GEOPHYSICS®

Frequency domain multi-parameter acoustic inversion for transversely isotropic media with a vertical axis of symmetry

Journal:	<i>Geophysics</i>
Manuscript ID	GEO-2017-0564.R2
Manuscript Type:	Technical Paper
Keywords:	anisotropy, full-waveform inversion, multiparameter, VTI
Area of Expertise:	Anisotropy, Seismic Inversion

SCHOLARONE™
Manuscripts

Frequency domain multi-parameter acoustic inversion for transversely isotropic media with a vertical axis of symmetry

Ramzi Djebbi* and Tariq Alkhalifah†

**Formerly King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.*

Presently PGS, Weybridge, UK. E-mail: ramzi.djebbi@pgs.com

†King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.

E-mail: Tariq.alkhalifah@kaust.edu.sa

(August 30, 2018)

Running head: **Multi-parameter inversion for VTI media**

ABSTRACT

Multi-parameter full waveform inversion (FWI) for transversely isotropic (TI) media with a vertical axis of symmetry (VTI) suffers from the trade-off between the parameters. The trade-off results in the leakage of one parameter’s update into the other. It affects the accuracy and convergence of the inversion. The sensitivity analyses suggested a parameterization using the horizontal velocity v_h , Thomsen’s parameter ϵ and the an-elliptic parameter η to reduce the trade-off for surface recorded seismic data. We aim to invert for this parameterization using the scattering integral (SI) method. The available Born sensitivity kernels, within this approach, can be used to calculate additional inversion information. We mainly compute the diagonal of the approximate Hessian, used as a conjugate-gradient preconditioner, and the gradients step lengths. We consider modeling in the frequency domain. The large computational cost of the scattering integral method can be avoided with direct Helmholtz equation solvers. We apply the proposed method to the VTI Marmousi II model for various inversion strategies. We show that we can invert the v_h accurately. For the

ϵ parameter, only the short wavelengths are well recovered. On the other hand, the η parameter impact is weak on the inversion results and can be fixed. However, a good background η , with accurate long wavelengths, is needed to correctly invert for v_h . Furthermore, we invert a real data set acquired by CGG from offshore Australia. We invert simultaneously all three parameters using the proposed inversion approach. The velocity model is improved and additional layers are recovered. We confirm the accuracy of the results by comparing them with well-log information, as well as, looking at the data and angle gathers.

INTRODUCTION

Full waveform inversion aims to invert the full recorded data content to recover an accurate Earth model (Virieux and Operto, 2009). It is based on the minimization of data residuals, measured between the real data and the data generated numerically using an approximate Earth model. The FWI objective function is highly non-linear with respect to the model parameters. Therefore, the minimization process using gradient methods can converge to a local minimum resulting in a wrong inverted model. FWI is often implemented in the time domain (Tarantola, 1984; Mora, 1988; Bunks et al., 1995). Bunks et al. (1995) proposed a multiscale method to reduce the FWI nonlinearity. The inversion problem, in this case, is decomposed into scales starting from a long wavelength scale to the shorter ones. The low-enough frequency content of data mitigates the nonlinearity and permits converging to a global minimum low-resolution model. This model serves as a good starting point for the next scale of frequencies. Another approach to deal with the cycle-skipping problem is to window the data, so we focus the inversion on specific arrivals (Shipp and Singh, 2002). A more natural multiscale inversion scheme is realized in frequency domain (Pratt and Worthington, 1990; Zhou et al., 1995; Liao and McMechan, 1996) where a discrete number of increasing frequencies are selected. Sirgue and Pratt (2004) proposed a frequency selection algorithm which fills the wavenumber spectrum with a limited number of frequencies. For multi-parameter inversion, inverting a group of frequencies is necessary to resolve more than one parameter. In fact, the number of inverted parameters and their inversion resolution limits highly depend on the data redundancy criterion (Podgornova et al., 2015; Kazei and Alkhalifah, 2018).

Conventionally, the adjoint-state method (Tromp et al., 2005; Plessix, 2006) is used for FWI. The main advantage is the reasonable efficiency in the gradient computation. Its essence is to circumvent the explicit evaluation of the Fréchet derivatives. The scattering integral (SI) method is an

alternative approach, where the gradient is computed explicitly using the sensitivity kernels (Chen et al., 2007; Tao and Sen, 2013). The gradient is given by the product of the Fréchet derivatives matrix with the complex conjugate of the residual vector. The large computational cost and memory requirements to compute the gradient made the SI method less popular than the adjoint-state in seismic inversion. Chen et al. (2007) compared the two methods, and showed that the SI method is less expensive when the number of sources is comparable to the number of receivers (such as OBS exploration and earthquake seismology). Furthermore, the availability of the sensitivity information can be used to estimate the gradient step size and/or approximations of the inverse Hessian matrix (Gauss-Newton direction). Liu et al. (2015) introduced a method to avoid storing the Fréchet derivatives. The Green's functions are stored in memory, and a vector-matrix product is used to compute the gradient on the fly. When modeling is performed in the frequency domain, direct solvers are used to reduce the computational cost of the SI method. The main cost of the modeling is in the decomposition of the impedance matrix into a lower-upper (LU) structure and solving the wave equation, for a large number of sources and receivers, is small compared to time domain solvers.

Most of the existing inversion schemes consider only an isotropic Earth model to reduce the computational cost. However, the recent improvements in data acquisition made anisotropic effects more prominent and inverting for anisotropy parameters is getting more attention. Several studies showed that the possibility to recover anisotropy in a multiparameter inversion set-up is limited by the trade-off between the parameters (Gholami et al., 2013; Alkhalifah and Plessix, 2014; Alkhalifah, 2016; Masmoudi and Alkhalifah, 2016; Djebbi et al., 2017). Understanding the sensitivity of the data to the anisotropy parameters is essential to choose a parameterization with a minimum trade-off. For transversely isotropic (TI) media with a vertical axis of symmetry (VTI), Plessix and Cao (2011) studied the trade-off problem using an eigenvalue-eigenvector decomposition of the

Hessian matrix. Gholami et al. (2013) and Alkhalifah and Plessix (2014) analyzed the radiation patterns for several parameterizations to understand the sensitivity to anisotropy parameters at the scattering point level. Recently, Alkhalifah (2016) studied the short and long wavelength influences using perturbations in the model parameters. The author perturbed the model parameters and investigated the effects on the recorded data. In a more qualitative study, Djebbi et al. (2017) used the sensitivity kernels for anisotropy parameter as a tool of analysis. Compared to the radiation patterns, the kernels show additional trade-off information along the wave-path.

The VTI wave equation parameterization using the normal move-out velocity v_n , Thomsen's parameter δ (Thomsen, 1986) and the an-elliptic parameter η (Alkhalifah and Tsvankin, 1995) is adapted for an inversion that includes simultaneously reflections and diving waves (Plessix and Cao, 2011; Alkhalifah and Plessix, 2014). For this parameterization, the δ parameter is considered as a secondary parameter, which is used to fit the inaccuracy in the amplitude caused by the acoustic assumption. However, the trade-off between v_n and η parameter for large scattering angles (diving waves) is significant and results in the leakage of η parameter into v_n during the inversion. Another widely used parameterization is based on the vertical velocity v_v , δ , and ϵ parameters (Vigh et al., 2014). It suffers from the same type of trade-off for large scattering angles between v_v and ϵ parameter. Alkhalifah (2016) suggested to use the horizontal velocity v_h , η and ϵ . This parameterization is considered optimal for conventional surface seismic experiments with the least trade-off. To construct the model's long wavelengths, diving waves are the dominant source of information at least for the first stages of FWI. With v_h , η , and ϵ parameterization, diving waves are sensitive to the horizontal velocity v_h only. Therefore, v_h long wavelengths can be recovered with minimum cross-talk with the other two parameters. For small scattering angles, coming mainly from reflections, data are sensitive to ϵ and v_h . With the acoustic approximation, seismic data amplitudes are inaccurate and the ϵ parameter helps to absorb these errors. The sensitivity to the η parameter is principally

in the mid-range scattering angles (around 45°) with weak magnitude. Provided that a good long wavelength η model is available (generally from tomography), the impact of η on the inversion is negligible (Alkhalifah, 2016). Guitton and Alkhalifah (2016) showed that using v_h , η and ϵ parameterization the leakage from density and shear waves into v_h is reduced. Wu and Alkhalifah (2016) used the same parameterization for reflection waveform inversion (RWI). They assumed $\epsilon = \eta$ to reduce the number of variables and used ϵ as a perturbation parameter. The inversion strategy performed well, and both the background and perturbation models are reconstructed accurately.

