Inversion gradients for acoustic VTI wavefield tomography*

Vladimir Li[1], Hui Wang[1], Ilya Tsvankin[1], Esteban Díaz[1] & Tariq Alkhalifah[2]
[1]*Center for Wave Phenomena, Colorado School of Mines*
[2]*King Abdullah University of Science and Technology*

ABSTRACT
Wavefield tomography can handle complex subsurface geology better than ray-based techniques and, ultimately, provide a higher resolution. Here, we implement forward and adjoint wavefield extrapolation for VTI (transversely isotropic with a vertical symmetry axis) media using a generalized pseudospectral operator based on a separable approximation for the P-wave dispersion relation. This operator is employed to derive the gradients of the differential semblance optimization (DSO) and modified image-power objective functions. We also obtain the gradient expressions for a data-domain objective function that can more easily incorporate borehole information necessary for stable VTI velocity analysis. These gradients are similar to the ones obtained with a space-time finite-difference (FD) scheme for a system of coupled wave equations but the pseudospectral method is not hampered by the imprint of the shear-wave artifact. Numerical examples also show the potential advantages of the modified image-power objective function in estimating the anellipticity parameter η.

Key words: anisotropy, wave equation, data domain, image domain, tomography, gradient

1 INTRODUCTION

Wavefield tomography can be implemented in the data or image domain depending on the way of formulating the objective function. Data-domain methods enforce the similarity between the predicted and observed seismic wavefields. The image-domain approach requires an additional migration step and relies, in accordance with the semblance principle, on the consistency of migrated images for different experiments (Al-Yahya, 1989; Sattlegger, 1975; Perrone and Sava, 2012). There are various modifications of image-domain tomography that employ different migration operators, imaging conditions, and types of image gathers (e.g. Sava, 2014). The objective function in either domain is typically minimized using gradient-based techniques, with the gradients obtained by the adjoint-state method (ASM) (Tarantola, 1984; Tromp et al., 2005; Plessix, 2006). Despite the difference in their objective functions, both data- and image-domain methods use the same wave equation and observed wavefields (Sava, 2014).

In this paper, we focus on wavefield extrapolation and gradient derivation, which are common key steps for both groups of methods. Our algorithm is designed for transversely isotropic models with a vertical symmetry axis (VTI) and can be easily extended to tilted TI (TTI) media. Both VTI and TTI models are widely used to improve the results of time and depth imaging and reflection tomography. Optimally, anisotropic inversion requires elastic wavefield extrapolation and benefits from including shear and mode-converted waves. However, incorporating shear-wave information into wavefield-based inversion remains challenging due to the high cost and complexity of elastic modeling, imaging, and inversion, as well as the limited availability of multicomponent data. Therefore, anisotropic wavefield tomography is typically implemented under the pseudoacoustic assumption originally proposed by Alkhalifah (1998, 2000).

P-wave kinematics in VTI media is controlled by the vertical velocity V_{P0} and Thomsen parameters ϵ and δ (Tsvankin and Thomsen, 1994; Tsvankin, 2012). Alternative parameter combinations for acoustic VTI media also involve the P-wave horizontal velocity $\left[V_{\rm hor} = V_{P0}\sqrt{1+2\epsilon}\right]$, the anellipticity parameter $\eta = (\epsilon-\delta)/(1+2\delta)$, and the normal-moveout (NMO) velocity for a horizontal interface $\left[V_{\rm nmo} = V_{P0}\sqrt{1+2\delta}\right]$. The main challenge in anisotropic wavefield-based inversion is the trade-off between model parameters, which strongly depends on the chosen parameterization.

Acoustic modeling in TI media is based either on differential or intergral wave-equation solutions. The first group of methods operates with coupled second-order partial differential equations (Duveneck et al., 2008; Fletcher et al., 2009;

*This paper has been accepted for publication in Geophysics. It represents a significantly revised and expanded version of Research Report CWP-858.

Fowler et al., 2010; Zhang et al., 2011). Because of the coupling of P- and SV- modes, the differential methods propagate shear-wave "artifacts" caused by setting the shear-wave symmetry-direction velocity V_{S0} to zero (Alkhalifah, 1998, 2000; Grechka et al., 2004). These artifacts can contaminate migrated images and hamper the acoustic inversion. The simplest way to suppress the artifact is to place sources and receivers in an elliptic ($\epsilon = \delta$, $\eta = 0$) or purely isotropic medium (Alkhalifah, 2000; Duveneck et al., 2008). However, this strategy can be legitimately applied only in the case of the data-domain waveform inversion of surface data when the physical sources and receivers, as well as the adjoint sources, reside in the near-surface layer, which can be made elliptic. More elaborate methods for suppressing the artifact involve using a finite V_{S0}, wave-mode separation, or introducing a damping term into the wave equation (Fletcher et al., 2009; Le and Levin, 2014; Suh, 2014; Fowler and King, 2011). Another issue with the differential solutions is their numerical instability for models with $\eta < 0$.

Here, we focus on integral-solution methods, which are designed to propagate only P-waves by producing decoupled modes in the wavenumber domain (Etgen and Brandsberg-Dahl, 2009; Crawley et al., 2010; Pestana and Stoffa, 2010; Song and Alkhalifah, 2013; Fomel et al., 2013b; Sun et al., 2016). A comprehensive review and classification of these methods can be found in Du et al. (2014). Separable P-mode dispersion-relation approximations for TI and orthorhombic media are described in Pestana et al. (2011), Zhan et al. (2012), Du et al. (2014), and Schleicher and Costa (2015).

Anisotropic waveform inversion has drawn considerable attention in the literature, but it is usually implemented in the data domain (Warner et al., 2013; Gholami et al., 2013; Plessix et al., 2014; Wang and Sava, 2015; Kamath and Tsvankin, 2016). Compared to the data-domain inversion, image-domain methods are less sensitive to the amplitude and shape of reflected arrivals. Whereas data-domain FWI is based on the direct trace-by-trace comparison of the observed and simulated data, image-domain inversion involves such smoothing operations as wavefield correlations and summation over the experiments, as well as the summation over image extensions for the adjoint-source computation. This property of image-domain methods is highly beneficial for acoustic inversion that cannot produce accurate reflection amplitudes.

The most common approach to image-domain tomography involves evaluating the energy focusing in the extended images (Rickett and Sava, 2002; Sava and Fomel, 2006; Sava and Vasconcelos, 2011), which can be done with differential semplance optimization (DSO) (Symes and Carazzone, 1991; Shen and Symes, 2008) or image-power estimates (Chavent and Jacewitz, 1995; Soubaras and Gratacos, 2007). The DSO and image-power objective functions can be combined to use both zero-lag and residual energy, which poses the challenge of optimal balancing of the corresponding terms. Determination of optimal weights using such inversion-theory methods as the L-curve (Nocedal and Wright, 2006) is not computationally affordable, so the balancing is commonly done empirically. Zhang and Shan (2013) propose a "partial" image-power objective function that combines the DSO and image-power criteria without a need to determine the optimal weights. Still, robust parameter estimation for complicated anisotropic velocity models may require using both the partial image-power and DSO operators.

In general, P-wave reflection moveout must be supplemented with borehole (Wang and Tsvankin, 2013a,b) or other information to resolve the VTI parameters. Y. Li et al. (2016a) build an algorithm for image-domain tomography in acoustic VTI media that operates with angle-domain common-image gathers (Sava and Fomel, 2003; Biondi, 2007; Sava and Alkhalifah, 2013). They use prior rock-physics information and structure-guided steering filters to precondition the gradient of the objective function in order to mitigate the dominant contribution of the NMO velocity to the gradient. This technique, however, requires an accurate estimate of the covariance between model parameters at each subsfurface location. A realistic error in the covariance matrix may result in the suppresion of the updates in the anisotropy coefficients. Y. Li et al. (2016b) test the algorithm on field data using the image-power objective function, but their approach does not produce sufficient updates in ϵ and δ. Weibull and Arntsen (2014) use elastic P-wave extended images to estimate V_{P0}, ϵ, and δ. However, their imaging condition is based on a purely isotropic wave-mode separation technique.

V. Li et al. (2016) analyze the defocusing in the extended domain caused by errors in the VTI parameters and show that the coefficient δ could be constrained only if it strongly varies laterally. As is the case for conventional moveout analysis, the sensitivity to the anellipticity parameter η in the image domain is higher for dipping interfaces than for horizontal reflectors.

In this paper, we derive the gradients of the data- and image-domain objective functions for acoustic VTI media using a wave-equation operator based on the separable P-mode approximation. After reviewing parameterization and wavefield extrapolation for acoustic VTI models, we discuss the objective functions for wavefield tomography, with the main focus on the image-domain approach. For data-domain tomography, the analysis is restricted to the conventional objective function that represents the ℓ_2-norm of the data-difference. Then we obtain the corresponding gradients of the objective function in both domains using the adjoint-state method. Finally, the gradients are computed and analyzed for typical VTI models.

2 PARAMETERIZATION FOR ACOUSTIC VTI MEDIA

In general, VTI acoustic wavefield tomography in either domain cannot simultaneously constrain all three relevant model parameters due to the parameter trade-offs in surface P-wave data. For data-domain inversion, an optimal parameter choice depends on the directions in which the source and receiver wavefields interact to produce a model update. Alkhalifah and Plessix (2014) analyze the radiation (sensitivity) patterns for horizontal reflectors in acoustic VTI media. They conclude that if the inversion is driven primarily by waves traveling in

near-horizontal directions (e.g., diving waves recorded at long offsets), then the optimal parameter set includes $V_{\rm hor}$, η, and ϵ. For near-vertical propagation, better results can be obtained with $V_{\rm nmo}$, η, and δ.

For image-domain inversion, parameter trade-offs stem from the properties of P-wave reflection moveout. Alkhalifah and Tsvankin (1995) demonstrate that P-wave reflection moveout for a laterally homogeneous VTI medium above the target horizon (which could be dipping or curved) is controlled by the velocity $V_{\rm nmo}$ and parameter η. For layer-cake VTI media, η contributes only to the nonhyperbolic (long-offset) portion of the P-wave moveout. If the reflector is dipping, however, η influences the NMO velocity and, therefore, conventional-spread moveout. P-wave reflection traveltimes are not sensitive to the coefficient δ, unless it varies laterally above the target reflector (Alkhalifah et al., 2001; Tsvankin and Grechka, 2011).

3 WAVEFIELD EXTRAPOLATION METHODS

Pseudoacoustic modeling operators are widely used in imaging and tomography because of their simplicity and computational efficiency. Acoustic algorithms, however, cannot accurately predict P-wave amplitudes and often have to rely on the phase of recorded arrivals or use a "dummy" model parameter that absorbs unphysical model updates (e.g., Alkhalifah and Plessix, 2014). As mentioned above, image-domain algorithms are less sensitive to the amplitude and shape of the reflection arrivals and may not require "dummy" variables.

