2014-02

2d Frequency-Domain Elastic Full-Waveform Inversion using Time-Domain Modeling and a Multistep-Length Gradient Approach

UTD AUTHOR(S): Kun Xu and George A. McMechan

GEOPHYSICS, VOL. 79, NO. 2 (MARCH-APRIL 2014); P. R41–R53, 21 FIGS., 2 TABLES.
10.1190/GEO2013-0134.1

2D frequency-domain elastic full-waveform inversion using time-domain modeling and a multistep-length gradient approach

Kun Xu[1] and George A. McMechan[1]

ABSTRACT

To decouple the parameters in elastic full-waveform inversion (FWI), we evaluated a new multistep-length gradient approach to assign individual weights separately for each parameter gradient and search for an optimal step length along the composite gradient direction. To perform wavefield extrapolations for the inversion, we used parallelized high-precision finite-element (FE) modeling in the time domain. The inversion was implemented in the frequency domain; the data were obtained at every subsurface grid point using the discrete Fourier transform at each time-domain extrapolation step. We also used frequency selection to reduce cycle skipping, time windowing to remove the artifacts associated with different source spatial patterns between the test and predicted data, and source wavelet estimation at the receivers over the full frequency spectrum by using a fast Fourier transform. In the inversion, the velocity and density reconstructions behaved differently; as a low-wavenumber tomography (for velocities) and as a high-wavenumber migration (for density). Because velocities and density were coupled to some extent, variations were usually underestimated (smoothed) for V_P and V_S and correspondingly overestimated (sharpened) for ρ. The impedances I_P and I_S from the products of the velocity and density results compensated for the under- or overestimations of their variations, so the recovered impedances were closer to the correct ones than V_P, V_S, and ρ were separately. Simultaneous reconstruction of V_P, V_S, and ρ was robust on the FE and finite-difference synthetic data (without surface waves) from the elastic Marmousi-2 model; satisfactory results are obtained for V_P, V_S, ρ, and the recovered I_P and I_S from their products. Convergence is fast, needing only a few tens of iterations, rather than a few hundreds of iterations that are typical in most other elastic FWI algorithms.

INTRODUCTION

Multiparameter elastic full-waveform inversion (FWI) is more suitable for wide-aperture, multicomponent seismic data than monoparameter scalar-wave velocity (or acoustic) FWI because the acoustic solution is not able to fit P-S-mode conversions well at wide apertures (Mulder and Plessix, 2008). As more parameters are obtained in elastic FWI, it is increasingly useful for geologic interpretation of rock properties (e.g., reservoir characterization and monitoring). However, because different parameters are more or less coupled as functions of the incident angles in elastic FWI (Tarantola, 1986; Virieux and Operto, 2009), it is more ill-conditioned than acoustic velocity-only FWI. Especially, among all elastic parameters in the inversion, ρ has no impact on the phases of seismic wavefields and has a radiation pattern similar to that of V_P at short apertures. Hence, ρ is the most difficult parameter to be inverted, so it is typically fixed, or recovered from very low frequencies (e.g., below 1 Hz) along with other parameters in the elastic inversion.

To date, only a few studies of elastic FWIs are conducted in the time or frequency domains. In the time domain, Tarantola (1986) and Mora (1987) expand 2D FWI from the monoparameter scalar, to the multiparameter elastic, inversions with different combinations of elastic parameters; they found that ρ cannot be well resolved because of the ambiguity of the second diffraction patterns between V_P and ρ. Therefore, they suggest that probably only two velocities (V_P and V_S) or their impedances (I_P and I_S) are suitable parameters for the elastic FWI. Crase et al. (1990) reconstruct the short-wavelength impedances I_P and I_S using a 2D time-domain elastic FWI with different minimization criteria. The impedance results are similar to migration profiles, which lack low-wavenumber information. Shipp and Singh (2002) and Sears et al. (2008) perform 2D

Manuscript received by the Editor 31 March 2013; revised manuscript received 9 August 2013; published online 17 February 2014.
[1]The University of Texas at Dallas, Center for Lithospheric Studies, Richardson, Texas, USA. E-mail: stonekingxk@hotmail.com; mcmec@utdallas.edu.

time-domain elastic FWIs to recover velocities using wide-aperture marine data. Using a time-windowing strategy, they design hierarchical multistep approaches to mitigate the nonlinearity in the data-to-parameter coupling, which makes the inversion convergence slow or difficult. Nevertheless, not all three parameters V_P, V_S, and ρ are recovered during the inversions in their work; V_S or ρ is reconstructed using empirical relation equations to predict from V_P. Köhn et al. (2012) show successful reconstructions of V_P, V_S, and ρ for the elastic Marmousi-2 model (Martin et al., 2002) with a multiscale strategy using different frequency bands in the time domain. However, they conclude that low frequencies are essential for the reconstruction of ρ, so very low frequencies (close to 0 Hz) in the full frequency band are included in the time-domain inversion of their synthetic data.

In the frequency domain, Pratt et al. (1998) describe least-squares optimization schemes in detail using different approximate forms of the Hessian matrix for the monoparameter scalar-wave velocity FWI, which can be generalized into elastic FWIs. Xu et al. (2006) discuss different source-independent methods used in a frequency-domain elastic FWI, on 2D crosswell synthetic data. Gelis et al. (2007) linearize the inverse problem using Born and Rytov approximations in a 2D frequency-domain elastic FWI. Brossier et al. (2009) conduct a 2D frequency-domain elastic FWI on a 2D slice of the SEG/EAGE overthrust model by using frequencies above 1.7 Hz; to mitigate the nonlinearity and thus to improve the convergence, they apply frequency selection and multiple damping terms in the frequency domain. All the above frequency-domain elastic FWIs are limited to solving only for V_P and V_S by fixing ρ in the inversion. Abubakar et al. (2012) propose an elastic FWI scheme using a compressed implicit Jacobian matrix to invert V_P, V_S, and ρ simultaneously, but fail to give a good estimation of ρ. Choi et al. (2008) successfully recover V_P, V_S, and ρ simultaneously on the elastic Marmousi-2 model in a marine acoustic-elastic environment. To improve the reconstruction of ρ, Jeong et al. (2012) propose a hierarchical strategy to estimate the velocities V_P and V_S initially, and then subsequently to estimate the density and velocities simultaneously. They show successful reconstructions of V_P, V_S, and ρ on the elastic Marmousi-2 model and on a modified 2D SEG elastic salt model. However, they (Choi et al., 2008; Jeong et al., 2012) perform the frequency-domain inversions from very low frequencies (e.g., 0.167 Hz), which are usually not available in field data.

Because the time and frequency domains have their own advantages and disadvantages for the inversion, some researchers (Sirgue et al., 2008; Etienne et al., 2010) recently propose to combine both for FWI. Source and receiver wavefield extrapolations are performed using time-domain finite-difference (FD) or finite-element (FE) modeling, which is fast, memory-efficient, and easily parallelized, especially for 3D computations. Time windowing can be applied to the inversion to mitigate the nonlinearity, by gradually extending the window length or removing strong interference waves (e.g., surface waves or direct waves). The inversion is implemented in the frequency domain to calculate the objective function and gradients. The frequency data are calculated at each extrapolation time step using a discrete Fourier transform (DFT) (Sirgue et al., 2008; Etienne et al., 2010), or phase sensitive detection (Nihei and Li, 2007). The storage of the whole time-domain source wavefields at every subsurface grid point is greatly reduced to that for only a few frequency snapshots. Most importantly, the gradual low-to-high frequency-selection strategy can be naturally applied in the inversion to avoid cycle skipping.

In this paper, we use two strategies to improve the computational efficiency, to reduce the inversion nonlinearity, and to decouple V_P, V_S, and ρ in elastic inversion. First, a time-domain high-precision FE extrapolation is parallelized over shots with the message-passing interface (MPI), for forward and backward elastic wavefield propagations, and the inversion is implemented in the frequency domain. This improves the computational efficiency. Frequency-selection and time-windowing techniques are applied to mitigate the nonlinearity in the elastic FWI.

In the second strategy, to decouple V_P, V_S, and ρ in the inversion, we propose a multistep-length gradient approach to assign different optimal weights to each of the parameter gradients, and then we search for an overall optimal steplength along the composite gradient at each iteration. Because different model parameters have different contributions to the elastic multicomponent wavefields, it is reasonable to treat them with different weights (individual step lengths) in elastic FWI. For the elastic Marmousi-2 model, which has complicated structures, the elastic FWI scheme gives satisfactory results even using frequencies above 3 Hz. Because the approach always searches for a minimum using parabolic fits at each iteration, the convergence is stable and fast.

In the gradients and inverted results of different parameters, the long-wavenumber velocity results that contribute to propagation times (phases) and amplitudes of the full wavefields are often underestimated or smoothed in their variations, while the short-wavelength density result that contributes only to reflection amplitudes is overestimated or sharpened in its variations, correspondingly. Mora (1989) reveals the insight that inversion can be interpreted as a composite of a tomography and a migration to recover the low- and high-wavenumber parts of the model, respectively. Here, we distinguish between full-waveform inversion (FWI) and full-waveform tomography (FWT), which are often confused as being the same (Xu and Greenhalgh, 2010; Köhn et al., 2012). FWI tries to recover the complete wavenumber information using the full wavefields including phases and amplitudes, whereas FWT reconstructs the low-wavenumber velocities using predominantly phase (or traveltime) information rather than amplitude information. The recovered impedances (I_P and I_S) from the velocity and density products compensate for the under- or overestimations of the velocity and density variations, respectively, because the amplitudes in the data that are fitted to recover impedances depend on velocities and density. Therefore, the impedance results are closer to the correct ones than V_P, V_S, and ρ are separately.

TIME-DOMAIN FE MODELING

In the time domain, we use a high-precision elastic-wave FE method to simulate forward and backward wavefield propagations. The elastic-wave FE equation can be expressed in matrix form (Marfurt, 1984):

$$\mathbf{K}\hat{\mathbf{u}}(t) + \mathbf{M}\frac{\partial^2 \hat{\mathbf{u}}(t)}{\partial t^2} = \hat{\mathbf{f}}(t), \tag{1}$$

where $\mathbf{K}$ is the stiffness matrix, $\mathbf{M}$ is the mass matrix, $\hat{\mathbf{u}}(t)$ is the 2D 2C time-domain elastic wavefield, and $\hat{\mathbf{f}}(t)$ is the source vector.

