (CSI) (see van den Berg and Kleinman, 1997; Abubakar et al., 2003) is one method that employs this concept. Originally, CSI was based on the IE approach. For efficiency, the background medium is a simple homogeneous or layered medium. To handle an inhomogeneous background medium, we use a finite-difference version of CSI as introduced in Abubakar et al. (2008b).

In CSI, instead of eliminating the contrast-source quantities w_j in equations 13 and 15, the inverse scattering problem is treated as an optimization problem to reconstruct both contrast function χ and contrast sources w_j. The cost function to be minimized is the normalized sum of the data and object errors

$$F_n(\chi, w_j) = \frac{\sum_j \|f_j - \mathcal{M}^S\{\mathcal{L}_{\mathrm{b}}[w_j]\}\|_S^2}{\sum_j \|f_j\|_S^2} + \frac{\sum_j \|\chi u_j^{\mathrm{inc}} - w_j + \chi\mathcal{M}^D\{\mathcal{L}_{\mathrm{b}}[w_j]\}\|_D^2}{\sum_j \|\chi_n u_j^{\mathrm{inc}}\|_D^2} \tag{21}$$

or

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

$$F_n(\chi,w_j) = F^S(w_j) + F_n^D(\chi,w_j), \tag{22}$$

where $F^S(w_j)$ represents the data equation error, $F^D(\chi,w_j)$ represents the object or domain equation error. We have utilized $f_j \approx u_j^{\mathrm{sct}}$ to emphasize that the measurement of the scattered field is corrupted unavoidably with noise. When we deal with real data, f_j is built by subtracting the simulated background field from the measured total field. The L_2-norm on the object domain D is defined as

$$\|v_j\|_D^2 = \langle v_j,v_j\rangle_D = \int_D v_j(\mathbf{r})\overline{v_j(\mathbf{r})}\mathrm{d}\mathbf{r}. \tag{23}$$

Normalization factors in the cost function are chosen so we weight errors in the data and object equations equally. For each frequency, finite-difference operator $\mathcal{H}_\mathrm{b}$ needs to be inverted only once because the background medium does not change throughout the inversion process. In the 2D case, our FDFD code (see Hu et al., 2009) employs an LU decomposition technique, hence the cost of applying operator $\mathcal{L}_\mathrm{b}$ on a function is relatively cheap because the LU decomposition is done only once and then stored and used later in each step of the inversion iteration.

We define the data error at the nth iteration to be

$$\rho_{j,n} = f_j - \mathcal{M}^S\{\mathcal{L}_\mathrm{b}[w_{j,n}]\}, \tag{24}$$

and the object error to be

$$r_{j,n} = \chi_n u_j^{\mathrm{inc}} - w_{j,n} + \chi_n \mathcal{M}^D\{\mathcal{L}_\mathrm{b}[w_{j,n}]\}. \tag{25}$$

The CSI method minimizes the objective function of equation 21 by constructing two interlaced sequences $\{w_{j,n}\}$ and $\{\chi_n\}$ via the alternate application of the conjugate gradient (CG) minimization method on each unknown, χ and w_j. As each interlaced sequence is updated, the other is assumed to be constant.

The first step of the algorithm is to update the contrast sources with the formula

$$w_{j,n} = w_{j,n-1} + \alpha_{j,n}^w v_{j,n}, \tag{26}$$

where $\alpha_{j,n}^w$ is an update-step size, and, $v_{j,n}$ represents the Polak-Ribière search directions given by

$$v_{j,0} = 0 \tag{27}$$

and

$$v_{j,n} = g_{j,n}^w + \frac{\sum_k \langle g_{k,n}^w, g_{k,n}^w - g_{k,n-1}^w\rangle_D}{\sum_k \|g_{k,n-1}^w\|_D^2} v_{j,n-1}, n>0, \tag{28}$$

where $g_{j,n}^w$ is the gradient (or Fréchet derivative) of the cost function with respect to w_j evaluated at the $(n-1)$th iteration. The gradient can be shown to be (see Abubakar et al., 2008b)

$$\begin{aligned} g_{j,n}^w = \left.\frac{\partial F(\chi_{n-1} w_j)}{\partial w_j}\right|_{w_j = w_{j,n-1}} &= -\eta^S \mathcal{L}_\mathrm{b}^*\{\mathcal{M}^{S*}[\gamma_j \rho_{j,n-1}]\} \\ &- \eta_n^D(r_{j,n-1} - \mathcal{L}_\mathrm{b}^*\{\mathcal{M}^{D*}[\overline{\chi_{n-1}} r_{j,n-1}]\}). \end{aligned} \tag{29}$$

Here, γ_j appears because of our norm definition in equation 18 and the normalization factors are given by

$$\eta^S = \left(\sum_j \|f_j\|_S^2\right)^{-1} \tag{30}$$

and

$$\eta_n^D = \left(\sum_j \|\chi_{n-1} u_j^{\mathrm{inc}}\|_D^2\right)^{-1}. \tag{31}$$

Operator $\mathcal{L}_\mathrm{b}^*$ represents the adjoint of operator $\mathcal{L}_\mathrm{b}$, and is given by

$$\mathcal{L}_\mathrm{b}^*[\cdot] = -\overline{k_\mathrm{b}^2}(\mathcal{H}_\mathrm{b}^*)^{-1}[\cdot], \tag{32}$$

where

$$\mathcal{H}_\mathrm{b}^*[\cdot] = [\nabla^2 + \overline{k_\mathrm{b}^2}(\mathbf{r})][\cdot]. \tag{33}$$

We employed the same approach to calculate $\mathcal{L}_\mathrm{b}[\cdot]$ for solving equation 32. Adjoint operator $\mathcal{M}^{D*}$ simply takes fields from the inversion domain and maps them into total domain T. Adjoint operator $\mathcal{M}^{S*}$ takes the fields from the measurement domain S, divides them by the area of the FD cell, and maps them into the total domain T. Division by the area of the cell is required by the definition of the adjoint operator (see Abubakar et al., 2008b).

After the search directions $v_{j,n}$ are obtained, the minimizer $\alpha_{j,n}$ of the cost function in equation 21 is found:

$$\alpha_{j,n}^w = \frac{-\sum_j \langle g_{j,n}^w, v_{j,n}\rangle_D}{\eta_S \sum_j \|\mathcal{M}^S\{\mathcal{L}_\mathrm{b}[v_{j,n}]\}\|_S^2 + \eta_n^D \sum_j \|v_{j,n} - \chi_{n-1}\mathcal{M}^D\{\mathcal{L}_\mathrm{b}[v_{j,n}]\}\|_D^2}. \tag{34}$$

This minimizer is obtained by substituting equation 26 in the cost function in equation 21 and minimizing this equation with respect to $\alpha_{j,n}^w$.

Once the contrast sources w_j have been updated with a single step of the CG algorithm, the contrast χ is updated by again minimizing the cost function in equation 21 with the updated value of $w_{j,n}$. The updating formula is given in a closed form as

$$\chi_n = \frac{\sum_j \mathrm{Re}(w_{j,n}\bar{u}_{j,n})}{\sum_j |u_{j,n}|^2}, \tag{35}$$

where

$$u_{j,n} = u_{j,n}^{\mathrm{inc}} + \mathcal{M}^D\{\mathcal{L}_\mathrm{b}[w_{j,n}]\}. \tag{36}$$

The derivation of this updating formula can be found in Habashy et al. (1994) and van den Berg and Kleinman (1997).