3.1 Differential solution of the pseudoacoustic wave equation

Here, we use the formulation proposed by Fletcher et al. (2009) and Fowler et al. (2010). The 2D version of their equations for VTI media can be written as:

$$\begin{aligned}\frac{\partial^2 u^p}{\partial t^2} &= V_{\rm hor}^2(x,z)\ \frac{\partial^2 u^p}{\partial x^2} + V_{P0}^2(x,z)\ \frac{\partial^2 u^q}{\partial z^2} + f^p\ ,\\ \frac{\partial^2 u^q}{\partial t^2} &= V_{\rm nmo}^2(x,z)\ \frac{\partial^2 u^p}{\partial x^2} + V_{P0}^2(x,z)\ \frac{\partial^2 u^q}{\partial z^2} + f^q\ ,\end{aligned} \tag{1}$$

where $u^p(\mathbf{x},t)$ and $u^q(\mathbf{x},t)$ are the solutions of the fourth-order acoustic VTI equation (Alkhalifah, 2000), and f^p and f^q are the source functions. Thus, this wave-equation operator propagates the two-component vector wavefield $\mathbf{u}$. In the matrix-vector notation equation 1 can be expressed as:

$$\mathbf{L}_{\rm FD}\begin{bmatrix}u^p\\ u^q\end{bmatrix} + \begin{bmatrix}f^p\\ f^q\end{bmatrix} = 0, \tag{2}$$

where $\mathbf{L}_{\rm FD}$ is the following operator:

$$\mathbf{L}_{\rm FD} = \begin{bmatrix}V_{\rm hor}^2\,\partial_{xx} - \partial_{tt} & V_{P0}^2\,\partial_{zz}\\ V_{\rm nmo}^2\,\partial_{xx} & V_{P0}^2\,\partial_{zz} - \partial_{tt}\end{bmatrix}. \tag{3}$$

For gradient computation, we use the system of equations adjoint to equation 1 (Wang and Sava, 2015).

3.2 Integral solution of the pseudoacoustic wave equation

The integral solutions use the P-wave dispersion relation to obtain the phase shift for extrapolating (time-stepping) the wavefield (Du et al., 2014). The general integral wave-equation solutions can be written as follows:

$$\begin{aligned}U(\mathbf{x},t\pm\Delta t) &= \int \widehat{U}(\mathbf{k},t)e^{\pm i\phi(\mathbf{x},\mathbf{k},\Delta t)}d\mathbf{k},\\ \widehat{U}(\mathbf{k},t) &= \left(\frac{1}{2\pi}\right)^n \int U(\mathbf{x},t)e^{-i\mathbf{kx}}d\mathbf{x},\end{aligned} \tag{4}$$

where Δt is the time step, $\widehat{U}(\mathbf{k},t)$ is the spatial Fourier transform of the wavefield $U(\mathbf{x},t)$, $\mathbf{k}$ is the wave vector, n is the dimension of the Fourier transform, and the phase function $\phi = \Delta t\sqrt{\mathbf{A}}$, where $\mathbf{A}$ is the right-hand side of a dispersion relation (e.g., see equation 5 below).

Application of this approach to anisotropic wave equations may involve the generalized pseudospectral methods (Du et al., 2014), which require approximate dispersion relations with separable wavenumber and model-parameter terms. In other words, the contribution of the spatial wavefield variation should be decoupled from the spatial variation of medium parameters (Du et al., 2014). In the pseudoacoustic approximation, the 2D P-wave dispersion relation for VTI media can be written as (Alkhalifah, 1998):

$$\begin{aligned}\omega^2 = &\frac{1}{2}\Big[(1+2\epsilon)V_{P0}^2\,k_x^2 + V_{P0}^2\,k_z^2\Big]\\ &\left\{1+\sqrt{1-\frac{8\,(\epsilon-\delta)\,k_x^2\,k_z^2}{\Big[(1+2\epsilon)\,k_x^2+k_z^2\Big]^2}}\right\},\end{aligned} \tag{5}$$

where k_x and k_z are the horizontal and vertical wavenumbers. However, equation 5 is not suitable for pseudospectral methods because it contains the radical term. Assuming that the term containing $\epsilon-\delta$ under the radical is small, a Taylor series expansion in that term yields:

$$\omega^2 = (1+2\epsilon)V_{P0}^2k_x^2 + V_{P0}^2k_z^2 - 2(\epsilon-\delta)V_{P0}\,\frac{k_x^2\,k_z^2}{k_z^2+Fk_x^2}, \tag{6}$$

where $F = 1+2\epsilon$. Pestana et al. (2011) set F to a constant to achieve separable formulas suitable for pseudospectral methods. Physically, the Taylor series expansion produces a weak-anellipticity approximation for the dispersion relation (the medium is elliptic if $\epsilon = \delta$).

A more accurate dispersion relation can be obtained from Padé's expansion in the same term that contains $\epsilon-\delta$ in equation 5. With the first-order Padé expansion, the separable dispersion relation takes the form (Schleicher and Costa, 2015):

$$\begin{aligned}\omega^2 = &(1+2\epsilon)V_{P0}^2\,k_x^2 + V_{P0}^2\,k_z^2 - 2(\epsilon-\delta)V_{P0}^2\,\frac{k_x^2k_z^2}{k_x^2+k_z^2}\\ &\times\left[1-2\epsilon\,\frac{k_x^2}{k_x^2+k_z^2}+2(\epsilon-\delta)\,\frac{k_x^2k_z^2}{(k_x^2+k_z^2)^2}\right].\end{aligned} \tag{7}$$

Here, the Padé coefficients α and β in equation 17 of Schleicher and Costa (2015) are set to $1/2$ and $1/4$ respectively,

and their coefficient f is set to unity according to the acoustic assumption. Equation 7 can be referred to as the "separable strong-anellipticity approximation," which is suitable for implementation with pseudospectral methods. Therefore, the integral wave-equation operator can be written as:

$$\begin{aligned}\mathbf{L}_{\rm INT} = &- \frac{\partial^2}{\partial t^2} - (1+2\epsilon)V_{P0}^2\, k_x^2 - V_{P0}^2\, k_z^2 \\ &+ 2(\epsilon-\delta)V_{P0}^2\, \frac{k_x^2 k_z^2}{k_x^2+k_z^2} \\ &\times\left[1 - 2\epsilon\, \frac{k_x^2}{k_x^2+k_z^2} + 2(\epsilon-\delta)\, \frac{k_x^2 k_z^2}{(k_x^2+k_z^2)^2}\right].\end{aligned} \tag{8}$$

At each time step all terms containing the wavenumbers are computed separately as follows:

(i) Compute the spatial Fourier transform $\widehat{U}(\mathbf{k},t)$.

(ii) Multiply $\widehat{U}(\mathbf{k},t)$ by the corresponding wavenumbers (e.g., k_x^2).

(iii) Compute the inverse Fourier transform of that product [e.g., $k_x^2\, \widehat{U}(\mathbf{k},t)$].

(iv) Multiply the result by the corresponding medium parameters [e.g., $(1+2\epsilon)V_{P0}^2$].

An extension to TTI media can be obtained by locally applying the appropriate rotation matrix to the wavenumbers because equation 7 remains valid for $\widehat{k}_x$ and $\widehat{k}_z$ in the rotated coordinates. However, the rotation matrix makes the equations more complex, and the resulting wavefield simulation involves additional Fourier transforms (Zhan et al., 2012).

4 OBJECTIVE FUNCTIONS FOR WAVEFIELD TOMOGRAPHY

4.1 Data domain

Data-domain methods enforce the similarity between the observed and modeled data. The objective function is typically defined as the ℓ_2-norm data difference:

$$\mathcal{J} = \frac{1}{2}\left\|\mathbf{K}(r)\, u - d^{\,\rm obs}\right\|_2^2, \tag{9}$$

where the action of the operator $\mathbf{K}(r)$ on the modeled wavefield u produces the predicted data, and $d^{\,\rm obs}$ is the observed data for fixed receiver coordinates and time. However, because acoustic wavefield extrapolation cannot adequately predict P-wave amplitudes, application of equation 7 to field data might be problematic. Acoustic data-domain tomography is often implemented with the objective functions that rely mostly on phase information and, therefore, are less prone to get trapped in local minima (Luo and Schuster, 1991; Alkhalifah, 2015; Choi and Alkhalifah, 2015; Díaz and Sava, 2015). Alternatively, one could use a "dummy" model parameter to absorb the model updates caused by unphysical amplitudes produced by acoustic equations.

4.2 Image domain

Image-domain tomography uses migrated reflection data as the input for the inversion with the goal of updating the background velocty model (note that parameter updates are smeared along the reflection wavepaths). Our treatment is restricted to the residual energy minimization in the so-called extended domain. Extended images are produced by retaining correlation lags between the source and receiver wavefields in the output of wave-equation migration. The general imaging condition can be formulated as follows (Sava and Vasconcelos, 2011):

$$I\left(\mathbf{x},\boldsymbol{\lambda},\tau\right) = \sum_{e,t} W_s\left(\mathbf{x}-\boldsymbol{\lambda},t-\tau\right) W_r\left(\mathbf{x}+\boldsymbol{\lambda},t+\tau\right), \tag{10}$$

where $I(\mathbf{x},\boldsymbol{\lambda},\tau)$ is the extended image, W_s and W_r denote the source and receiver wavefields, respectively, $\boldsymbol{\lambda}$ is the space lag, τ is the time lag, and e indicates summation over experiments. To reduce computational cost, one can compute only extended common-image-gathers (CIG), which are space-lag or time-lag extensions at fixed horizontal coordinates (Rickett and Sava, 2002; Sava and Fomel, 2006), or common-image-point (CIP) gathers, which represent multiple extensions computed at sparse points in the image space (Sava and Vasconcelos, 2009). Residual energy at nonzero lags can be used to update the migration velocity model and is most commonly measured with differential semblance optimization (DSO) (Symes and Carazzone, 1991; Shen and Symes, 2008). The DSO objective function for a horizontal space-lag extended image I has the form:

$$J_{DSO} = \frac{1}{2}\|\lambda_x\, I(x,z,\lambda_x)\|_2^2, \tag{11}$$

where the horizontal lag λ_x plays the role of the penalty operator. Another commonly used (image-power) objective function measures zero-lag energy:

$$J_{ST} = -\frac{1}{2}\|I(x,z,\lambda_x=0)\|_{\ell_2}^2\,. \tag{12}$$

Zhang and Shan (2013) propose a "partial" image-power objective function that combines the criteria in equations 11 and 12:

$$J_{PST} = -\frac{1}{2}\|H(\lambda_x)\, I(x,z,\lambda_x)\|_{\ell_2}^2, \tag{13}$$

where H is a Gaussian operator centered at zero lag.