Because the standard FE method has only second-order spatial accuracy, we apply a dispersion correction (Krenk, 2001) to reduce

the wavefield's spatial dispersion using an average mass matrix combined with the lumped and consistent mass matrices (Marfurt, 1984) at each explicit time step. The FE scheme is parallelized with MPI to distribute shot simulations over cluster nodes.

To calculate the frequency wavefields at every subsurface grid point, we apply the DFT equation at each step of time-domain wavefield extrapolations:

$$u(\omega) = \sum_{t=0}^{T} \hat{u}(t) e^{i\omega t}, \tag{2}$$

where $u(\omega)$ is the monofrequency snapshot at the circular frequency ω and T is the total recording time. Thus, a few selected frequency snapshots are calculated at each subsurface grid point for the gradient calculation. The gradual low-to-high frequency-selection strategy (Mora, 1987; Xu et al., 1995) can be applied. The whole time-domain source wavefields are recorded only at receivers to calculate the residual wavefields for backward propagation. The time-windowing strategy can also be applied in the time-domain residual wavefields at the receivers.

FREQUENCY-DOMAIN GRADIENT EQUATION

The frequency-domain FE modeling equation (Marfurt, 1984) corresponding to equation 1 is

$$(\mathbf{K} - \omega^2 \mathbf{M})\mathbf{u}(\omega) = \mathbf{S}\mathbf{u}(\omega) = \mathbf{f}(\omega), \tag{3}$$

where $\mathbf{S}$ is the complex-valued impedance matrix. In the FWI algorithm, the same stiffness $\mathbf{K}$ and mass $\mathbf{M}$ matrices in the time-domain equation 1 and frequency-domain equation 3 are defined piecewise and assembled from their element matrices $\mathbf{K}^e$ and $\mathbf{M}^e$ for each 2D rectangular element, respectively. According to Hamilton's principle in elastodynamic media (Nikishkov, 2007),

$$\mathbf{K}^e = \iint_e \mathbf{B}^T \mathbf{D} \mathbf{B} \mathrm{d}s, \quad \text{and} \quad \mathbf{M}^e = \iint_e \rho \mathbf{N}^T \mathbf{N} \mathrm{d}s, \tag{4}$$

where $\mathbf{N}$ is the Lagrangian bilinear rectangular function matrix; $\mathbf{B}$ is the displacement differentiation matrix, which is obtained with the spatial differentiation of $\mathbf{N}$; $\iint_e \mathrm{d}s$ represents the 2D spatial integral over each rectangular element; the superscript T indicates the transpose of the matrix; and $\mathbf{D}$ is the matrix of elastic coefficients (Nikishkov, 2007), which can be expressed using V_P, V_S, and ρ as

$$\mathbf{D} = \begin{pmatrix} V_\mathrm{P}^2\rho & (V_\mathrm{P}^2 - 2V_\mathrm{S}^2)\rho & 0 \\ (V_\mathrm{P}^2 - 2V_\mathrm{S}^2)\rho & V_\mathrm{P}^2\rho & 0 \\ 0 & 0 & V_\mathrm{S}^2\rho \end{pmatrix}. \tag{5}$$

The impedance matrix $\mathbf{S}$ is also defined piecewise and assembled from its element matrices; $\mathbf{S}_e = \mathbf{K}^e - \omega^2 \mathbf{M}^e$.

Based on equation 3, the least-squares objective misfit function for the elastic FWI is

$$\min(E) = \frac{1}{2} \sum_{\omega} \sum_{s} [\mathbf{u}_f(\mathbf{x}_R, \omega) - \mathbf{d}(\mathbf{x}_R, \omega)]^T \times [\mathbf{u}_f(\mathbf{x}_R, \omega) - \mathbf{d}(\mathbf{x}_R, \omega)]^*, \tag{6}$$

where $\mathbf{u}_f(\mathbf{x}_R, \omega)$ contains the predicted frequency data at the receivers $\mathbf{x}_R$, $\mathbf{d}(\mathbf{x}_R, \omega)$ contains the measured frequency data at the receivers $\mathbf{x}_R$, $\sum_\omega$ indicates summation over all the selected frequencies in the inversion, $\sum_s$ indicates summation over all shots, and the superscript * indicates the complex conjugate of the matrix.

Each element of the gradient vectors is the derivative of the objective function 6 with respect to a corresponding elastic parameter p_j at the subsurface grid element j. According to Shin and Cha (2009), the gradient can be calculated using a virtual-source normalization as

$$\frac{\partial E}{\partial p_j} \approx \sum_{\omega} \frac{\mathrm{Re}\left[\sum_s \mathbf{u}_f^T(\mathbf{x}, \omega) \frac{\partial \mathbf{S}^T}{\partial p_j} \mathbf{u}_b^*(\mathbf{x}, \omega)\right]}{\sum_s \mathbf{u}_f^T(\mathbf{x}, \omega) \frac{\partial \mathbf{S}^T}{\partial p_j} \frac{\partial \mathbf{S}^*}{\partial p_j} \mathbf{u}_f^*(\mathbf{x}, \omega) + \gamma}, \tag{7}$$

where $\mathbf{u}_f(\mathbf{x}, \omega)$ are the frequency data of the forward-propagated source wavefields at each subsurface grid point $\mathbf{x}$, $\mathbf{u}_b(\mathbf{x}, \omega)$ are the frequency data of the backward-propagated residual wavefields from receivers at each subsurface grid point $\mathbf{x}$, γ is a small damping term, and Re represents the operation to take the real part. Hence, we construct the gradient vectors $\mathbf{g}_{V_\mathrm{P}}$, $\mathbf{g}_{V_\mathrm{S}}$, and $\mathbf{g}_\rho$ independently for V_P, V_S, and ρ, respectively. In equation 7, $\frac{\partial \mathbf{S}}{\partial p_j}$ has nonzero coefficients only for the corresponding elastic parameters V_{P_j}, V_{S_j}, and ρ_j at the jth element. The element matrices $\frac{\partial \mathbf{S}_e}{\partial V_{\mathrm{P}_j}}$, $\frac{\partial \mathbf{S}_e}{\partial V_{\mathrm{S}_j}}$, and $\frac{\partial \mathbf{S}_e}{\partial \rho_j}$ at the jth element are solved as

$$\begin{aligned} \frac{\partial \mathbf{S}_e}{\partial V_{\mathrm{P}_j}} &= \frac{\partial \mathbf{K}_e}{\partial V_{\mathrm{P}_j}} = \iint_e \mathbf{B}^T \frac{\partial \mathbf{D}}{\partial V_{\mathrm{P}_j}} \mathbf{B} \mathrm{d}s, \\ \frac{\partial \mathbf{S}_e}{\partial V_{\mathrm{S}_j}} &= \frac{\partial \mathbf{K}_e}{\partial V_{\mathrm{S}_j}} = \iint_e \mathbf{B}^T \frac{\partial \mathbf{D}}{\partial V_{\mathrm{S}_j}} \mathbf{B} \mathrm{d}s, \quad \text{and} \\ \frac{\partial \mathbf{S}_e}{\partial \rho_j} &= \frac{\partial \mathbf{K}_e}{\partial \rho_j} - \omega^2 \frac{\partial \mathbf{M}_e}{\partial \rho_j} = \iint_e \mathbf{B}^T \frac{\partial \mathbf{D}}{\partial \rho_j} \mathbf{B} \mathrm{d}s \\ &\quad - \omega^2 \iint_e \mathbf{N}^T \mathbf{N} \mathrm{d}s, \end{aligned} \tag{8}$$

respectively.

MULTIPLE STEP LENGTHS

Theory

To update the model parameters, we implement the steepest-descent gradient approach by assigning different weights (individual step lengths) to the gradients of three elastic parameters V_P, V_S, and ρ. The update equation is

$$\mathbf{P}^{k+1}(V_\mathrm{P}, V_\mathrm{S}, \rho) = \mathbf{P}^k(V_\mathrm{P}, V_\mathrm{S}, \rho) - \alpha^k(\beta_1^k \mathbf{g}_{V_\mathrm{P}}^k + \beta_2^k \mathbf{g}_{V_\mathrm{S}}^k + \beta_3^k \mathbf{g}_\rho^k), \tag{9}$$

where $\mathbf{p}(V_\mathrm{P}, V_\mathrm{S}, \rho)$ is the model parameter vector including three parameters V_P, V_S, and ρ at every element; $k+1$ and k indicate the iterations; α^k is the overall step length at the kth iteration; β_1^k, β_2^k, and β_3^k are the individual step lengths (weights) at the kth iteration for the individual elastic parameters V_P, V_S, and ρ, respectively.

Similar to Pica et al. (1990) and Meles et al. (2010), we evaluate the optimal values of the individual weights β_j, for $j = 1, 2, 3$ by giving a small perturbation of the corresponding parameter to the current model along their own gradient directions $\mathbf{g}_{V_P}$, $\mathbf{g}_{V_S}$, and $\mathbf{g}_\rho$, respectively. Because the parameters V_P, V_S, and ρ are coupled more or less in the elastic FWI, the weights β_j that are estimated independently cannot guarantee the convergence of the inversion. We further use the overall step-length factor α to adjust the length of the composite gradient $(\beta_1^k \mathbf{g}_{V_P}^k + \beta_2^k \mathbf{g}_{V_S}^k + \beta_3^k \mathbf{g}_\rho^k)$, to guarantee the updated objective function to be the minimum at each iteration.

In the best situation, in which the parameters are completely uncoupled, and the individual weights β_j for different parameters are estimated very well, the value of α is, or very close to, 1. Normally, we give two trial values of α near 1 such as 0.7 (smaller than 1) and 1.3 (larger than 1) and solve for the objective functions at these two trial values. Then we perform (at most) two parabolic fits (Press et al., 1992; Vigh and Starr, 2008) to search for the optimal value of α using three values at each fitting, picked from 0, 0.7, 1.3, a new estimated one, or the maximum one. The maximum value of α is set far away from 1 (e.g., 2 or 3), in case the parabolic fits cannot locate the minimum value of the objective function below 1.3.

