



Original Paper

Truncated Gauss-Newton full-waveform inversion of pure quasi-P waves in vertical transverse isotropic media

Zhi-Ming Ren ^{a, b, *}, Lei Wang ^a, Qian-Zong Bao ^{a, b}^a College of Geological Engineering and Geomatics, Chang'an University, Xi'an, 710064, Shaanxi, China^b National Engineering Research Center of Offshore Oil and Gas Exploration, Beijing, 100028, China

ARTICLE INFO

Article history:

Received 3 January 2024

Received in revised form

23 January 2024

Accepted 29 April 2024

Available online 4 May 2024

Edited by Meng-Jiao Zhou

Keywords:

Full waveform inversion

Anisotropy

Pure quasi-P wave

Gauss-Newton

Hessian

ABSTRACT

Full-waveform inversion (FWI) uses the full information of seismic data to obtain a quantitative estimation of subsurface physical parameters. Anisotropic FWI has the potential to recover high-resolution velocity and anisotropy parameter models, which are critical for imaging the long-offset and wide-azimuth data. We develop an acoustic anisotropic FWI method based on a simplified pure quasi P-wave (qP-wave) equation, which can be solved efficiently and is beneficial for the subsequent inversion. Using the inverse Hessian operator to precondition the functional gradients helps to reduce the parameter tradeoff in the multi-parameter inversion. To balance the accuracy and efficiency, we extend the truncated Gauss-Newton (TGN) method into FWI of pure qP-waves in vertical transverse isotropic (VTI) media. The inversion is performed in a nested way: a linear inner loop and a nonlinear outer loop. We derive the formulation of Hessian-vector products for pure qP-waves in VTI media based on the Lagrange multiplier method and compute the model update by solving a Gauss-Newton linear system via a matrix-free conjugate gradient method. A suitable preconditioner and the Eisenstat and Walker stopping criterion for the inner iterations are used to accelerate the convergence and avoid prohibitive computational cost. We test the proposed FWI method on several synthetic data sets. Inversion results reveal that the pure acoustic VTI FWI exhibits greater accuracy than the conventional pseudoacoustic VTI FWI. Additionally, the TGN method proves effective in mitigating the parameter crosstalk and increasing the accuracy of anisotropy parameters.

© 2024 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Full-waveform inversion (FWI) iteratively updates subsurface parameter models by a data-fitting procedure minimizing the distance between predicted and observed data (Tarantola, 1984; Virieux and Operto, 2009). Using the full wave equation to govern seismic wave propagation in the Earth, FWI can provide accurate physical parameters and quantitative images of the subsurface. Owing to the ability of obtaining P- and S-velocities, density, attenuation and anisotropic parameters, the multi-parameter FWI in elastic, anisotropic and viscoelastic media attracts an increasing interest in the past decades (Tarantola, 1986; Sheen et al., 2006; Sears et al., 2008; Prioux et al., 2013; Alkhalifah and Plessix, 2014;

Ren et al., 2021; Gao et al., 2021; Yong et al., 2022).

The multi-parameter FWI suffers from the parameter tradeoff because different parameters can have a coupled effect on the seismic response. There are mainly two kinds of methods to analyze the tradeoff effect: the radiation pattern and the Hessian operator (Operto et al., 2013). The radiation pattern indicates the amplitude variations of the partial derivative wavefield as a function of the scattering angle and shows the sensitivity of the scattered wave to a parameter perturbation along different direction. The crosstalk is serious when two parameter classes exhibit similar scattering patterns. It is also noted that the model update wavenumber depends on the scattering angle and the modulus of the local wavenumber. The wide and small scattering angles, associated with the diving and short-spread reflected waves, image the long and short wavelengths of the subsurface, respectively. The radiation patterns can be analytically derived based on the framework of the asymptotic ray + Born approximation (Forgues and Lambaré, 1997) or numerically computed by the subtraction of two

* Corresponding author. College of Geological Engineering and Geomatics, Chang'an University, Xi'an, 710064, Shaanxi, China.

E-mail address: rzm-213@163.com (Z.-M. Ren).

snapshots for a homogeneous model with or without a certain parameter perturbation (Malinowski et al., 2011; Zhou et al., 2015; Ren and Liu, 2016). The Hessian operator is used in the Newton-based method and can be divided into several blocks in the multi-parameter inversion. The diagonal blocks, formed by the autocorrelation of the partial derivative wavefield associated with the parameters of the same nature, determine the partial resolution of the estimation. The off-diagonal blocks, formed by the cross-correlation of the partial derivative wavefield associated with the parameters of different nature, show the parameter tradeoff effect. The interparameter crosstalk disappears when the elements in the off-diagonal blocks of the Hessian matrix are zero. Incorporating the inverse Hessian operator on the gradients helps to remove the blurring effects, caused by the uneven illumination and the limited bandwidth of seismic data, improve the spatial resolution, and remove the interparameter crosstalk (Pratt et al., 1998; Fichtner and Trampert, 2011; Métivier et al., 2013; Yang et al., 2018; Liu et al., 2020; Wang et al., 2022). The explicit computation of the Hessian and its inverse is unaffordable for current computational capability. The quasi-Newton methods, using an approximate inverse Hessian, can dramatically improve the efficiency at the cost of the inversion accuracy. The L-BFGS approximation is the most popular quasi-Newton method in FWI (Nocedal and Wright, 1999; Virieux and Operto, 2009). Afterward, the truncated Gauss-Newton (TGN) and truncated Newton (TN) methods, using the linear and full Hessian operator, respectively, are proposed to accelerate the convergence rate and minimize the interparameter tradeoff (Métivier et al., 2013, 2014). The methods are implemented based on the computation of Hessian-vector products and a matrix-free conjugate gradient solution of a Newton linear system. Up to now, the TGN and TN methods have been applied in elastic (Liu et al., 2020), anisotropic (Pan et al., 2016), viscoacoustic (Yang et al., 2018) and viscoelastic (Gao et al., 2021) FWI in the time or frequency domain. The TGN and TN methods require a large number of forward modeling in the inner linear iterations. The source encoding (Castellanos et al., 2015), shot subsampling (Matharu and Sacchi, 2019; Wang et al., 2022) and Fourier-domain full-scattered-field approximation (Yong et al., 2022) are effective strategies to lower the computational burden. The parameter tradeoff can further be reduced by applying a suitable parameterization (Gholami et al., 2013; Alkhalifah and Plessix, 2014), the P- and S-wave separation (Ren and Liu, 2016; Wang et al., 2017) or a hierarchical inversion workflow (Jeong et al., 2012; Ren et al., 2019).

Accounting for the anisotropic effect in FWI is crucial for quantitative imaging of the subsurface. Researchers have explored elastic anisotropic FWI in vertical (Kamath and Tsvankin, 2016; Pan et al., 2019; Singh et al., 2020), horizontal (Pan et al., 2016), tilted (Rusmanugroho et al., 2017; Singh et al., 2021) transverse isotropic (TI), and orthorhombic (Oh and Alkhalifah, 2016) media. Based on the radiation patterns, He et al. (2018) analyzed the sensitivity of seismic data to a certain elastic parameter in the land elastic vertical TI (VTI) FWI. Pan et al. (2016) analytically computed the ra-

tilt angles in elastic tilted TI FWI. Compared with elastic anisotropic FWI, acoustic anisotropic FWI is computationally efficient (Plessix and Cao, 2011; Prieux et al., 2011; Gholami et al., 2013; Alkhalifah and Plessix, 2014; Hadden et al., 2019). Plessix and Cao (2011) employed the eigenanalysis of the Hessian operator to analyze the tradeoff between parameters in acoustic VTI FWI. Gholami et al. (2013) performed the sensitivity and crosstalk analyses for several combinations of parameters and concluded that the velocity had more dominant influence on seismic data than the anisotropy parameters. Alkhalifah and Plessix (2014) derived the radiation patterns for six parameter classes and revealed that the parameterization with the normal moveout (NMO) velocity and anisotropy parameters exhibited limited interparameter crosstalk. Wang and Dong (2015) carried out TN FWI in VTI media and proved the advantage of the TN method over the L-BFGS method. The work in this field is focused on discussing which parameterization is suitable for FWI in VTI media and which parameters can be reliably inverted for. Also, the present acoustic anisotropic FWI methods are based on the pseudoacoustic wave equations. Two problems arise when the pseudoacoustic solvers are used in anisotropic FWI. The spurious S-wave artifacts greatly reduce the accuracy and the inversion becomes unstable if the anisotropy parameter $\varepsilon < \delta$ (Alkhalifah, 2000; Duveneck et al., 2008). The pure quasi-P (qP) wave equations have been further proposed for modeling seismic wave propagation in TI media (Chu et al., 2011; Li and Zhu, 2018; Mu et al., 2023). The pure qP-wave equations are free from the spurious S-waves and can be solved stably for any ε and δ . The radiation patterns and the Hessian operator for the pure qP-wave FWI are different from those for the pseudoacoustic and elastic anisotropic FWI. Hence, it is necessary to investigate anisotropic FWI methods under the pure acoustic approximation. Ren et al. (2024) performed FWI of pure qP-waves in TTI media based on different parameterizations and concluded that the parameterization of the NMO velocity and anisotropy parameters helped to enhance the inversion accuracy. However, the improvement is relatively small.