5 GRADIENT COMPUTATION USING THE ADJOINT-STATE METHOD

The adjoint-state method (Tarantola, 1984; Tromp et al., 2005; Plessix, 2006) is designed to efficiently evaluate the gradient of the objective function with respect to the model parameters. For seismic wavefield tomography, general gradient expressions for acoustic wavefields written in matrix-vector notation can be found in Sava (2014). In addition to the objective function, application of the adjoint-state method involves state and adjoint equations. Minimization of the objective function $\mathcal{J}$ is

subject to the constraints $\mathcal{F}_s$ and $\mathcal{F}_r$:

$$\begin{bmatrix}\mathcal{F}_s\\ \mathcal{F}_r\end{bmatrix} = \begin{bmatrix}\mathbf{L} & \mathbf{0}\\ \mathbf{0} & \mathbf{L}^\dagger\end{bmatrix}\begin{bmatrix}\mathbf{u}_s\\ \mathbf{u}_r\end{bmatrix} - \begin{bmatrix}\mathbf{d}_s\\ \mathbf{d}_r\end{bmatrix} = \mathbf{0}, \tag{14}$$

where $\mathbf{L}$ and $\mathbf{L}^\dagger$ are the forward and adjoint wave-equation operators, respectively, $\mathbf{d}_s$ is the source function, $\mathbf{d}_r$ is the observed data, and $\mathbf{u}_s$ and $\mathbf{u}_r$ are the source and receiver wavefields, respectively. The zero matrix $\mathbf{0}$ has the same dimensions as the wave-equation matrix (operator) $\mathbf{L}$. These constraints indicate that the wavefields $\mathbf{u}_s$ and $\mathbf{u}_r$ used in the minimization problem should be solutions of the wave equation:

$$\begin{bmatrix}\mathbf{L} & \mathbf{0}\\ \mathbf{0} & \mathbf{L}^\dagger\end{bmatrix}\begin{bmatrix}\mathbf{u}_s\\ \mathbf{u}_r\end{bmatrix} = \begin{bmatrix}\mathbf{d}_s\\ \mathbf{d}_r\end{bmatrix}. \tag{15}$$

The method of Lagrange multipliers can be used to formulate the minimization as an unconstrained problem:

$$\mathcal{H} = \mathcal{J} - \begin{bmatrix}\mathcal{F}_s^{\mathrm{T}} & \mathcal{F}_r^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\mathbf{a}_s\\ \mathbf{a}_r\end{bmatrix}; \tag{16}$$

where $\mathcal{H}$ is Lagrangian and T denotes a transpose. The Lagrange multipliers $\mathbf{a}_s$ and $\mathbf{a}_r$ are referred to as the "adjoint-state variables," which are found from the following adjoint equations that involve the source terms $\mathbf{g}_s$ and $\mathbf{g}_r$:

$$\begin{bmatrix}\mathbf{L}^\dagger & \mathbf{0}\\ \mathbf{0} & \mathbf{L}\end{bmatrix}\begin{bmatrix}\mathbf{a}_s\\ \mathbf{a}_r\end{bmatrix} = \begin{bmatrix}\mathbf{g}_s\\ \mathbf{g}_r\end{bmatrix}, \tag{17}$$

The magnitude and spatial distribution of $\mathbf{g}_s$ and $\mathbf{g}_r$ are obtained from the derivatives of the objective function:

$$\begin{bmatrix}\mathbf{g}_s\\ \mathbf{g}_r\end{bmatrix} = \begin{bmatrix}\dfrac{\partial\mathcal{J}}{\partial\mathbf{u}_s}\\ \dfrac{\partial\mathcal{J}}{\partial\mathbf{u}_r}\end{bmatrix}. \tag{18}$$

Finally, the gradient of the augmented function $\mathcal{H}$ with respect to the vector $\mathbf{m}$ of the model parameters is found as

$$\frac{\partial\mathcal{H}}{\partial\mathbf{m}} = \frac{\partial\mathcal{J}}{\partial\mathbf{m}} + \sum_e\left[\left(\frac{\partial\mathcal{F}_s}{\partial\mathbf{m}}\right)^T\left(\frac{\partial\mathcal{F}_r}{\partial\mathbf{m}}\right)^T\right]\begin{bmatrix}\mathbf{a}_s\\ \mathbf{a}_r\end{bmatrix} \tag{19}$$

Additionally, the summation on the right-hand side is performed not just over experiments, but also over time, which is equivalent to the zero time-lag correlation (Sava, 2014):

$$\frac{\partial\mathcal{H}}{\partial\mathbf{m}} = \frac{\partial\mathcal{J}}{\partial\mathbf{m}} + \sum_{e,\tau}\delta(\tau)\left(\frac{\partial\mathcal{F}_s}{\partial\mathbf{m}}\star\mathbf{a}_s + \frac{\partial\mathcal{F}_r}{\partial\mathbf{m}}\star\mathbf{a}_r\right), \tag{20}$$

where τ is the correlation lag, $\delta(\tau)$ is the Dirac delta function, and '$\star$' denotes cross-correlation. Overall, application of the adjoint-state method involves computing the following quantities:

(i) The state variables $\mathbf{u}_s$ and $\mathbf{u}_r$ by solving the state equations 15.

(ii) The adjoint sources $\mathbf{g}_s$ and $\mathbf{g}_r$ that depend on the chosen objective function in equation 18.

(iii) The adjoint-state variables $\mathbf{a}_s$ and $\mathbf{a}_r$ by solving the adjoint equations 17.

(iv) The gradient of the objective function, which depends on the wave-equation operator and chosen parameterization.

Here, we apply the adjoint-state method to the pseudoacoustic operators $\mathbf{L}_{\mathrm{FD}}$ and $\mathbf{L}_{\mathrm{INT}}$ discussed above and obtain gradient expressions for the objective functions in equations 9, 11, and 13.

5.1 Differential-solution operator

For VTI media, the forward (state) wave-equation operator $\mathbf{L}$ is defined as (equation 3):

$$\mathbf{L} = \begin{bmatrix}\mathbf{L}_{11} & \mathbf{L}_{12}\\ \mathbf{L}_{21} & \mathbf{L}_{22}\end{bmatrix} = \begin{bmatrix}V_{\mathrm{hor}}^2\partial_{xx} - \partial_{tt} & V_{P0}^2\partial_{zz}\\ V_{\mathrm{nmo}}^2\partial_{xx} & V_{P0}^2\partial_{zz} - \partial_{tt}\end{bmatrix}. \tag{21}$$

As shown by Wang and Sava (2015), the corresponding adjoint operator $\mathbf{L}^T$ is:

$$\mathbf{L}^T = \begin{bmatrix}\mathbf{L}_{11}^T & \mathbf{L}_{21}^T\\ \mathbf{L}_{12}^T & \mathbf{L}_{22}^T\end{bmatrix} = \begin{bmatrix}\partial_{xx}V_{\mathrm{hor}}^2 - \partial_{tt} & \partial_{xx}V_{\mathrm{nmo}}^2\\ \partial_{zz}V_{P0}^2 & \partial_{zz}V_{P0}^2 - \partial_{tt}\end{bmatrix}. \tag{22}$$

5.1.1 Data domain

For the data-domain objective function (equation 9), the gradients can be found in Wang and Sava (2015). For 2D models, they define the data residual as $\mathbf{K}_r(\mathbf{u}^p + \mathbf{u}^q) - d^{\mathrm{obs}}$, the model parameters as $\mathbf{m} = \{V_{P0}^2, V_{\mathrm{nmo}}^2, V_{\mathrm{hor}}^2\}$, and obtain the following expressions:

$$\begin{aligned}\frac{\partial\mathcal{J}}{\partial\mathbf{m}} &= \begin{bmatrix}\dfrac{\partial\mathcal{J}}{\partial V_{P0}^2}\\ \dfrac{\partial\mathcal{J}}{\partial V_{\mathrm{nmo}}^2}\\ \dfrac{\partial\mathcal{J}}{\partial V_{\mathrm{hor}}^2}\end{bmatrix} = \sum_{e,\tau}\delta(\tau)\begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix},\\ b_1 &= \partial_{zz}u^q\star(a^p + a^q),\\ b_2 &= \partial_{xx}u^p\star a^q,\\ b_3 &= \partial_{xx}u^p\star a^p.\end{aligned} \tag{23}$$

where a^p and a^q are the components of the adjoint wavefield. Application of the chain rule yields the gradient expressions for the vector $\widehat{\mathbf{m}} = \{V_{\mathrm{hor}}, \eta, \epsilon\}$:

$$\begin{aligned}\frac{\partial\mathcal{J}}{\partial\widehat{\mathbf{m}}} &= \begin{bmatrix}\dfrac{\partial\mathcal{J}}{\partial\epsilon}\\ \dfrac{\partial\mathcal{J}}{\partial\eta}\\ \dfrac{\partial\mathcal{J}}{\partial V_{\mathrm{hor}}}\end{bmatrix} = \sum_{e,\tau}\delta(\tau)\\ &\times\begin{bmatrix}\dfrac{-2V_{\mathrm{hor}}^2}{(1+2\epsilon)^2} & 0 & 0\\ 0 & \dfrac{-2V_{\mathrm{hor}}^2}{(1+2\eta)^2} & 0\\ 0 & 0 & 2V_{\mathrm{hor}}\end{bmatrix}\begin{bmatrix}f_1\\ f_2\\ f_3\end{bmatrix},\\ f_1 &= \partial_{zz}u^q\star(a^p + a^q),\\ f_2 &= \partial_{xx}u^p\star a^q,\\ f_3 &= \partial_{xx}u^p\star a^p + \frac{\partial_{xx}u^p}{1+2\eta}\star a^q + \frac{\partial_{zz}u^q}{1+2\epsilon}\star(a^p + a^q).\end{aligned} \tag{24}$$

5.1.2 *Image domain*

We define the space-lag common-image gather through the sum of the p and q components of the source and receiver wavefields:

$$I(\mathbf{x},\boldsymbol{\lambda}) = \sum_{e,t} W_s(e,\mathbf{x}-\boldsymbol{\lambda},t)\, W_r(e,\mathbf{x}+\boldsymbol{\lambda},t), \tag{25}$$

where

$$W_i(e,\mathbf{x},t) = u_i^p(e,\mathbf{x},t) + u_i^q(e,\mathbf{x},t),\ i=s,r\,. \tag{26}$$

As a result, for the objective function in equation 11, equations 18 for the adjoint sources take the following form:

$$\begin{aligned}
\begin{bmatrix} g_s^p \\ g_s^q \end{bmatrix} &= \sum_{\lambda_x} \lambda_x^2 \begin{bmatrix} I(x+\lambda_x,\lambda_x)\, W_r(x+2\lambda_x,t) \\ I(x+\lambda_x,\lambda_x)\, W_r(x+2\lambda_x,t) \end{bmatrix}, \\
\begin{bmatrix} g_r^p \\ g_r^q \end{bmatrix} &= \sum_{\lambda_x} \lambda_x^2 \begin{bmatrix} I(x-\lambda_x,\lambda_x)\, W_s(x-2\lambda_x,t) \\ I(x-\lambda_x,\lambda_x)\, W_s(x-2\lambda_x,t) \end{bmatrix}.
\end{aligned} \tag{27}$$