In this paper, we present a multi-parameter acoustic VTI inversion using v_h , η and ϵ parameterization. The conjugate gradient (CG) method is used for the optimization within the scattering integral framework. The step length as well as the diagonal approximate Hessian, which is used as CG preconditioners, are directly estimated using the sensitivity kernels. The first section presents the frequency domain modeling approach. Then, we present the scattering integral method for multi-parameter FWI. In the numerical examples section, we consider various multiscale inversion tests to show which strategy is best suited for anisotropic multiparameter FWI. The aim is to recover a good velocity field v_h with minimum cross-talk with ϵ and η . Considering the complexity of the model, we show accurate inverted v_h . For ϵ and η , only short wavelengths are recoverable as for these two parameters the sensitivity to large scattering angles data is zero. We show a realistic example, where the source signature is unknown, and the observed data are modeled with a different wave equation solver. We show that we can achieve reasonable inversion results using this parameterization. Finally, we apply the inversion method to a real data set acquired by CGG from offshore Australia. We invert simultaneously all three parameters using the proposed inversion approach. We use the well-log profile and the final modeled data to confirm the accuracy of the inversion. The migrated images and angle gathers are also used as a quality-control (QC) of the final model.

FREQUENCY DOMAIN VTI MODELING

The key element in inversion is to efficiently solve the wave equation that represents the correct physics of wave propagation. We consider the VTI system of wave equations (Duveneck et al., 2008) in frequency domain,

$$\begin{aligned}\frac{\omega^2}{v_v^2}p + (1 + 2\epsilon)\frac{\partial^2 p}{\partial x^2} + \sqrt{1 + 2\delta}\frac{\partial^2 q}{\partial z^2} &= f_x, \\ \frac{\omega^2}{v_v^2}q + \sqrt{1 + 2\delta}\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 q}{\partial z^2} &= f_z,\end{aligned}\tag{1}$$

where p and q are the wavefield components, f_x and f_z are the source components and ω is the angular frequency. The wave equations are parameterized using the vertical velocity v_v and Thomsen anisotropy parameters ϵ and δ . We denote the total wavefield by $U(\mathbf{x}, \mathbf{x}_s, \omega)$ and it can be calculated as $U(\mathbf{x}, \mathbf{x}_s, \omega) = \frac{q+2p}{3}$. Considering that parameterization v_h , η and ϵ is used for the inversion, we re-parameterize the system of wave equations as,

$$\begin{aligned}\frac{\omega^2(1+2\epsilon)}{v_h^2}p + (1 + 2\epsilon)\frac{\partial^2 p}{\partial x^2} + \sqrt{\frac{1+2\epsilon}{1+2\eta}}\frac{\partial^2 q}{\partial z^2} &= f_x, \\ \frac{\omega^2(1+2\epsilon)}{v_h^2}q + \sqrt{\frac{1+2\epsilon}{1+2\eta}}\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 q}{\partial z^2} &= f_z.\end{aligned}\tag{2}$$

In a compact form, this system can be written as,

$$\mathbf{B}(\mathbf{x}, \omega)\mathbf{u}(\mathbf{x}, \mathbf{x}_s, \omega) = \mathbf{f}(\mathbf{x}_s, \omega),\tag{3}$$

where $\mathbf{B}(\mathbf{x}, \omega)$ is the impedance matrix. $\mathbf{u}(\mathbf{x}, \mathbf{x}_s, \omega) = [p \quad q]^T$ the wavefield vector and $\mathbf{f}(\mathbf{x}_s, \omega) = [f_x \quad f_z]^T$ is the source function vector.

For 2-D problems, solving equation (3) is generally performed using direct solvers (Pratt et al., 1998). The sparse impedance matrix $\mathbf{B}(\mathbf{x}, \omega)$ is decomposed into a lower-upper (LU) decomposition. The main advantage of this approach is the non-dependence of $\mathbf{B}(\mathbf{x}, \omega)$ on the source location.

Thus, the same LU can be used for all sources. Considering the low computational cost of the forward and backward substitutions, compared to the LU decomposition, the wave equation can be solved efficiently for a large number of source locations. We consider 4th-order finite-differences to discretize the VTI wave equations system. We also use SuiteSparse package to perform the LU decomposition efficiently (Davis, 2004). We consider perfectly-matched layers (PML) boundary conditions to absorb the wavefields on the side and bottom boundaries (Berenger, 1994; Lu, 2006).

For the 3-D case, the LU factorization was restricted by the large memory requirements. However, with the new developments of multifrontal methods, where low-rank blocks are used to represent dense matrices, the cost and memory requirements for the LU factorization are largely reduced (Wang et al., 2011; Amestoy et al., 2015). Furthermore, Operto et al. (2014) proposed a finite-difference stencil for visco-acoustic VTI wave equation to control the memory and efficiency. Operto et al. (2015) inverted for the vertical velocity with a wavefield modeled under the visco-acoustic VTI approximation to analyze the performance of the proposed 3-D direct solver.

MULTI-PARAMETER FULL WAVEFORM INVERSION USING THE SCATTERING INTEGRAL METHOD

Full waveform inversion is based on the minimization of the residuals between the modeled data $\mathbf{U}(\mathbf{m})$ and the observed data $\mathbf{d}$. We minimize the l_2 norm misfit function,

$$E(\mathbf{m}) = \frac{1}{2} \Delta \mathbf{d}^T \Delta \mathbf{d}^*, \quad (4)$$

where $\Delta \mathbf{d} = \mathbf{U}(\mathbf{m}) - \mathbf{d}$ is the data residuals vector and $\mathbf{m}$ is the model. The superscripts T and * denote the transpose and complex conjugate, respectively.

We use the preconditioned conjugate gradient (CG) method to update the model where the ap-

proximate inverse Hessian is considered as a preconditioner to improve the convergence properties.

The update is given as,

$$\Delta \mathbf{m}_{CG} = -\gamma \mathbf{P}(\mathbf{m}) \frac{\partial E(\mathbf{m})}{\partial \mathbf{m}} = -\gamma \mathbf{p}_2, \quad (5)$$

where γ is the step size and $\mathbf{P}(\mathbf{m})$ is the preconditioning operator. We denote the preconditioned CG direction as $\mathbf{p}_2$. Taking the derivative of the misfit function with respect to the model parameters, the gradient reads,

$$\mathbf{g}(\mathbf{m}) = \mathbf{p}_1 = \frac{\partial E(\mathbf{m})}{\partial \mathbf{m}} = \Re(\mathbf{F}^T \Delta \mathbf{d}^*), \quad (6)$$

where $\Re$ represents the real part. We denote the gradient update direction as $\mathbf{p}_1$. $\mathbf{F}$ is an $(n \times m)$ matrix, representing the Fréchet derivatives where n and m are the data and model sizes, respectively.

The Fréchet derivatives show the sensitivity of the data to the model parameters. The elements of this matrix are given as,

$$k_{ij} = \frac{\partial U_i}{\partial m_j} \quad i = (1, \dots, n); \quad j = (1, \dots, m). \quad (7)$$

The scattering integral method

We consider the scattering integral (SI) method to compute the gradient. The gradient is given by the product of the Fréchet derivatives matrix with the complex conjugate of the residual vector as shown with equation (6). For a single source and receiver, each row of the Fréchet derivatives matrix $\mathbf{F}$ corresponds to the Born wavefield sensitivity kernel $\mathbf{K}(\mathbf{x}_r, \mathbf{x}, \mathbf{x}_s, \omega)$. For the isotropic case, the Born kernel for the velocity model is given as (Beydoun and Tarantola, 1988),

$$\mathbf{K}^{(v)}(\mathbf{x}_r, \mathbf{x}, \mathbf{x}_s, \omega) = \frac{2\omega^2}{v_0^3} U_0(\mathbf{x}, \mathbf{x}_s, \omega) G_0(\mathbf{x}, \mathbf{x}_r, \omega), \quad (8)$$

where $U_0(\mathbf{x}, \mathbf{x}_s, \omega)$ is the source wavefield and $G_0(\mathbf{x}, \mathbf{x}_r, \omega)$ is the receiver's Green's function in the background medium v_0 .