In this paper, we develop an acoustic VTI FWI method based on a simplified pure qP-wave equation. The approximate Hessian operator is used to analyze the tradeoff between parameters. We derive the formulations of Hessian-vector products based on the Lagrange multiplier method and present an efficient matrix-free Gauss-Newton method. Synthetic data on a simple model and the modified Hess VTI model are applied to test the proposed method. The pseudoacoustic FWI, pure acoustic FWI, and pure acoustic TGN FWI are compared.

2. Methodology

2.1. The forward problem

The pure qP-wave equation in VTI media is expressed as (Liu et al., 2009)

$$\left(\omega^2 - \frac{v^2}{2} (k_x^2(1+2\varepsilon) + k_z^2) - \frac{v^2}{2} \sqrt{k_x^4(1+2\varepsilon)^2 + 2(1-2\varepsilon+4\delta)k_x^2k_z^2 + k_z^4} \right) \hat{P} = 0 \quad (1)$$

diation patterns in general anisotropic media and implemented elastic TN FWI in horizontal TI media. Rusmanugroho et al. (2017) showed that the Voigt parameters and the Chen and Tromp parameters contributed to the recovery of anisotropy parameters and

where $\hat{P}$ denotes the pressure field in the frequency-wavenumber domain, ω is the angular frequency, $k_x = k \cos \theta$ and $k_z = k \sin \theta$, k is the wavenumber, θ is the angle between the x -axis and the propagation direction, v is the propagation velocity along the

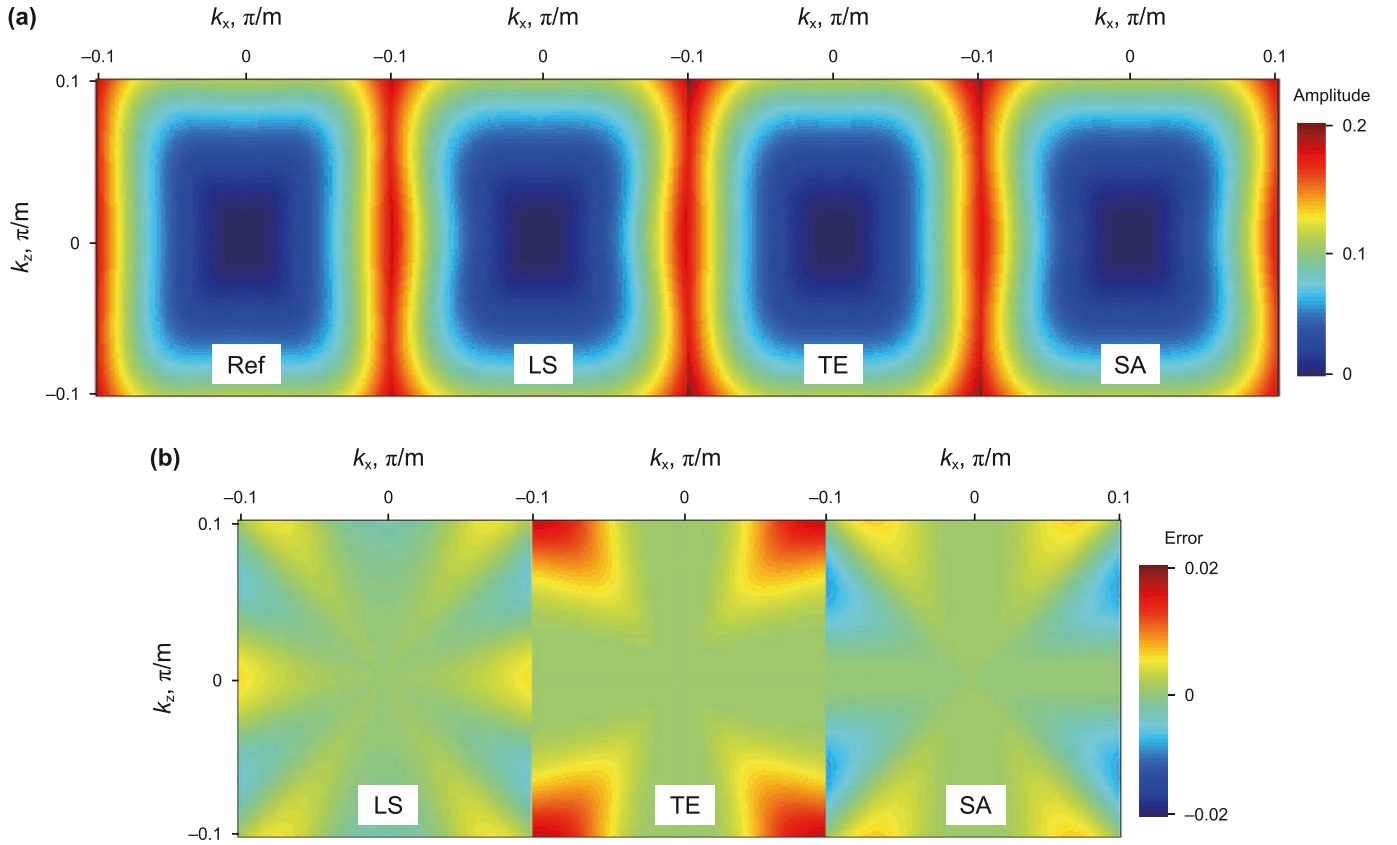


Fig. 1. (a) Different approximations of the pseudo-differential operator and (b) the approximation errors in the wavenumber domain, $\varepsilon = 0.3$ and $\delta = -0.2$. Ref, LS, TE and SA represent the exact pseudo-differential operator and the approximations based on LS, TE and SA, respectively.

symmetry axis, and ε and δ are the Thomsen's anisotropy parameters (Thomsen, 1986). It is cumbersome to cope with the pseudo-differential operator in Eq. (1). Chu et al. (2011) used the Taylor series expansion (TE) to approximate the square root term:

$$\sqrt{k_x^4(1+2\varepsilon)^2 + 2(1-2\varepsilon+4\delta)k_x^2k_z^2 + k_z^4} \approx (1+2\varepsilon)k_x^2 + k_z^2 + \frac{4(\delta-\varepsilon)k_x^2k_z^2}{(1+2\varepsilon)k_x^2 + k_z^2} \quad (2)$$

Based on Eqs. (1) and (2), the pure qP-wave equation in the time-space domain is derived:

$$\frac{\partial^2 p}{\partial t^2} = v^2 \left((1+2\varepsilon) \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} + 2(\delta-\varepsilon) \frac{\partial^4 q}{\partial x^2 \partial z^2} \right) \quad (3)$$

$$(1+2\varepsilon) \frac{\partial^2 q}{\partial x^2} + \frac{\partial^2 q}{\partial z^2} = p \quad (4)$$

where p is the time-space-domain pressure and q is the auxiliary field. Eq. (3) can be solved by the finite-difference (FD) or pseudo-spectral method. Discretizing Eq. (4) by FD leads to a sparse linear equation system, which can be solved by an iterative linear solver.

Li and Zhu (2018) applied a different algebraic expression to approximate the square root term:

$$\sqrt{k_x^4(1+2\varepsilon)^2 + 2(1-2\varepsilon+4\delta)k_x^2k_z^2 + k_z^4} \approx c_1 k_x^2 + c_2 k_z^2 + c_3 \frac{k_x^2 k_z^2}{k_x^2 + k_z^2} \quad (5)$$

and optimized the coefficients $c_1 - c_3$ by minimizing the absolute error between both sides of Eq. (5) based on the least-squares (LS). For details on the computation of $c_1 - c_3$, please see Li and Zhu (2018). The LS-based pure qP-wave equation is written as

$$\frac{\partial^2 p}{\partial t^2} = \frac{v^2}{2} \left(d_1 \frac{\partial^2 p}{\partial x^2} + d_2 \frac{\partial^2 p}{\partial z^2} + d_3 \frac{\partial^4 q}{\partial x^2 \partial z^2} \right) \quad (6)$$

$$\frac{\partial^2 q}{\partial x^2} + \frac{\partial^2 q}{\partial z^2} = p \quad (7)$$

where

$$d_1 = 1 + 2\varepsilon + c_1, d_2 = 1 + c_2, d_3 = c_3 \quad (8)$$

Eq. (7) is a Poisson's equation with constant coefficients and can be solved by the fast Poisson solver (Strang, 2007). From the point of view of computational efficiency, the LS-based pure qP-wave equation is superior to the TE-based one. Note that the coefficients $c_1 - c_3$ depend on the Thomsen's anisotropy parameters