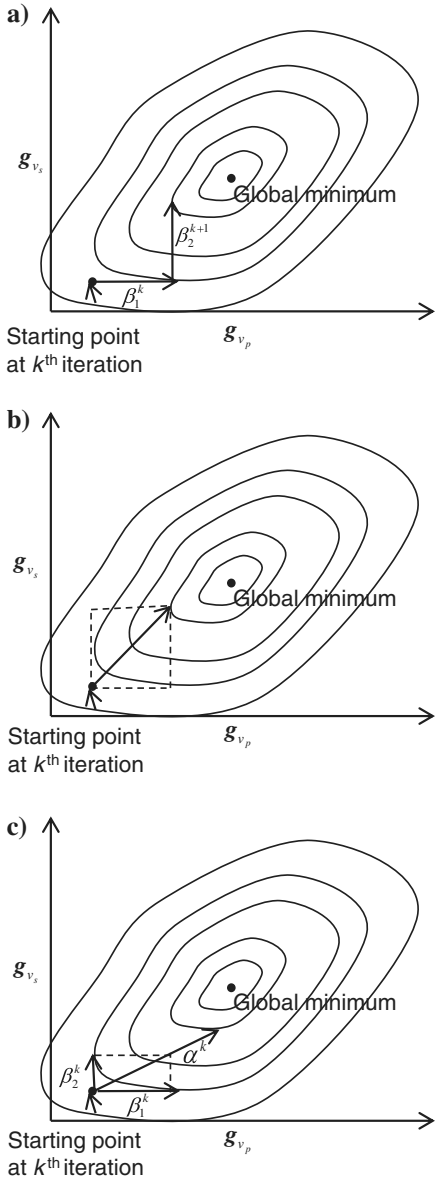

Figure 1. Three procedures are used to update two model parameters V_P and V_S: (a) one by one in turn, (b) simultaneously, all with the same weights, and (c) simultaneously with different weights.

Physical meaning

Here, we give a simple example to show the implementation and advantages of the proposed multistep-length gradient approach for multiparameter inversion. If we invert for two elastic parameters V_P and V_S, there are usually three possible procedures to update the model using the gradient method (see Figure 1). In Figure 1a, as Tarantola (1986) suggests, a minimum is searched along the V_P gradient $\mathbf{g}_{V_P}$, then only V_P is updated. Next, only V_S is updated in the same way. V_P and V_S are updated in turn until convergence. This approach is relatively stable but converges slowly. In Figure 1b, the difference between parameter types is ignored, and all parameters are inverted together with the same weights. By using more expensive optimization approaches such as Gauss-Newton or quasi-Newton algorithms (Pratt et al., 1998; Virieux and Operto, 2009), the convergence can be accelerated. However, it still has difficulties in inverting ρ along with V_P and V_S simultaneously because ρ has a very different behavior from V_P and V_S.

In Figure 1c, the gradients $\mathbf{g}_{V_P}$ and $\mathbf{g}_{V_S}$ are assigned different weights using our new scheme (or a subspace method) (Kennett et al., 1988; Sambridge et al., 1991). Using our new scheme, we treat them with different weights β_1 and β_2 at each iteration step; the weights are evaluated as the optimal individual step lengths to reach a minimum along their own gradient directions, respectively. Then, the overall optimal step length α is evaluated along the composite gradient direction using parabolic fits. This scheme guarantees a suitable composite gradient direction (adjusted by β_1 and β_2) with an optimal overall step length α at each step. Therefore, its convergence is fast and stable for simultaneous reconstruction of different parameters (including ρ); each only needs to be assigned a different weight.

The subspace method (Kennett et al., 1988; Sambridge et al., 1991) is similar to our multistep-length gradient approach to assign different weights for different parameter gradients, but without the overall step length α. Because the cross variances between different model parameter types are generally impossible to estimate at each subsurface grid point, to resolve different parameters simultaneously in the inversion, the subspace approach normally assumes that the subsurface parameter types are completely independent (Sambridge et al., 1991; Sakai, 2011). The assumption is opposed to most real cases that subsurface parameters are correlated to some extent. Brossier et al. (2009) and Sakai (2011) conclude that the subspace approach is not better than the standard quasi-Newton or preconditioning conjugate gradient methods in elastic inversion for V_P and V_S. In comparison, our new scheme uses one more overall step length α to always determine the "best" length for the composite gradient direction, so it does not have the limitation of the independent-parameter assumption. Without this limitation, the multistep-length gradient approach can be expanded into the inversion for any combinations of multiple parameters.

NUMERICAL TESTS

Elastic Marmousi-2 model and experiment setup

We test our elastic FWI on a modified elastic Marmousi-2 model. The original Marmousi-2 model (Martin et al., 2002) has a water layer ($V_S = 0$) and very low V_S values (i.e., high Poisson's ratios), which requires a large number of grid points with a small grid size in wavefield simulations. To simulate an onshore environment and avoid the computational burden, we remove the upper water layer and retain only the central portion of the V_P and ρ models from the original Marmousi-2 model. We also regenerate the V_S model from the original V_P model using two V_S/V_P ratios:

$$V_S = \begin{cases} 0.6V_P & V_P < 3.5 \text{ km/S}, \\ \frac{1}{\sqrt{3}}V_P & V_P \geq 3.5 \text{ km/S}. \end{cases} \quad (10)$$

The nonlinearity of the inversion is increased by using two V_S/V_P ratios, rather than only one constant ratio. Figure 2 shows the modified Marmousi-2 V_P, V_S, and ρ models with the dimensions of 9.2×3.04 km and with a uniform grid interval of 0.01 km. Hence, the total number of the grid elements for the FE modeling is $nxe * nze = 920 \times 304$, where nxe is the number of grid elements along the horizontal direction and nze is along the vertical direction. The total number of the unknowns in the inversion is $920 \times 304 \times 3$ for V_P, V_S, and ρ. To reduce the inversion nonlinearity, each gradient vector (e.g, $\mathbf{g}_{V_P}$, $\mathbf{g}_{V_S}$, or $\mathbf{g}_\rho$) to update the corresponding model parameter is discretized into a coarser grid mesh as 460×152, with a larger grid interval of 0.02 km. In Figure 3, the corresponding correct P- and S-impedance (I_P and I_S) models are also calculated from the products of the velocities and density (Figure 2).

To remove unwanted model edge reflections, for our FE modeling and inversion, we put a hybrid absorbing boundary condition composed of a damping zone (Shin, 1995), and a second-order one-way wave-equation boundary condition (Higdon, 1991), along the bottom, left, and right edges of the models. To further remove free-surface effects (i.e., surface waves and body-wave reflections from the free surface) that increase the nonlinearity of the inversion (Brossier et al., 2009), we also place the hybrid absorbing condition on the top of model. To simulate onshore synthetic data on the Marmousi-2 model (Figure 2), 111 explosive sources along the top surface of the model are spaced every 0.08 km, and are excited in turn with a Ricker wavelet with peak frequency of 6 Hz and maximum frequency of about 13 Hz. For every shot, 231 2C receivers are spaced every 0.04 km along the top surface. The survey line is long enough to record all wave types (without surface waves) and P-S conversions, which are all accounted for in the elastic FWI. For comparison, the acoustic FWI has many problems with long-offset elastic data (Mulder and Plessix, 2008).

To provide a starting model (Figure 4) for the following FWI tests, the correct model (Figure 2) is smoothed using a 2D Gaussian filter with a correlation distance of 1.6 km. Thus, all the small-scale structures vanish, and only the large-scale (low-wavenumber)

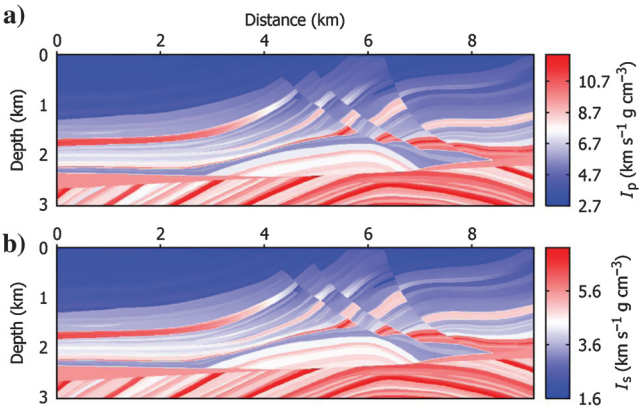

Figure 3. The target model impedances for (a) I_P and (b) I_S from the products of the velocities and density in Figure 2.

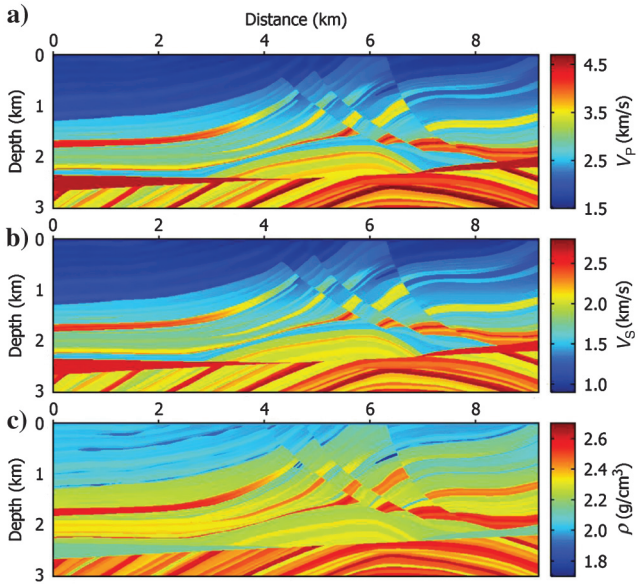

Figure 2. The target elastic Marmousi-2 models for (a) V_P, (b) V_S, and (c) ρ.

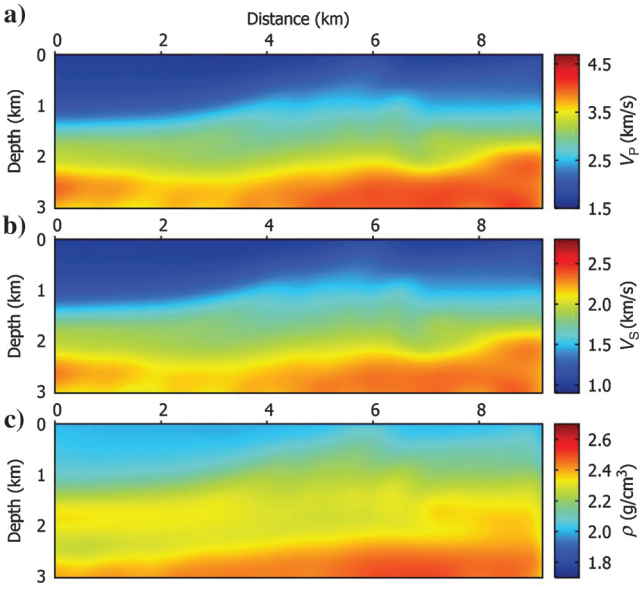

Figure 4. The smooth starting models for (a) V_P, (b) V_S, and (c) ρ for the elastic FWI.

structures are present. Because very low frequencies are usually unavailable in field data, we use frequencies above 3 Hz for all inversion tests. The requirement here is that the initial predicted data must match the measured data to within a half-cycle at the lowest available frequency (e.g., here, 3.09 Hz) (Mora, 1987). More highly smoothed starting models are also feasible for the inversion algorithm, if we can start the inversion from lower frequencies (e.g., <3 Hz) to retrieve the corresponding lower wavenumber velocity information, which is missing in the starting model. For field data, the starting model can be estimated using migration velocity analysis or kinematic tomography for V_P, and empirical geologic relations for V_S and ρ, or Laplace-domain inversion (Chung et al., 2008) for all V_P, V_S, and ρ.