After the adjoint wavefields are computed, the source- and receiver-side gradients with respect to the vector $\mathbf{m} = \{V_{\rm nmo}, \eta, \delta\}$ are found as:

$$\begin{aligned}
\left[\frac{\partial \mathcal{J}}{\partial \mathbf{m}}\right]_i &= \begin{bmatrix} \dfrac{\partial \mathcal{J}}{\partial \delta} \\ \dfrac{\partial \mathcal{J}}{\partial V_{\rm nmo}} \\ \dfrac{\partial \mathcal{J}}{\partial \eta} \end{bmatrix}_i = \sum_{e,\tau} \delta(\tau) \\
&\times \begin{bmatrix} \dfrac{-2V_{\rm nmo}^2}{(1+2\delta)^2} & 0 & 0 \\ 0 & 2V_{\rm nmo} & 0 \\ 0 & 0 & 2V_{\rm nmo}^2 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix},
\end{aligned} \tag{28}$$

$$\begin{aligned}
f_1 &= \partial_{zz} u_i^q \star (a_i^p + a_i^q)\,, \\
f_2 &= (1+2\eta)\partial_{xx} u_i^p \star a_i^p + \partial_{xx} u_i^p \star a_i^q \\
&\quad + \frac{\partial_{zz} u_i^q}{1+2\delta} \star (a_i^p + a_i^q)\,, \\
f_3 &= \partial_{xx} u_i^p \star a_i^p\,,\ i=s,r\,,
\end{aligned}$$

where i denotes either the source or receiver side.

5.2 Integral-solution operator

For most TI models (with the exception of uncommonly strong anisotropy), sufficient accuracy can be provided by the three leading terms of the separable dispersion relation in equation 7, which simplifies the gradient expressions. However, we truncate equation 7 only for deriving the gradient expressions but not for wavefield extrapolation. For VTI media, the forward (state) wave-equation operator $\mathbf{L}$ can be defined as

$$\begin{aligned}
\mathbf{L}_{\rm INT} = &-\frac{\partial^2}{\partial t^2} - V_{\rm hor}^2\, k_x^2 - \frac{V_{\rm hor}^2}{1+2\epsilon}\, k_z^2 \\
&+ 2\eta\, \frac{V_{\rm hor}^2}{1+2\eta} \frac{k_x^2 k_z^2}{k_x^2+k_z^2},
\end{aligned} \tag{29}$$

or, equivalently,

$$\begin{aligned}
\mathbf{L}_{\rm INT} = &-\frac{\partial^2}{\partial t^2} - V_{\rm nmo}^2\, k_x^2 - \frac{V_{\rm nmo}^2}{1+2\delta}\, k_z^2 \\
&- 2\eta\, V_{\rm nmo}^2 \frac{k_x^4}{k_x^2+k_z^2}.
\end{aligned} \tag{30}$$

The corresponding adjoint operator $\mathbf{L}^\dagger$ is:

$$\begin{aligned}
\mathbf{L}_{\rm INT}^\dagger = &-\frac{\partial^2}{\partial t^2} - k_x^2\, V_{\rm hor}^2 - k_z^2 \frac{V_{\rm hor}^2}{1+2\epsilon} \\
&+ \frac{2\, k_x^2\, k_z^2}{k_x^2+k_z^2}\, \eta\, \frac{V_{\rm hor}^2}{1+2\eta},
\end{aligned} \tag{31}$$

or

$$\begin{aligned}
\mathbf{L}_{\rm INT}^\dagger = &-\frac{\partial^2}{\partial t^2} - k_x^2\, V_{\rm nmo}^2 - k_z^2\, \frac{V_{\rm nmo}^2}{1+2\delta} \\
&- \frac{2\, k_x^4}{k_x^2+k_z^2}\, \eta\, V_{\rm nmo}^2.
\end{aligned} \tag{32}$$

5.2.1 *Data domain*

Below, we obtain the gradient expressions for the data-domain objective function in equation 9. The data residual is defined as $\mathbf{K}_r u - d^{\,\rm obs}$. Therefore, equation 18 for the adjoint sources g_s and g_r becomes:

$$\begin{bmatrix} g_s \\ g_r \end{bmatrix} = \begin{bmatrix} \mathbf{K}_r^{\rm T}\, (\mathbf{K}_r u - d^{\rm obs}) \\ -\mathbf{K}_r^{\rm T}\, (\mathbf{K}_r u - d^{\rm obs}) \end{bmatrix}. \tag{33}$$

For data-domain methods, only the adjoint source wavefield a is relevant (Sava, 2014), and the gradient with respect to the model parameter $\widehat{\mathbf{m}} = \{V_{\rm hor}, \eta, \epsilon\}$ is given by the following expression:

$$\begin{aligned}
\frac{\partial \mathcal{J}}{\partial \widehat{\mathbf{m}}} &= \begin{bmatrix} \dfrac{\partial \mathcal{J}}{\partial \epsilon} \\ \dfrac{\partial \mathcal{J}}{\partial \eta} \\ \dfrac{\partial \mathcal{J}}{\partial V_{\rm hor}} \end{bmatrix} = -\sum_{e,\tau} \delta(\tau) \\
&\times \begin{bmatrix} \dfrac{-2V_{\rm hor}^2}{(1+2\epsilon)^2} & 0 & 0 \\ 0 & \dfrac{-2V_{\rm hor}^2}{(1+2\eta)^2} & 0 \\ 0 & 0 & 2V_{\rm hor} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
f_1 &= k_z^2 u \star a\,, \\
f_2 &= \frac{k_x^2 k_z^2}{k_x^2+k_z^2} u \star a\,, \\
f_3 &= k_x^2 u \star a + \frac{k_z^2 u}{1+2\epsilon} \star a\ - \frac{2\eta}{1+2\eta}\, \frac{k_x^2 k_z^2}{k_x^2+k_z^2} u \star a\,.
\end{aligned} \tag{34}$$

5.2.2 Image domain

The image residual can be defined as:

$$\boldsymbol{\lambda}\, I(\mathbf{x},\boldsymbol{\lambda}) = \boldsymbol{\lambda}\Big(\sum_{e,t} u_s(e,\mathbf{x}-\boldsymbol{\lambda},t)\, u_r(e,\mathbf{x}+\boldsymbol{\lambda},t)\Big). \tag{35}$$

Hence, for the objective function in equation 11, equation 18 for the adjoint sources becomes:

$$\begin{bmatrix} g_s \\ g_r \end{bmatrix} = \sum_{\lambda_x} \lambda_x^2 \begin{bmatrix} I(x+\lambda_x,\lambda_x)\, u_r(x+2\lambda_x,t) \\ I(x-\lambda_x,\lambda_x)\, u_s(x-2\lambda_x,t) \end{bmatrix}. \tag{36}$$

Similarly, for the partial image-power objective function in equation 13, equation 18 for the adjoint sources becomes:

$$\begin{bmatrix} g_s \\ g_r \end{bmatrix} = -\sum_{\lambda_x} H(\lambda_x)^2 \begin{bmatrix} I(x+\lambda_x,\lambda_x)\, u_r(x+2\lambda_x,t) \\ I(x-\lambda_x,\lambda_x)\, u_s(x-2\lambda_x,t) \end{bmatrix}. \tag{37}$$

Then the source- and receicer-side gradients with respect to the model vector $\mathbf{m} = \{V_{\rm nmo}, \eta, \delta\}$ are given by:

$$\begin{aligned}
\left[\frac{\partial \mathcal{J}}{\partial \mathbf{m}}\right]_i &= \begin{bmatrix} \dfrac{\partial \mathcal{J}}{\partial \delta} \\ \dfrac{\partial \mathcal{J}}{\partial V_{\rm nmo}} \\ \dfrac{\partial \mathcal{J}}{\partial \eta} \end{bmatrix}_i = -\sum_{e,\tau} \delta(\tau) \\
&\times \begin{bmatrix} \dfrac{-2V_{\rm nmo}^2}{(1+2\delta)^2} & 0 & 0 \\ 0 & 2V_{\rm nmo} & 0 \\ 0 & 0 & 2V_{\rm nmo}^2 \end{bmatrix} \begin{bmatrix} f1 \\ f2 \\ f3 \end{bmatrix}, \\
f_1 &= k_z^2 u_i \star a_i\ , \\
f_2 &= k_x^2 u_i \star a_i + \frac{k_z^2}{1+2\delta} u_i \star a_i + 2\eta \frac{k_x^4}{k_x^2+k_z^2} u_i \star a_i\ , \\
f_3 &= \frac{k_x^4}{k_x^2+k_z^2}\ u_i \star a_i\ ,\ i = s, r\ .
\end{aligned} \tag{38}$$

6 SYNTHETIC EXAMPLES

Below, we test the gradient expressions derived above on several VTI models. The medium parameters are specified on a rectangular grid, and the density is assumed to be constant. For forward and adjoint wavefield extrapolation, we use both the differential (operators $\mathbf{L}_{\rm FD}$ and $\mathbf{L}_{\rm FD}^{\dagger}$) and integral methods (operators $\mathbf{L}_{\rm INT}$ and $\mathbf{L}_{\rm INT}^{\dagger}$) described above. The gradients obtained with the integral operator are compared with the ones for the differential operator alrorithm (equations 21 and 22).

6.1 Model 1

First, we compute the gradients in the data domain for a model that includes a constant $V_{\rm hor}$-field and Gaussian anomalies in the parameters η (reaching 0.2 at the center; the background η = 0.05) and ϵ (reaching 0.15; the background ϵ = 0) (Figure 1). Only transmitted waves are employed to generate parameter updates. The source function is a Ricker wavelet with a central frequency of 2 Hz. Using the actual η-field, we compute the gradients for understated and overstated peak values of the ϵ-anomaly (ϵ = 0 and 0.3; the background ϵ = 0 is correct). Note that for the peak frequency of the source signal (2 Hz) and the model size, the time shifts caused by errors in ϵ do not exceed half a cycle.

For the chosen parameterization ($V_{\rm hor}$, η, ϵ), the coefficient ϵ should be constrained for near-vertical propagation, if $V_{\rm hor}$ has been estimated from long-offset data (Alkhalifah and Plessix, 2014). We compute the gradients using the vertical ("borehole") receiver array shown in Figure 1d. In general, P-wave reflection moveout must be supplemented with borehole (Wang and Tsvankin, 2013a) or other information to resolve the VTI parameters. The gradients generated by both operators are similar and, as expected, change sign depending on the sign of the ϵ-error (Figure 2). Because the background η-field is positive, the differential extrapolator produces a pronounced shear-wave artifact. In the data domain, the gradient for the actual ϵ-field goes to zero. However, the data-difference estimate may be questionable for field-data applications because the acoustic approximation does not accurately model reflection amplitudes.