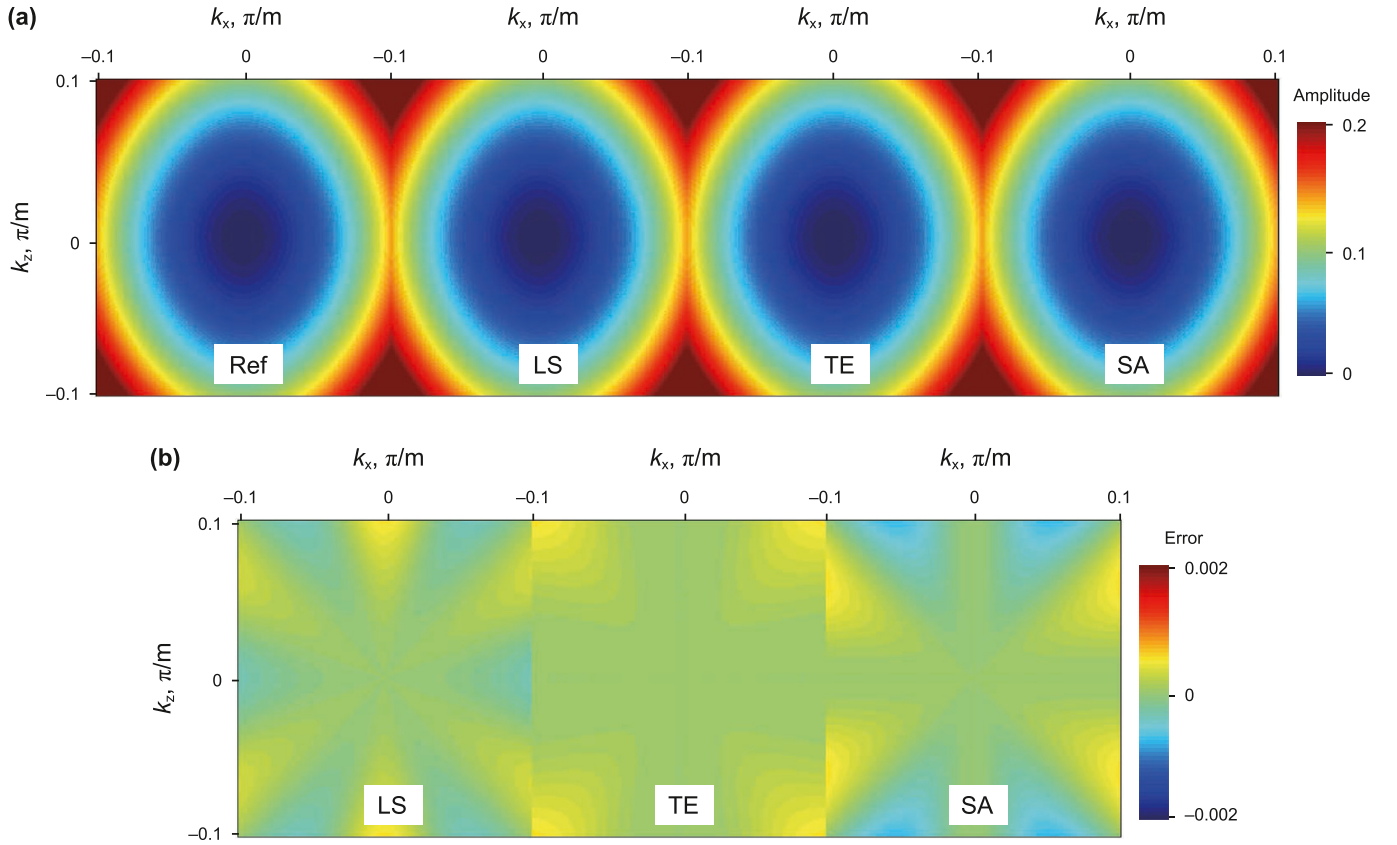


Fig. 2. (a) Different approximations of the pseudo-differential operator and (b) the approximation errors in the wavenumber domain, $\varepsilon = 0.2$ and $\delta = 0.3$.

and need to be recomputed in each iteration because ε and δ keep changed during the inversion.

We then use a sampling approximation (SA) (Finkelstein and Kastner, 2007; Yang et al., 2016) to approximate the pseudo-differential operator. Setting $k_x^2 = 0$, $k_z^2 = 0$ and $k_x^2 = k_z^2$, respectively in Eq. (5), gives

$$c_1 = 1 + 2\varepsilon, c_2 = 1, c_1 + c_2 + \frac{c_3}{2} = 2\sqrt{\varepsilon^2 + 2\delta + 1} \quad (9)$$

which is equivalent to

$$c_1 = 1 + 2\varepsilon, c_2 = 1, c_3 = 4(\sqrt{\varepsilon^2 + 2\delta + 1} - \varepsilon - 1) \quad (10)$$

The SA approximation guarantees that the errors along eight propagation directions ($\theta = \frac{n\pi}{4}$; $n = 0, 1, \dots, 7$) are equal to zero for the 2D case. Substituting Eq. (10) into Eqs. (6)–(8) results in

$$\frac{\partial^2 p}{\partial t^2} = v^2 \left((1 + 2\varepsilon) \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} + 2(\sqrt{\varepsilon^2 + 2\delta + 1} - \varepsilon - 1) \frac{\partial^4 q}{\partial x^2 \partial z^2} \right) \quad (11)$$

$$\frac{\partial^2 q}{\partial x^2} + \frac{\partial^2 q}{\partial z^2} = p \quad (12)$$

Eqs. (11) and (12) represent the simplified pure qP-wave equation, which can still be solved by FD and the fast Poisson solver

(Strang, 2007). The coefficients of the spatial derivatives in Eq. (11) are expressed explicitly by velocity and anisotropy parameters, which is beneficial for the subsequent inversion. The functional gradients with respect to model parameters can be analytically derived.

In Figs. 1 and 2, two homogenous models are used to evaluate the accuracy of different approximations of the pseudo-differential operator. The case that ε is close and far from to δ is called as weak and strong anelliptical anisotropy, respectively. It is observed that the TE-based approximation exhibits larger errors than the other two approximations for the anelliptical anisotropic case ($\varepsilon = 0.3$ and $\delta = -0.2$). The accuracy of the SA approximation is very close to that of the LS-based approximation for either weak or strong anelliptical anisotropy. Figs. 3 and 4 illustrate the modeled snapshots using different pure qP-wave equations. The time step is 1.0 ms, the recording time is 1.0 s, the grid spacing is 10 m, and the grid dimensions are 201×201 . A Ricker wavelet with the peak frequency of 30 Hz is used as a source. The approximate equations are solved by FD with the second-order accuracy in time and the tenth-order accuracy in space. The equation based on the exact pseudo-differential operator is solved by the pseudo-spectral method. All the three approximate equations can obtain reasonable wavefield for the weak anelliptical anisotropy, whereas the amplitude errors appear in the simulated results with the anellipticity increasing. However, the solutions of the approximate equations have correct kinematics of qP-waves for the strong anelliptical anisotropic case. In addition, the pure qP-wave equations can be solved stably for the