The explicit computation of the Fréchet derivatives requires wavefields computed at the shot locations, as well as, the Green's functions from the receivers locations, which makes it computationally expensive. In the frequency domain, the computational cost is largely reduced as a single LU factorization is required to solve the wave equation for all sources and receivers locations. Furthermore, to reduce the memory needed to store the Fréchet derivative (sensitivity) matrix, Liu et al. (2015) proposed to save the Green's functions into memory and use matrix decomposition to compute the gradient. Equation (9) shows the gradient computation procedure,

$$\mathbf{g}(\mathbf{m}) = \begin{pmatrix} k_{11} & \dots & k_{i1} & \dots & k_{n1} \\ \vdots & \ddots & \ddots & \vdots & \ddots \\ k_{1j} & \dots & k_{ij} & \dots & k_{nj} \\ \vdots & \ddots & \ddots & \vdots & \ddots \\ k_{1m} & \dots & k_{im} & \dots & k_{nm} \end{pmatrix} \begin{pmatrix} \Delta d_1^* \\ \vdots \\ \Delta d_j^* \\ \vdots \\ \Delta d_n^* \end{pmatrix} = \sum_{i=0}^n \begin{pmatrix} k_{i1} \\ \vdots \\ k_{ij} \\ \vdots \\ k_{im} \end{pmatrix} \Delta d_i^*. \quad (9)$$

The diagonal approximate Hessian is given as,

$$\mathbf{H}_d(\mathbf{m}) = \text{diag}(\mathbf{F}^\dagger \mathbf{F}) = \sum_{i=0}^n \begin{pmatrix} k_{i1}^* k_{i1} & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & \ddots \\ 0 & \dots & k_{ij}^* k_{ij} & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & \ddots \\ 0 & \dots & 0 & \dots & k_{im}^* k_{im} \end{pmatrix}. \quad (10)$$

$\mathbf{F}^\dagger$ stands for the complex conjugate transpose (Hermitian) of the Fréchet derivatives matrix.

Considering a second order approximation of the objective function $E_1(\gamma) = E(\mathbf{m} + \gamma\mathbf{p})$, the time step size can be directly computed using the sensitivity kernels as (Hu et al., 2011),

$$\gamma = \frac{\mathbf{p}^\dagger \mathbf{p}_1}{\mathbf{p}^\dagger \mathbf{H}_a \mathbf{p}_1} = \frac{\mathbf{p}^\dagger \mathbf{p}_1}{(\mathbf{F}\mathbf{p})^\dagger (\mathbf{F}\mathbf{p})}. \quad (11)$$

Depending on the optimization method, $\mathbf{p}$ can be either the gradient direction $\mathbf{p}_1$ or the preconditioned gradient direction $\mathbf{p}_2$. $\mathbf{H}_a = \mathbf{F}^\dagger \mathbf{F}$ is the approximate Hessian. For an adjoint state based FWI, an additional modeling step is required to approximate the denominator in equation (11). However, using the sensitivity kernels, within the scattering integral approach, the gradient time step can be accurately computed with a minimal cost.

The Born sensitivity kernels for VTI media

For a VTI medium parameterized using v_h , ϵ and η , the Fréchet derivatives matrix is given as:

$$\mathbf{F} = [\mathbf{F}_{v_h} \quad \mathbf{F}_\epsilon \quad \mathbf{F}_\eta]. \quad (12)$$

The total size of the Fréchet matrix is $3 \times (n \times m)$. Each row of the three sub-matrices, given in equation (12), is the sensitivity kernel for v_h , ϵ and η , respectively, for a specific source-receiver pair.

To derive the sensitivity kernels, the VTI system of equations (2) is combined into a single 4th order equation (Djebbi et al., 2017). Then, the three anisotropy parameters are perturbed to obtain the corresponding single-frequency Born sensitivity kernels (Djebbi et al., 2017). Equation (13)

shows the resulting kernels,

$$\begin{aligned}
 \mathbf{K}^{(v_h)}(\mathbf{x}_r, \mathbf{x}_s, \mathbf{x}, \omega) &= \frac{2\omega^2}{v_0^3} U_0(\mathbf{x}, \mathbf{x}_s, \omega) G_0(\mathbf{x}, \mathbf{x}_r, \omega); \\
 \mathbf{K}^{(\epsilon)}(\mathbf{x}_r, \mathbf{x}_s, \mathbf{x}, \omega) &= - (U_0(\mathbf{x}, \mathbf{x}_s, \omega) \partial_{zz} G_0(\mathbf{x}, \mathbf{x}_r, \omega) \\
 &\quad + \partial_{zz} U_0(\mathbf{x}, \mathbf{x}_s, \omega) G_0(\mathbf{x}, \mathbf{x}_r, \omega)); \\
 \mathbf{K}^{(\eta)}(\mathbf{x}_r, \mathbf{x}_s, \mathbf{x}, \omega) &= - \frac{v_0^2}{\omega^2} (\partial_{xx} U_0(\mathbf{x}, \mathbf{x}_s, \omega) \partial_{zz} G_0(\mathbf{x}, \mathbf{x}_r, \omega) \\
 &\quad + \partial_{zz} U_0(\mathbf{x}, \mathbf{x}_s, \omega) \partial_{xx} G_0(\mathbf{x}, \mathbf{x}_r, \omega));
 \end{aligned} \tag{13}$$

where $U_0(\mathbf{x}, \mathbf{x}_s, \omega)$ and $G_0(\mathbf{x}, \mathbf{x}_r, \omega)$ are the source wavefield and the receiver Green's function using the background model, respectively.

[Figure 1 about here.]

We show in Figure 1 the Born sensitivity kernels for an increasing velocity with depth model ($v(z) = 1.5 + 0.8z$ km/s) for a source located at 0.5 km and a receiver located at 7.5 km. The kernels are normalized to the maximum value and plotted at the same scale. The sensitivity kernel is composed of a central region, which corresponds to the diving wave update and the isochrones corresponding to reflection information. It can be observed, for this conventional seismic experiment, that the horizontal velocity sensitivity is large for both diving waves and reflections. The main contribution for the ϵ parameter sensitivity is due to reflections. η kernel looks similar to ϵ in this set-up, although it is scattering at intermediate angles only.

Multi-parameter step estimation

The step size is estimated for each parameter using a second-order approximation of the objective function $E_1(\gamma) = E(\mathbf{m} + \gamma^T \mathbf{p}_2)$. For a VTI medium, the vector $\mathbf{m}$ is composed of three vectors

corresponding to: v_h , ϵ and η . The vector $\mathbf{p}_2$ is the preconditioned CG update. Finally, γ is the step size vector given as: $\gamma = \begin{bmatrix} \gamma_{v_h} & \gamma_{\epsilon} & \gamma_{\eta} \end{bmatrix}^T$. We perturb simultaneously the anisotropy parameters. Then, using the sensitivity kernels, the CG time steps are computed efficiently by solving the following 3×3 system of equations:

$$\begin{bmatrix} \mathbf{p}_{2v_h}^\dagger (\mathbf{F}_{v_h}^\dagger \mathbf{F}_{v_h}) \mathbf{p}_{2v_h} & \mathbf{p}_{2v_h}^\dagger (\mathbf{F}_{v_h}^\dagger \mathbf{F}_{\epsilon}) \mathbf{p}_{2\epsilon} & \mathbf{p}_{2v_h}^\dagger (\mathbf{F}_{v_h}^\dagger \mathbf{F}_{\eta}) \mathbf{p}_{2\eta} \\ \mathbf{p}_{2\epsilon}^\dagger (\mathbf{F}_{\epsilon}^\dagger \mathbf{F}_{v_h}) \mathbf{p}_{2v_h} & \mathbf{p}_{2\epsilon}^\dagger (\mathbf{F}_{\epsilon}^\dagger \mathbf{F}_{\epsilon}) \mathbf{p}_{2\epsilon} & \mathbf{p}_{2\epsilon}^\dagger (\mathbf{F}_{\epsilon}^\dagger \mathbf{F}_{\eta}) \mathbf{p}_{2\eta} \\ \mathbf{p}_{2\eta}^\dagger (\mathbf{F}_{\eta}^\dagger \mathbf{F}_{v_h}) \mathbf{p}_{2v_h} & \mathbf{p}_{2\eta}^\dagger (\mathbf{F}_{\eta}^\dagger \mathbf{F}_{\epsilon}) \mathbf{p}_{2\epsilon} & \mathbf{p}_{2\eta}^\dagger (\mathbf{F}_{\eta}^\dagger \mathbf{F}_{\eta}) \mathbf{p}_{2\eta} \end{bmatrix} \cdot \begin{bmatrix} \gamma_{v_h} \\ \gamma_{\epsilon} \\ \gamma_{\eta} \end{bmatrix} = \begin{bmatrix} -\mathbf{p}_{1v_h}^\dagger \mathbf{p}_{2v_h} \\ -\mathbf{p}_{1\epsilon}^\dagger \mathbf{p}_{2\epsilon} \\ -\mathbf{p}_{1\eta}^\dagger \mathbf{p}_{2\eta} \end{bmatrix} \quad (14)$$

The CG approach, with the gradient scaling factors given by equation (14), is an affordable method to deal with the cross-talk between anisotropy parameters. The accumulation of the inverse Hessian matrix effects provides a proper scaling between anisotropy parameters. The method is related to the subspace method (Kennett et al., 1988; Baumstein et al., 2014) where the model's perturbation is represented by a combination of basis vectors. Baumstein et al. (2014) extended the subspace method to invert for anisotropy. Using CG frequency domain SI, the model is approached progressively with small update steps. In this situation, the proposed method can efficiently resolve the cross-talk between the parameters. However, compared to the truncated Gauss-Newton (Métivier et al., 2013) method, the action of the Hessian is not well treated.