After updating this contrast function, and if the value of the cost function $F_n(\chi_n,w_{j,n})$ is not smaller than the prescribed error criterion and the number of inversion iteration is still less than the prescribed maximum number of iterations, we repeat these two update steps until convergence is achieved. A flowchart of this FDCSI method is given in Figure 1.

The CSI algorithm might be improved significantly with the inclusion of a multiplicative-regularization (MR) term as shown in van den Berg et al. (1999). Here we also utilize the weighted L^2-norm regularizer introduced in Abubakar et al. (2002). Details of FDCSI with the MR approach can be found in Appendix A.

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

Algorithm initialization

When we have an inhomogeneous P-wave velocity distribution (a priori information) available from other independent measurements or from other seismic data-processing techniques, we can initialize FDCSI in two different ways:

1) We use the known inhomogeneous P-wave velocity distribution as the background medium ($c_b(\mathbf{r})$). Then, the algorithm is initialized with the contrast sources obtained by back-propagation, multiplied by a weight that ensures minimizing the data equation error. The expression for this initial contrast source is given by (see Habashy et al., 1994)

$$w_{j,0} = \frac{\|\mathcal{L}_b^*(\mathcal{M}^{S*}[f_j])\|_D^2}{\|\mathcal{M}^S\{\mathcal{L}_b[\mathcal{L}_b^*(\mathcal{M}^{S*}[f_j])]\}\|_S^2}\mathcal{L}_b^*(\mathcal{M}^{S*}[f_j]). \tag{37}$$

Contrast χ_0 is calculated using equation 35. With this option, we do need to provide the method with an initial model.

2) The other option is to employ a homogeneous medium as the background and then put the known inhomogeneous P-wave velocity distribution as $c(\boldsymbol{r})$ to calculate χ_0. Then, we calculate the contrast sources from

$$w_{j,0} = \chi_0 u_{j,0}, \tag{38}$$

where fields $u_{j,0}$ are obtained by solving equation 1 for $k = k_b\sqrt{\chi_0 + 1}$. This means we are solving a forward problem for the initial model χ_0.

We will show that the first way to initialize FDCSI produces a better result than the second. When no known inhomogeneous P-wave velocity a priori information is available, we can use any homogeneous medium as the background and initialize FDCSI using the first option.

Computational complexity

Operation of operators $\mathcal{H}_b^{-1}$ and $(\mathcal{H}_b^*)^{-1}$ on their arguments are computed via FDFD, using a fourth-order FD scheme, which incorporates the PML boundary conditions. In FDFD, both fields and acoustic velocities are collocated at the center of grid points. For high efficiency of the FDCSI algorithm, it is important that operators $\mathcal{H}_b^{-1}$ and $(\mathcal{H}_b^*)^{-1}$ are computed only once, at the beginning of the inversion process. This is possible because these operators depend only on the background medium, which does not change throughout the inversion process. These operators are calculated via an efficient LU decomposition of the resultant discretized operator (see Hu et al., 2009). Decomposition is computed once at the beginning of the inversion process, then stored and utilized at each subsequent step of the inversion process. The decomposition process takes $O(N^{1.5})$ operations (see Davis and Duff, 1997), where N is the number of grid cells in the finite-difference computation. The solution (back-substitution) of the decomposed matrix system takes $O(N \log N)$ for each source position. Solution of the $\mathcal{H}_b$ system is required once per iteration of the FDCSI algorithm. Solution of the adjoint system $\mathcal{H}_b^*$ is required twice per iteration. Hence, computation complexity of FDCSI for each frequency is given approximately by

$$\text{Computations}_{\text{FDCSI}} \approx O(2N^{1.5}) + N^{\text{iter}}O(3N^{\text{S}}N \log N). \tag{39}$$

The computational effort for updating contrast χ is not included in equation 39 because it is negligible in comparison to the effort of updating the contrast sources w_j.

NUMERICAL EXAMPLES

Marmousi model inversion

We investigate the algorithm performance using the 2D Marmousi model. The true model is given in Figure 2a. The velocity in this Marmousi model varies from 1500 m/s to 5500 m/s. The model is 384 × 126 grids in the x and z directions with a grid size of

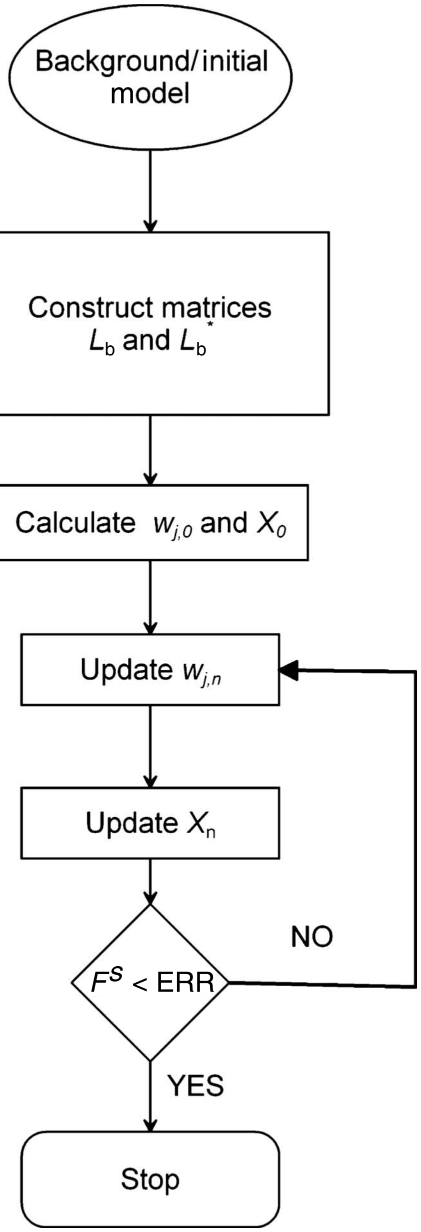


Figure 1. Flowchart of the FDCSI method. In this flowchart, F^S denotes the first term of the cost function in equation 21.

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

24 × 24 m. The synthetic data were generated using a finite-difference time-domain forward (FDTD) code described in Hu et al. (2007). The source wavelet used in data generation is a Ricker wavelet with the dominant frequency of 7.5 Hz. Because our inversion approach is formulated in the context of a frequency-domain framework, we transformed the time-domain data using a fast Fourier transform (FFT) and selected four frequencies (3, 7.5, 12, and 16.5 Hz) for the inversion. After that, we added 5% random noise to each frequency component data using the formula

$$d_j^{\text{noise}}(\mathbf{r}^{\text{R}}) = d_j(\mathbf{r}^{\text{R}}) + 2\times 0.05(\xi^{\text{Re}} + \mathrm{i}\xi^{\text{Im}})\text{Max}(\forall\, \mathbf{r}^{\text{R}} \forall\, j |d_j(\mathbf{r})|), \qquad (40)$$

where d_j^{noise} is the total field data with noise and d_j is the noiseless data. Symbols ξ^{Re} and ξ^{Im} denote pseudorandom numbers uniformly distributed on the interval $(-0.5, 0.5)$. To illustrate the noise effects in Figures 3a and b, we show the data without and with 5% noise as functions of the source and receiver indices. In these figures, we plot the amplitude of the frequency-domain data at 16.5 Hz. The relative difference between the clean and noisy data is given in Figure 3c. We employ 48 sources and 96 receivers located at surface $z = 0$ m and distributed uniformly from $x = 0$ to $x = 9216$ m. For all the sources, we used the full aperture of the model. The number of sources, receivers, and frequencies used in our simulation are significantly less than the ones employed in the literature (Sirgue and Pratt, 2004; Vigh and Starr, 2008; Mulder and Plessix, 2008).