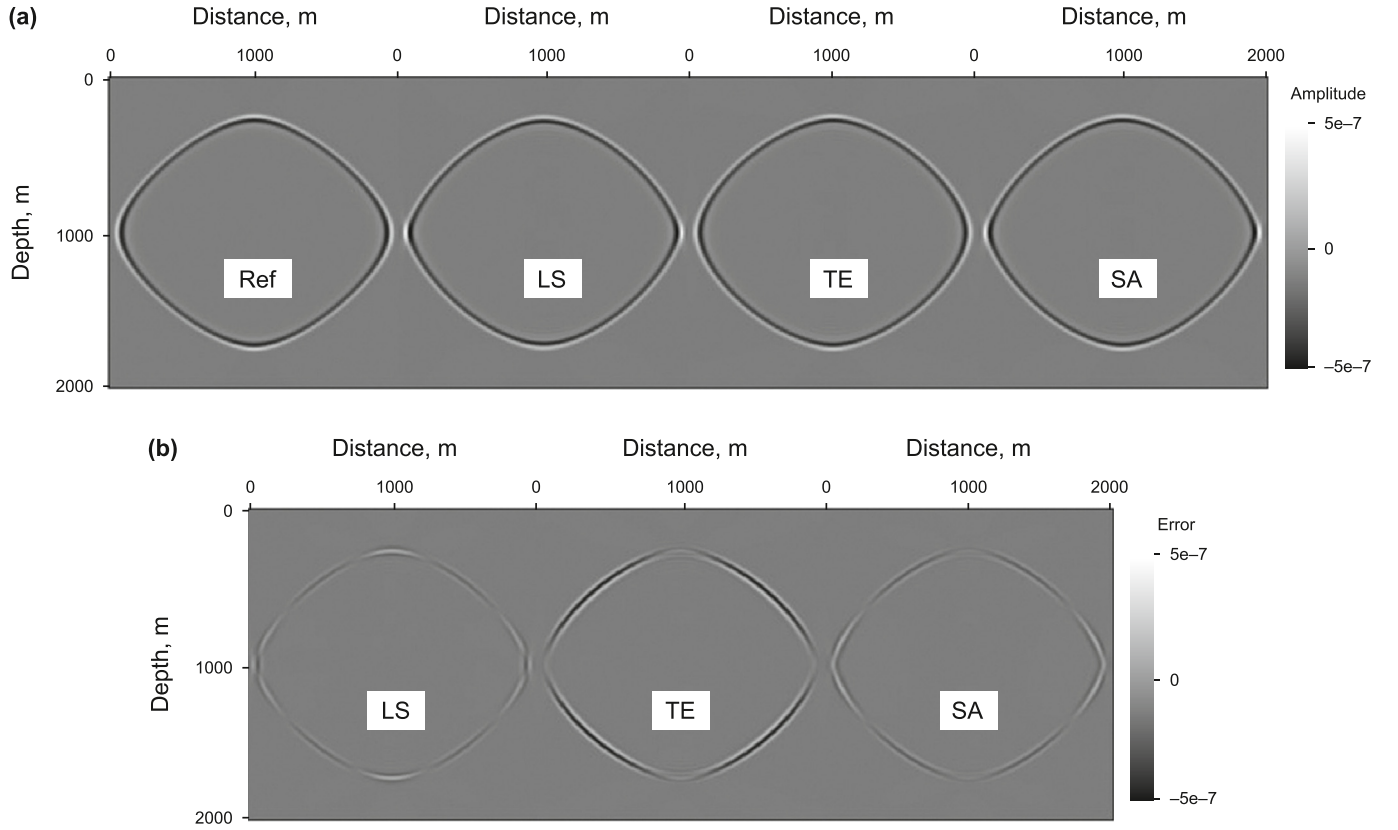


Fig. 3. (a) Snapshots at 0.4 s modeled by using different pure qP-wave equations based on the pseudo-differential operator and the approximations and (b) the errors of different approximations for a homogeneous VTI model, $v = 2000$ m/s, $\varepsilon = 0.3$ and $\delta = -0.2$.

case of $\varepsilon < \delta$ ($\varepsilon = 0.2$ and $\delta = 0.3$). The simplified and LS-based pure qP-wave equations are preferable to the TE-based equation for modeling the wave propagation in the media with strong anelliptical anisotropy. The computational time of solving the TE-based, LS-based and simplified equations is 41.6, 30.4 and 30.2 s, respectively on the workstation with two AMD EPYC 7542 processors and 384 GB of memory.

As mentioned above, the simplified pure qP-wave equation has very close accuracy, efficiency and stability to the LS-based equation. We apply the simplified equation in the following FWI because the coefficients can be expressed explicitly by the velocity and anisotropy parameters.

2.2. The inverse problem

FWI updates subsurface models by minimizing the distance between the predicted and observed data. The misfit function in the L2-norm criterion is constructed:

$$\min_{v, \varepsilon, \delta} E(v, \varepsilon, \delta) = \frac{1}{2} \int_{(x,z) \in H} \int_0^T (\mathbf{R}\mathbf{P} - \mathbf{P}_{\text{obs}})^T (\mathbf{R}\mathbf{P} - \mathbf{P}_{\text{obs}}) dt dx dz \quad (13)$$

where $\mathbf{P}$ and $\mathbf{P}_{\text{obs}}$ represent the modeled and observed data,

respectively, H is the computational domain, T is the computational time, the superscript T is the matrix transport operator, and $\mathbf{R}$ is a mapping of the wavefield to the receivers' locations. A summation over all sources is omitted for the sake of clarity. The wavefield $\mathbf{P}$ is modeled by the forward problem:

$$\mathbf{A}\mathbf{P} = \mathbf{s} \quad (14)$$

where $\mathbf{P} = [p \ q]^T$, $\mathbf{s}$ denotes the source item,

$$\mathbf{A} = \begin{bmatrix} \frac{\partial^2}{\partial t^2} - v^2(1+2\varepsilon)\frac{\partial^2}{\partial x^2} - v^2\frac{\partial^2}{\partial z^2} - 2v^2\left(\sqrt{\varepsilon^2 + 2\delta + 1} - \varepsilon - 1\right)\frac{\partial^4}{\partial x^2 \partial z^2} \\ -1 & \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \end{bmatrix} \quad (15)$$

Eq. (14), the same as Eqs. (11) and (12) except for the source term, is treated as a constraint condition, and we employ the Lagrange multiplier method (Bunks et al., 1995) to solve the minimization problem. The misfit function is rewritten as

$$\min_{v, \varepsilon, \delta} J(v, \varepsilon, \delta) = \min_{v, \varepsilon, \delta} E(v, \varepsilon, \delta) + \int_{(x,z) \in H} \int_0^T \lambda^T (\mathbf{A}\mathbf{P} - \mathbf{s}) dt dx dz \quad (16)$$

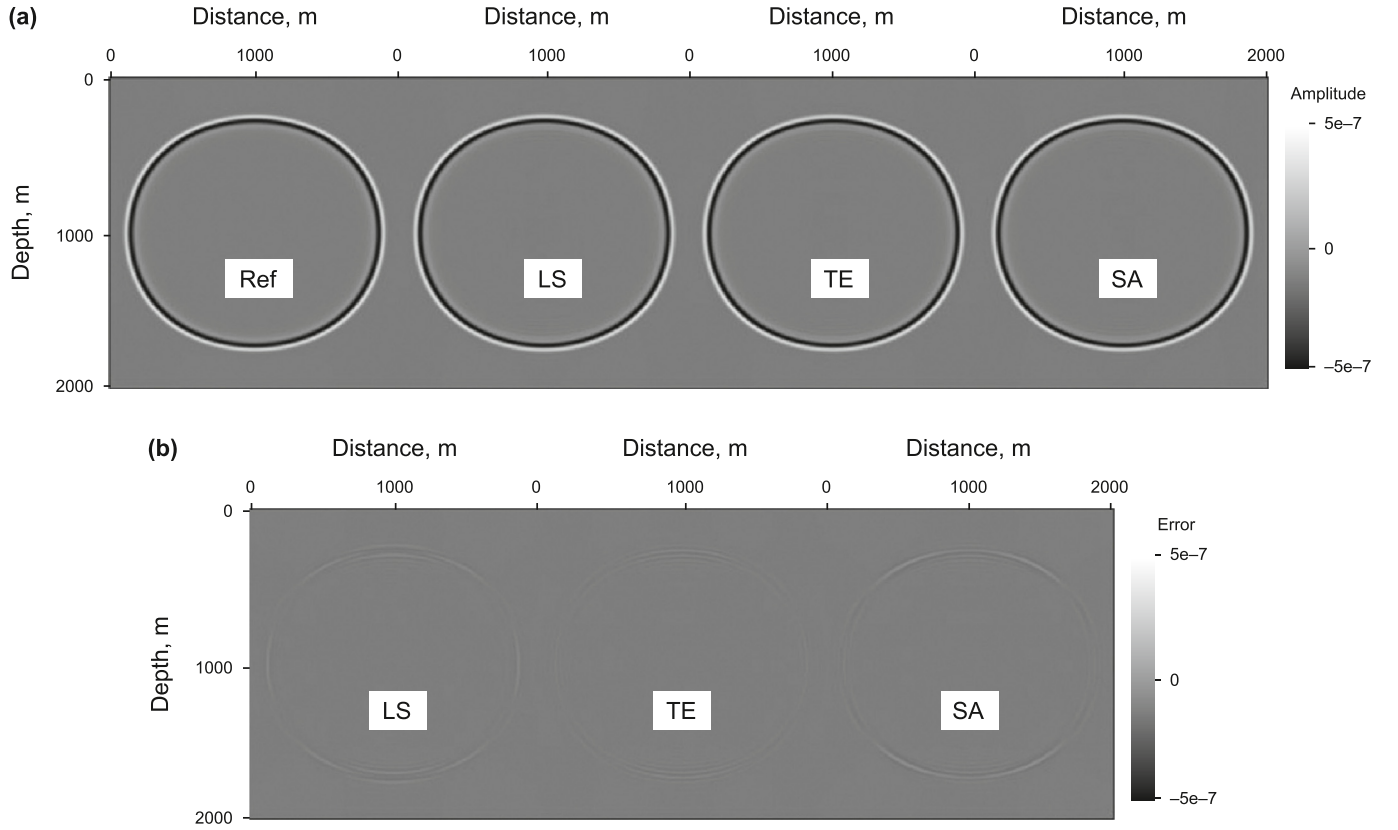


Fig. 4. (a) Snapshots at 0.4 s modeled by using different pure qP-wave equations based on the pseudo-differential operator and the approximations and (b) the errors of different approximations for a homogeneous VTI model, $v = 2000$ m/s, $\varepsilon = 0.2$ and $\delta = 0.3$.

where $\lambda = [\lambda_1 \ \lambda_2]^T$ stands for the Lagrange multiplier function. Integrating Eq. (16) by parts gives

$$\min_{v, \varepsilon, \delta} J(v, \varepsilon, \delta) = \min_{v, \varepsilon, \delta} E(v, \varepsilon, \delta) + \int_{(x,z) \in H} \int_0^T (\mathbf{P}^T \mathbf{A}^T \lambda - \lambda^T \mathbf{s}) dt dx dz \quad (17)$$