For the adjoint-state based FWI, the elements of the matrix on the left-hand side in equation (14) require additional modeling steps (needed to approximate the matrix elements by finite-difference). For the scattering integral (SI) approach only vector-matrix multiplications are required. As a result, no additional modeling steps are needed, and the main computation occurs during the construction of the sensitivity matrix and its decomposition. The scattering integral method can be extended to compute a truncated Gauss-Newton update through an inner minimization loop. In this paper, we

limit ourselves to the preconditioned CG method.

SYNTHETIC EXAMPLES

Based on the sensitivity analysis, we consider v_h , ϵ and η parameterization for the inversion. We invert for the VTI Marmousi II model, shown in Figure 2. The anisotropy parameters are computed using the Marmousi II density and P-wave velocity models. We set the near surface to be isotropic (with a velocity value equal to the water velocity) to avoid shear waves artifacts. The resulting models are complex with large anisotropy magnitudes. Therefore, we expect strong trade-off between the parameters, which complicates the inversion procedure. We intent to investigate which inversion strategy is best suited for multiparameter anisotropic frequency domain FWI. Therefore, we show in this section, inversion tests with various strategies and initial models. The final test is performed with an independent modeling code to avoid the inverse crime when using the same modeling algorithm for both the observed and modeled data.

[Figure 2 about here.]

Inversion strategy

The initial models, shown in Figure 3, are obtained by smoothing the exact anisotropy parameters. We smooth the exact models using a triangular smoothing function with 1.5 km smoothing radius. The modeling is performed in the frequency domain using the acoustic VTI Helmholtz solver. Aiming to study the impact of the inversion strategy on trade-off, we use the same modeling for both the observed and the modeled data. We consider a multiscale inversion approach where 4 frequency bands are used. Table 1 shows the inversion parameters for each frequency band. The minimum frequency for the inversion is 2 Hz.

[Table 1 about here.]

Simultaneous inversion of v_h , ϵ and η

In this test, v_h , η and ϵ are updated, simultaneously. We aim to invert for v_h with a minimum trade-off as it is the main parameter. Seismic data are sensitive to all scattering angles for v_h perturbation. Therefore, the resulting inverted v_h model should contain all wavelengths. The final v_h model, shown in Figure 4a, confirms the sensitivity analysis. Most of the features are well recovered. The reservoir area around $x = 8$ km and $z = 3$ km location is well resolved. ϵ perturbation scatters data only at small scattering angles. Therefore, it plays a secondary role in the fitting process focused on amplitudes, for example, for elastic data fitting. In this case, ϵ absorbs the inaccuracies in the inversion of short wavelengths when the acoustic approximation is used to invert elastic data (Alkhalifah, 2016; Guitton and Alkhalifah, 2016). In the final model (Figure 4b), the short to intermediate wavelengths are well reconstructed. However, the long wavelengths are not recovered, which is in good agreement with the radiation patterns analysis (Alkhalifah and Plessix, 2014). The recorded data are sensitive to η only for mid-range scattering angles (around 45°) and the sensitivity magnitude is small compared to the data sensitivity to v_h and ϵ . The final η model, Figure 4c, contains some features of the exact model. However, the trade-off with the other parameters affects the inversion in some areas of the model. For instance, the scattering point at $x=3$ km and $z=1$ km and the low η magnitude layer around $z=3$ km are altered with the cross-talk.

Figure 5 shows the vertical profile at 12 km. The horizontal velocity is well reconstructed and the main features are recovered. Furthermore, the inverted ϵ parameter is reasonable and the short wavelength features are well recovered for parts of the model deeper than 1.5 km. From the ϵ profile, the variations in the shallow part of ϵ model are not well reconstructed. In fact, in the shallow parts of the model, the scattering angles are large compared to the deeper parts. The same observation

holds for the η parameter. To improve these models, higher frequencies should be used. However, initial models for ϵ and η with accurate long wavelengths are essential to achieve good inversion results.

Updating the long wavelengths of the model can be achieved using the low frequencies. However, ultra-low frequencies are required to recover ϵ and η long wavelength smooth component. For reasonable low frequencies, the cross-talk between v_h and ϵ and η parameters is limited and we can reconstruct a good long wavelength v_h . The redundancy of seismic data can help in mitigating the trade-off between v_h and ϵ for small scattering angles. Therefore, acoustic frequency domain multi-parameter inversion using a parameterization based on v_h , ϵ and η can provide good inverted models which is inline with the sensitivity analysis by Alkhalifah (2016) and Djebbi et al. (2017).

Figure 6 shows the data misfit history for the 4 frequency bands. The misfit is the time domain data misfit. The misfit is largely reduced, which confirms the convergence of the inversion. In Figure 7 we compare a shot gather modelled using the inverted model and the observed one. We show the data residual in Figure 7c presented at the same scale as the modelled and observed shot gathers. In general, the data residual is weak and the diving waves and most of the reflections are perfectly modelled using the inverted model.

[Figure 3 about here.]

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]

Hierarchical inversion of v_h , ϵ and η

In this example, we consider inverting hierarchically for v_h , ϵ and η . For low frequencies, given by the first frequency band, we only invert for v_h . Long wavelengths of the horizontal velocity are updated without leakage from the remaining parameters. Starting from the second frequency band, we include ϵ parameter for the inversion. Finally, for frequency bands 3 and 4, we invert the three parameters simultaneously.

The final inverted models are shown in Figure 8. v_h model is well recovered. Also, short wavelengths for ϵ and η are well reconstructed in the deep parts of the model. In the shallow parts, ϵ and η parameters are not recovered accurately due to sensitivity properties of these parameters. η parameter is affected by the leakage from the other parameters. Therefore, many artifacts are present. For instance, the low-velocity layers around 2.2 and 3 km deep are mainly caused by leakage from ϵ parameter.

Figure 9 shows the vertical profiles at $x = 12$ km. Compared to the previous example, the quality of the inverted horizontal velocity is comparable. The main features are recovered. The inverted ϵ parameter is reasonable and the short wavelength features are well inverted. The long wavelengths features are missing in the η parameter inversion. Also, it is contaminated with trade-off artifacts from the other anisotropy parameters.

[Figure 7 about here.]

Simultaneous inversion of v_h and ϵ , η maintained as the initial η_0

As the sensitivity of the recorded data to η parameter is weak, Alkhalifah (2016) suggested inverting only for v_h and ϵ . With this approach, the main objective is to recover a good horizontal velocity model. The ϵ parameter is useful for inverting elastic data using the acoustic assump-

tion (Guitton and Alkhalifah, 2016). As the ϵ radiation pattern is similar to the radiation patterns of density, it can absorb the errors in amplitude caused by ignoring density during the acoustic inversion.

We test this inversion strategy for the same initial model. η parameter is fixed as the initial η_0 model. We perform a simultaneous inversion of both v_h and ϵ parameter. Figure 10 shows the inverted models. The vertical profiles at $x = 12$ km for v_h and ϵ are shown in Figure 11. The horizontal velocity is accurately reconstructed. The short wavelengths of ϵ parameter are also recovered, especially in the deep parts of the model. In these areas, the data are mainly governed by reflections with small scattering angles. A good background η model is required to minimize the errors in v_h inversion. Such models can be recovered using tomographic methods. Therefore, ignoring η short wavelengths, will not affect the inversion results.

[Figure 8 about here.]

[Figure 9 about here.]