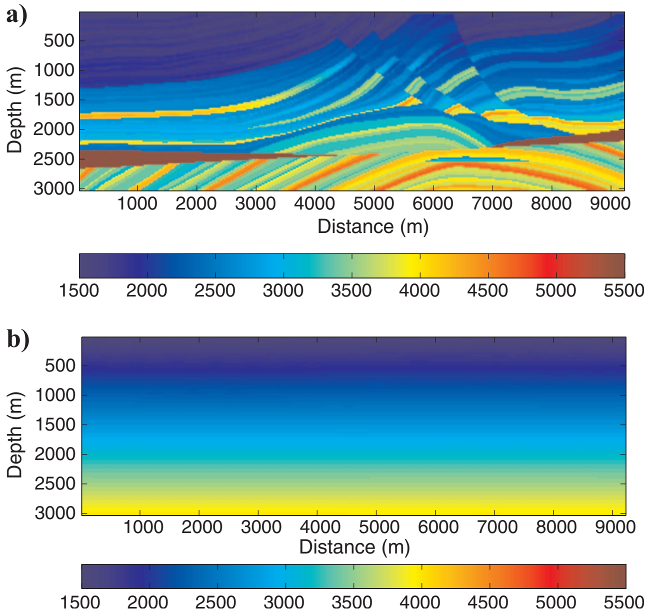


Figure 2. True Marmousi velocity model (a) and initial/background model used in the inversion (b). Color bars are in m/s.

We carried out the inversion employing the L_2-norm regularization given in equations A-2 and A-3. For an initial model, we took a linearly increasing velocity model, where the velocity varies from 1500 m/s to 4200 m/s as shown in Figure 2b. In FDCSI, we used this initial model also as the background model in the inversion. In carrying out the inversion, we employed the frequency-hopping approach, i.e., we inverted the data one frequency at a time from the lowest to the highest frequency. Each time, we used the lower frequency model as the background model for the inversion of the next higher frequency data set. The inverted velocity model using the data at 3 Hz is given in Figure 4a. The error in contrast ERR_χ has been reduced from 28.57% to 23.48% and the data misfit F^S defined in equation 22 has been reduced from 58.3% to 8.4%. The error in contrast ERR_χ is defined as

$$\text{ERR}_\chi = \frac{\|\chi - \chi^{\text{true}}\|_D}{\|\chi^{\text{true}}\|_D}, \qquad (41)$$

where χ^{true} is the contrast of the true model. Next, we used the inversion results of the 3-Hz data as the background model for the 7.5-Hz data inversion. The inverted velocity model using the data at 7.5 Hz is given in Figure 4b. The data misfit F^S has now reduced from 89.8% to 6.3% and the error in contrast value is reduced to 21.30%. We repeated the process for inverting the data at 12 Hz and 16.5 Hz. Results are given in Figure 4c and d. The final data-misfit value is 2.2% for the 12-Hz data inversion and 2.3% for the 16.5-Hz data inversion. The final error in contrast value is 17.67%. These results are summarized in Table 1. In Figure 5, we present the data misfit value F^S as a function of iteration number for inversion of the 16.5-Hz data. We observe that FDCSI has a well-behaved convergence curve. Note that for this Marmousi model, we also ran the inversion using the weighted L_2 regularization, however we did not obtain a significant improvement in the quality of the inversion result.

An alternate way to implement frequency hopping is to use any model (in our inversion run, we used the linearly increasing velocity model) as the background model in the inversion of each data frequency and to use the previous frequency inversion results as the initial model of the inversion of the next data frequency. This is achieved by initializing the algorithm using equation 38. This is the sequential multifrequency inversion approach that one employs usually in the Gauss-Newton or nonlinear conjugate gradient method. The final inversion result using this approach is given in Figure 6. The final error in contrast value of this approach is 19.57%. We observe that in the deep region the inversion results in Figure 4d are slightly better than those in Figure 6. Hence, for sequential multifrequency inversion using FDCSI, it seems that it is better to

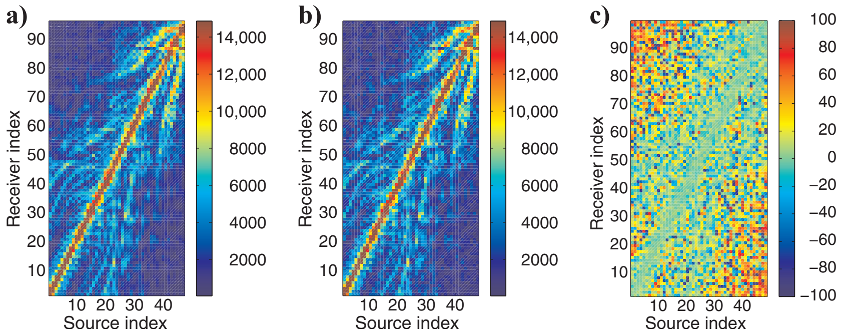


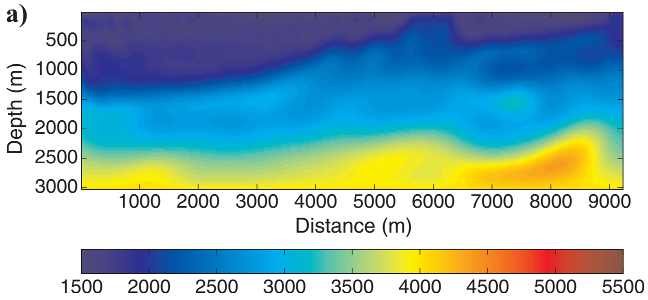
Figure 3. Amplitudes of the frequency-domain data at 16.5 Hz without (a) and with (b) 5% random noise and their relative difference (c).

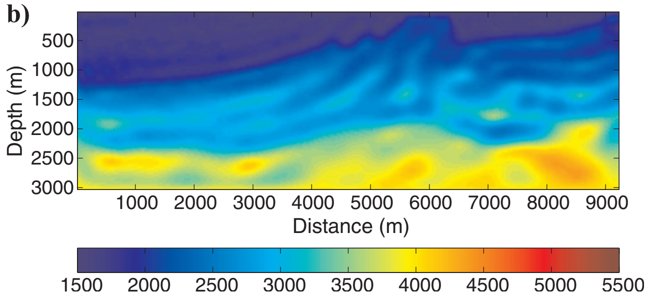
Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

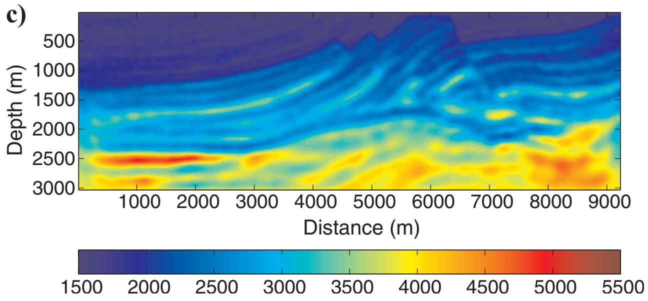
use the previous frequency-inversion result as the background model than to use it only as the initial model.

Crosswell time-lapse inversion example

In the second example, we apply FDCSI to a crosswell seismic monitoring problem. In this time-lapse imaging application (see Nur et al., 1984; Lumley, 1995), our goal is to reconstruct the changes in the model over time. We have the so-called baseline and monitor models as given in Figure 7a and b. The baseline model is the model at the initial time and is assumed to be either known or reconstructed from data collected at initial times. Hence, the goal is to reconstruct the difference between the baseline model and the monitor model.