Setting $\partial J / \partial \mathbf{P} = \mathbf{0}$ leads to

$$\mathbf{A}^T \lambda = -(\mathbf{R}\mathbf{P} - \mathbf{P}_{\text{obs}}) \quad (18)$$

Eq. (18) denotes the adjoint equation of pure qP-waves in VTI media and can also be solved by FD and the fast Poisson solver (Strang, 2007). The functional gradients with respect to velocity and anisotropy parameters are computed by the chain rule $\frac{\partial E}{\partial m} = \frac{\partial J}{\partial \mathbf{P}} \frac{\partial \mathbf{P}}{\partial m} + \frac{\partial J}{\partial m}$, $m \in \{v, \varepsilon, \delta\}$.

$$\frac{\partial E}{\partial m} = \int_0^T \lambda^T \frac{\partial \mathbf{A}}{\partial m} \mathbf{P} dt \quad (19)$$

where

$$\frac{\partial \mathbf{A}}{\partial v} = \begin{bmatrix} -2v(1+2\varepsilon) \frac{\partial^2}{\partial x^2} - 2v \frac{\partial^2}{\partial z^2} & -4v(\sqrt{\varepsilon^2 + 2\delta + 1} - \varepsilon - 1) \frac{\partial^4}{\partial x^2 \partial z^2} \\ 0 & 0 \end{bmatrix} \quad (20)$$

$$\frac{\partial \mathbf{A}}{\partial \varepsilon} = \begin{bmatrix} -2v^2 \frac{\partial^2}{\partial x^2} & -2v^2 \left(\frac{\varepsilon}{\sqrt{\varepsilon^2 + 2\delta + 1}} - 1 \right) \frac{\partial^4}{\partial x^2 \partial z^2} \\ 0 & 0 \end{bmatrix} \quad (21)$$

$$\frac{\partial \mathbf{A}}{\partial \delta} = \begin{bmatrix} 0 & \frac{-2v^2}{\sqrt{\varepsilon^2 + 2\delta + 1}} \frac{\partial^4}{\partial x^2 \partial z^2} \\ 0 & 0 \end{bmatrix} \quad (22)$$

We carry out FWI of pure qP-waves for VTI media as follows: forward propagating the source wavefield by Eq. (14); backward propagating the adjoint wavefield by Eq. (18); obtaining the gradients with the velocity and anisotropy parameters by Eq. (19); updating the velocity and anisotropy parameter models using a proper optimization algorithm.

The preconditioning of the functional gradients helps to accelerate the convergence and mitigate the parameter tradeoff. The

inverse Hessian operator is generally used in the Newton-based methods

$$\Delta \mathbf{m}_k = -\mathbf{H}_k^{-1} \mathbf{grad}_k \quad (23)$$

where

$$\mathbf{m} = [v_1 \ v_2 \ \dots \ v_N \ \varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_N \ \delta_1 \ \delta_2 \ \dots \ \delta_N]^T \quad (24)$$

$$\mathbf{grad} = \left[\frac{\partial E}{\partial v_1} \ \frac{\partial E}{\partial v_2} \ \dots \ \frac{\partial E}{\partial v_N} \ \frac{\partial E}{\partial \varepsilon_1} \ \frac{\partial E}{\partial \varepsilon_2} \ \dots \ \frac{\partial E}{\partial \varepsilon_N} \ \frac{\partial E}{\partial \delta_1} \ \frac{\partial E}{\partial \delta_2} \ \dots \ \frac{\partial E}{\partial \delta_N} \right]^T \quad (25)$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{vv} & \mathbf{H}_{v\varepsilon} & \mathbf{H}_{v\delta} \\ \mathbf{H}_{\varepsilon v} & \mathbf{H}_{\varepsilon\varepsilon} & \mathbf{H}_{\varepsilon\delta} \\ \mathbf{H}_{\delta v} & \mathbf{H}_{\delta\varepsilon} & \mathbf{H}_{\delta\delta} \end{bmatrix}^T \quad (26)$$

$$\mathbf{H}_{mm'}[i][j] = \frac{\partial^2 E}{\partial m_i \partial m'_j}, m, m' \in \{v, \varepsilon, \delta\}, i, j = 1, 2, \dots, N \quad (27)$$

where $\Delta \mathbf{m}$ is the model update, $\mathbf{grad}$ is the gradient vector, $\mathbf{H}$ is the Hessian operator, N is the grid dimensions, and k denotes the k th iteration. The relation between first- and second-order derivatives of the misfit function with respect to model parameters is

$$\mathbf{H} = \frac{\partial \mathbf{grad}^T}{\partial \mathbf{m}} \quad (28)$$

The TN methods can be performed without the explicit computation of the Hessian and its inverse operator. We then derive the formulation of Hessian-vector products for pure qP-waves in VTI media. A new misfit function is constructed:

$$F = \mathbf{grad}^T \mathbf{x} = \sum_{i=1}^N \left(\frac{\partial E}{\partial v_i} x_{v_i} + \frac{\partial E}{\partial \varepsilon_i} x_{\varepsilon_i} + \frac{\partial E}{\partial \delta_i} x_{\delta_i} \right) \quad (29)$$

where $\mathbf{x}$ is an arbitrary vector with the dimension of $3N$ and can be expressed as

$$\mathbf{x} = [x_{v_1} \ x_{v_2} \ \dots \ x_{v_N} \ x_{\varepsilon_1} \ x_{\varepsilon_2} \ \dots \ x_{\varepsilon_N} \ x_{\delta_1} \ x_{\delta_2} \ \dots \ x_{\delta_N}]^T \quad (30)$$

Based on Eqs. (28) and (29), we have

$$\frac{\partial F}{\partial \mathbf{m}} = \frac{\partial \mathbf{grad}^T}{\partial \mathbf{m}} \mathbf{x} = \mathbf{H} \mathbf{x} \quad (31)$$

Note that the Hessian-vector products ($\mathbf{H} \mathbf{x}$) are equivalent to the functional gradients with respect to model parameters. We use the Lagrange multiplier method instead of the second-order adjoint-state method to formulate the Hessian-vector products for pure qP-waves in VTI media. The continuous formulation is derived for the sake of clarity. The misfit function becomes

$$F = \int_{(x,z) \in H} \left(\frac{\partial E}{\partial v} x_v + \frac{\partial E}{\partial \varepsilon} x_\varepsilon + \frac{\partial E}{\partial \delta} x_\delta \right) dx dz \quad (32)$$

The gradient $\partial E / \partial m$, $m \in \{v, \varepsilon, \delta\}$ is computed by Eq. (19) and the forward and adjoint wavefield satisfy Eqs. (14) and (18), respectively. We treat Eqs. (14), (18) and (19) as a constraint condition and rewrite the misfit function as

$$\begin{aligned} G = & F \\ & + \int_{(x,z) \in H} \left(\mu_v \left(\frac{\partial E}{\partial v} - \int_0^T \lambda^T \frac{\partial \mathbf{A}}{\partial v} \mathbf{P} dt \right) + \mu_\varepsilon \left(\frac{\partial E}{\partial \varepsilon} - \int_0^T \lambda^T \frac{\partial \mathbf{A}}{\partial \varepsilon} \mathbf{P} dt \right) \right. \\ & \left. + \mu_\delta \left(\frac{\partial E}{\partial \delta} - \int_0^T \lambda^T \frac{\partial \mathbf{A}}{\partial \delta} \mathbf{P} dt \right) \right) dx dz \\ & + \int_{(x,z) \in H} \int_0^T \kappa^T (\mathbf{A} \mathbf{P} - \mathbf{s}) dt dx dz + \int_{(x,z) \in H} \int_0^T \Delta \mathbf{P}^T (\mathbf{A}^T \lambda + \mathbf{R} \mathbf{P} \\ & - \mathbf{P}_{\text{obs}}) dt dx dz \end{aligned} \quad (33)$$

where μ_m , $m \in \{v, \varepsilon, \delta\}$, and $\Delta \mathbf{P} = [\Delta p \ \Delta q]^T$ and $\kappa = [\kappa_1 \ \kappa_2]^T$ are the Lagrange multiplier functions. Eq. (33) can be simplified as

$$\begin{aligned} G = & \int_{(x,z) \in H} \left((x_v + \mu_v) \frac{\partial E}{\partial v} + (x_\varepsilon + \mu_\varepsilon) \frac{\partial E}{\partial \varepsilon} + (x_\delta + \mu_\delta) \frac{\partial E}{\partial \delta} \right) dx dz \\ & + \int_{(x,z) \in H} \int_0^T \left(\kappa^T (\mathbf{A} \mathbf{P} - \mathbf{s}) + \Delta \mathbf{P}^T (\mathbf{A}^T \lambda + \mathbf{R} \mathbf{P} - \mathbf{P}_{\text{obs}}) \right. \\ & \left. - \lambda^T \left(\mu_v \frac{\partial \mathbf{A}}{\partial v} + \mu_\varepsilon \frac{\partial \mathbf{A}}{\partial \varepsilon} + \mu_\delta \frac{\partial \mathbf{A}}{\partial \delta} \right) \mathbf{P} \right) dt dx dz \end{aligned} \quad (34)$$