6.2 Model 2

Next, we compute the η-gradient in the image domain using reflection data. The model includes a horizontal interface 8 km long beneath a homogeneous VTI layer with $V_{\rm nmo}$ = 2 km/s, η = δ = 0.15, and a thickness of 2 km. The near-surface layer, which is 0.2 km thick, is assumed to be elliptic (ϵ = δ) to suppress the shear-wave artifact produced by the differential extrapolator. We generate horizontal-space-lag extended images (Figure 3) and obtain the η-gradients for understated and overstated values of η. The η-errors induce residual energy in extended images (Figure 3) that has a linear ("V"-like) shape, which is typical for near-horizontal interfaces (V. Li et al., 2016; Sava and Alkhalifah, 2012). For both extrapolators, the extended images computed with the understated and even actual η-fields also contain considerable residual energy that spreads from the image point up to the surface. These kinematic artifacts, caused by the aperture truncation, may introduce bias in the image-domain objective function and lead to false model updates.

The DSO gradients computed using surface acquisition geometry and the entire extended image are shown in Figure 4. With either extrapolation operator, the gradient of the DSO objective function (equation 11) for the understated η-field is strongly influenced by the kinematic artifacts in the extended image. The contribution of the artifact is even larger than that of the residual induced by the η-error because the artifact is located closer to the physical sources and receivers. For this model, the partial image-power objective function (equation 13) significantly reduces the artifact (Figure 5). Nevertheless, robust anisotropic inversion may require additional suppression of kinematic artifacts by proper accounting for il-

lumination in the imaging or DSO operators (Lameloise et al., 2015; Hou and Symes, 2015; Yang and Sava, 2015).

6.3 Model 3

In this test, we compare the sensitivity of the DSO and partial image-power estimates to errors in the background $V_{\rm nmo}$-field. The actual $V_{\rm nmo}$ field consists of the constant background equal to 2 km/s and perturbations located at 1, 2, and 3 km depth. The anisotropy coefficients η and δ are taken constant (equal to 0.15 and 0.1, respectively) throughout the model. We compute the DSO and partial image-power objective functions for several models with different background $V_{\rm nmo}$-values ranging from 1.8 to 2.2 km/s. Figure 7 shows the space-lag CIGs for the understated, actual, and overstated background $V_{\rm nmo}$-values. Similarly to model 2, the gathers include defocused energy due to both velocity errors and the aperture-truncation effect.

Figures 8 and 9 show the same gathers after applying the DSO and partial-power operators. The DSO operator (Figure 8) is biased towards understated background models, whereas the partial image-power focuses most energy for the actual backround model. Figure 10 shows the difference in the behavior of the DSO and partial image-power objective functions. The DSO objective function amplifies the aperture-truncation artifacts and is not sensitive to the negative velocity errors for this model. In contrast, the partial image-power objective function is symmetric with the minimum at the actual background $V_{\rm nmo}$-value. However, the difference between the DSO and partial image-power objective functions needs to be studied further for more complicated models. The relative performance of these two functions is likely to depend on such factors as the accuracy of the initial model and the type of input data.

6.4 Image-domain tomography

In the last test, we apply the partial image-power gradients discussed above to perform image-domain tomography for model 3. The data are generated by 41 evenly distributed shots with the spacing equal to 0.2 km. The initial model is elliptic ($\eta = 0$) with the NMO velocity equal to 1.8 km/s (10% lower than the actual value). We assume that the parameter δ is known because it does not vary laterally, and, therefore, cannot be constrained by P-wave reflection data (V. Li et al., 2016). The model update is computed with the following equation:

$$m_{k+1} = m_k + \alpha_k \, \nabla J_k, \tag{39}$$

where α_k is the steplength and ∇J_k is the gradient of the partial image-power objective function. Given the simplicity of the model, we use the steepest-descent method (Nocedal and Wright, 2006), which relies only on the inversion gradient at the current iteration k.

Extended images and the inversion gradients are tapered in the top part of the section to reduce the influence of the aperture-truncation artifacts. Figure 11a,b shows the gradients for the parameters $V_{\rm nmo}$ and η computed for the initial model. We also apply Gaussian smoothing to the gradients, as shown in Figure 11 (c,d). After three iterations, the updated parameters $V_{\rm nmo} \approx 2.05$ and $\eta \approx 0.17$ (Figure 12) are close to the actual values (2 km/s and 0.15, respectively).

7 CONCLUSIONS

Wavefield extrapolation and gradient computation are key steps of wave-equation-based inversion algorithms. We implemented forward and adjoint integral extrapolation operators for acoustic VTI media based on a separable dispersion-relation approximation and derived the corresponding gradient expressions. This work is mostly focused on image-domain wavefield tomography, which is less susceptible to amplitude distortions produced by acoustic algorithms. However, because estimation of all three relevant VTI parameters (e.g., V_{P0}, ϵ, and δ) is seldom feasible using only P-wave reflection moveout, we also derived data-domain gradients, which are more suitable for incorporating borehole information.

The gradients of the image- and data-domain objective functions were computed for several VTI models and different acquisition geometries. The similarity between the gradients obtained with the integral and differential operators validates our analytic results. However, the gradients computed with these two operators do not exhibit the same spatial distribution, which can be explained by the difference in amplitude variation along the simulated wavefronts. This difference becomes larger with an increase in the parameter η. For a model where the sources and receivers were placed in a layer with $\eta > 0$, the gradients computed with the pseudospectral algorithm do not contain the imprint of the shear-wave artifact that contaminates the FD results.

The space-lag common-image gathers (CIGs) reveal illumination-related issues with the DSO objective function applied to cross-correlation extended images. Kinematic artifacts caused by insufficient illumination substantially distort the gradients and should be suppressed prior to updating the model. The partial image-power objective function may help reduce the false updates caused by these artifacts. However, the DSO and partial image-power objective functions need to be compared for more realistic, structurally complex models. Ongoing work involves implementing the imaging and inversion steps of anisotropic image-domain tomography and an extension of the algorithm to tilted TI media.

8 ACKNOWLEDGMENTS

We thank Paul Fowler, Jörg Schleicher, and the A(nisotropy)- and i(maging)-Teams at CWP for fruitful discussions. We also thank associate editor Igor Ravve and reviewers Alexey Stovas, Jiubing Cheng, and Junzhe Sun for their helpful comments. This work was supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP and the competitive research funding from King Abdullah University of Science and Technology (KAUST). The reproducible

numeric examples in this paper are generated with the Madagascar open-source software package (Fomel et al., 2013a) freely available from http://www.ahay.org.

REFERENCES

Al-Yahya, K., 1989, Velocity analysis by iterative profile migration: Geophysics, **54**, 718–729.

Alkhalifah, T., 1998, Acoustic approximations for processing in transversely isotropic media: Geophysics, **63**, 623–631.

———, 2000, An acoustic wave equation for anisotropic media: Geophysics, **65**, 1239–1250.

———, 2015, Efficient scattering angle filtering for full waveform inversion: SEG Technical Program Expanded Abstracts, 1138–1142.

Alkhalifah, T., S. Fomel, and B. Biondi, 2001, The space-time domain: theory and modelling for anisotropic media: Geophysical Journal International, **144**, 105–113.

Alkhalifah, T., and R.-E. 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.

Biondi, B., 2007, Angle-domain common-image gathers from anisotropic migration: Geophysics, **72**, S81–S91.

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

Choi, Y., and T. Alkhalifah, 2015, Unwrapped phase inversion with an exponential damping: Geophysics, **80**, R251–R264.

Crawley, S., S. Brandsberg-Dahl, and J. McClean, 2010, 3D TTI RTM using the pseudoanalytic method: SEG Technical Program Expanded Abstracts, 3216–3220.

Díaz, E., and P. Sava, 2015, Data domain wavefield tomography using local correlation functions: SEG Technical Program Expanded Abstracts, 1361–1365.

Du, X., P. J. Fowler, and R. P. Fletcher, 2014, Recursive integral time-extrapolation methods for waves: A comparative review: Geophysics, **79**, T9–T26.

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

Etgen, J. T., and S. Brandsberg-Dahl, 2009, The pseudoanalytical method: Application of pseudolaplacians to acoustic and acoustic anisotropic wave propagation: SEG Technical Program Expanded Abstracts, 2552–2556.

Fletcher, R. P., X. Du, and P. J. Fowler, 2009, Reverse time migration in tilted transversely isotropic (TTI) media: Geophysics, **74**, WCA179–WCA187.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013a, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1(1)**, e8.

Fomel, S., L. Ying, and X. Song, 2013b, Seismic wave extrapolation using lowrank symbol approximation: Geophysical Prospecting, **61**, 526–536.

Fowler, P. J., X. Du, and R. P. Fletcher, 2010, Coupled equations for reverse time migration in transversely isotropic media: Geophysics, **75**, S11–S22.

Fowler, P. J., and R. King, 2011, Modeling and reverse time migration of orthorhombic pseudoacoustic P-waves: SEG Technical Program Expanded Abstracts 2011, presentation, 190–195.

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.

Grechka, V., L. Zhang, and I. James W. Rector, 2004, Shear waves in acoustic anisotropic media: Geophysics, **69**, 576–582.

Hou, J., and W. W. Symes, 2015, An approximate inverse to the extended born modeling operator: Geophysics, **80**, R331–R349.

Kamath, N., and I. Tsvankin, 2016, Elastic full-waveform inversion for VTI media: Methodology and sensitivity analysis: Geophysics, **81**, C53–C68.

Lameloise, C.-A., H. Chauris, and M. Noble, 2015, Improving the gradient of the image-domain objective function using quantitative migration for a more robust migration velocity analysis: Geophysical Prospecting, **63**, 391–404.

Le, H., and S. A. Levin, 2014, Removing shear artifacts in acoustic wave propagation in orthorhombic media: SEG Technical Program Expanded Abstracts, 486–490.

Li, V., I. Tsvankin, and T. Alkhalifah, 2016, Analysis of RTM extended images for VTI media: Geophysics, **81**, S139–S150.

Li, Y., B. Biondi, R. Clapp, and D. Nichols, 2016a, Integrated VTI model building with seismic data, geologic information, and rock-physics modeling part 1: Theory and synthetic test: Geophysics, **81**, C177–C191.

———, 2016b, Integrated VTI model building with seismic data, geologic information, and rock-physics modeling part 2: Field data test: Geophysics, **81**, C205–C218.

Luo, Y., and G. T. Schuster, 1991, Waveequation traveltime inversion: Geophysics, **56**, 645–653.

Nocedal, J., and S. J. Wright, 2006, Numerical optimization, 2 ed.: Springer-Verlag New York, Inc.

Perrone, F., and P. Sava, 2012, Wavefield tomography based on local image correlations: CWP Project Review Report, 51–76.