FWI test on FE synthetic data

In this subsection, we use our high-precision dispersion-corrected elastic FE modeling to generate time-domain 2C (vertical and horizontal) particle-displacement wavefields for 111 explosive shots, recorded at 4096 time samples with the increment of 0.0015 s. Figure 5 shows the 2C elastic particle-displacement whole wavefields (without surface waves) for the 50th shot. For the inversion in this subsection, the source wavelet is assumed to be known a priori (in the later example, it is not). The elastic FWI using the multistep-length gradient approach straightforwardly works well on the whole test wavefields (Figure 5), partly because the extrapolations for the test data and the predicted data in the inversion are performed using the same time-domain FE modeling engine (in the later example, it is not). Hence, we do not apply the time-windowing strategy here.

Before inversion, we calculate the scaled (i.e., multiplied by step lengths $\alpha\beta_i$, $i = 1, 2, 3$) gradients $\mathbf{g}_{V_P}$, $\mathbf{g}_{V_S}$, and $\mathbf{g}_\rho$ (Figure 6) using all FE-synthetic shot gathers, at the first iteration of the elastic FWI and using 29 frequencies with an equal interval of 0.326 Hz in the frequency band of 3.09–12.2 Hz. The frequency increment of 0.326 Hz is double that defined by Nyquist sampling of the time-domain synthetic input shot gathers (e.g., Figure 5). The frequency increment is small enough to prevent frequency leakage, as long as the time length of the shot profiles (Figure 5) is long enough to record all relevant waves. We do not use the uneven frequency sampling strategy (Sirgue and Pratt, 2004) because it is not suitable for the fast Fourier transform (FFT), which is used for the later source wavelet estimation at the receivers. The gradient calculation can be considered as an elastic least-squares migration (Chung et al., 2011), but the velocity and density gradients have quite different characteristics. The velocity gradients (Figure 6a and 6b) recover the low-wavenumber spatial information with lower resolutions; they have higher amplitudes at shallower depths because the main contributions are from direct waves and refractions. The density gradient (Figure 6c) recovers the high-wavenumber spatial information with high resolution. It has a balanced amplitude (fairly independent of depth) background mainly from reflections; the shallow part is not affected by the strong energy of direct waves and refractions. This conforms to the elastic wavefield theory that the density variations affect only the reflection and transmission coefficients between layers to change the amplitudes of the propagated wavefields (mainly reflections recorded at the surface receivers), but they do not affect the phases or traveltimes of the whole wavefields except for the phase shifts of postcritical reflections.

Then, we apply the frequency-selection strategy in two phases (for two frequency ranges) to avoid waveform cycle-skipping fitting in the elastic FWI. In phase I, 15 lower frequencies with an equal interval of 0.326 Hz from 3.09 to 7.65 Hz are used in the elastic FWI for 30 iterations, which uses the smooth model (Figure 4) as the starting model. In phase II, 15 higher frequencies from 7.65 to 12.2 Hz are used in the elastic FWI for 30 iterations, which uses the final results in phase I as the initial model. Figure 7 shows the final results after the two phases; V_P, V_S, and ρ are all well resolved. The velocity results (Figure 7a and 7b) are better resolved at shallow depths, because the strong shallow direct waves and refractions dominate their shallow reconstructions. In contrast, the density result (Figure 7c) has the highest resolution with a balanced amplitude background, which is similar to a depth migration. This is because it reconstructs the high-wavenumber part of the whole model using the amplitude information mainly from reflections.

Figure 8 shows the comparisons between the target (red lines), the starting (green lines), and the inverted (blue lines) parameter

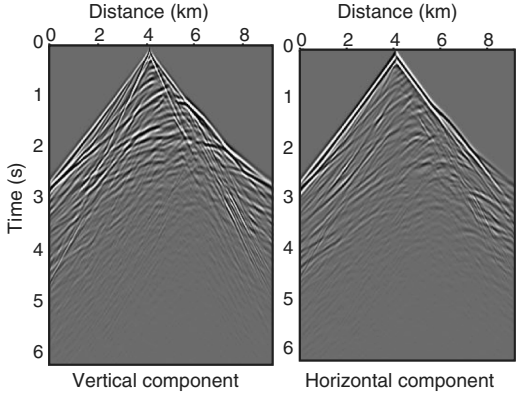

Figure 5. (a) The synthetic vertical-component and (b) horizontal-component elastic particle-displacement wavefields (without surface waves) for the 50th shot, using FE modeling.

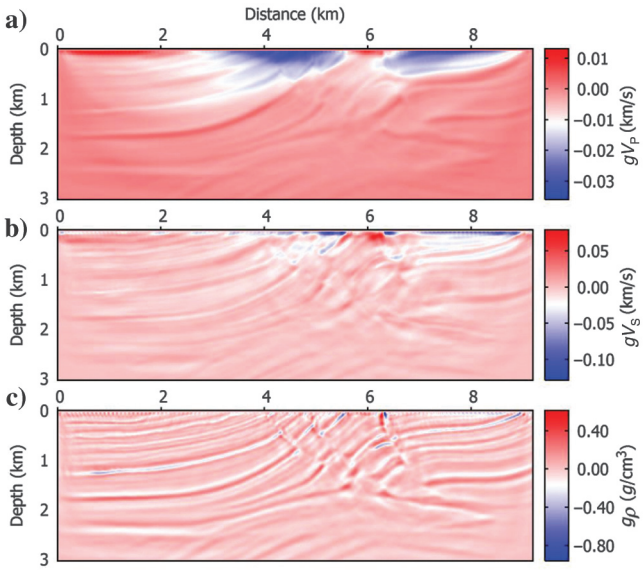

Figure 6. The scaled gradients (multiplied by step lengths $\alpha\beta_i$ $i = 1, 2, 3$) using all FE-synthetic shot gathers, at the first iteration and using 29 frequencies in the range of 3.09–12.2 Hz.

values in the depth profiles at the horizontal positions $x = 3.5$ km and 6.5 km in the models in Figures 2, 4, and 7. The V_P and V_S variations are underestimated (smoothed), especially in depth, while the density variations are overestimated (sharpened) correspondingly; for example, the under- and overestimations are obvious at 2.5 km depth, in the depth profile at $x = 3.5$ km. The density result (Figure 7c) is completely different from that which would be approximated from the recovered V_P (Figure 7a) using an empirical relation; it does not have the same smooth (or underestimated) variations that the V_P result does. Because the density and velocities are partly coupled (in their influence on the reflection amplitudes), the variations in the inverted density (Figure 7c) are overestimated to compensate for the underestimation of the variations in the inverted velocities (Figure 7a and 7b) to make the amplitudes of the predicted data a better match to those of the input data.

As the velocity deviations are often underestimated and the density deviations are correspondingly overestimated, we calculate the impedances I_P and I_S (Figure 9) from their products. The recovered impedances (Figure 9) are closer to those of the target model (Figure 3) in structural and physical meanings, than V_P, V_S, and ρ (Figure 7) are separately. Figure 10 contains the comparisons between the target (the red lines), the starting (the green lines), and the recovered (the blue lines) impedances in the depth profiles at the horizontal positions $x = 3.5$ km and 6.5 km in the models in Figures 3 and 9. Note, the starting impedance models are not shown before because the impedances are calculated from the velocities and density. The recovered impedances (Figure 10) are better matched to the correct ones than V_P, V_S, and ρ (Figure 8) are, separately.

To further evaluate the inverted results quantitatively, we calculate the posterior model errors separately for the elastic parameters V_P, V_S, ρ, I_P, and I_S, using

$$\text{error}_{\text{model}} = \sqrt{\frac{1}{nxe * nze}\left\| \frac{\mathbf{p}_{\text{inverted}} - \mathbf{p}_{\text{true}}}{\mathbf{p}_{\text{true}}} \right\|}, \qquad (11)$$

where $\mathbf{p}_{\text{true}}$ and $\mathbf{p}_{\text{inverted}}$ are the true and inverted parameter vectors with the dimension of $nxe * nze$ for each of the elastic parameters, respectively. If the error is smaller, the corresponding parameter is closer to that of the target. To judge which parameter is reconstructed better in inversion, we also calculate the error decrease (the initial model error minus the inverted model error). Table 1 gives the initial and inverted model errors, and the error decreases for V_P, V_S, ρ, I_P, and I_S. The model error of ρ is increased during the elastic FWI because of the overestimation of the density

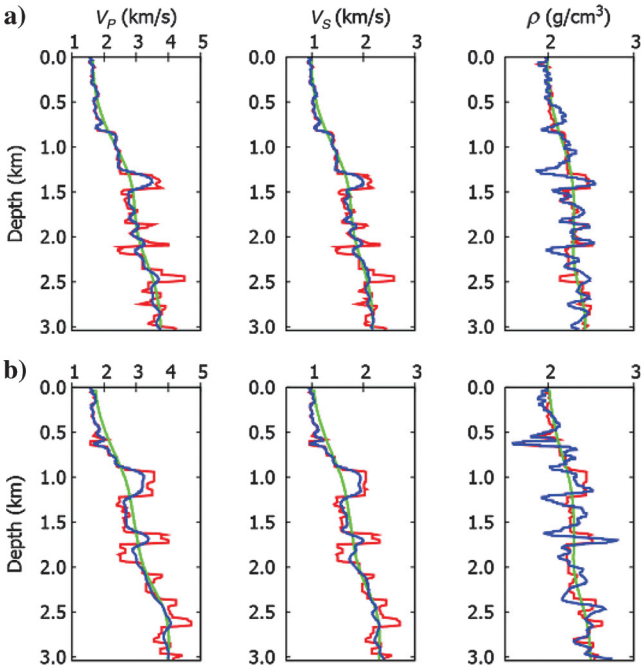

Figure 8. Comparisons between the target (red lines), starting (green lines), and inverted (blue lines) parameter values in the depth profiles at (a) $x = 3.5$ km, and (b) $x = 6.5$ km in the models in Figures 2, 4, and 7, respectively.