Realistic initial models

In this example, we consider more realistic initial models. Thomsen parameter δ , relating ϵ and η , is not inverted using surface recorded seismic data. It is obtained by matching well logs. As a result, when well information is not available, a realistic initial model is to consider δ equals zero. Conventionally, the initial NMO velocity v_{n0} and η_0 models are obtained using tomographic methods. Therefore, we construct the initial models for v_h , ϵ and η parameters as follows: we smooth the exact v_h and η parameters using 1.5 km radius triangular smoothing. The resulting

models are equivalent to the ones obtained from tomography. Then, we consider $\delta = 0$ which gives $\epsilon_0 = \eta_0$. The models are shown in Figure 12.

[Figure 10 about here.]

The inversion parameters are the same as in the previous examples. We use a frequency continuation approach with the same frequency parameters listed in Table 1. The inversion results are shown in Figure 13. The v_h model contains the exact model structures. However, looking into the vertical profiles in Figure 14, the reconstructed models are shifted in depth. These shifts are caused by ambiguity between the δ parameter and the depth. In imaging, inaccurate δ models produce a vertical misplacement of the reflectors. For the ϵ parameter, only the short wavelengths are recovered. The difference between the initial and exact models is large. Due to the scattering properties of ϵ , this large difference can not be retrieved. The η parameter is also affected by ϵ model errors. Leakage from v_h and ϵ affects the inverted η parameter. We conclude that to accurately invert for v_h , good background ϵ and η parameters are required.

[Figure 11 about here.]

[Figure 12 about here.]

An independent modelling code

We model the observed data with a time domain acoustic VTI solver. We consider a free-surface boundary condition. Also, the absorbing boundary conditions are implemented differently compared to the frequency domain solver. To further complicate the inversion task and mimic a real data inversion situation, the source wavelet is unknown. The source wavelet is estimated at every

iteration. The l_2 misfit function is used for the inversion. In this example, we consider initial models obtained by smoothing the exact ones.

Figure 15 shows the inverted models. Figure 16 shows the vertical profiles $x = 12$ km. Most of the features are well recovered, especially in the shallow parts of the model. Details in the deep are not well reconstructed. For ϵ and η parameters, the short wavelengths are well resolved. Considering the complexity of the Marmousi II model, and the problems associated with multi-parameter inversion, the final inversion result is acceptable.

[Figure 13 about here.]

REAL DATA EXAMPLE

We consider a 2-D marine data set acquired by CGG from the North-Western Australia Continental shelf. The data are BroadSeis data acquired with variable depth streamers. This type of acquisition improves the signal to noise ratio for low frequencies (Soubaras and Dowle, 2010). The data set consists of 1824 shots with 648 receivers per shot. We display one shot gather and its frequency content in Figures 17a and 17b, respectively. The data minimum frequency is 2.5 Hz and the maximum recording offset is 8.3 km. The number of receivers is reduced to 324 receivers per shot, covering the full offset, to reduce the computational cost.

[Figure 14 about here.]

We mute the noise prior to the first arrivals and filter the data to low frequencies appropriate for full waveform inversion (Figure 18). We consider only 300 shots covering a part of the model to reduce the computational cost. The lateral extent of the inverted model is 12.5 km. The initial model is isotropic and is estimated using reflection waveform inversion (RWI) (Wu and Alkhalifah,

2015). The initial velocity model is smooth and represents well the kinematics of wave propagation. Figure 19 shows the initial velocity model. We estimate the water depth by migrating the data using the water velocity. The source wavelet, shown in Figure 20, is estimated using the direct arrivals in the water layer. As the water velocity is known, the source signature can be well inverted (Kim et al., 2011; Kalita and Alkhalifah, 2017).

[Figure 15 about here.]

[Figure 16 about here.]

[Figure 17 about here.]

The acoustic approximation, used to model the synthetic data, suffers from inaccurate estimation of data amplitudes. Therefore, the conventional least squares misfit function fails to properly invert for the Earth's model. To deal with this problem, we use the logarithmic misfit function (Shin and Min, 2006) to invert for the phase only. The data residuals for the logarithmic misfit function are given as,

$$\Delta \mathbf{d} = \Delta \Phi = \Im \left(\log \left(\frac{\mathbf{U}(\mathbf{m})}{\mathbf{d}} \right) \right), \tag{15}$$

where $\Delta \Phi$ is the phase difference, $\log$ is the complex logarithm function, $\mathbf{U}(\mathbf{m})$ is the modeled data and $\mathbf{d}$ is the observed data. $\Im$ denotes the imaginary part. To avoid dividing by zero, a small constant (compared to the data average value) is added to the denominator in equation (15).

The logarithmic misfit function is given by,

$$E(\mathbf{m}) = \frac{1}{2} \Im \left(\log \left(\frac{\mathbf{U}(\mathbf{m})}{\mathbf{d}} \right) \right)^2. \tag{16}$$

The gradient is given as,

$$\mathbf{g}(\mathbf{m}) = \mathbf{p}_1 = \frac{\partial E(\mathbf{m})}{\partial \mathbf{m}} = \Im(\mathbf{F}^T \Delta \mathbf{d}_b), \quad (17)$$

where $\Delta \mathbf{d}_b = \frac{\Im(\log(\frac{\mathbf{U}(\mathbf{m})}{\mathbf{d}}))}{\mathbf{U}(\mathbf{m})}$ is the complex back-propagated residual.

We consider a preconditioned conjugate gradient method for the inversion. The preconditioning operator is the diagonal of the approximate Hessian matrix. We use the frequency domain scattering integral method for inversion and we invert for two frequency bands, given in Table 2,

[Table 2 about here.]

We show in Figure 21 the inverted anisotropic models. Compared to the initial model, additional high-velocity layers can be observed in the reconstructed v_h . We compare a v_h vertical profile at $x = 10.5$ km to the well-log filtered to the inversion resolution in Figure 22. The final v_h vertical profile contains a low-resolution high-velocity layer which corresponds, most likely, to the 2.1 km high velocity in the well-log. The well-log velocity corresponds to the vertical velocity, therefore, compared to the inverted v_h we can observe a shift which might be the result of the under/over estimation of anisotropy. The same geological area was the focus of Kalita and Alkhalifah (2017) in the context of time domain acoustic isotropic FWI. The inverted velocities are similar but with a lower resolution in our study. To recover a more accurate and higher resolution model we need to consider higher frequencies and good long-wavelength initial ϵ and η anisotropy models. These initial models will help in placing the velocity layers at their correct depth. The inverted ϵ and η parameters contain several high magnitude areas. As the background model is not available, only high wavenumber features are recovered. The resulting models are in good agreement with the radiations patterns. The long-wavelengths features of ϵ and η can not be reconstructed using FWI,

unless lower frequencies and larger offsets are available. We finally show the phase misfit for the two frequency bands in Figure 23. The data misfits are largely reduced.

[Figure 18 about here.]

[Figure 19 about here.]

[Figure 20 about here.]

To confirm the accuracy of the inverted models, we generate the time domain modeled data using the initial and final anisotropic models. The data modeled using the initial and inverted models are compared to the filtered real data. We show two shot gathers in Figures 24 and 25. The data generated with the inverted models contain additional reflections, which correspond to the real data reflections. However, we notice a small phase shift of the modeled data using the final model. This phase shift is the result of under-estimating the velocity in some shallow parts of the model.

[Figure 21 about here.]

[Figure 22 about here.]

Finally, we perform reverse time migration (RTM) using the initial and inverted models. We show the images, and the resulting angle gathers in Figures 26 and 27, respectively. As the initial model already contains the required low wavenumbers for migration, slight changes are observed in the RTM image with the final model. In the angle gathers, multiple locations are improved, and the gathers are flattened. These locations are indicated by the green arrows. At the well location ($x = 10.5$ km) we observe an improvement in the gathers. However, there is still errors in the velocity which explains the shifts observed in data comparison Figure 25.

[Figure 23 about here.]

[Figure 24 about here.]