Pestana, R. C., and P. L. Stoffa, 2010, Time evolution of the wave equation using rapid expansion method: Geophysics, **75**, T121–T131.

Pestana, R. C., B. Ursin, and P. L. Stoffa, 2011, Separate P- and SV-wave equations for VTI media: SEG Technical Program Expanded Abstracts 2011, 163–167.

Plessix, R.-E., 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.-E., A. Stopin, P. Milcik, and K. Matson, 2014, Acoustic and anisotropic multi-parameter seismic full waveform inversion case studies: SEG Technical Program Expanded Abstracts 2014, 1056–1060.

Rickett, J., and P. Sava, 2002, Offset and angle–domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889.

Sattlegger, J. W., 1975, Migration velocity determination: Part I. Philosophy: Geophysics, **40**, 1–5.

Sava, P., 2014, A comparative review of wavefield tomography methods: CWP Project Review Report, 119–144.

Sava, P., and T. Alkhalifah, 2012, Anisotropy signature in extended images from reverse-time migration: SEG, Technical Program Expanded Abstracts, 1–6.

———, 2013, Wide-azimuth angle gathers for anisotropic wave-equation migration: Geophysical Prospecting, **61**, 75–91.

Sava, P., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, S209–S217.

Sava, P., and I. Vasconcelos, 2009, Efficient computation of extended images by wavefield-based migration: SEG, Technical Program Expanded Abstracts, 2824–2828.

———, 2011, Extended imaging conditions for wave-equation migration: Geophysical Prospecting, **59**, 35–55.

Sava, P. C., and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Schleicher, J., and J. C. Costa, 2015, A separable strong-anisotropy approximation for pure qp wave propagation in TI media: SEG Technical Program Expanded Abstracts, 3565–3570.

Shen, P., and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.

Song, X., and T. Alkhalifah, 2013, Modeling of pseudoacoustic P-waves in orthorhombic media with a low-rank approximation: Geophysics, **78**, C33–C40.

Soubaras, R., and B. Gratacos, 2007, Velocity model building by semblance maximization of modulated-shot gathers: Geophysics, **72**, U67–U73.

Suh, S. Y., 2014, A finite-difference method for orthorhombic reverse time migration: SEG Technical Program Expanded Abstracts, 3915–3919.

Sun, J., S. Fomel, and L. Ying, 2016, Low-rank one-step wave extrapolation for reverse time migration: Geophysics, **81**, S39–S54.

Symes, W. W., and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

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

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

Tsvankin, I., 2012, Seismic signatures and analysis of reflection data in anisotropic media, third edition: Society of Exploration Geophysicists.

Tsvankin, I., and V. Grechka, 2011, Seismology of azimuthally anisotropic media and seismic fracture characterization: Society of Exploration Geophysicists.

Tsvankin, I., and L. Thomsen, 1994, Nonhyperbolic reflection moveout in anisotropic media: Geophysics, **59**, 1290–1304.

Wang, H., and P. Sava, 2015, Pseudo-acoustic wavefield tomography with model constraints: CWP Project Review Report, 95–106.

Wang, X., and I. Tsvankin, 2013a, Multiparameter TTI tomography of P-wave reflection and VSP data: Geophysics, **78**, WC51–WC63.

———, 2013b, Ray–based gridded tomography for tilted transversely isotropic media: Geophysics, **78**, C11–C23.

Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. tekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: Geophysics, **78**, R59–R80.

Weibull, W. W., and B. Arntsen, 2014, Anisotropic migration velocity analysis using reverse-time migration: Geophysics, **79**, R13–R25.

Yang, T., and P. Sava, 2015, Image-domain wavefield tomography with extended common-image-point gathers: Geophysical Prospecting, 1086–1096.

Zhan, G., R. C. Pestana, and P. L. Stoffa, 2012, Decoupled equations for reverse time migration in tilted transversely isotropic media: Geophysics, **77**, T37–T45.

Zhang, Y., and G. Shan, 2013, Wave-equation migration velocity analysis using partial stack-power maximization: SEG Technical Program Expanded Abstracts, 4847–4852.

Zhang, Y., H. Zhang, and G. Zhang, 2011, A stable TTI reverse time migration and its implementation: Geophysics, **76**, WA3–WA11.

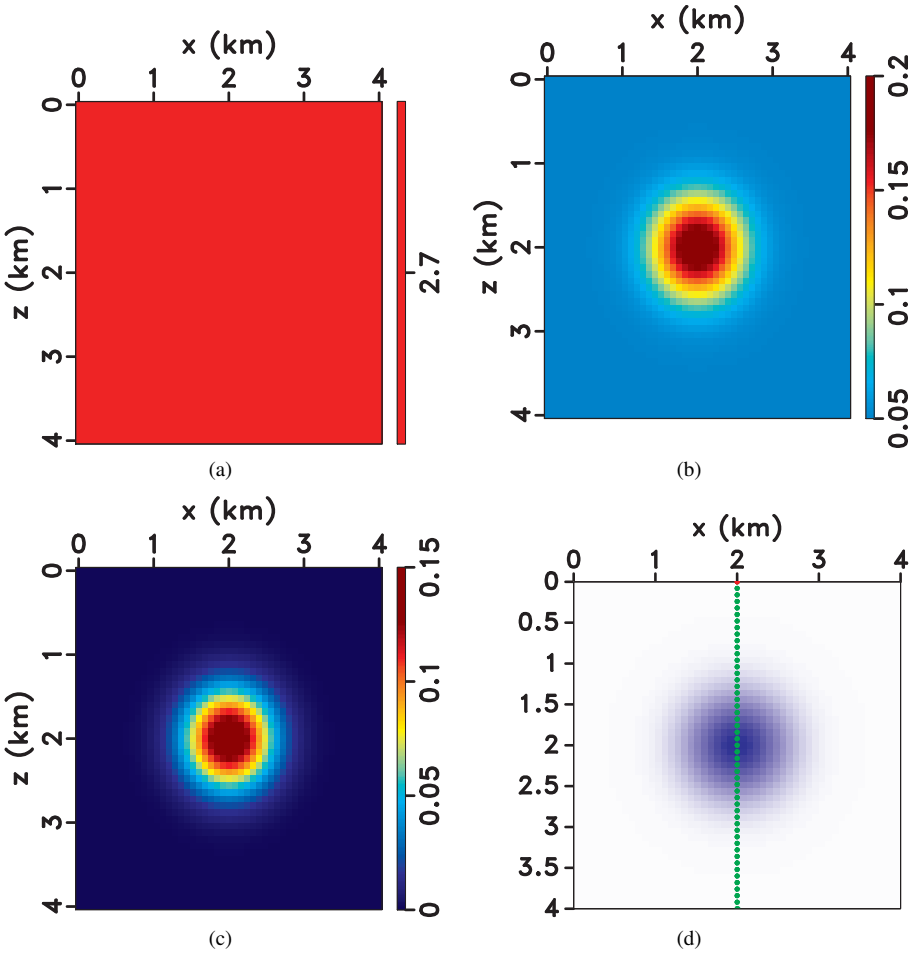

Figure 1. VTI model with Gaussian anomalies in the parameters η and ϵ: (a) V_{hor}, (b) η, and (c) ϵ (model 1). (d) Source (red dot)-receiver (green dots) geometry.

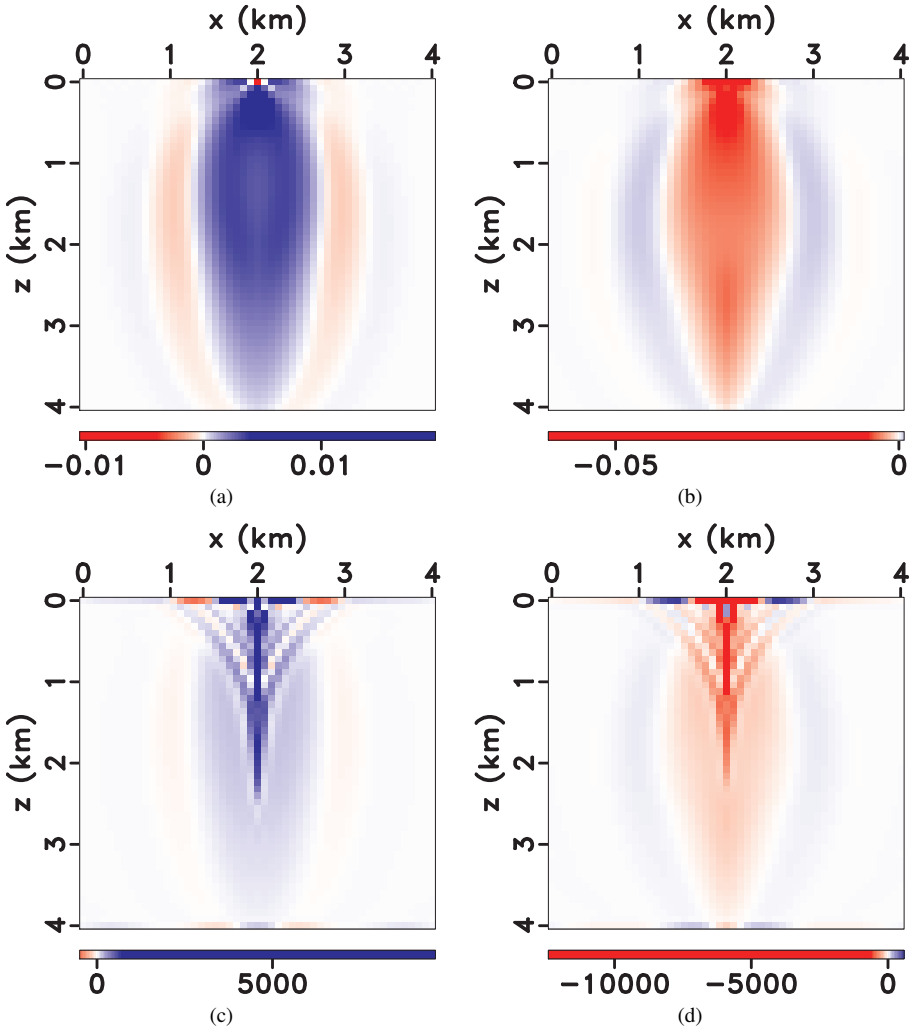

Figure 2. Gradients for model 1 computed using the (a,b) integral and (c,d) differential extrapolators with different peak values of ϵ: (a,c) $\epsilon = 0$ and (b,d) $\epsilon = 0.3$ (the actual peak $\epsilon = 0.15$). The differential operator produces a strong artifact at $x = 2$ km.