Setting

$$\mu_m = -x_m, m \in \{v, \varepsilon, \delta\} \quad (35)$$

we have

$$G = \int_{(x,z) \in H} \int_0^T \left(\kappa^T (\mathbf{A} \mathbf{P} - \mathbf{s}) + \Delta \mathbf{P}^T (\mathbf{A}^T \lambda + \mathbf{R} \mathbf{P} - \mathbf{P}_{\text{obs}}) \right. \\ \left. - \lambda^T \left(\mu_v \frac{\partial \mathbf{A}}{\partial v} + \mu_\varepsilon \frac{\partial \mathbf{A}}{\partial \varepsilon} + \mu_\delta \frac{\partial \mathbf{A}}{\partial \delta} \right) \mathbf{P} \right) dt dx dz \quad (36)$$

Integrating Eq. (36) by parts results in

$$G = \int_{(x,z) \in H} \int_0^T \left(\kappa^T (\mathbf{A} \mathbf{P} - \mathbf{s}) + \lambda^T \mathbf{A} \Delta \mathbf{P} + \Delta \mathbf{P}^T (\mathbf{R} \mathbf{P} - \mathbf{P}_{\text{obs}}) \right. \\ \left. - \lambda^T \left(\mu_v \frac{\partial \mathbf{A}}{\partial v} + \mu_\varepsilon \frac{\partial \mathbf{A}}{\partial \varepsilon} + \mu_\delta \frac{\partial \mathbf{A}}{\partial \delta} \right) \mathbf{P} \right) dt dx dz \quad (37)$$

or

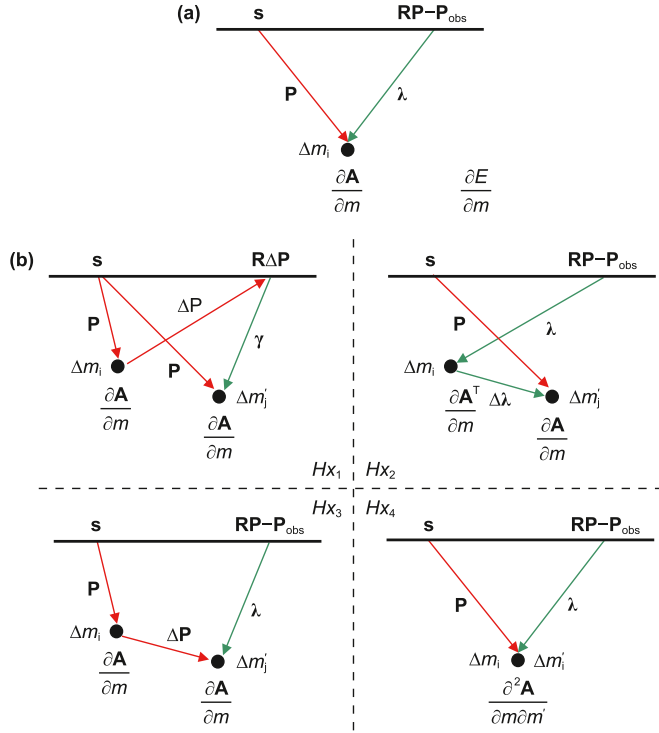


Fig. 5. Schematic diagrams of the computation of (a) functional gradients and (b) Hessian-vector products.

$$G = \int_{(x,z) \in H} \int_0^T \left(\mathbf{P}^T \mathbf{A}^T \kappa - \kappa^T \mathbf{s} + \Delta \mathbf{P}^T (\mathbf{A}^T \lambda - \mathbf{P}_{\text{obs}}) + \mathbf{P}^T \mathbf{R} \Delta \mathbf{P} \right. \\ \left. - \mathbf{P}^T \left(\mu_v \frac{\partial \mathbf{A}^T}{\partial v} + \mu_\varepsilon \frac{\partial \mathbf{A}^T}{\partial \varepsilon} + \mu_\delta \frac{\partial \mathbf{A}^T}{\partial \delta} \right) \lambda \right) dt dx dz \quad (38)$$

where $\partial \mathbf{A}^T / \partial m = (\partial \mathbf{A} / \partial m)^T$, $m \in \{v, \varepsilon, \delta\}$. Setting $\partial G / \partial \lambda = \mathbf{0}$ or $\partial G / \partial \mathbf{P} = \mathbf{0}$ gives

$$\mathbf{A} \Delta \mathbf{P} = \left(\mu_v \frac{\partial \mathbf{A}}{\partial v} + \mu_\varepsilon \frac{\partial \mathbf{A}}{\partial \varepsilon} + \mu_\delta \frac{\partial \mathbf{A}}{\partial \delta} \right) \mathbf{P} \quad (39)$$

$$\mathbf{A}^T \kappa = \left(\mu_v \frac{\partial \mathbf{A}^T}{\partial v} + \mu_\varepsilon \frac{\partial \mathbf{A}^T}{\partial \varepsilon} + \mu_\delta \frac{\partial \mathbf{A}^T}{\partial \delta} \right) \lambda - \mathbf{R} \Delta \mathbf{P} \quad (40)$$

Eqs. (39) and (40) can be regarded as the perturbed forward and adjoint equations and are still solved by FD and the fast Poisson solver (Strang, 2007). The gradients of the misfit function F with respect to model parameters are calculated by the chain rule $\frac{\partial F}{\partial m} = \frac{\partial G}{\partial m} = \frac{\partial G}{\partial \mathbf{P}} \frac{\partial \mathbf{P}}{\partial m} + \frac{\partial G}{\partial \lambda} \frac{\partial \lambda}{\partial m} + \frac{\partial G}{\partial m}$, $m \in \{v, \varepsilon, \delta\}$.

$$\frac{\partial F}{\partial m} = \int_0^T \kappa^T \frac{\partial \mathbf{A}}{\partial m} \mathbf{P} dt + \int_0^T \lambda^T \frac{\partial \mathbf{A}}{\partial m} \Delta \mathbf{P} dt$$

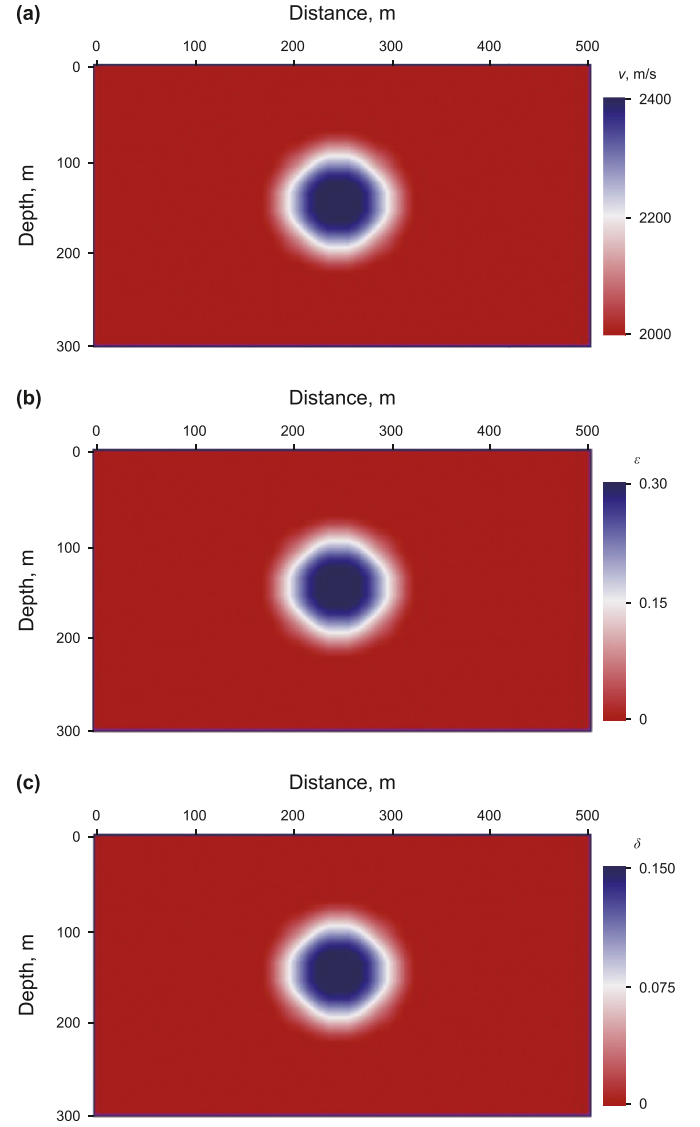


Fig. 6. A VTI model with correlated (a) v , (b) ε and (c) δ structures.