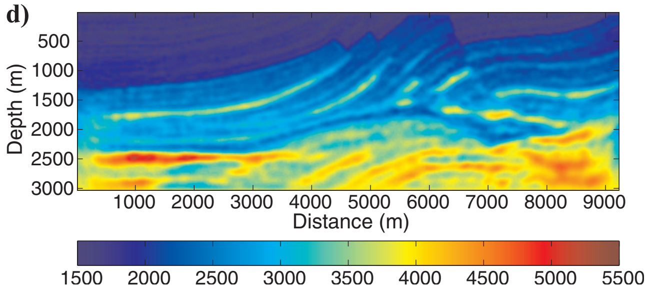


Figure 4. Reconstructed Marmousi velocity model using FDCSI method: (a) 3 Hz, $\mathrm{ERR}_\chi = 23.48\%$; (b) 7.5 Hz, $\mathrm{ERR}_\chi = 21.3\%$; (c) 12 Hz, $\mathrm{ERR}_\chi = 18.5\%$; and (d) 16.5 Hz, $\mathrm{ERR}_\chi = 17.67\%$.

The true baseline model in Figure 7a consists of seven layers with velocities of 2560, 2830, 2560, 2830, 2560, 2830, and 3000 m/s from top to bottom. In the monitor model (Figure 7b), we have placed an anomaly in the third layer and another in the bottom layer to mimic fluid movements in an injection experiment. The velocity of these anomalous regions is 2200 m/s. The models are discretized into 45 grids in the x-direction and 120 grids in the z-direction with a grid size of 1×1 m.

Synthetic data were generated using a FDTD algorithm presented in Hu et al. (2007). The 30 sources and 30 receivers are spaced equally on the left and right edges of the inversion domain. Well spacing is 45 m. After applying a FFT to the synthetic time-domain data, we extracted data at three frequencies: 50, 150, and 250 Hz. We also

Table 1. The FDCSI data misfit and error in contrast for each frequency inversion.

	3 Hz	7.5 Hz	12 Hz	16.5 Hz
$F^S(\chi_0)$	58.3	89.8	86.8	77.7
$F^S(\chi_N)$	8.4	6.3	2.2	2.2
ERR_{χ_N}	23.48	21.30	18.59	17.67

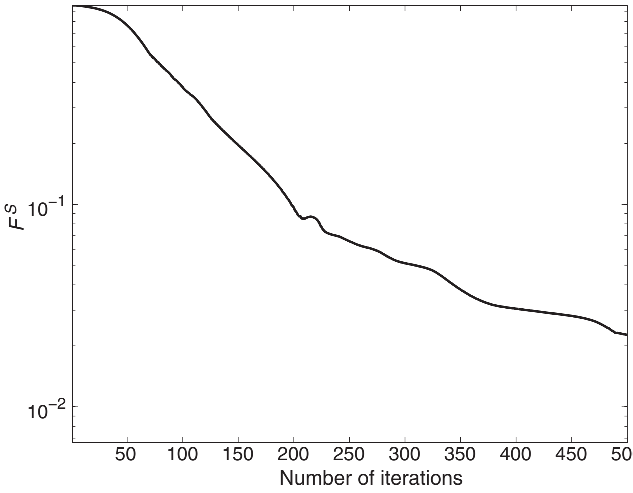


Figure 5. The data misfit value F^S as a function of iteration number for the inversion of 16.5 Hz data using FDCSI.

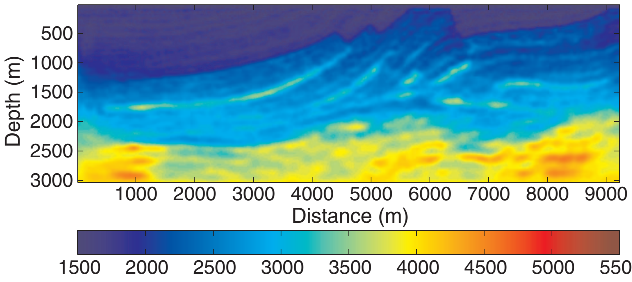


Figure 6. Reconstructed Marmousi velocity model using FDCSI ($\mathrm{ERR}_\chi = 19.57\%$). The linearly increasing velocity model is used as the background model in the inversion of each data frequency and the previous frequency-inversion result is used as the initial model of inversion for the next data frequency.

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

added 5% random noise to the data. In this example, we inverted all data simultaneously by replacing the cost function in equation 21 with

$$F_n(\chi,w_{k,j}) = \frac{1}{N^{\mathrm{f}}}\sum_{k=1}^{N^{\mathrm{f}}}\left[\frac{\sum_j \|f_{k,j} - \mathcal{M}^S\{\mathcal{L}_{\mathrm{b};k}[w_{k,j}]\}\|_S^2}{\sum_j \|f_{k,j}\|_S^2} + \frac{\sum_j \|\chi u_{k,j}^{\mathrm{inc}} - w_{k,j} + \chi\mathcal{M}^D\{\mathcal{L}_{\mathrm{b};k}[w_{j,k}]\}\|_D^2}{\sum_j \|\chi_n u_{k,j}^{\mathrm{inc}}\|_D^2}\right], \tag{42}$$

where N^{f} denotes the number of frequencies employed in the inversion and subscript k represents the frequency index. Because the changes between the baseline and monitor models exhibit sharp boundaries, we employed the weighted L_2-norm regularization given in equations A-2 and A-4.

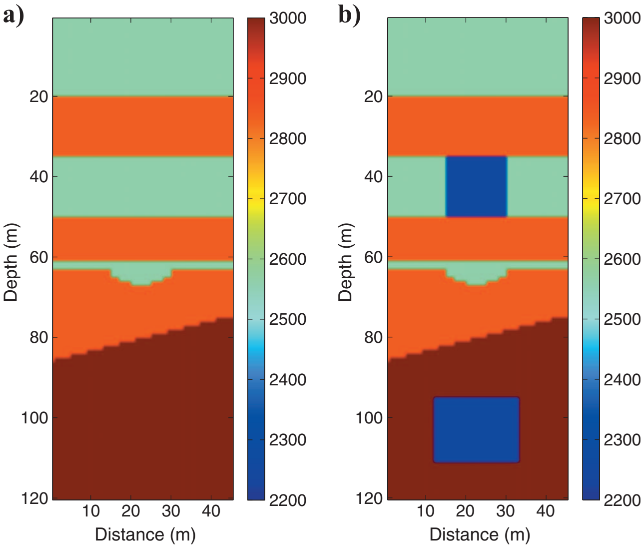


Figure 7. True models of the crosswell time-lapse example: (a) baseline model and (b) monitor model. Color bars are in m/s.

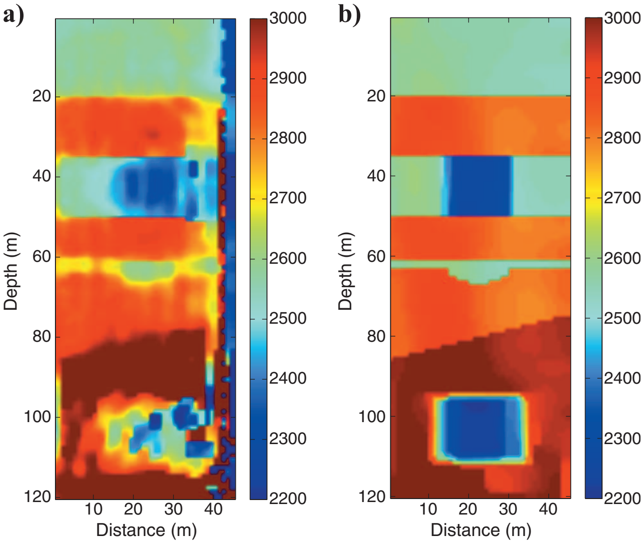


Figure 8. Inversion results of the crosswell time-lapse example using the true baseline model as the initial estimate: (a) using a homogeneous background medium, ERR_χ=40.7%; and (b) using an inhomogeneous background medium, ER_χ=5.1%.