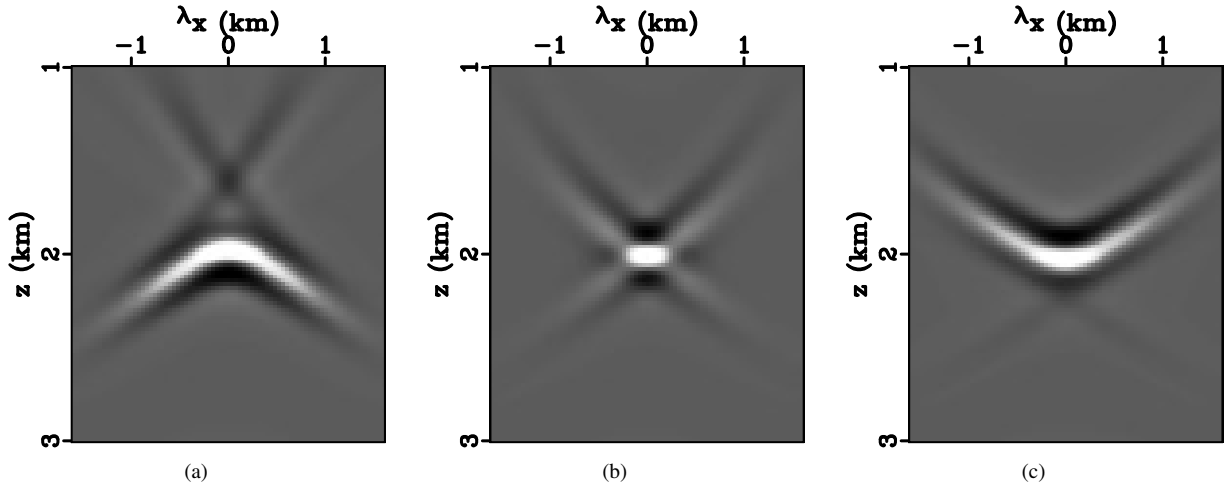

Figure 3. Space-lag CIGs for a horizontal VTI layer (model 2) computed in the middle of the model (x = 4 km) using the integral extrapolator with (a) $\eta = 0$, (b) $\eta = 0.15$ (actual value), and (c) $\eta = 0.3$.

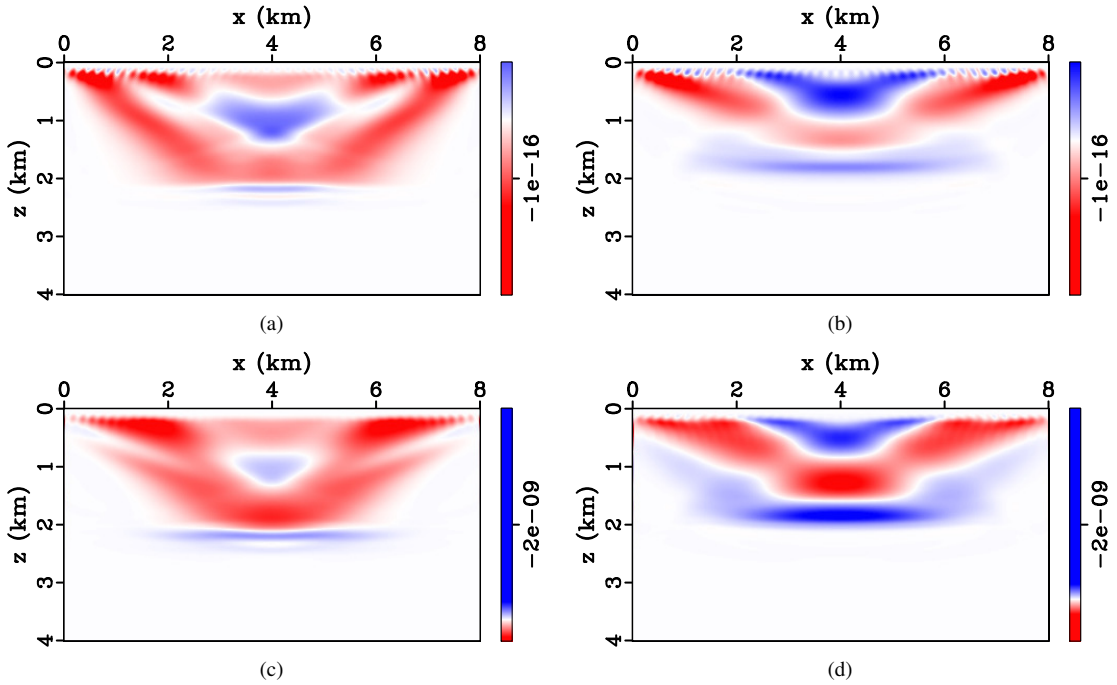

Figure 4. Gradients of the DSO objective function (equation 11) for model 2 computed using the (a,b) integral and (c,d) differential extrapolators for (a,c) $\eta = 0$ and (b,d) $\eta = 0.3$.

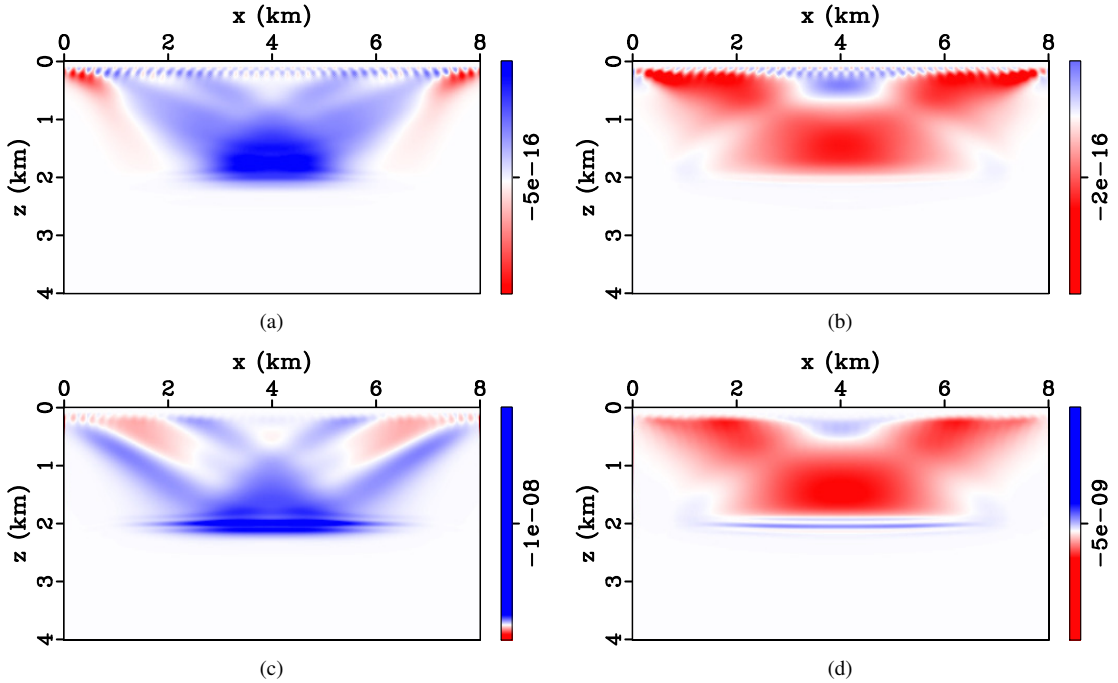

Figure 5. Gradients of the partial image-power function (equation 13) for model 2 computed using the (a,b) integral and (c,d) differential extrapolators for (a,c) $\eta = 0$ and (b,d) $\eta = 0.3$.

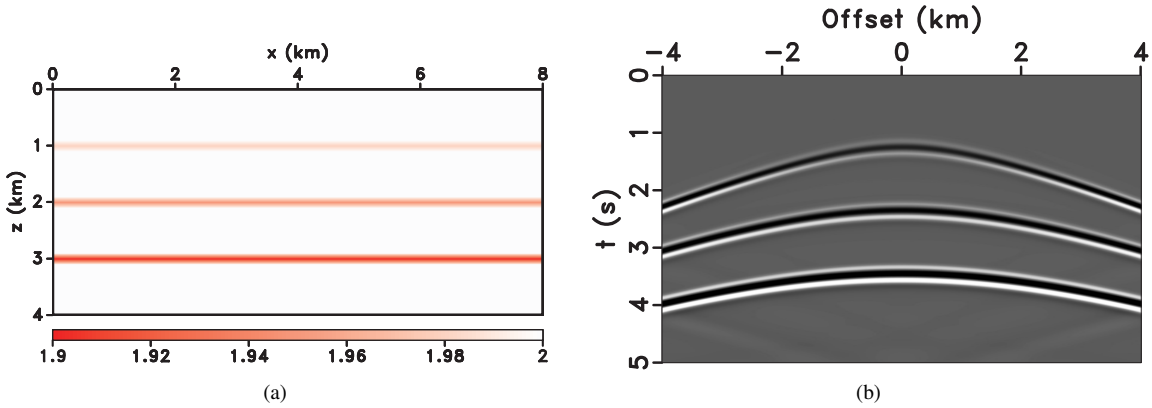

Figure 6. (a) VTI model with perturbations in the V_{nmo}-field (model 3); the anisotropy coefficients η and δ are constant throughout the model. (b) Shot gather for the source located in the middle of the model.

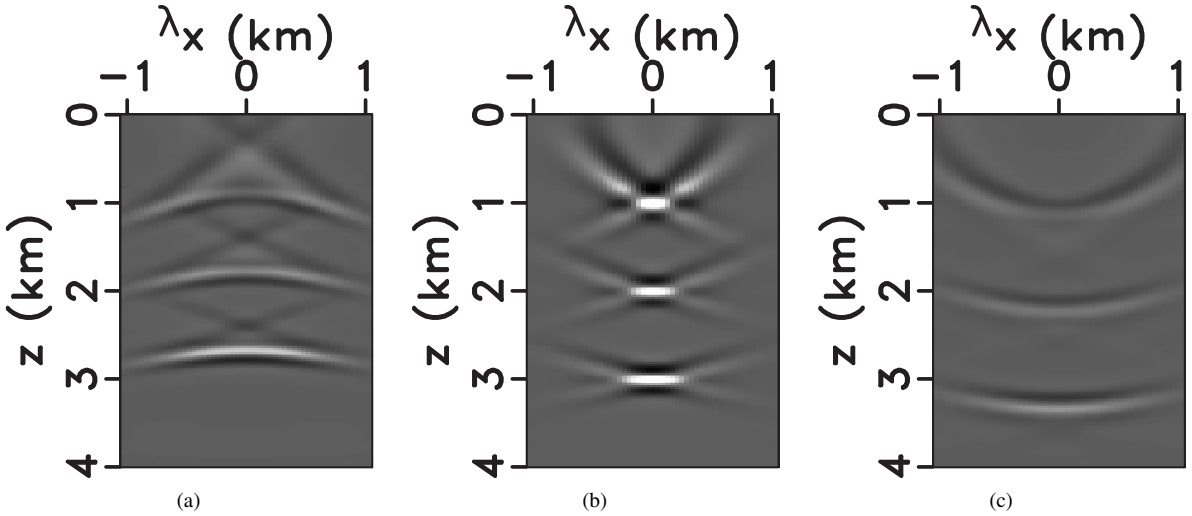

Figure 7. Space-lag CIGs for model 3 at x = 4 km using the integral extrapolator with (a) $V_{\rm nmo}$ = 1.8, (b) $V_{\rm nmo}$ = 2.0 (actual value), and (c) $V_{\rm nmo}$ = 2.2 km/s.