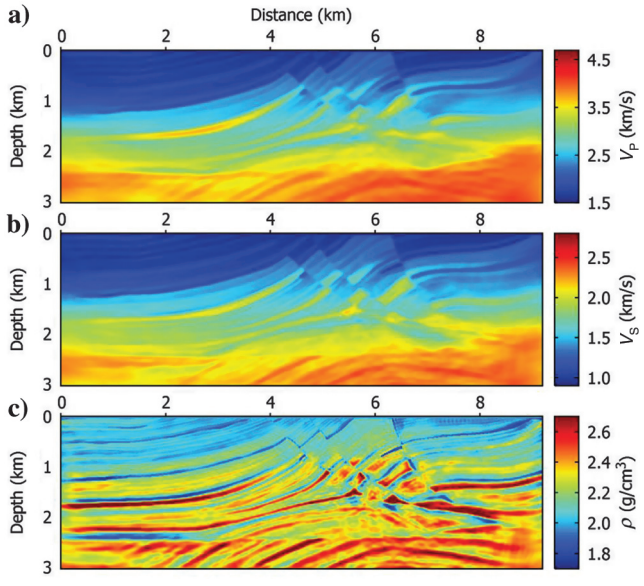

Figure 7. The inverted results for (a) V_P, (b) V_S, and (c) ρ using the whole FE-synthetic shot gathers (see Figure 5), after phases I and II (using 15 frequencies from 3.09 to 7.65 Hz and from 7.65 to 12.2 Hz, with 30 iterations each).

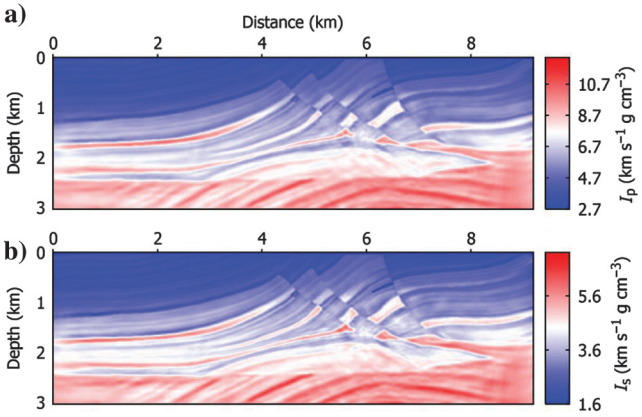

Figure 9. The recovered impedances for (a) I_P and (b) I_S, from the products of the velocity and density results in Figure 7.

variations, which is the opposite of the other parameters' behaviors. We consider the density reconstruction to be more like a depth migration than an inversion (Mora, 1989). Noticeably, among all the parameters, the impedances have the largest error decreases; their results are the best matched to the correct ones.

In Figure 11, the 2C (vertical and horizontal) elastic whole wavefields for the 50th shot are simulated for the starting model (Figure 4) and the final model (Figure 7) using our FE modeling. The corresponding initial and final residuals between themselves and the measured shot gather (Figure 5) are also displayed. All the figures are plotted using the same amplitude level. The final time-domain data residuals (Figure 11d) are significantly decreased by the inversion.

Figure 12 shows the normalized least-squares errors of the frequency data residuals (i.e., the objective function in equation 6 with a normalization that makes the error at iteration 1 to be 1) for phases I (the dashed line) and II (the dotted line), with 30 iterations each. At each iteration, the elastic FWI performs two extrapolations for the gradient calculation, three for the estimation of the individual weights β_i, $i = 1, 2, 3$, and three or four for the estimation of the overall step length α (equation 9). Thus, the elastic FWI needs three more extrapolations than other elastic FWIs using the standard steepest-descent gradient approach, which do not assign different weights for different parameter gradients. However, our elastic FWI has improved convergence and stability by using optimal step lengths; the total number of iterations is decreased to a few tens from the several hundreds that most other elastic FWIs need (Jeong et al., 2012; Köhn et al., 2012). The computation for all 60 iterations for the elastic FWI is done in parallel over sources using MPI on 56 CPU cores of an Intel Xeon X5650 cluster, using about 115 hours elapsed time.

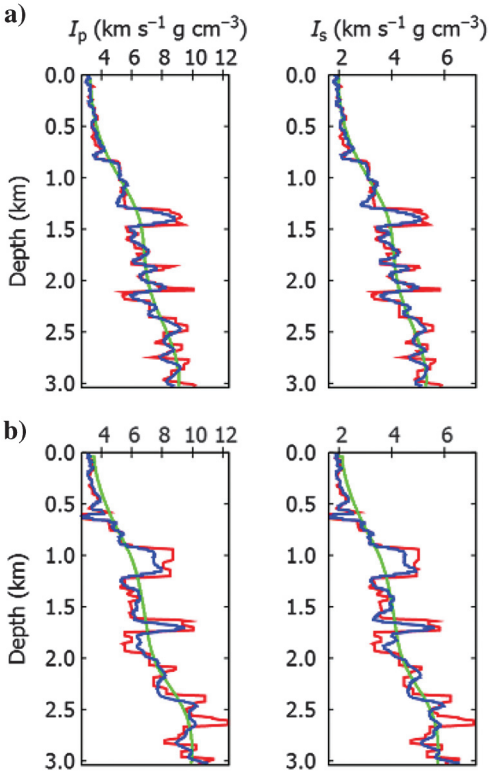

Figure 10. Comparisons between the target (the red lines), starting (the green lines), and recovered (the blue lines) impedances in the depth profiles at (a) $x = 3.5$ km and (b) $x = 6.5$ km in the models in Figures 3 and 9. Note that the starting impedance models are not shown before because the impedances are calculated from the velocities and density.

FWI test on FD synthetic data

To verify the robustness of our elastic FWI, we use a 2D P-SV-wave fourth-order stress-velocity staggered-grid FD modeling algorithm (Virieux, 1986; Levander, 1988) to produce onshore time-domain elastic particle-velocity wavefields for 111 explosive shots, recorded at 4096 time samples with the increment of 0.0015 s. The perfectly matched layer absorbing boundary conditions (Berenger, 1994) are placed along the four edges of the model to remove unwanted model edge reflections and highly nonlinear free-surface effects at the top surface. After the FD simulation, to adjust the FD and FE modeled wavefields to the same amplitude level, all FD-synthetic shot gathers are scaled by a factor of −50. Figure 13 shows the 2C particle-velocity whole wavefields (without surface waves) for the 50th shot, using FD modeling.

The FD and FE modeling schemes are different. FE modeling that simulates particle-displacement wavefields can also generate the particle-velocity wavefields, if the source wavelet is changed to its time derivative. However, the source wavelet used for FE extrapolations in the inversion is not equal to the time derivative of that used for the FD synthetic data (see Figure 14), due to their different modeling engines. To match the FD synthetic wavefields (Figure 13) in our elastic FWI with the FE modeling engine, we estimate the time-domain source wavelet for each shot gather at each iteration. Using FFT, the source-wavelet estimation is performed at every discrete frequency in the full frequency spectrum (e.g., 0.326–29.3 Hz) at the receivers (Song et al., 1995; Xu et al., 2006), and they are then transformed back to a time-domain source wavelet for the later time-domain wavefield extrapolations. Because FFT runs much faster than DFT on the whole time traces at the receivers to solve for the full frequency spectra, the source wavelet estimation is fast and efficient. Figure 14 shows the estimations of the source wavelet for the 50th shot gather. At the first iteration, starting from the first derivative of a Gaussian function wavelet (the green line), the estimated wavelet (the blue line) is quite good, with some phase errors at the later times. At the final iteration (24th iteration in phase II), the phases (of the yellow line) match well, and the amplitudes are a little higher than those of the correct Ricker

Table 1. Initial and inverted errors, and error decreases in the elastic FWI, using the FE-synthetic data.

Parameter	Initial error (%)	Inverted error (%)	Error decrease (%)
V_P	11.9	9.49	2.37
V_S	11.1	8.67	2.41
ρ	3.70	5.65	−1.95
I_P	14.2	10.4	3.81
I_S	13.4	9.67	3.78

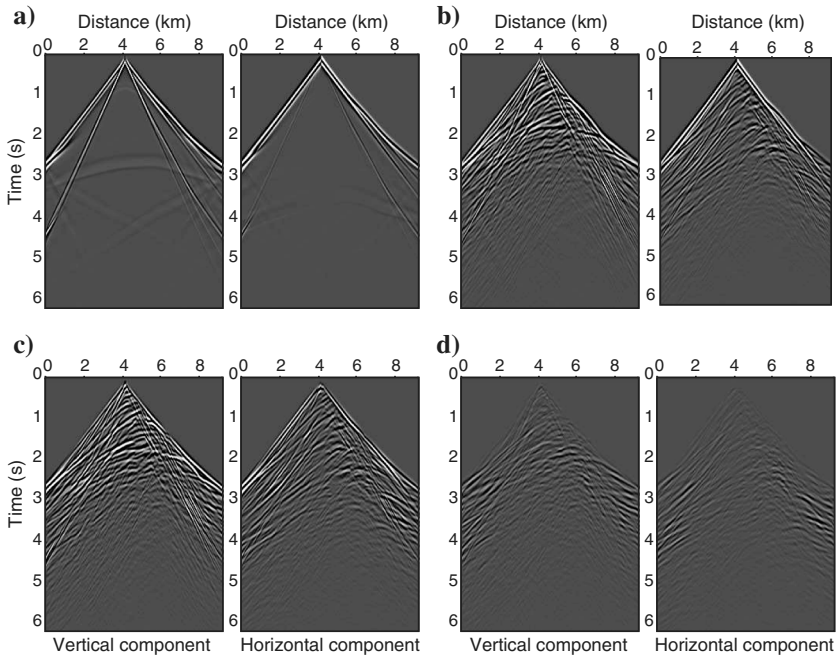

Figure 11. The predicted 2C elastic wavefields for (a) the starting model (Figure 4) and (b) the final result (Figure 7) for the 50th shot, using the FE modeling; and the corresponding (c) initial and (d) final residuals between themselves and the measured shot gather (Figure 5). Note that in each subfigure, the left panel is the vertical component wavefield and the right panel is the horizontal component. The amplitude scaling is the same as that in Figure 5.