In this example, we investigated the advantages of using an inhomogeneous background medium in FDCSI. First, we inverted the monitor data set using the true baseline model given in Figure 7a as the initial model and a homogeneous background medium of 1500 m/s. Inversion results are given in Figure 8a. Next, we employed the baseline model in Figure 7a as the inhomogeneous background medium. Figure 8b shows inversion results using the inhomogeneous background medium. The error in contrast ERR_χ at the initial model is 16.4%. After 128 iterations, its value becomes 40.7% for the reconstruction using the homogeneous background medium and 5.1% for the reconstruction using the inhomogeneous background medium. Initial data misfits for each frequency are 18.4%, 70.3%, and 86.3%. Final data misfits are 7.4%, 1.9%, and 1.4% for the reconstruction using the homogeneous background medium and 4.7%, 2.5%, and 1.2% for the reconstruction using the inhomogeneous background medium. Many artifacts are observed in the homogeneous background inversion result. Moreover, the velocity variations do not seem to be reconstructed satisfactorily. However, the reconstructed structure using the inhomogeneous background medium is very close to the true monitor model, which demonstrates that FDCSI has the tendency to preserve the inhomogeneous background and tries to reconstruct only anomalous regions that correspond to velocity changes.

To simulate a more realistic time-lapse experiment, we used the inverted baseline model as the background medium in the inversion of the monitor data set. The inverted baseline model obtained using FDCSI is given in Figure 9a. The initial model used to obtain the inverted baseline model is the back-propagation model in equation 37. For the inversion of this baseline model, both algorithm initializations might be used. In our present example, they have produced the same results. We were able to retrieve most of the features in this baseline model. Then we inverted the monitor data set using the inverted baseline model as the background medium. Inversion results

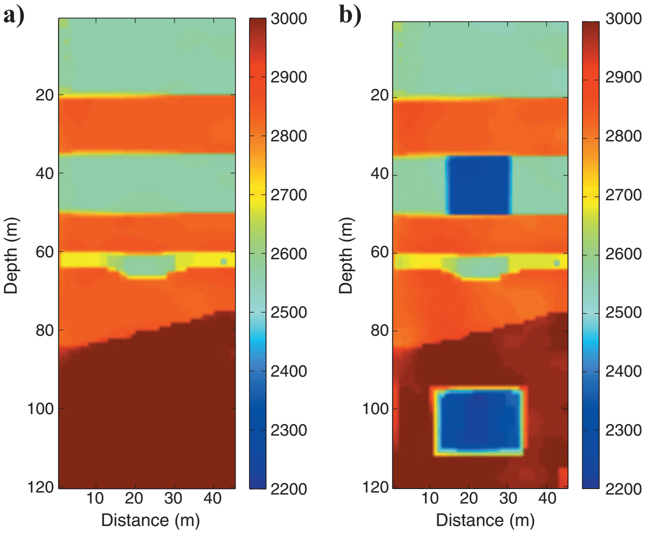


Figure 9. Inversion results of the crosswell time-lapse example using the reconstructed baseline model as the background medium: (a) inverted baseline model and (b) inverted monitor model.

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

are given in Figure 9b. Although we did not use the true baseline model, the reconstructed monitor model is quite comparable to the one obtained in Figure 8b. In Figure 10b, we also plot the difference between the reconstructed velocity model and the baseline model. The reconstructed velocity difference in Figure 10b is very close to the true one.

For the time-lapse application, the first initialization procedure in equation 37 would be the most suitable for most cases. In cases where we have a priori information in addition to the inversion results of the baseline data, the second initialization procedure in equation 38 also can be used. In this workflow, we might want to use the inversion results of the baseline data as the background model and other available a priori information in the initial model.

DISCUSSION

Comparison with other nonlinear inversion algorithms

To assess the quality of the inversion results of the FDCSI method, we also inverted the same data sets using the Gauss-Newton method and the nonlinear conjugate gradient method described in Hu et al. (2009). We employed the sequential frequency inversion in both methods. The inversion results are given in Figure 11a for the nonlinear conjugate gradient method and in Figure 11b for the Gauss-Newton method. The resolution of the inversion results of the nonlinear conjugate gradient method is lower than that of the FDCSI method. However, in the nonlinear conjugate gradient method, we do not employ any preconditioning operator. One can improve this nonlinear conjugate-gradient-method inversion result as described in Choi et al. (2007) at an additional computation cost. Inversion results of the Gauss-Newton method and the FDCSI method are comparable in the zones where we have good sensitivity in the data (up to a depth of $z = 2000$ m).

Finite-difference contrast-source inversion uses conjugate gradient methods to update both the contrast sources and the contrast function. However, the minimization process is carried out by alternately minimizing the contrast sources and the contrast function. All the updating parameters in each step are found in closed form. The latter helps in reducing the numerical errors in the minimization process. In the closed-form update of the contrast function of equation 35, the numerator represents the gradient direction and the denominator acts as a Jacobi preconditioner. In principle, we have reformulated the minimization process in a well-posed updating process for the contrast sources, whereas the ill-posed part of the updating process is solved using a closed-form equation.

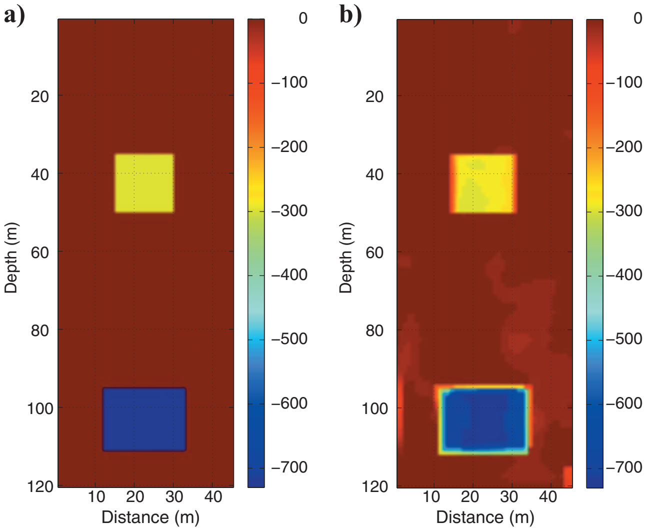


Figure 10. Velocity differences between the monitor and the baseline models: (a) true difference and (b) reconstructed difference.

For more detail on these comparisons in Figure 12, we present the extracted P-wave velocity vertical profiles at $x = 4596$ m for the three inversion methods. From Figure 12, we observe that FDCSI produces results that are close to the Gauss-Newton method.