CONCLUSIONS

We applied a frequency domain FWI using the scattering integral approach to invert for a transversely isotropic medium with a vertical axis of symmetry. The VTI wave equation is parameterized using the horizontal velocity v_h , ϵ and η anisotropy parameters to reduce the trade-off between different parameters. We used a preconditioned conjugate gradient method to update the model. The gradient step lengths are estimated through a second order approximation of the objective function. It takes into account perturbations in the three anisotropy parameters and can reduce the inversion errors caused by the trade-off between the parameters. We tested various inversion strategies: first, v_h , ϵ and η are inverted simultaneously. In a second test, we considered a hierarchical inversion approach. Finally, we kept η fixed to its initial values and inverted for v_h and ϵ . We showed that the inversion gives an accurate inverted v_h model for all three tests. Ignoring η in the inversion reduces the cost and the accuracy of v_h inversion is not affected. However, a good initial η model is required. For ϵ parameter, we recovered only its short wavelengths which confirms the radiation patterns predictions. We also considered a realistic initial model where $\epsilon_0 = \eta_0$. In this situation, the difference between the exact and initial ϵ models is large. The structure of the inverted models is accurate; however, the positions of the velocity layers are shifted. Furthermore, we tested the inversion on data generated with a different solver. The inversion results are also acceptable.

Finally, the proposed inversion method is applied to a real data set provided by CGG. The inversion results show additional features not visible in the initial model. Data comparison, RTM images,

and angle gathers show improvements in the inverted models compared to the initial velocity. To conclude, v_h , ϵ and η parameterization is suitable for multi-parameter inversion within a frequency domain multi-scale framework.

ACKNOWLEDGEMENTS

We are grateful to King Abdullah University of Science and Technology (KAUST) for the financial support. We thank the members of seismic wave analysis group (SWAG), namely Yunseok Choi, Juwon Oh and Nabil Masmoudi for the useful discussions. Also, we are grateful to Zedong Wu and Mahesh Kalita for their pieces of advice on the real data example. The real data shown in this study are provided courtesy of CGG. The well-log information is provided by Geoscience Australia. We thank Stéphane Operto and anonymous reviewers for their fruitful suggestions.

REFERENCES

- Alkhalifah, T., 2016, Research note: Insights into the data dependency on anisotropy: An inversion prospective: *Geophysical Prospecting*, **64**, 505–513.
- Alkhalifah, T., and R.-É. Plessix, 2014, A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study: *Geophysics*, **79**, R91–R101.
- Alkhalifah, T., and I. Tsvankin, 1995, Velocity analysis for transversely isotropic media: *Geophysics*, **60**, 1550–1566.
- Amestoy, P., C. Ashcraft, O. Boiteau, A. Buttari, J.-Y. L'Excellent, and C. Weisbecker, 2015, Improving multifrontal methods by means of block low-rank representations: *SIAM Journal on Scientific Computing*, **37**, A1451–A1474.
- Baumstein, A., et al., 2014, Extended subspace method for attenuation of crosstalk in multi-parameter full wavefield inversion: Presented at the 2014 SEG Annual Meeting, Society of Exploration Geophysicists.
- Berenger, J.-P., 1994, A perfectly matched layer for the absorption of electromagnetic waves: *Journal of computational physics*, **114**, 185–200.
- Beydoun, W. B., and A. Tarantola, 1988, First Born and Rytov approximations: Modeling and inversion conditions in a canonical example.: *Acoust. Soc. Am.*, 1045–1055.
- Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: *Geophysics*, **60**, 1457–1473.
- Chen, P., T. H. Jordan, and L. Zhao, 2007, Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods: *Geophysical Journal International*, **170**, 175–181.
- Davis, T. A., 2004, Algorithm 832: Umfpack v4. 3—an unsymmetric-pattern multifrontal method: *ACM Transactions on Mathematical Software (TOMS)*, **30**, 196–199.

Djebbi, R., R.-É. Plessix, and T. Alkhalifah, 2017, Analysis of the traveltime sensitivity kernels for an acoustic transversely isotropic medium with a vertical axis of symmetry: *Geophysical Prospecting*, **65**, 22–34.

Duveneck, E., P. Milcik, P. M. Bakker, and C. Perkins, 2008, Acoustic VTI wave equations and their application for anisotropic reverse-time migration, *in* SEG Technical Program Expanded Abstracts 2008: Society of Exploration Geophysicists, 2186–2190.

Gholami, Y., R. Brossier, S. Operto, A. Ribodetti, and J. Virieux, 2013, Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 1: Sensitivity and trade-off analysis: *Geophysics*, **78**, R81–R105.

Guittou, A., and T. Alkhalifah, 2016, Full-waveform Inversion in an Anisotropic Elastic Earth-Can we isolate the role of density and shear wave velocity?, *in* 78th EAGE Conference and Exhibition 2016: European Association of Geoscientists and Engineers.

Hu, W., A. Abubakar, T. Habashy, and J. Liu, 2011, Preconditioned non-linear conjugate gradient method for frequency domain full-waveform seismic inversion: *Geophysical Prospecting*, **59**, 477–491.

Kalita, M., and T. Alkhalifah, 2017, Efficient full waveform inversion using the excitation representation of the source wavefield: *Geophysical Journal International*, **210**, 1581–1594.

Kazei, V., and T. Alkhalifah, 2018, Waveform inversion for orthorhombic anisotropy with P-waves: feasibility & resolution: *Geophysical Journal International*.

Kennett, B., M. Sambridge, and P. Williamson, 1988, Subspace methods for large inverse problems with multiple parameter classes: *Geophysical Journal International*, **94**, 237–247.

Kim, Y., D.-J. Min, and C. Shin, 2011, Frequency-domain reverse-time migration with source estimation: *Geophysics*, **76**, S41–S49.

Liao, Q., and G. A. McMechan, 1996, Multifrequency viscoacoustic modeling and inversion: *Geo-*

- physics, **61**, 1371–1378.
- Liu, Y., J. Yang, B. Chi, and L. Dong, 2015, An improved scattering-integral approach for frequency-domain full waveform inversion: *Geophysical Journal International*, **202**, 1827–1842.
- Lu, Y. Y., 2006, Minimizing the discrete reflectivity of perfectly matched layers: *IEEE photonics technology letters*, **18**, 487–489.
- Masmoudi, N., and T. Alkhalifah, 2016, A new parameterization for waveform inversion in acoustic orthorhombic media: *Geophysics*, **81**, R157–R171.
- Métivier, L., R. Brossier, J. Virieux, and S. Operto, 2013, Full waveform inversion and the truncated Newton method: *SIAM Journal on Scientific Computing*, **35**, B401–B437.
- Mora, P., 1988, Elastic wave-field inversion of reflection and transmission data: *Geophysics*, **53**, 750–759.
- Operto, S., R. Brossier, L. Combe, L. Métivier, A. Ribodetti, and J. Virieux, 2014, Computationally efficient three-dimensional acoustic finite-difference frequency-domain seismic modeling in vertical transversely isotropic media with sparse direct solver: *Geophysics*, **79**, T257–T275.
- Operto, S., A. Miniussi, R. Brossier, L. Combe, L. Métivier, V. Monteiller, A. Ribodetti, and J. Virieux, 2015, Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: Application to Valhall in the visco-acoustic vertical transverse isotropic approximation: *Geophysical Journal International*, **202**, 1362–1391.
- Plessix, R.-É., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: *Geophysical Journal International*, **167**, 495–503.
- Plessix, R.-É., and Q. Cao, 2011, A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium: *Geophysical Journal International*, **185**, 539–556.
- Podgornova, O., S. Leaney, L. Liang, et al., 2015, Analysis of resolution limits of VTI anisotropy

with full waveform inversion: Presented at the 2015 SEG Annual Meeting, Society of Exploration Geophysicists.

Pratt, R. G., C. Shin, and G. Hick, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362.

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

Shin, C., and D.-J. Min, 2006, Waveform inversion using a logarithmic wavefield: *Geophysics*, **71**, R31–R42.

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.

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

Soubaras, R., and R. Dowle, 2010, Variable-depth streamer- a broadband marine solution: First break, **28**.

Tao, Y., and M. K. Sen, 2013, Frequency-domain full waveform inversion with a scattering-integral approach and its sensitivity analysis: *Journal of Geophysics and Engineering*, **10**, 065008.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.

Thomsen, L., 1986, Weak elastic anisotropy: *Geophysics*, **51**, 1954–1966.

Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: *Geophysical Journal International*, **160**, 195–216.

Vigh, D., X. Cheng, K. Jiao, D. Sun, and J. Kapoor, 2014, Multiparameter TTI full waveform inversion on long-offset broadband acquisition: a case study, *in* SEG Technical Program Expanded Abstracts 2014: Society of Exploration Geophysicists, 1061–1065.

- Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, WCC1–WCC26.
- Wang, S., M. V. de Hoop, and J. Xia, 2011, On 3D modeling of seismic wave propagation via a structured parallel multifrontal direct helmholtz solver: *Geophysical Prospecting*, **59**, 857–873.
- Wu, Z., and T. Alkhalifah, 2015, Full waveform inversion based on scattering angle enrichment with application to real dataset, *in* SEG Technical Program Expanded Abstracts 2015: Society of Exploration Geophysicists, 1258–1262.
- , 2016, Waveform inversion for acoustic VTI media in frequency domain, *in* SEG Technical Program Expanded Abstracts 2016: Society of Exploration Geophysicists, 1184–1189.
- Zhou, C., W. Cai, Y. Luo, G. T. Schuster, and S. Hassanzadeh, 1995, Acoustic wave-equation traveltime and waveform inversion of crosshole seismic data: *Geophysics*, **60**, 765–773.

LIST OF FIGURES

1. Figure 1: Single frequency Born sensitivity kernels for a VTI with increasing velocity with depth model: $v(z) = 1.5 + 0.8z$ km/s: (a) v_h , (b) ϵ and (c) η . The kernels amplitudes are normalized to the maximum value and plotted at the same scale.
2. Figure 2: The VTI Marmousi II model: (a) the horizontal velocity v_h , (b) ϵ parameter, (c) η parameter.
3. Figure 3: The VTI initial models: (a) the horizontal velocity v_h , (b) ϵ parameter, (c) η parameter.
4. Figure 4: FWI inverted models: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter. All three parameters are simultaneously inverted.
5. Figure 5: Vertical profiles for the simultaneous inversion at $x = 12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.
6. Figure 6: Convergence history for the simultaneous inversion. The misfit is the $l-2$ time domain misfit, therefore it represents the misfit for all frequencies between 2 and 12 Hz.
7. Figure 7: Shot gather comparison for a source located at $x = 8.25$ km and $z = 0.025$ km. (a) the exact observed data, (b) the modelled data with the inverted model, (c) the data residual. All figures are plotted at the same scale.
8. Figure 8: FWI inverted models: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter. The three parameters are hierarchically inverted: for frequency band 1, only v_h is inverted, then for frequency band 2, v_h and ϵ are inverted. Finally all three parameters are inverted for frequency bands 3 and 4.

9. Figure 9: Vertical profiles for the hierarchical inversion at $x = 12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.
10. Figure 10: FWI inverted models: (a) the horizontal velocity v_h , (b) ϵ parameter. η parameter is fixed as the initial η_0 . The two parameters are inverted simultaneously.
11. Figure 11: Vertical profiles for the inversion with η fixed as the initial η_0 at $x = 12$ km. (a) v_h and (b) ϵ parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.
12. Figure 12: Realistic VTI initial models, corresponding to δ parameter equals zero: (a) the horizontal velocity v_h , (b) ϵ parameter, (c) η parameter.
13. Figure 13: FWI inverted models for inversion with realistic initial models: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter.
14. Figure 14: Vertical profiles for the inversion with realistic initial models at $x = 12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.
15. Figure 15: FWI inverted models for inversion without inverse crime: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter. The observed and synthetic data are modeled with two separate modeling codes to avoid the inverse crime (the same modeling code is used to generate both data).
16. Figure 16: Vertical profiles for the inversion without inverse crime at $x = 12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

17. Figure 17: Shot gather example: (a) Shot gather for a source located at $(x, z) = (3.75, 0.005)$ km, (b) data frequency spectrum.
18. Figure 18: Bandpass filtered shot gather between 3 and 12 Hz: (a) Shot gather for a source located at $(x, z) = (3.75, 0.005)$ km, (b) data frequency spectrum.
19. Figure 19: Initial velocity model.
20. Figure 20: The inverted source wavelet. (a) the wavelet and (b) the wavelet spectrum.
21. Figure 21: Inverted models: (a) v_h model (b) ϵ parameter model and (c) η parameter model.
22. Figure 22: Vertical profiles at $x = 10.5$ km compared with well-log.
23. Figure 23: Convergence history. The misfit is given by the phase difference.
24. Figure 24: Data comparison for shot gather 50: (a) initial model data (b) final model data.
25. Figure 25: Data comparison for shot gather 100: (a) initial model data (b) final model data.
26. Figure 26: Image using reverse time migration (RTM) for (a) the initial model (b) the final model.
27. Figure 27: Angle gathers for (a) the initial model (b) the final model. The angle gathers are shown for multiple lateral positions. The angles are between 0 and 45° . The green arrows show locations where there is improvement.

LIST OF TABLES

1. Table 1: VTI Marmousi II inversion parameters. 4 bands of multiples frequencies are used.

itmax is the maximum number of iterations.

2. Table 2: CGG real data inversion parameters. 2 bands of multiples frequencies are used.

itmax is the maximum number of iterations.

Figure1a: Single frequency Born sensitivity kernels for a VTI with increasing velocity with depth model: $v(z)=1.5+0.8z$ km/s: (a) v_{hr} , (b) ϵ and (c) η . The kernels amplitudes are normalized to the maximum value and plotted at the same scale.

44x19mm (300 x 300 DPI)

Figure 1b: Single frequency Born sensitivity kernels for a VTI with increasing velocity with depth model: $v(z) = 1.5 + 0.8z$ km/s: (a) v_h , (b) ϵ and (c) η . The kernels amplitudes are normalized to the maximum value and plotted at the same scale.

101x44mm (300 x 300 DPI)

Figure1c: Single frequency Born sensitivity kernels for a VTI with increasing velocity with depth model: $v(z)=1.5+0.8z$ km/s: (a) v_h , (b) ϵ and (c) η . The kernels amplitudes are normalized to the maximum value and plotted at the same scale.

101x44mm (300 x 300 DPI)

Figure 2a: The VTI Marmousi II model: (a) the horizontal velocity v_h , (b) ϵ parameter, (c) η parameter.

220x70mm (300 x 300 DPI)

Figure2b: The VTI Marmousi II model: (a) the horizontal velocity v_h , (b) ϵ parameter, (c) η parameter.
220x71mm (300 x 300 DPI)

Figure 2c: The VTI Marmousi II model: (a) the horizontal velocity v_h , (b) ϵ parameter, (c) η parameter.

220x71mm (300 x 300 DPI)

Figure3a: The VTI initial models: (a) the horizontal velocity v_h , (b) ϵ parameter, (c) η parameter.
220x70mm (300 x 300 DPI)

Figure 3b: The VTI initial models: (a) the horizontal velocity v_h , (b) ϵ parameter, (c) η parameter.

220x71mm (300 x 300 DPI)

Figure 3c: The VTI initial models: (a) the horizontal velocity v_h , (b) ϵ parameter, (c) η parameter.
220x71mm (300 x 300 DPI)

Figure 4a: FWI inverted models: (a) the horizontal velocity v_h , (b) ε parameter and (c) η parameter. All three parameters are simultaneously inverted.

221x71mm (300 x 300 DPI)

Figure4b: FWI inverted models: (a) the horizontal velocity v_h , (b) ε parameter and (c) η parameter. All three parameters are simultaneously inverted.

220x71mm (300 x 300 DPI)

Figure 4c: FWI inverted models: (a) the horizontal velocity v_h , (b) ε parameter and (c) η parameter. All three parameters are simultaneously inverted.

220x71mm (300 x 300 DPI)

Figure5a: Vertical profiles for the simultaneous inversion at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

191x83mm (300 x 300 DPI)

Figure 5b: Vertical profiles for the simultaneous inversion at $x=12$ km. (a) v_h , (b) ε parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

191x80mm (300 x 300 DPI)

Figure 5c: Vertical profiles for the simultaneous inversion at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

191x78mm (300 x 300 DPI)

Figure 6: Convergence history for the simultaneous inversion. The misfit is the l-2 time domain misfit, therefore it represents the misfit for all frequencies between 2 and 12 Hz.

238x109mm (300 x 300 DPI)

Figure 7a: Shot gather comparison for a source located at $x=8.25$ km and $z=0.025$ km. (a) the exact observed data, (b) the modelled data with the inverted model, (c) the data residual. All figures are plotted at the same scale.

224x119mm (300 x 300 DPI)

Figure 7b: Shot gather comparison for a source located at $x=8.25$ km and $z=0.025$ km. (a) the exact observed data, (b) the modelled data with the inverted model, (c) the data residual. All figures are plotted at the same scale.

224x119mm (300 x 300 DPI)

Figure 7c: Shot gather comparison for a source located at $x=8.25$ km and $z=0.025$ km. (a) the exact observed data, (b) the modelled data with the inverted model, (c) the data residual. All figures are plotted at the same scale.