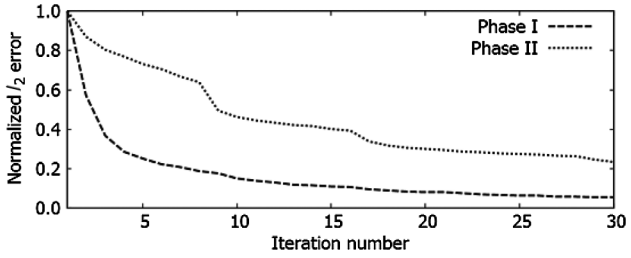

Figure 12. The normalized least-squares (l_2) errors of the frequency data residuals for the inversion using the FE-synthetic shot gathers.

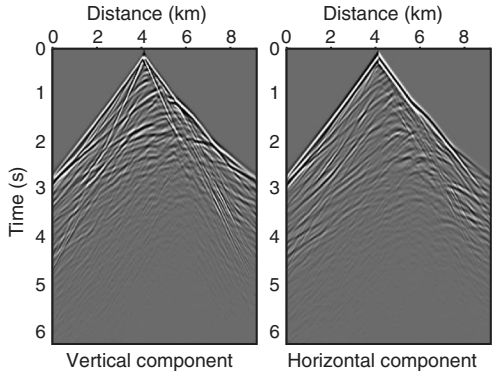

Figure 13. The (a) synthetic vertical-component and (b) horizontal-component elastic particle-velocity wavefields (without surface waves) for the 50th shot, using a staggered-grid FD modeling and scaled by a factor of −50. Compare with Figure 5.

wavelet (the red line) to compensate for errors in the estimated model velocities and density.

From tests, the inversion still converges very slowly if using the whole FD synthetic shot gathers because of strong artifacts around the source locations associated with their inconsistent source spatial patterns (combinations of source vectors in space that are implemented for numerical calculations). Therefore, we apply the time-windowing strategy to mute out the strong direct waves or refractions from the inversion, for which the waveforms are severely affected by source spatial patterns, especially around source locations.

Because the density initial model for inversion is difficult to estimate for field data, we consider the elastic FWI by using a constant density of 2.25 g/cm^3 (the average value of the density model) (Figure 2c) as the density starting model and the same velocity starting models and frequencies as those in the previous test. Before inversion, we calculate the scaled gradients $\mathbf{g}_{V_P}$, $\mathbf{g}_{V_S}$, and $\mathbf{g}_\rho$ (Figure 15) using all time-windowed FD-synthetic shot gathers, at the first iteration and using 29 frequencies from 3.09 to 12.2 Hz. The density gradient (Figure 15c) is similar to the previous one (Figure 6c) because the density variations are mainly related to the amplitudes of reflections. Compared with the previous velocity gradients (Figure 6a and 6b), the current gradients (Figure 15a and 15b) have higher resolutions and are less influenced by strong direct waves and refractions because most of them are muted out in the inversion.

We perform the inversion in the same two phases (frequency groups) as the previous ones. In phase II, the inversion stops at the 24th iteration as the convergence becomes slow. The final results of V_P, V_S, and ρ after two phases of inversion are shown in Figure 16. Even for the FD particle-velocity synthetic data,

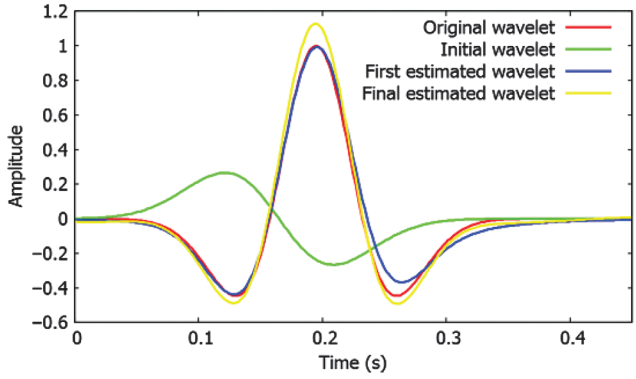

Figure 14. The time-domain source wavelets. The original Ricker source wavelet (the red line) used for the FD-synthetic wavefields (Figure 13); the initial source wavelet of the first derivative of the Gaussian function (the green line) as the input for the first iteration of the inversion, the estimated Ricker wavelet (the blue line) after the first iteration of the inversion, and the estimated Ricker wavelet (the yellow line) after the final iteration (the 24th iteration in phase II) of the inversion.

by using the frequency-domain source wavelet estimation (Song et al., 1995; Xu et al., 2006) and the time windowing to remove direct waves and refractions, the elastic FWI still gives the satisfactory velocity results (Figure 16a and 16b) similar to the previous one (Figure 7a and 7b). Because of the removal of direct waves and refractions, the qualities of the velocity results (Figure 16a and 16b) become somewhat worse near both sides of the model, while the resolutions in depths are enhanced. In addition, because the velocity reconstructions try to retrieve the long-wavelength spatial information, their results (Figure 16a and 16b) are not significantly affected by using a constant density initial model. The density result (Figure 16c) still gives a correct structural image with high resolution, even though it has an inaccurate low-wavenumber background. The density reconstruction is similar to a depth migration to reveal the short-wavelength density perturbation information, as long as the long-wavelength velocities are adequately recovered.

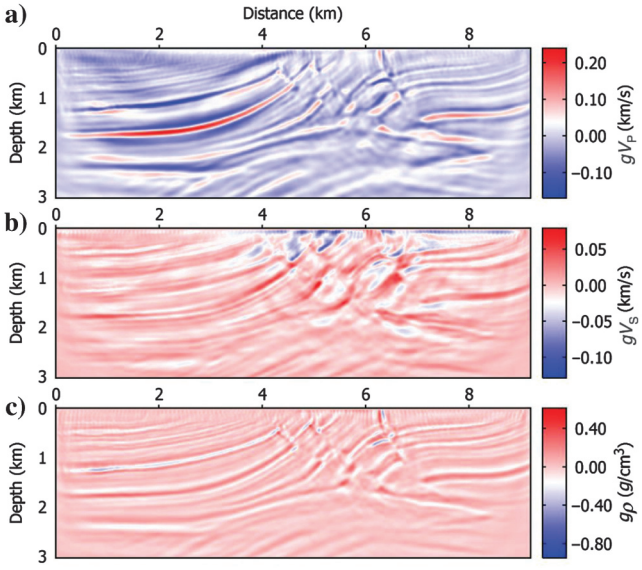

Figure 15. The scaled gradients (multiplied by step lengths $\alpha\beta_i$ $i = 1, 2, 3$) using all time-windowed FD-synthetic shot gathers, at the first iteration and using 29 frequencies in the range of 3.09–12.2 Hz.

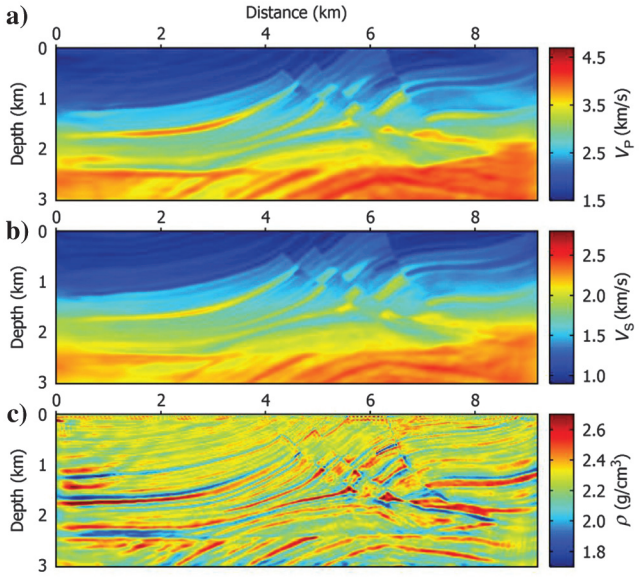

Figure 16. The inverted results for (a) V_P, (b) V_S, and (c) ρ by time-windowing the FD-synthetic shot gathers (see Figure 13) to remove direct waves and refractions, after phase I (using 15 frequencies from 3.09 to 7.65 Hz, with 30 iterations) and phase II (using 15 frequencies from 7.65 to 12.2 Hz, with 24 iterations).

Figure 17 shows the comparisons between the target (red lines), the previously inverted (green lines), and the currently inverted (blue lines) parameter values in the depth profiles at the horizontal positions of $x = 3.5$ km and $x = 6.5$ km in the models in Figures 2, 7, and 16, respectively. The velocity results (the left two panels) for both tests (green and blue lines) are very similar. The density results (the right panels) for both tests (green and blue lines) have different backgrounds, while their short-wavelenegth variations are well matched and close to the true ones (red lines).

The recovered impedances I_P and I_S (using the constant density starting model) in Figure 18 are again closer to the correct ones (Figure 3) in structural and physical meanings, than V_P, V_S, and ρ (Figure 16) are separately. Figure 19 shows the comparisons between the correct (the red lines), the previously recovered (the green lines), and the currently recovered (the blue lines) impedances in the depth profiles at the horizontal positions of $x = 3.5$ km and 6.5 km in the models in Figures 3, 9, and 18, respectively. Because the errors of the density results are smaller than those of the velocity results, the density errors have lower impacts on the impedances. The currently recovered impedances (the blue lines) are close to the previously recovered ones (the green lines), and they are only a little worse matched to the target impedances (the red lines).