As an extra check, we compare the responses of the true model and the reconstructed model obtained using the FDCSI, Gauss-New-

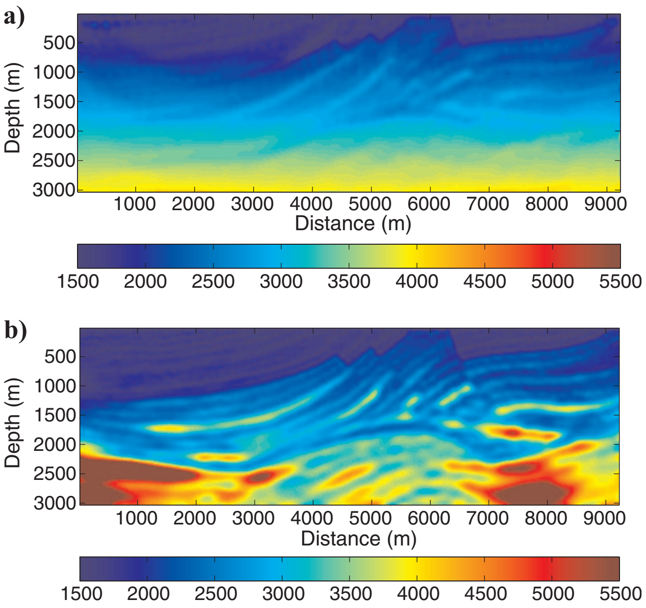


Figure 11. Reconstructed Marmousi velocity model using the standard methods: (a) nonlinear conjugate gradient inversion result and (b) Gauss-Newton inversion result.

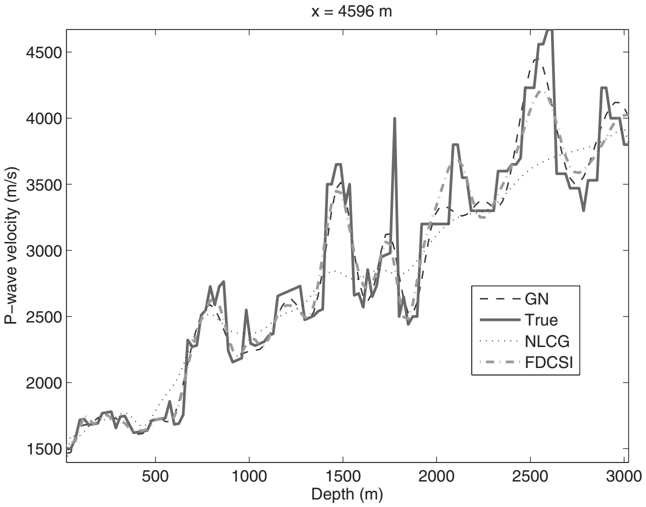


Figure 12. P-wave velocity vertical profile at $x = 4596$ m of true model (solid line) and the inversion results using the FDCSI (dashed dotted line), Gauss-Newton (dashed line) and nonlinear conjugate gradient (dotted line) method.

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

ton, and nonlinear conjugate gradient methods. The time-waveform responses in Figure 13a-d and Figure 14 are calculated using an FDTD code that uses the same discretization grids as the frequency-domain simulator used in the inversions. We use a source with a Ricker wavelet with a dominant frequency of 7.5 Hz that is located at $x = 4620$ m. The time waveforms are recorded at three receivers located at $x = 2292$ m, $x = 4596$ m, and $x = 6900$ m. As shown in Figure 14, the time waveforms of the true model and the reconstructed models are very close, except for the late time signals of the responses produced by the nonlinear conjugate gradient inversion results. These late time results carry mainly the information of the deep region of the Marmousi model.

On average, the number of iterations in each inversion run is five for the Gauss-Newton method, 30 for nonlinear conjugate gradient inversion, and 128 for FDCSI. Total CPU time for all four frequency inversions is 831,560 seconds for the Gauss-Newton method, 21,975 seconds for the nonlinear conjugate gradient method, and 8801 seconds for FDCSI using a personal computer with a PIV 3.04 GHz processor.

Extension to 3D geometries

Because FDCSI does not require the explicit computation of the full forward solution in each of the inversion iterations, we expect the method to have a great potential for solving 3D problems. However, more research on efficient computation of the CSI finite-difference operator for the 3D configuration is needed. It will not be efficient for FDCSI to employ a standard 3D iterative frequency-domain forward solver. We speculate that there are a few options for FDCSI, such as constructing a basis for the CSI finite-difference operator and using it in each FDCSI iteration, calculating the CSI finite-difference operator using a limited number of iterations of the iterative solver, or even using a direct solver because we calculate LU decompositions only twice in FDCSI. All these options need to be investigated.

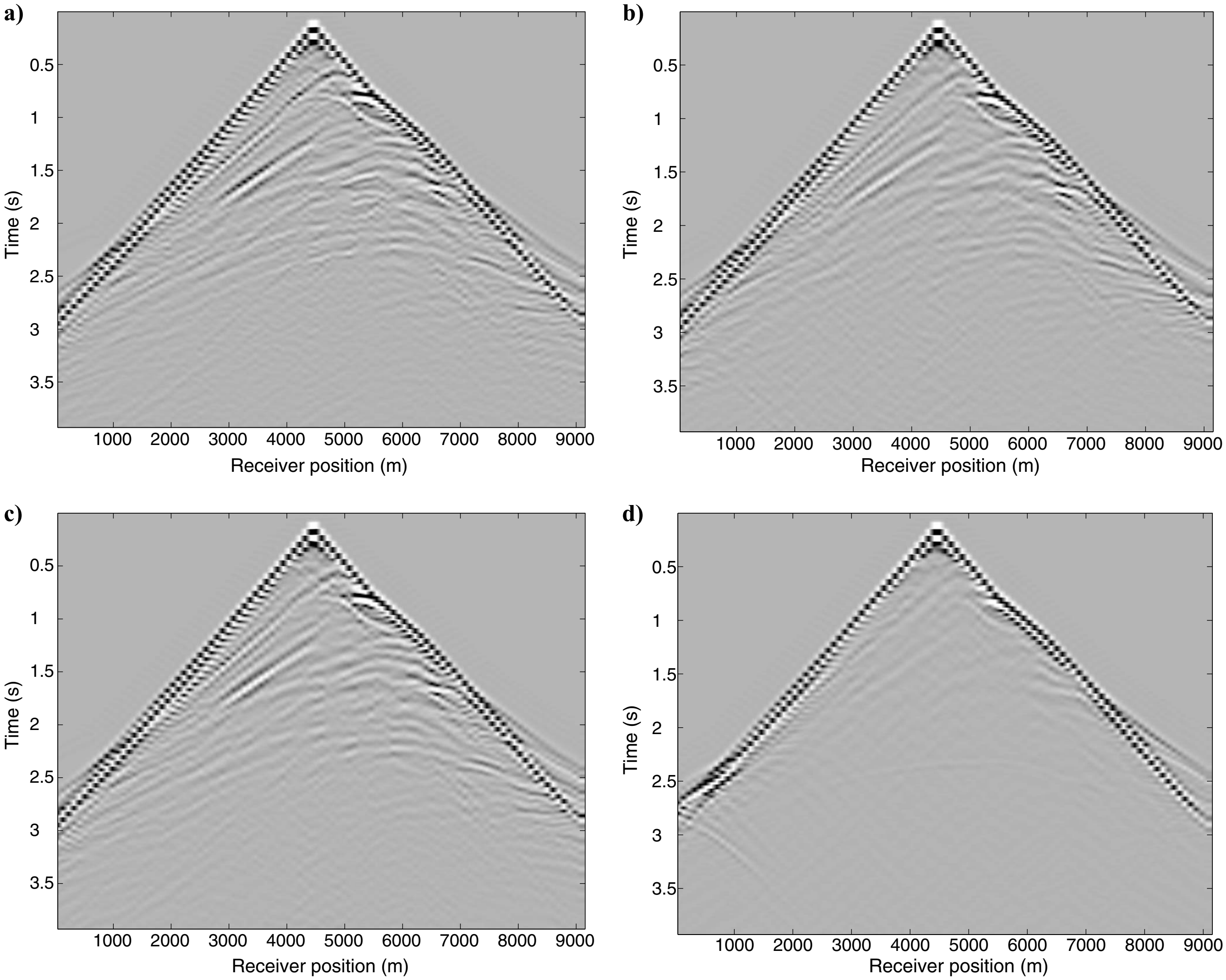


Figure 13. The responses to the true model and the reconstructed model using the FDCSI, Gauss-Newton, and nonlinear conjugate gradient methods. The source is a Ricker wavelet with dominant frequency of 7.5 Hz, located at (4620 m, 0 m).