224x119mm (300 x 300 DPI)

Figure 8a: FWI inverted models: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter. The three parameters are hierarchically inverted: for frequency band 1, only v_h is inverted, then for frequency band 2, v_h and ϵ are inverted. Finally all three parameters are inverted for frequency bands 3 and 4.

224x72mm (300 x 300 DPI)

Figure 8b: FWI inverted models: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter. The three parameters are hierarchically inverted: for frequency band 1, only v_h is inverted, then for frequency band 2, v_h and ϵ are inverted. Finally all three parameters are inverted for frequency bands 3 and 4.

224x73mm (300 x 300 DPI)

Figure 8c: FWI inverted models: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter. The three parameters are hierarchically inverted: for frequency band 1, only v_h is inverted, then for frequency band 2, v_h and ϵ are inverted. Finally all three parameters are inverted for frequency bands 3 and 4.

224x73mm (300 x 300 DPI)

Figure 9a: Vertical profiles for the hierarchical inversion at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x85mm (300 x 300 DPI)

Figure 9b: Vertical profiles for the hierarchical inversion at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x82mm (300 x 300 DPI)

Figure 9c: Vertical profiles for the hierarchical inversion at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x80mm (300 x 300 DPI)

Figure 10a: FWI inverted models: (a) the horizontal velocity v_h , (b) ϵ parameter. η parameter is fixed as the initial η_0 . The two parameters are inverted simultaneously.

224x72mm (300 x 300 DPI)

Figure 10b: FWI inverted models: (a) the horizontal velocity v_h , (b) ϵ parameter. η parameter is fixed as the initial η_0 . The two parameters are inverted simultaneously.

224x73mm (300 x 300 DPI)

Figure 11a: Vertical profiles for the inversion with η fixed as the initial η_0 at $x=12$ km. (a) v_h , (b) ε parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x85mm (300 x 300 DPI)

Figure 11b: Vertical profiles for the inversion with η fixed as the initial η_0 at $x=12$ km. (a) v_h , (b) ϵ parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x82mm (300 x 300 DPI)

Figure 12a: Realistic VTI initial models, corresponding to δ parameter equals zero: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter.

224x72mm (300 x 300 DPI)

Figure 12b: Realistic VTI initial models, corresponding to δ parameter equals zero: (a) the horizontal velocity v_{hr} , (b) ϵ parameter and (c) η parameter.

224x73mm (300 x 300 DPI)

Figure 12c: Realistic VTI initial models, corresponding to δ parameter equals zero: (a) the horizontal velocity v_h , (b) ε parameter and (c) η parameter.

224x73mm (300 x 300 DPI)

Figure 13a: FWI inverted models for inversion with realistic initial models: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter.

224x72mm (300 x 300 DPI)

Figure 13b: FWI inverted models for inversion with realistic initial models: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter.

224x73mm (300 x 300 DPI)

Figure 13c: FWI inverted models for inversion with realistic initial models: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter.

224x73mm (300 x 300 DPI)

Figure 14a: Vertical profiles for the inversion with realistic initial models at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x85mm (300 x 300 DPI)

Figure 14b: Vertical profiles for the inversion with realistic initial models at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x82mm (300 x 300 DPI)

Figure 14c: Vertical profiles for the inversion with realistic initial models at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x80mm (300 x 300 DPI)

Figure 15a: FWI inverted models for inversion without inverse crime: (a) the horizontal velocity v_h , (b) ϵ parameter and (c) η parameter. The observed and synthetic data are modeled with two separate modeling codes to avoid the inverse crime (the same modeling code is used to generate both data).

128x51mm (300 x 300 DPI)

Figure 15b: FWI inverted models for inversion without inverse crime: (a) the horizontal velocity v_h , (b) ε parameter and (c) η parameter. The observed and synthetic data are modeled with two separate modeling codes to avoid the inverse crime (the same modeling code is used to generate both data).

127x50mm (300 x 300 DPI)

Figure 15c: FWI inverted models for inversion without inverse crime: (a) the horizontal velocity v_h , (b) ε parameter and (c) η parameter. The observed and synthetic data are modeled with two separate modeling codes to avoid the inverse crime (the same modeling code is used to generate both data).

128x51mm (300 x 300 DPI)

Figure 16a: Vertical profiles for the inversion without inverse crime at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x85mm (300 x 300 DPI)

Figure 16b: Vertical profiles for the inversion without inverse crime at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x82mm (300 x 300 DPI)

Figure 16c: Vertical profiles for the inversion without inverse crime at $x=12$ km. (a) v_h , (b) ϵ parameter and (c) η parameter. The blue, black and red curves are the exact, the initial and the inverted models, respectively.

195x80mm (300 x 300 DPI)

Figure 17a: Shot gather example: (a) Shot gather for a source located at $(x, z)=(3.75, 0.005)$ km, (b) data frequency spectrum.

250x223mm (300 x 300 DPI)

Figure 17b: Shot gather example: (a) Shot gather for a source located at $(x, z) = (3.75, 0.005)$ km, (b) data frequency spectrum.

225x129mm (300 x 300 DPI)

Figure 18a: Bandpass filtered shot gather between 3 and 12 Hz: (a) Shot gather for a source located at (x, z)=(3.75, 0.005) km, (b) data frequency spectrum.

250x223mm (300 x 300 DPI)

Figure 18b: Bandpass filtered shot gather between 3 and 12 Hz: (a) Shot gather for a source located at $(x, z) = (3.75, 0.005)$ km, (b) data frequency spectrum.

225x127mm (300 x 300 DPI)

Figure 19: Initial velocity model.

188x54mm (300 x 300 DPI)

Figure 20a: The inverted source wavelet. (a) the wavelet and (b) the wavelet spectrum.

220x119mm (300 x 300 DPI)

Figure 20b: The inverted source wavelet. (a) the wavelet and (b) the wavelet spectrum.

225x127mm (300 x 300 DPI)

Figure 21a: Inverted models: (a) v_h model (b) ϵ parameter model and (c) η parameter model.

188x52mm (300 x 300 DPI)

Figure 21b: Inverted models: (a) v_h model (b) ϵ parameter model and (c) η parameter model.

188x53mm (300 x 300 DPI)

Figure 21c: Inverted models: (a) v_h model (b) ϵ parameter model and (c) η parameter model.

188x53mm (300 x 300 DPI)

Figure 22: Vertical profiles at $x=10.5$ km compared with well-log.

243x95mm (300 x 300 DPI)

Figure 23: Convergence history. The misfit is given by the phase difference.

240x116mm (300 x 300 DPI)

Figure 24a: Data comparison for shot gather 50: (a) initial model data (b) final model data.

138x62mm (300 x 300 DPI)

Figure 24b: Data comparison for shot gather 50: (a) initial model data (b) final model data.

138x62mm (300 x 300 DPI)

Figure 25a: Data comparison for shot gather 100: (a) initial model data (b) final model data.

138x62mm (300 x 300 DPI)

Figure 25b: Data comparison for shot gather 100: (a) initial model data (b) final model data.

138x62mm (300 x 300 DPI)

Figure 26a: Image using reverse time migration (RTM) for (a) the initial model (b) the final model.
147x69mm (300 x 300 DPI)

Figure 26b: Image using reverse time migration (RTM) for (a) the initial model (b) the final model.

147x69mm (300 x 300 DPI)

Figure 27a: Angle gathers for (a) the initial model (b) the final model. The angle gathers are shown for multiple lateral positions. The angles are between 0 and 45°. The green arrows show locations where there is improvement.

120x45mm (300 x 300 DPI)

Figure 27b: Angle gathers for (a) the initial model (b) the final model. The angle gathers are shown for multiple lateral positions. The angles are between 0 and 45° . The green arrows show locations where there is improvement.

120x45mm (300 x 300 DPI)

Band	f_{min} (Hz)	f_{max} (Hz)	Δf (Hz)	itmax
1	2.0	4.0	0.25	50
2	3.5	6.0	0.5	50
3	5.5	9.0	0.5	40
4	8.5	12.0	0.5	40

Table 1: VTI Marmousi II inversion parameters. 4 bands of multiples frequencies are used. itmax is the maximum number of iterations.

Band	f_{min} (Hz)	f_{max} (Hz)	Δf (Hz)	itmax
1	4.0	5.5	0.125	40
2	5.0	7.5	0.25	40

Table 2: CGG real data inversion parameters. 2 bands of multiples frequencies are used. itmax is the maximum number of iterations.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.