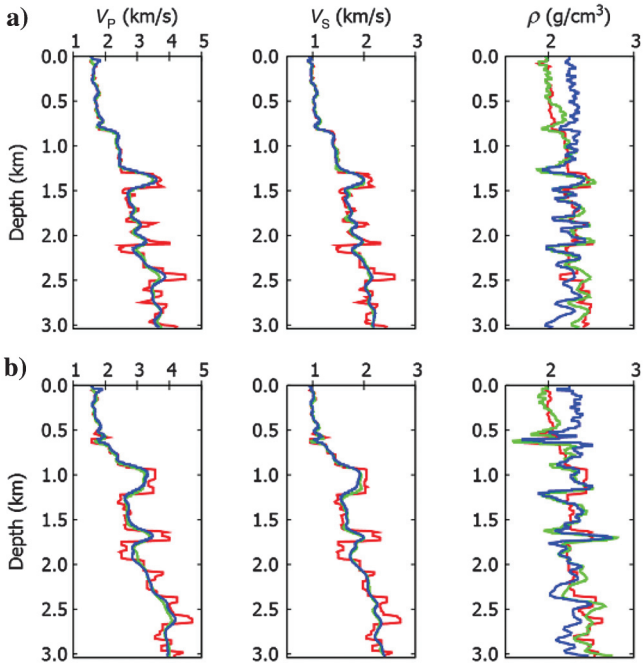

Figure 17. Comparisons between the target (red lines), the previously inverted (green lines), and the currently inverted (blue lines) parameter values in the depth profiles at (a) $x = 3.5$ km and (b) $x = 6.5$ km in the models in Figures 2, 7, and 16, respectively.

Table 2 gives the initial and inverted model errors and the error decreases for different parameters in this test. The model error of ρ is again increased in the elastic FWI because of the overestimation of density deviations and a constant initial model that is far from the correct one. Noticeably, among all the parameters, the impedances still have the largest error decreases. Compared with Table 1, V_P and V_s in Table 2 are better resolved (especially for V_P) with smaller inverted errors. This is because the velocity reconstructions are not dominated by strong shallow direct waves and refractions in this example because most of them are removed by using the time windowing.

The 2C elastic whole wavefields for the 50th shot are simulated using our FE modeling engine with the corresponding estimated source wavelets, for the starting model (Figure 4), and the final result (Figure 16). Because the synthetic wavefields are similar to those in the previous test (Figure 11a and 11b), we do not display them again. Compare the whole-wavefield initial (Figure 20a) and final (Figure 20b) residuals and the time-windowed initial (Figure 20c) and final (Figure 20d) residuals that are used in the inversion. Strong residuals exists for the direct waves and refractions (Figure 20a), and are not reduced after the inversion (Figure 20b). This is because the FD modeling for the synthetic measured data (Figure 13) and the FE modeling engine for the wavefield extrapolations of the inversion have different source spatial patterns. Using the time-windowing technique to mute out the direct waves and refractions, the residuals of the remaining reflections and diffractions are obviously reduced by the inversion (Figure 20d).

Figure 21 shows the normalized objective function of the frequency data residuals in the two phases; the convergence is still fast and stable. All 54 iterations are performed in parallel in the same cluster as in the previous test, using about 132 hours elapsed time. The reason for a longer run time even for fewer iterations (than 60) in this example is that the source-wavelet estimation and some additional wavefield simulations for the second parabolic fits at iterations to search for an optimal value of α are required at each iteration.

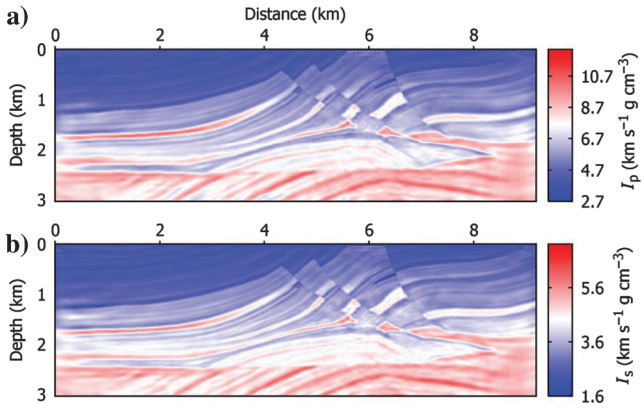

Figure 18. The recovered impedances for (a) I_P and (b) I_S, from the products of the velocity and density results in Figure 16.

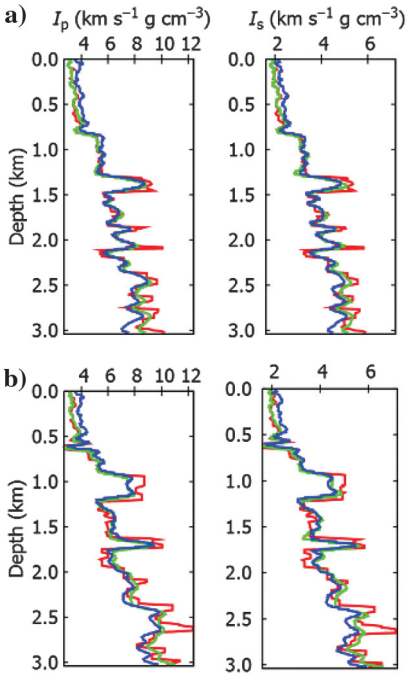

Figure 19. Comparisons between the target (red lines), the previously recovered (green lines), and the currently recovered (blue lines) impedances in the depth profiles at (a) $x = 3.5$ km and (b) $x = 6.5$ km in the models in Figures 3, 9, and 18, respectively.

DISCUSSION

Using a parallelized time-domain modeling engine for wavefield extrapolations, the FWI algorithm is more efficient than a pure frequency-domain FWI algorithm for multiple-frequency simulations. Because the inversion is implemented in the frequency domain, the FWI needs storage for only a few frequency snapshots, not for all the source time-domain wavefields. For comparison, the pure time-domain parallelized inversion conventionally saves the whole source wavefields at all subsurface grid points, so storage in each cluster node is usually unavoidable even when using optimal approaches. However, in our hybrid FWI scheme, the DFT calculation at each time step for all subsurface grid points is more expensive for more frequencies, which needs to be considered for improvement in future research.

In the above FE and FD synthetic tests, the elastic FWI is robust by using the time-windowing and frequency-selection strategies. Unlike the usual one single point in acoustic-wave modeling, the source function has different spatial patterns (different combinations of the source vectors in space) in elastic-wave modeling, which have an important impact on the success of the inversion. It is sometimes hard to match the source spatial pattern of the

Table 2. Initial and inverted errors, and error decreases in the elastic FWI, using the FD-synthetic data.

Parameter	Initial error (%)	Inverted error (%)	Error decrease (%)
V_P	11.9	8.79	3.08
V_S	11.1	8.57	2.51
ρ	7.63	9.13	−1.49
I_P	15.9	12.4	3.46
I_S	15.3	11.7	3.61

observed elastic data, but it is easy and feasible to mute out direct waves (or refractions) using time windowing to remove the unmatched effects in inversion. However, as long as the modeling engine can match the observed source patterns well (as in the first test), this step is not required, and more low-wavenumber information is then available for inversion in the shallow part of the model. The time windowing can also be used to remove surface waves for field data. Meanwhile, a stable source wavelet estimation is performed at each inversion iteration in the full frequency spectrum to obtain a complete time-domain wavelet via an FFT, for the time-domain source and receiver wavefield extrapolations. Here, the reason to choose FFT not DFT to perform the Fourier transformation is that FFT runs faster on the whole time traces at the receivers to solve for the full frequency spectra.

To decouple different parameters by assigning corresponding individual weights, the multistep-length gradient approach reconstructs multiple (velocities and density) parameters simultaneously. Using one more overall optimal steplength along the composite gradient direction, the multistep-length gradient approach always guarantees a stable convergence, and it does not have the limitation of the independent-parameter assumption that is normally present in the subspace method. Therefore, our multistep-length gradient approach is general, and it can be expanded, in a straightforward manner, for use with any multiple-parameter media such as acoustic-density, viscous, anisotropic, or porous. The multistep-length gradient approach is also feasible for any type of FWI such as 2D or 3D, in the frequency or time domains, and for land or marine environments, as long as a suitable FE or FD modeling engine is implemented for wavefield extrapolations.

In the inversion, the density reconstruction tries to image subsurface high-wavenumber density perturbations and has the highest resolution among all three reconstructions, using the amplitude information mainly from reflections. The density gradient imaging using all effective frequencies at the first iteration can be considered as an elastic least-squares migration to provide the subsurface structural information of the density distribution.

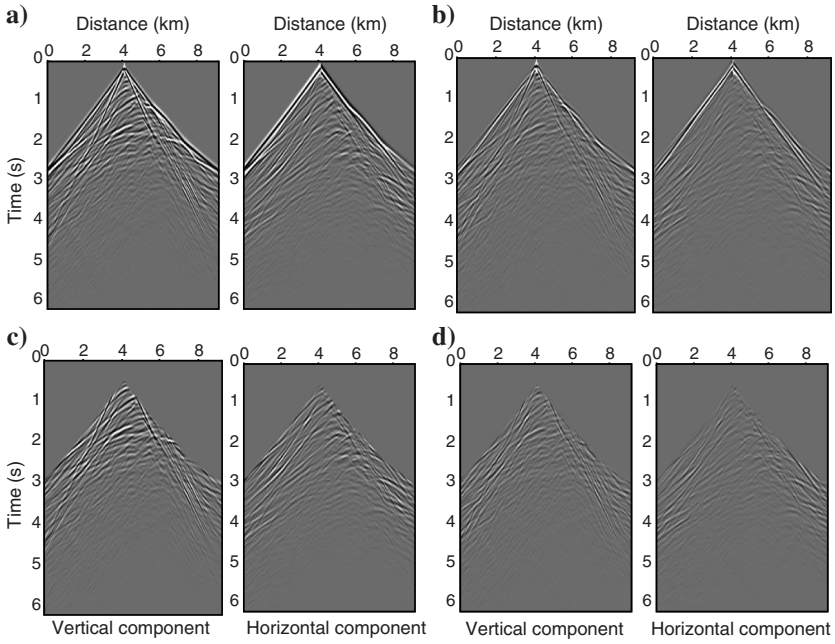

Figure 20. The whole (a) initial and (b) final residuals between the FE predicted and the FD measured wavefields (Figure 13), for the 50th shot and the corresponding time-windowed (c) initial and (d) final residuals. The amplitude scaling is the same as that in Figure 13.

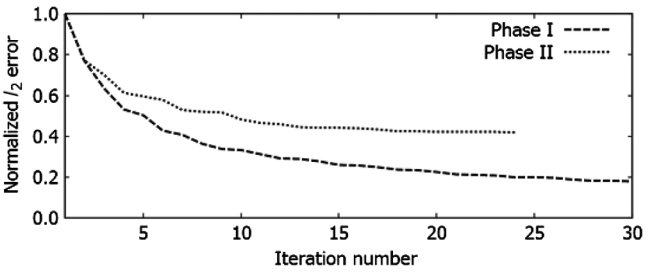

Figure 21. The normalized least-squares (l_2) errors of the frequency data residuals for the inversion using the FD-synthetic shot gathers.