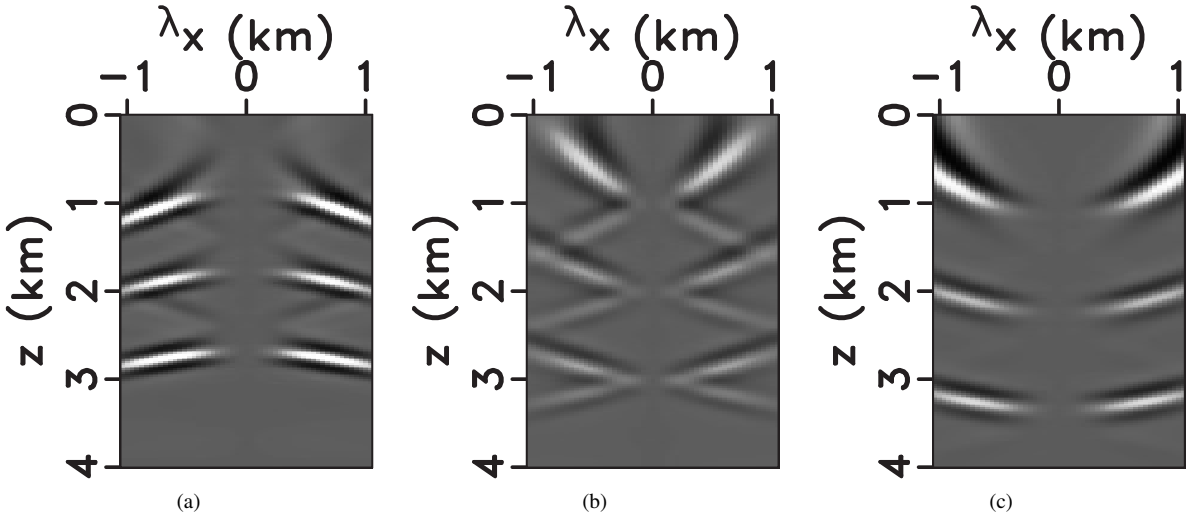

Figure 8. Space-lag CIGs for model 3 after applying the DSO operator. The gathers computed with the actual (plot b) and understated (plot a) value of $V_{\rm nmo}$ contain comparable residual energy.

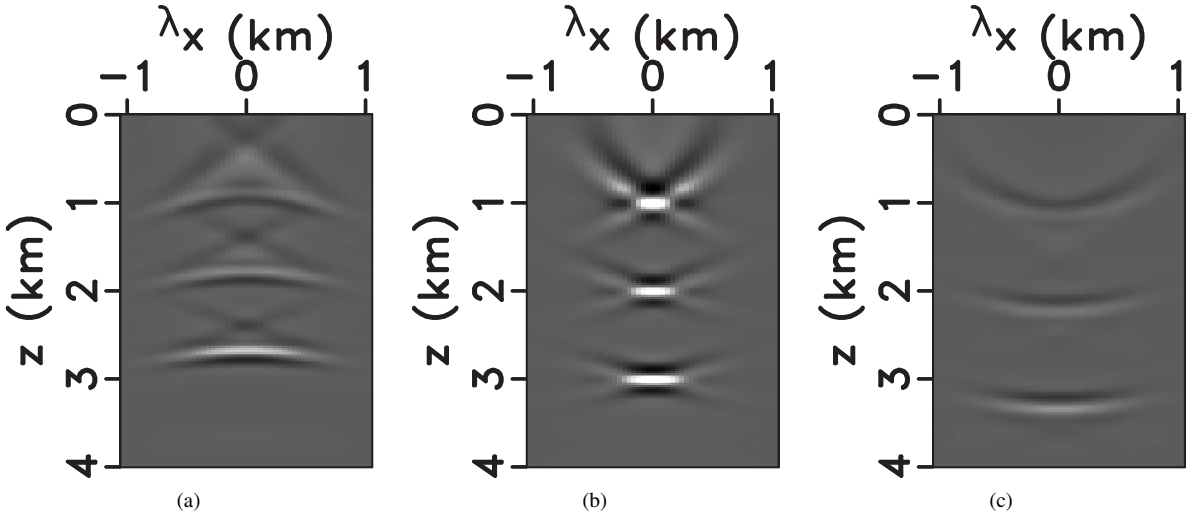

Figure 9. Space-lag CIGs for model 3 after applying the partial image-power operator. The gather computed with the actual model (plot b) features strong energy focusing at zero lag.

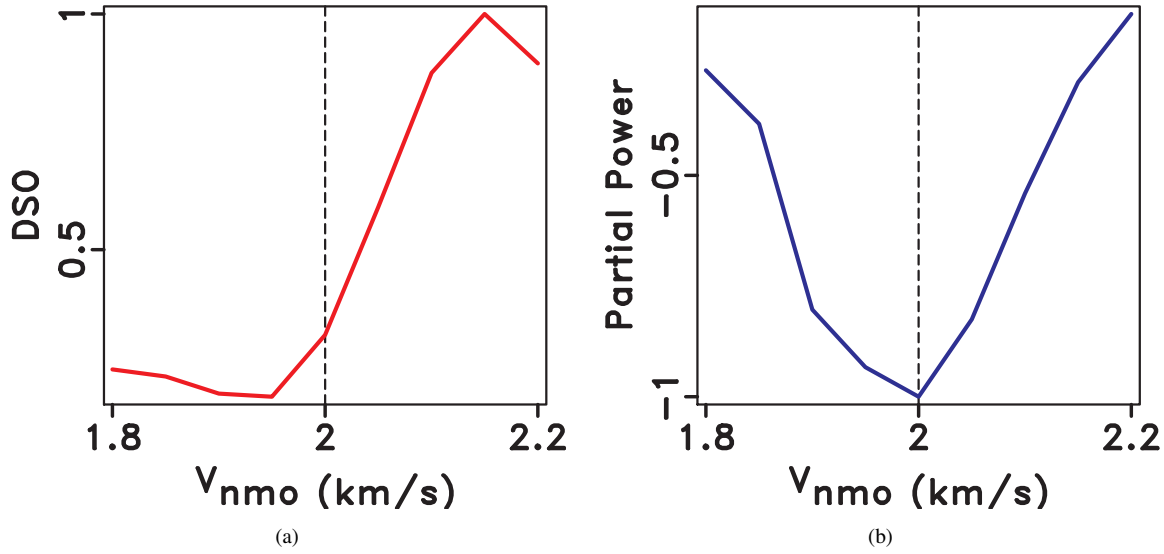

Figure 10. Dependence of the objective functions on the background $V_{\rm nmo}$-field: (a) DSO and (b) partial image-power. The actual background $V_{\rm nmo}$ = 2.0 km/s.

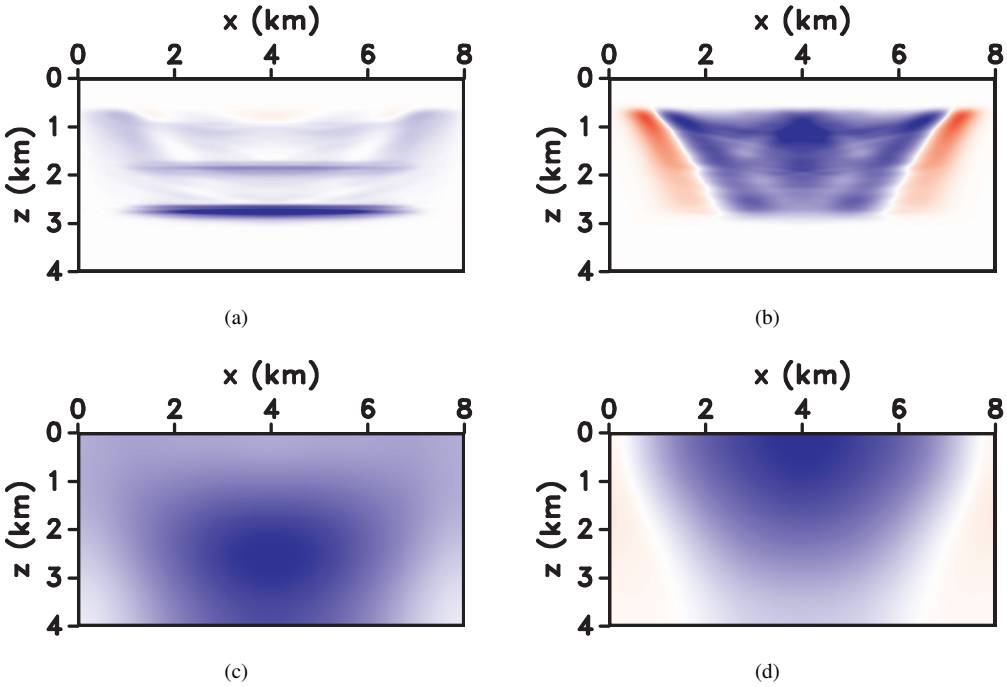

Figure 11. Gradients of the partial image-power objective function for model 3 (Figure 6): (a,c) V_{nmo} and (b,d) η. The gradients are computed before (a,b) and after (c,d) smoothing for the initial elliptic model ($\epsilon = \delta$).

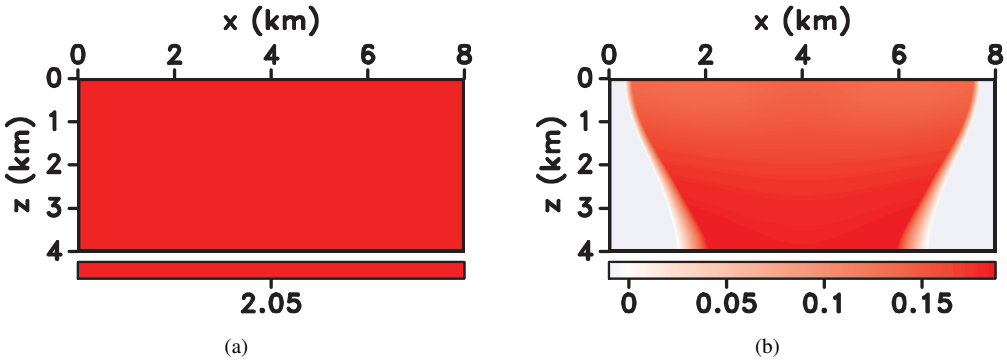

Figure 12. Estimated parameters V_{nmo} and η for model 3 after three iterations of image-domain tomography with the partial image-power objective function.