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

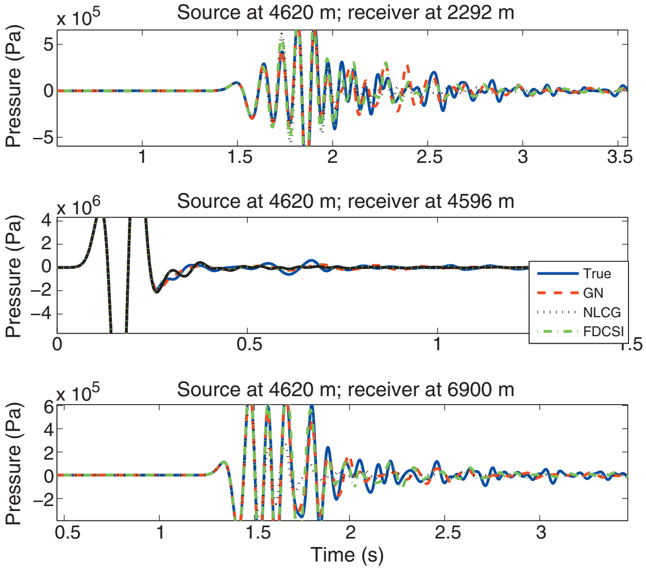


Figure 14. The responses to the true model (a) and the reconstructed model using the FDCSI (b), Gauss-Newton (c), and nonlinear conjugate gradient (d) method. The source is a Ricker wavelet with dominant frequency of 7.5 Hz, located at 4620 m, 0 m.

CONCLUSIONS

We have presented a finite-difference contrast-source inversion (FDCSI) method and compared its performance against the Gauss-Newton method and the nonlinear conjugate gradient method. Our numerical simulation results show that FDCSI is able to produce inversion results that are comparable to those obtained by the Gauss-Newton method, however with significantly less computation cost. The FDCSI inversion results are also better than those obtained by the unpreconditioned nonlinear conjugate gradient method.

When a priori knowledge is known about parts of the model to be inverted, FDCSI has the capability to incorporate this a priori information in the construction of the background medium, better preserving the known part of the model. This a priori knowledge might come from either other independent measurements or seismic data processing techniques. We also showed that FDCSI has great potential for time-lapse inversion applications. In such applications, FDCSI preserves the baseline model and can allow one to focus the inversion on a particular area of interest.

Future work will be focused on testing the method on more complicated models and on any available field data.

ACKNOWLEDGMENTS

The authors acknowledge valuable comments from F. Bleibinhaus, G. O. Ayeni, two anonymous reviewers, and the associate editor that improved the quality and readability of this manuscript significantly.

APPENDIX A

MULTIPLICATIVE REGULARIZATION

The CSI algorithm can be improved significantly with the inclusion of a multiplicative-regularization (MR) term shown in van den Berg et al. (1999). Here we utilize the weighted L_2-norm regularizer introduced in Abubakar et al. (2002). With the MR term, the CSI cost function then becomes

$$C_n(\chi,w_j) = [F^S(w_j) + F_n^D(\chi,w_j)]F_n^R(\chi), \qquad \text{(A-1)}$$

where the regularization cost function is given by

$$F_n^R(\chi) = \int_D b_n^2(\mathbf{r})(|\nabla\chi(\mathbf{r})|^2 + \delta_n^2)d\mathbf{r}, \qquad \text{(A-2)}$$

in which

$$b_n^2(\mathbf{r}) = \frac{1}{\int_D (|\nabla\chi_{n-1}(\mathbf{r})|^2 + \delta_n^2)d\mathbf{r}} \qquad \text{(A-3)}$$

for the L_2-norm regularization and

$$b_n^2(\mathbf{r}) = \frac{1}{A(|\nabla\chi_{n-1}(\mathbf{r})|^2 + \delta_n^2)} \qquad \text{(A-4)}$$

for the weighted L_2-norm regularization. Quantity δ_n^2 is defined as

$$\delta_n^2 = \frac{F_n^D(\chi_{n-1},w_{j,n-1})}{\Delta x \Delta z}, \qquad \text{(A-5)}$$

where $A = \int_D d\mathbf{r}$ is the area of the inversion domain D and $\Delta x \Delta z$ is the area of the inversion domain cell. The L_2-norm regularizer is known to favor smooth profiles, and the weighted L_2-norm regularizer is known for its ability to preserve edges (see Abubakar et al., 2002; van den Berg et al., 2003; Abubakar et al., 2008a). This weighted L_2-norm regularization factor belongs to the same class as the well-known total-variation regularization (see Rudin et al., 1992; Charbonnier et al., 1996; Dobson and Santosa, 1996; Vogel and Oman, 1996; Farquharson and Oldenburg, 1998). This cost function of a weighted L_2-norm regularization has all the advantages of the total-variation regularization function. Moreover, it is a quadratic function. It has a mathematically defined gradient so it is more suitable to be used with a gradient-based approach.

By introducing this extra regularization, there is no change in the updating procedure for the contrast sources w_j because $F_n^R(\chi)$ does not depend on the contrast sources and $F_n^R(\chi_{n-1}) = 1$. However, updating for contrast χ now must be done using a CG step

$$\chi_n = \chi_n^{\mathrm{c}} + \alpha_n^{\chi} d_n, \qquad \text{(A-6)}$$

where χ_n^{c} is the closed-form solution (which is the minimizer of $F_n^D(w_{j,n},\chi)$) given in equation 35, and d_n is the Polak-Ribière search direction

$$d_n = g_n^{\chi} + \frac{\langle g_n^{\chi}, g_n^{\chi} - g_{n-1}^{\chi}\rangle_D}{\|g_{n-1}^{\chi}\|_D^2} d_{n-1}, \qquad \text{(A-7)}$$

in which g_n^{χ} is the gradient of the cost function with respect to χ evaluated at the $(n-1)$th iteration. The gradient can be shown to be given by (see Abubakar et al., 2008b)

$$g_n^{\chi} = F_n^R(\chi_n^{\mathrm{c}}) 2\eta_n^D \sum_j (\chi_n^{\mathrm{c}} u_{j,n} - w_{j,n})\overline{u_{j,n}} - 2[F^S(w_{j,n})$$
$$+ F_n^D(\chi_n^{\mathrm{c}}, w_{j,n})]\nabla \cdot (b_n^2 \nabla \chi_n^{\mathrm{c}}) = -2[F^S(w_{j,n})$$

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/

$$+ F_n^D(\chi_n^c, w_{j,n})]\nabla \cdot (b_n^2 \nabla \chi_n^c). \qquad \text{(A-8)}$$

The first term in the right side of equation A-8 vanishes because gradient g_n^χ is evaluated at $\chi = \chi_n^c$. The real parameter minimizer α_n^χ is found from a line minimization of

$$\alpha_n^\chi = \min_{\alpha^\chi}[C_n(\chi_n^c + \alpha^\chi d_n, w_{j,n})]. \qquad \text{(A-9)}$$

Minimization of the multiplicative cost functional can be performed analytically because the cost functional is a fourth-degree polynomial in α^χ:

$$C_n(\alpha^\chi) = [A + B(\alpha^\chi)^2][X + 2Y\alpha^\chi + Z(\alpha^\chi)^2]. \qquad \text{(A-10)}$$

Here,

$$\begin{aligned} A &= F^S(w_{j,n}) + F_n^D(\chi_n^c, w_{j,n}), \\ B &= \eta_n^D \sum_j \|d_n u_{j,n}\|_D^2, \\ X &= \|b_n \nabla \chi_n^c\|_D^2 + \delta_n^2 \|b_n\|_D^2, \\ Y &= \mathrm{Re}\langle b_n \nabla \chi_n^c, b_n \nabla d_n\rangle_D, \\ Z &= \|b_n \nabla d_n\|_D^2. \end{aligned} \qquad \text{(A-11)}$$

Differentiation with respect to α^χ yields a cubic equation with one real root and two complex conjugate roots. The real root is the desired minimizer.