CONCLUSIONS

A composite of advances in inversion techniques leads to a performance improvement of approximately an order of magnitude in elastic FWI. Wavefield extrapolations are done in time-space with finite element; computations of the parameter gradients and the objective function are done in the frequency domain. Frequencies are divided into two ranges; lower frequencies are fitted first; the resulting elastic model is used as the starting model to fit the higher frequencies. A multistep-length gradient assigns optimal weights to each of the parameter gradients to decouple the parameters and to search for an optimal overall step length for the composite gradient to stabilize the nonlinearity of the solution. Even though requiring a few more simulations to search for the optimal values of the individual weights, the inversion algorithm converges stably and fast within a few tens of iterations, rather than several hundreds that most other elastic inversions need. By using effective strategies such as frequency selection, time windowing, and source wavelet estimation in the full frequency spectrum, frequency-domain elastic FWI based on the parallelized time-domain FE modeling engine is still successful on the elastic Marmousi-2 model, even when using a constant density model and for synthetic data that are generated by an FD modeling scheme.

The velocity and density reconstructions have different behaviors in the inversion, but they are also partly coupled. Velocities contribute to fitting of the phases (times) and amplitudes; density contributes mainly to amplitudes. Thus, the velocity component of the inversion behaves as a tomography to provide information inside layers as well as at layer boundaries; the density component of the inversion behaves as a least-squares migration to provide amplitudes only on layer boundaries. The inverted velocities have lower resolutions than the inverted density. However because velocities and density are coupled to some extent, variations are usually underestimated (smoothed) for velocities, and overestimated

(sharpened) for density correspondingly. The resulting P and S impedances are more accurate than velocities and density that are separately because the data amplitudes depend on the impedances rather than the velocities and density separately; errors in one are compensated by those in the other as they are fitted concurrently.

ACKNOWLEDGMENTS

The research in this paper was supported by the sponsors of the UT-Dallas Geophysical Consortium. S. Operto (SEISCOPE) provided a 2D P-SV-wave fourth-order staggered-grid FD modeling algorithm to verify our elastic FWI. K. Xu wishes to express his gratitude to F. Gao (at Total) for constructive discussions on the FWI algorithm. Finally, we would like to thank T. Alkhalifah, M. Sen, and two anonymous reviewers for their fruitful comments on the paper. The computations were performed on Linux clusters in the center for Lithospheric Studies at UT Dallas and at the Texas Advanced Computing Center. This paper is contribution no. 1255 from the Department of Geosciences at the University of Texas at Dallas.

REFERENCES

Abubakar, A., M. Li, Y. Lin, and T. M. Habashy, 2012, Compressed implicit Jacobian scheme for elastic full-waveform inversion: Geophysical Journal International, **189**, 1626–1634, doi: 10.1111/j.1365-246X.2012.05439.x.

Berenger, J. P., 1994, A perfectly matched layer for the absorption of electromagnetic waves: Journal of Computational Physics, **114**, 185–200, doi: 10.1006/jcph.1994.1159.

Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: Geophysics, **74**, no. 6, WCC105–WCC118, doi: 10.1190/1.3215771.

Choi, Y., D. Min, and C. Shin, 2008, Two-dimensional waveform inversion of multi-component data in acoustic-elastic coupled media: Geophysical Prospecting, **56**, 863–881, doi: 10.1111/j.1365-2478.2008.00735.x.

Chung, W., S. Pyun, H. S. Bae, and C. Shin, 2011, Frequency-domain elastic reverse-time migration using wave-field separation: 82nd Annual International Meeting, SEG, Expanded Abstracts, 3419–3424.

Chung, W., C. Shin, S. Pyun, and H. Calandra, 2008, 2-D elastic waveform inversion in the Laplace domain: 82nd Annual International Meeting, SEG, Expanded Abstracts, 1059–1064.

Crase, E., A. Pica, M. Noble, J. McDonald, and A. Tarantola, 1990, Robust elastic nonlinear waveform inversion: Application to real data: Geophysics, **55**, 527–538, doi: 10.1190/1.1442864.

Etienne, V., J. Virieux, S. Operto, and Y. Jia, 2010, Computational issues and strategies related to full waveform inversion in 3D elastic media: Methodological developments: 81st Annual International Meeting, SEG, Expanded Abstracts, 1050–1054.

Gelis, C., J. Virieux, and G. Grandjean, 2007, 2D elastic waveform inversion using Born and Rytov approximations in the frequency domain: Geophysical Journal International, **168**, 605–633, doi: 10.1111/j.1365-246X.2006.03135.x.

Higdon, R. L., 1991, Absorbing boundary conditions for elastic waves: Geophysics, **56**, 231–241, doi: 10.1190/1.1443035.

Jeong, W., H. Lee, and D. Min, 2012, Full waveform inversion strategy for density in the frequency domain: Geophysical Journal International, **188**, 1221–1242, doi: 10.1111/j.1365-246X.2011.05314.x.

Kennett, B. L. N., M. S. Sambridge, and P. R. Williamson, 1988, Subspace methods for large inverse problems with multiple parameter classes: Geophysical Journal, **94**, 237–247, doi: 10.1111/j.1365-246X.1988.tb05898.x.

Köhn, D., D. D. Nil, A. Kurzmann, A. Przebindowska, and T. Bohlen, 2012, On the influence of model parametrization in elastic full waveform tomography: Geophysical Journal International, **191**, 325–345, doi: 10.1111/j.1365-246X.2012.05633.x.

Krenk, S., 2001, Dispersion-corrected explicit integration of the wave equation: Computer methods in applied mechanics and engineering, **191**, 975–987, doi: 10.1016/S0045-7825(01)00297-3.

Levander, A. R., 1988, Fourth-order finite-difference P-SV seismograms: Geophysics, **53**, 1425–1436, doi: 10.1190/1.1442422.

Marfurt, K. J., 1984, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations: Geophysics, **49**, 533–549, doi: 10.1190/1.1441689.

Martin, G. S., K. J. Marfurt, and S. Larsen, 2002, Marmousi-2: An updated model for the investigation of AVO in structurally complex areas: 71st Annual International Meeting, SEG, Expanded Abstracts, 1979–1982.

Meles, G. A., J. V. Kruk, S. A. Greenhalgh, J. R. Ernst, H. Maurer, and A. G. Green, 2010, A new vector waveform inversion algorithm for simultaneously updating of conductivity and permittivity parameters from combination crosshole/borehole-to-surface GPR data: IEEE Transactions on Geoscience and Remote Sensing, **48**, 3391–3407, doi: 10.1109/TGRS.2010.2046670.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228, doi: 10.1190/1.1442384.

Mora, P., 1989, Inversion = migration + tomography: Geophysics, **54**, 1575–1586, doi: 10.1190/1.1442625.

Mulder, W. A., and R. E. Plessix, 2008, Exploring some issues in acoustic full waveform inversion: Geophysical Prospecting, **56**, 827–841, doi: 10.1111/j.1365-2478.2008.00708.x.

Nihei, K. T., and X. Li, 2007, Frequency response modeling of seismic waves using finite difference time domain with phase sensitive detection (TD-PSD): Geophysical Journal International, **169**, 1069–1078, doi: 10.1111/j.1365-246X.2006.03262.x.

Nikishkov, G. P., 2007, Introduction to the finite element method: University of Aizu.

Pica, A., J. P. Diet, and A. Tarantola, 1990, Nonlinear inversion of seismic reflection data in a laterally invariant medium: Geophysics, **55**, 284–292, doi: 10.1190/1.1442836.

Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362, doi: 10.1046/j.1365-246X.1998.00498.x.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992, Numerical recipes in Fortran 77: Cambridge University Press.

Sakai, A., 2011, A strategy of non-linear elastic wavefield inversion of marine seismic streamer data in lower seabed shear-wave velocity structure: Proceedings of the 10th SEGJ International Symposium, SEG, 108–111.

Sambridge, M. S., A. Tarantola, and B. L. N. Kennett, 1991, An alternative strategy for non-linear inversion of seismic waveforms: Geophysical Prospecting, **39**, 723–736, doi: 10.1111/j.1365-2478.1991.tb00341.x.

Sears, T. J., S. C. Singh, and P. J. Barton, 2008, Elastic full waveform inversion of multi-component OBC seismic data: Geophysical Prospecting, **56**, 843–862, doi: 10.1111/j.1365-2478.2008.00692.x.

Shin, C., 1995, Sponge boundary condition for frequency-domain modeling: Geophysics, **60**, 1870–1874, doi: 10.1190/1.1443918.

Shin, C., and Y. H. Cha, 2009, Waveform inversion in the Laplace-Fourier domain: Geophysical Journal International, **177**, 1067–1079, doi: 10.1111/j.1365-246X.2009.04102.x.

Shipp, R. M., and S. C. Singh, 2002, Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data: Geophysical Journal International, **151**, 325–344, doi: 10.1046/j.1365-246X.2002.01645.x.

Sirgue, L., J. T. Etgen, and U. Albertin, 2008, 3D frequency domain waveform inversion using time domain finite difference methods: 70th Annual International Conference and Exhibition, EAGE, Extended Abstracts, F022.

Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248, doi: 10.1190/1.1649391.

Song, Z. M., P. R. Williamson, and P. 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, doi: 10.1190/1.1443818.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903, doi: 10.1190/1.1442046.

Vigh, D., and E. W. Starr, 2008, 3D prestack plane-wave, full-waveform inversion: Geophysics, **73**, no. 5, VE135–VE144, doi: 10.1190/1.2952623.

Virieux, J., 1986, P-SV wave propagation in heterogeneous media, velocity-stress finite difference method: Geophysics, **51**, 889–901, doi: 10.1190/1.1442147.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, doi: 10.1190/1.3238367.

Xu, K., and S. A. Greenhalgh, 2010, Ore-body imaging by crosswell seismic waveform inversion: A case study from Kambalda, Western Australia: Journal of Applied Geophysics, **70**, 38–45, doi: 10.1016/j.jappgeo.2009.11.001.

Xu, K., S. A. Greenhalgh, and M. Wang, 2006, Comparison of source-independent methods of elastic waveform inversion: Geophysics, **71**, no. 6, R91–R100, doi: 10.1190/1.2356256.

Xu, T., G. A. McMechan, and R. Sun, 1995, 3D prestack full-wavefield inversion: Geophysics, **60**, 1805–1818, doi: 10.1190/1.1443913.