REFERENCES

Abubakar, A., T. Habashy, V. Druskin, L. Knizhnerman, and D. Alumbaugh, 2008a, 2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements: Geophysics, **73**, no. 4, F165–F177.

Abubakar, A., W. Hu, P. van den Berg, and T. M. Habashy, 2008b, A finite-diference contrast source inversion method: Inverse Problems, **24**, 1–17.

Abubakar, A., P. van den Berg, and J. T. Fokkema, 2003, Towards nonlinear inversion for characterization of time-lapse phenomena through numerical modelling: Geophysical Prospecting, **51**, 285–293.

Abubakar, A., P. van den Berg, and J. Mallorqui, 2002, Imaging of biomedical data using a multiplicative regularized contrast source inversion method: IEEE Transactions on Microwave Theory and Techniques, **50**, 1761–1771.

Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2008, Velocity model building by 3D frequency-domain, full-waveform inversion of wide-aperture seismic data: Geophysics, **73**, VE101-VE117.

Bérenger, J. P., 1994, A perfectly matched layer for the absorption of electromagnetic waves: Journal of Computational Physics, **114**, 185–200.

Bloemenkamp, R., and P. van den Berg, 2000, Time-domain profile inversion using contrast sources: Inverse Problems, **16**, 1173–1193.

Charbonnier, P., L. Blanc-Féraud, G. Aubert, and M. Barlaud, 1997, Deterministic edge-preserving regularization in computed imaging: IEEE Transactions on Image Processing, **6**, 298–311.

Chew, W., and W. Weedon, 1994, A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates: IEEE Transactions on Antennas and Propagation, **43**, 599–604.

Choi, Y., C. Shin, and D.-J. Min, 2007, Frequency-domain elastic full waveform inversion using the new pseudo-Hessian matrix: Elastic Marmousi-2 synthetic test: 77th Annual International Meeting, SEG, Expanded Abstracts, 1908–1912.

Dablain, M. A., 1986, The application of high-order differencing to the scalar equation: Geophysics, **51**, 54–66.

Davis, T. A., and I. S. Duff, 1997, An unsymmetric pattern multifrontal method for sparse LU factorization: SIAM Journal on Matrix Analysis and Applications, **18**, 140158.

Dobson, D. C., and F. Santosa, 1996, Recovery of blocky images for noisy and blurred data: SIAM Journal of Applied Mathematics, **56**, 1181–1198.

Farquharson, C. G., and D. W. Oldenburg, 1998, Nonlinear inversion using general measures of data misfit and model structure: Geophysical Journal International, **134**, 213–227.

Fokkema, J., and P. van den Berg, 1993, Seismic applications of acoustic reciprocity: Elsevier.

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

Habashy, T. M., M. L. Oristaglio, and A. T. de Hoop, 1994, Simultaneous nonlinear re-construction of two-dimensional permittivity and conductivity: Radio Science, **29**, 1101–1118.

Harari, I., and E. Turkel, 1995, Accurate finite difference methods for time-harmonic wave propagation: Journal of Computational Physics, **119**, 252–270.

Hu, W., A. Abubakar, and T. M. Habashy, 2007, Application of the nearly perfectly matched layer in acoustic wave modeling: Geophysics, **72**, no. 5, SM169–SM175.

——— 2009, Simultaneous multi-frequency inversion of full waveform seismic data: Geophysics, **74**, no. 2, R1–R14.

Hustedt, B., S. Operto, and J. Virieux, 2004, Mixed-grid and staggered-grid finite-difference methods for frequency-domain acoustic wave modelling: Geophysical Journal International, **157**, 1269–1296.

Liao, Q., and G. A. McMechan, 1996, Multifrequency viscoacoustic modeling and inversion: Geophysics, **61**, 1371–1378.

Lumley, D. E., 1995, 4D seismic monitoring of an active steamflood: 65th Annual International Meeting, SEG, Expanded Abstracts, 203–206.

Malinowski, M., and S. Operto, 2008, Quantitative imaging of the Permo-Mesozoic complex and its basement by frequency domain waveform tomography of wide-aperture seismic data from the Polish basin: Geophysical Prospecting, **56**, no. 6, 805–825.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **54**, 1211–1228.

Mulder, W., and R.-E. Plessix, 2008, Exploring some issues in acoustic full waveform inversion: Geophysical Prospecting, **56**, no. 6, 827–841.

Nur, A., C. Tosaya, and D. V-Thanh, 1984, Seismic monitoring of thermal enhanced oil recovery processes: 54th Annual International Meeting, SEG, Expanded Abstracts, 118–221.

Pratt, R., 1990, Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging: Geophysics, **55**, 626–632.

——— 1999, Seismic waveform inverison in the frequency domain, part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

Pratt, R., C. Shin, and G. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **13**, 341–362.

Pratt, R. G., and M. H. Worthington, 1990, Inverse theory applied to multi-source cross-hole tomography: Part 1. Acoustic wave-equation method: Geophysical Prospecting, **38**, 287–310.

Ravaut, C., S. Operto, L. Improta, J. Virieux, A. Herrero, and P. dell'Aversana, 2004, Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-wavefield inversions: Application to a thrust belt: Geophysical Journal International, **159**, 1032–1056.

Rudin, L., S. Osher, and C. Fatemi, 1992, Nonlinear total variation based noise removal algorithm: Physica, **60D**, 259–268.

Shipp, R., and S. C. Singh, 2002, Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data: Geophysical Journal International, **151**, 325–344.

Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

Song, Z.-M., P. R. Williamson, and R. G. Pratt, 1995, Frequency-domain acoustic-wave modeling and inversion of crosshole data: Part II — Inversion method, synthetic experiments, and real-data results: Geophysics, **60**, 796–809.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903.

van den Berg, P., A. Abubakar, and J. Fokkema, 2003, Multiplicative regularization for contrast profile inversion: Radio Science, **38**, 23.1–23.10.

van den Berg, P., A. L. Broekhoven, and A. Abubakar, 1999, Extended contrast source inversion: Inverse Problems, **15**, 1325–1344.

van den Berg, P. M., and R. E. Kleinman, 1997, A contrast source inversion method: Inverse Problems, **13**, 1607–1620.

van Dongen, K., and W. Wright, 2007, A full vectorial contrast source inversion scheme for 3D acoustic imaging of both compressibility and density profiles: Journal of the Acoustical Society of America, **121**, 1538–1549.

Vigh, D., and E. W. Starr, 2008, 3D prestack plane-wave, full-waveform inversion: Geophysics, **73**, no. 5, VE135–VE144.

Vogel, C. R., and M. E. Oman, 1996, Iterative methods for total variation denoising: SIAM Journal of Scientific Computing, **17**, 227–238.

Downloaded 01 Oct 2012 to 131.180.130.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://segdl.org/