$$- \int_0^T \lambda^T \left(\mu_v \frac{\partial^2 \mathbf{A}}{\partial v \partial m} + \mu_\varepsilon \frac{\partial^2 \mathbf{A}}{\partial \varepsilon \partial m} + \mu_\delta \frac{\partial^2 \mathbf{A}}{\partial \delta \partial m} \right) \mathbf{P} dt \quad (41)$$

where $\partial^2 \mathbf{A} / (\partial m \partial m') = \partial^2 \mathbf{A} / (\partial m' \partial m)$, $m, m' \in \{v, \varepsilon, \delta\}$,

$$\frac{\partial^2 \mathbf{A}}{\partial v^2} = \begin{bmatrix} -2(1+2\varepsilon) \frac{\partial^2}{\partial x^2} - 2 \frac{\partial^2}{\partial z^2} & -4(\sqrt{\varepsilon^2 + 2\delta + 1} - \varepsilon - 1) \frac{\partial^4}{\partial x^2 \partial z^2} \\ 0 & 0 \end{bmatrix} \quad (42)$$

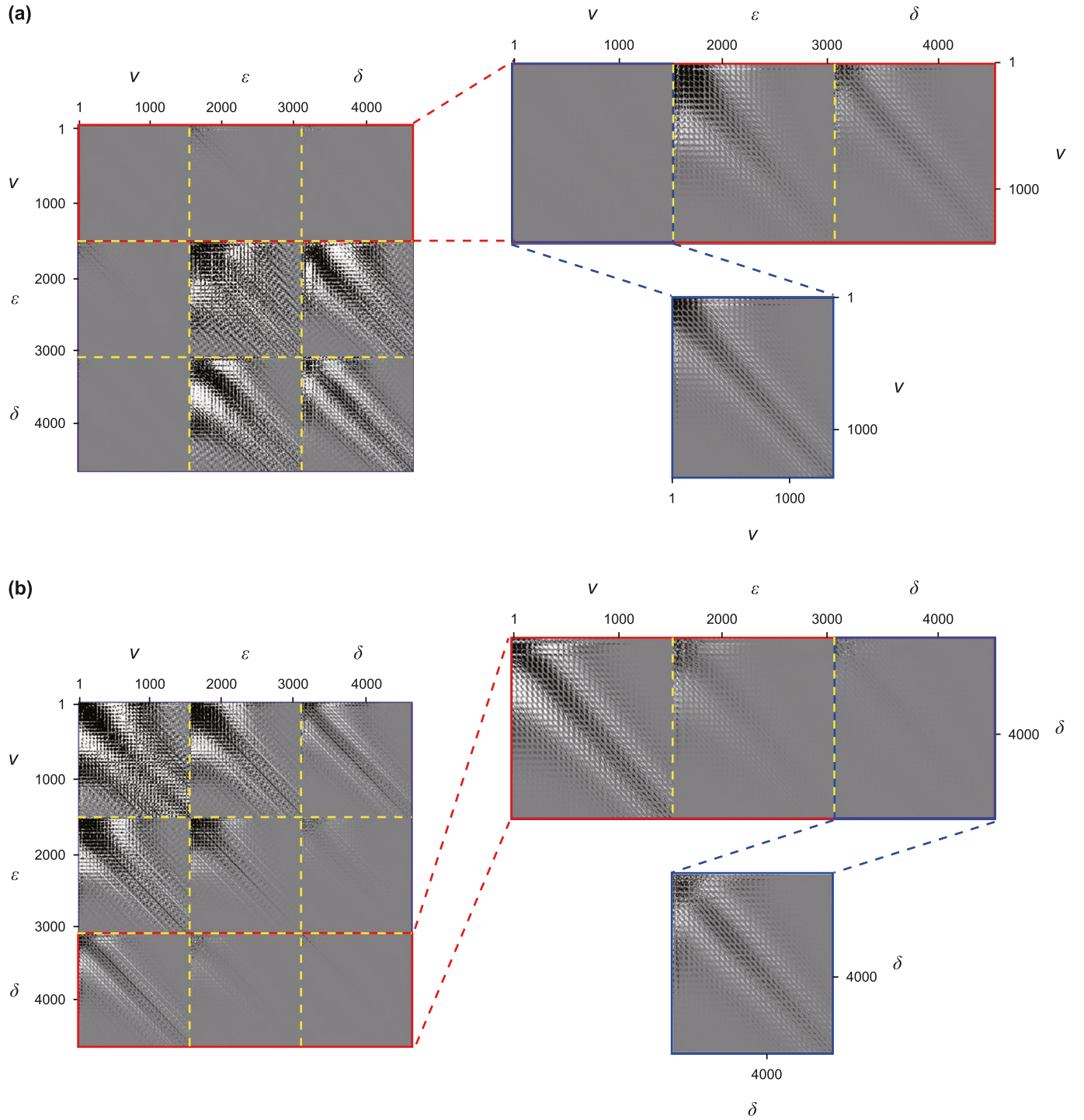


Fig. 7. Plots of the approximate Hessian **(a)** before and **(b)** after the velocity is adimensionalized for the VTI model with correlated v , ε and δ structures.

$$\frac{\partial^2 \mathbf{A}}{\partial v \partial \varepsilon} = \begin{bmatrix} -4v \frac{\partial^2}{\partial x^2} & -4v \left(\frac{\varepsilon}{\sqrt{\varepsilon^2 + 2\delta + 1}} - 1 \right) \frac{\partial^4}{\partial x^2 \partial z^2} \\ 0 & 0 \end{bmatrix} \quad (43)$$

$$\frac{\partial^2 \mathbf{A}}{\partial v \partial \delta} = \begin{bmatrix} 0 & \frac{-4v}{\sqrt{\varepsilon^2 + 2\delta + 1}} \frac{\partial^4}{\partial x^2 \partial z^2} \\ 0 & 0 \end{bmatrix} \quad (44)$$

$$\frac{\partial^2 \mathbf{A}}{\partial \varepsilon^2} = \begin{bmatrix} 0 & \frac{-2v^2(2\delta+1)}{(\sqrt{\varepsilon^2+2\delta+1})^3} \frac{\partial^4}{\partial x^2 \partial z^2} \\ 0 & 0 \end{bmatrix} \quad (45)$$

$$\frac{\partial^2 \mathbf{A}}{\partial \varepsilon \partial \delta} = \begin{bmatrix} 0 & \frac{2v^2 \varepsilon}{(\sqrt{\varepsilon^2+2\delta+1})^3} \frac{\partial^4}{\partial x^2 \partial z^2} \\ 0 & 0 \end{bmatrix} \quad (46)$$

$$\frac{\partial^2 \mathbf{A}}{\partial \delta^2} = \begin{bmatrix} 0 & \frac{2v^2}{(\sqrt{\varepsilon^2+2\delta+1})^3} \frac{\partial^4}{\partial x^2 \partial z^2} \\ 0 & 0 \end{bmatrix} \quad (47)$$

Based on Eqs. (14), (18), (35) and (39)–(41), we can obtain the gradients of F with respect to model parameters, which are the solution to the Hessian-vector products for the continuous problem. The discrete Hessian-vector products ($\mathbf{H}\mathbf{x}$) are estimated by numerically solving Eqs. (14), (18), (39) and (40) and calculating the gradients for each grid point.

Eq. (40) can be split into two sub-equations:

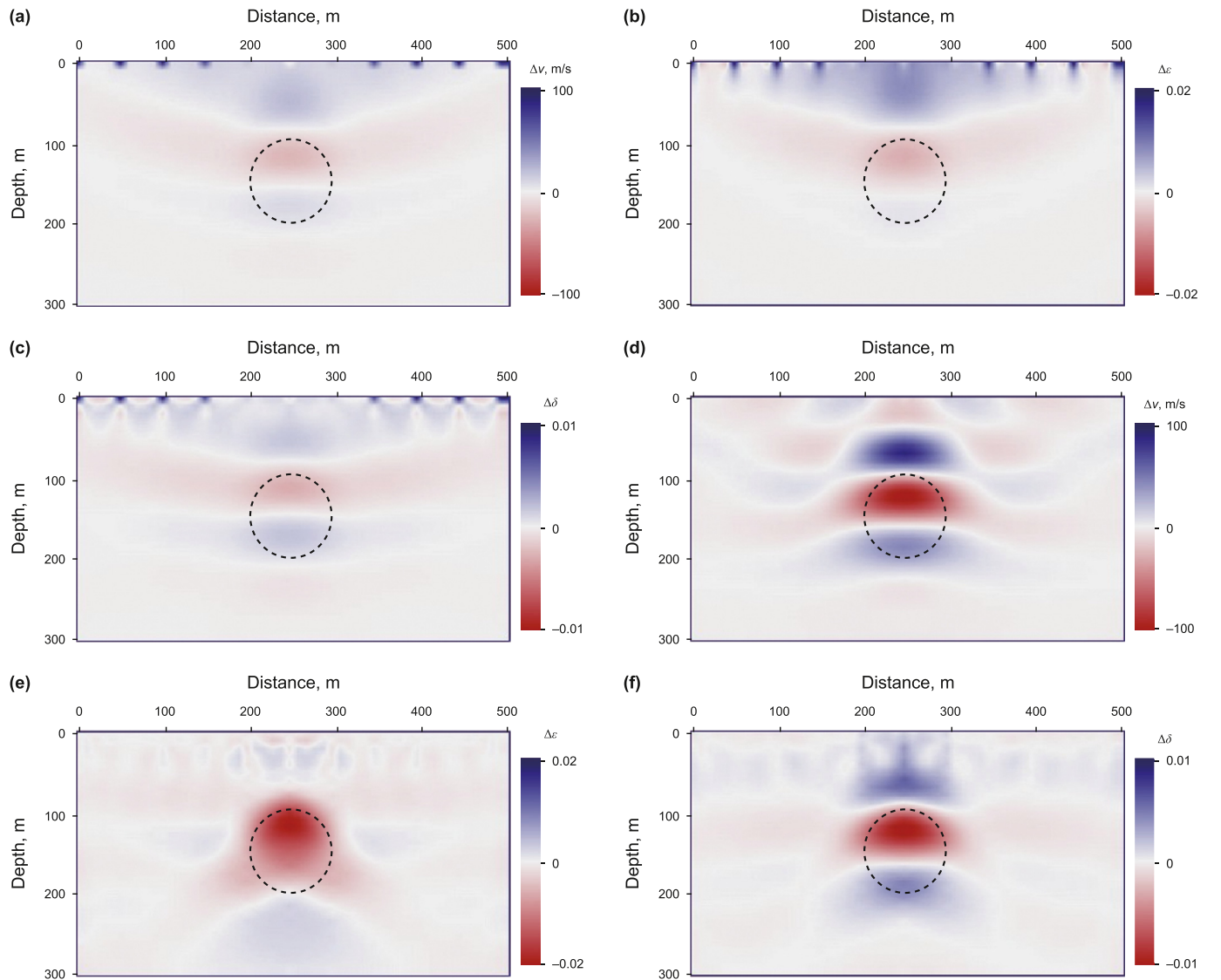


Fig. 8. FWI gradients (a)–(c) before and (d)–(f) after the inverse approximate Hessian preconditioning for the VTI model with correlated v , ε and δ structures. Gradients with respect to (a) and (d) v , (b) and (e) ε and (c) and (f) δ .

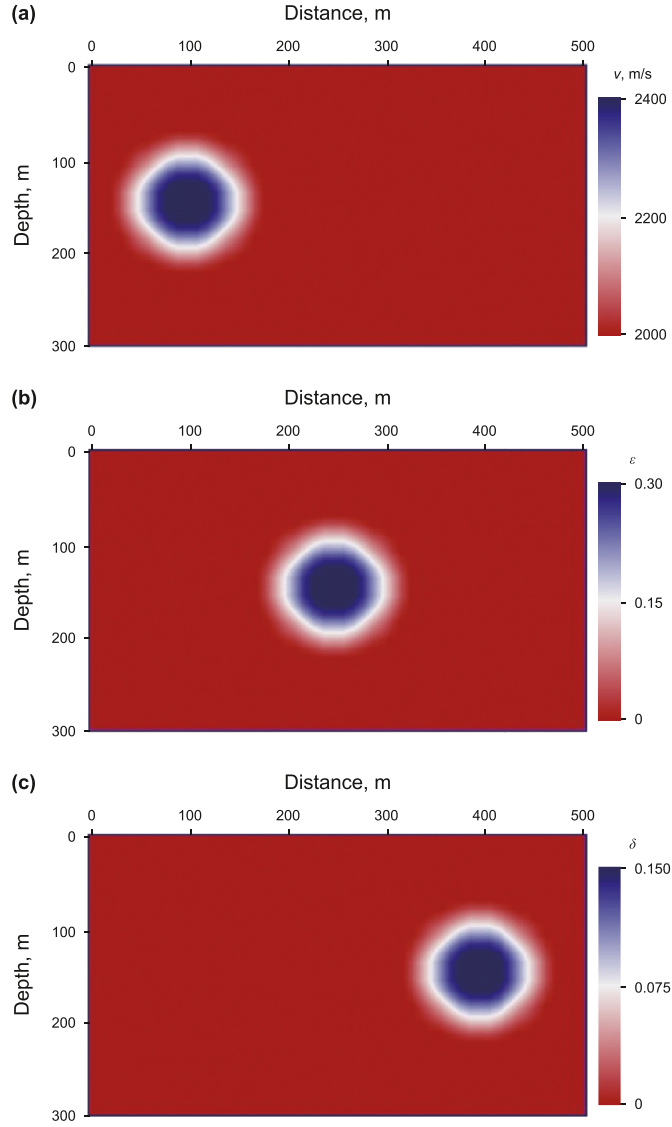


Fig. 9. A VTI model with uncorrelated (a) v , (b) ϵ and (c) δ structures.

$$\mathbf{A}^T \Delta \lambda = \left(\mu_v \frac{\partial \mathbf{A}^T}{\partial v} + \mu_\epsilon \frac{\partial \mathbf{A}^T}{\partial \epsilon} + \mu_\delta \frac{\partial \mathbf{A}^T}{\partial \delta} \right) \lambda \quad (48)$$

$$\mathbf{A}^T \gamma = -\mathbf{R} \Delta \mathbf{P} \quad (49)$$

where $\kappa = \Delta \lambda + \gamma$, $\Delta \lambda = [\Delta \lambda_1 \ \Delta \lambda_2]^T$ and $\Delta \lambda = [\gamma_1 \ \gamma_2]^T$. Eq. (41) is rewritten as a summation of four terms

$$\frac{\partial F}{\partial m} = Hx_1 + Hx_2 + Hx_3 + Hx_4 \quad (50)$$

where

$$Hx_1 = \int_0^T \gamma^T \frac{\partial \mathbf{A}}{\partial m} \mathbf{P} dt \quad (51)$$

$$Hx_2 = \int_0^T \Delta \lambda^T \frac{\partial \mathbf{A}}{\partial m} \mathbf{P} dt \quad (52)$$

$$Hx_3 = \int_0^T \lambda^T \frac{\partial \mathbf{A}}{\partial m} \Delta \mathbf{P} dt \quad (53)$$

$$Hx_4 = - \int_0^T \lambda^T \left(\mu_v \frac{\partial^2 \mathbf{A}}{\partial v \partial m} + \mu_\epsilon \frac{\partial^2 \mathbf{A}}{\partial \epsilon \partial m} + \mu_\delta \frac{\partial^2 \mathbf{A}}{\partial \delta \partial m} \right) \mathbf{P} dt \quad (54)$$

Fig. 5 illustrates the physical interpretation of different terms in Eq. (50). It is seen that the first term and the latter three terms are associated with the first- and second-order scattering effect, respectively. The full Hessian matrix $\mathbf{H}$ is composed of the Gauss-Newton term $\mathbf{B}$ and the second-order term $\Delta \mathbf{H}$. The TGN method only takes the term $\mathbf{B}$, determined by the first-order derivatives of the wavefield with respect to model parameters (the Jacobian matrix), into account. The Gauss-Newton Hessian-vector products ($\mathbf{B}\mathbf{x}$) can be obtained by ignoring the contribution associated with the second-order scattering. As a result, Eq. (50) becomes

$$\frac{\partial F}{\partial m} = Hx_1 = \int_0^T \gamma^T \frac{\partial \mathbf{A}}{\partial m} \mathbf{P} dt \quad (55)$$

The TGN and TN methods require 3 (Eqs. (14), (39) and (49)) and 4 (Eqs. (14), (18), (39) and (40)) wave propagation simulation, respectively, to compute the Hessian-vector products. The second-order term makes the Hessian operator symmetric indefinite and a negative curvature might be detected. The Gauss-Newton term is symmetric positive definite. Therefore, the Gauss-Newton method generally exhibits better stability. The second-order term of the Hessian matrix plays a key role in the inversion of multiples (prismatic waves, Liu et al., 2020). However, this is beyond the scope of this paper. We utilize the TGN method in FWI of pure qP-waves for the sake of efficiency and stability.

The Gauss-Newton linear system is written as

$$\mathbf{B}_k \Delta \mathbf{m}_k = -\mathbf{grad}_k \quad (56)$$

which can be solved stably and efficiently by a matrix-free conjugate gradient method based on the computation of Hessian-vector products ($\mathbf{B}\mathbf{x}$). A simple preconditioner $\mathbf{Q}$ is used to accelerate the convergence of the inner iterations. Eq. (56) becomes

$$\mathbf{Q}_k \mathbf{B}_k \Delta \mathbf{m}_k = -\mathbf{Q}_k \mathbf{grad}_k \quad (57)$$

where

$$\mathbf{Q}_k = \begin{bmatrix} s_v \bar{\mathbf{B}}_v & & \\ & s_\epsilon \bar{\mathbf{B}}_\epsilon & \\ & & s_\delta \bar{\mathbf{B}}_\delta \end{bmatrix}^{-1} \quad (58)$$

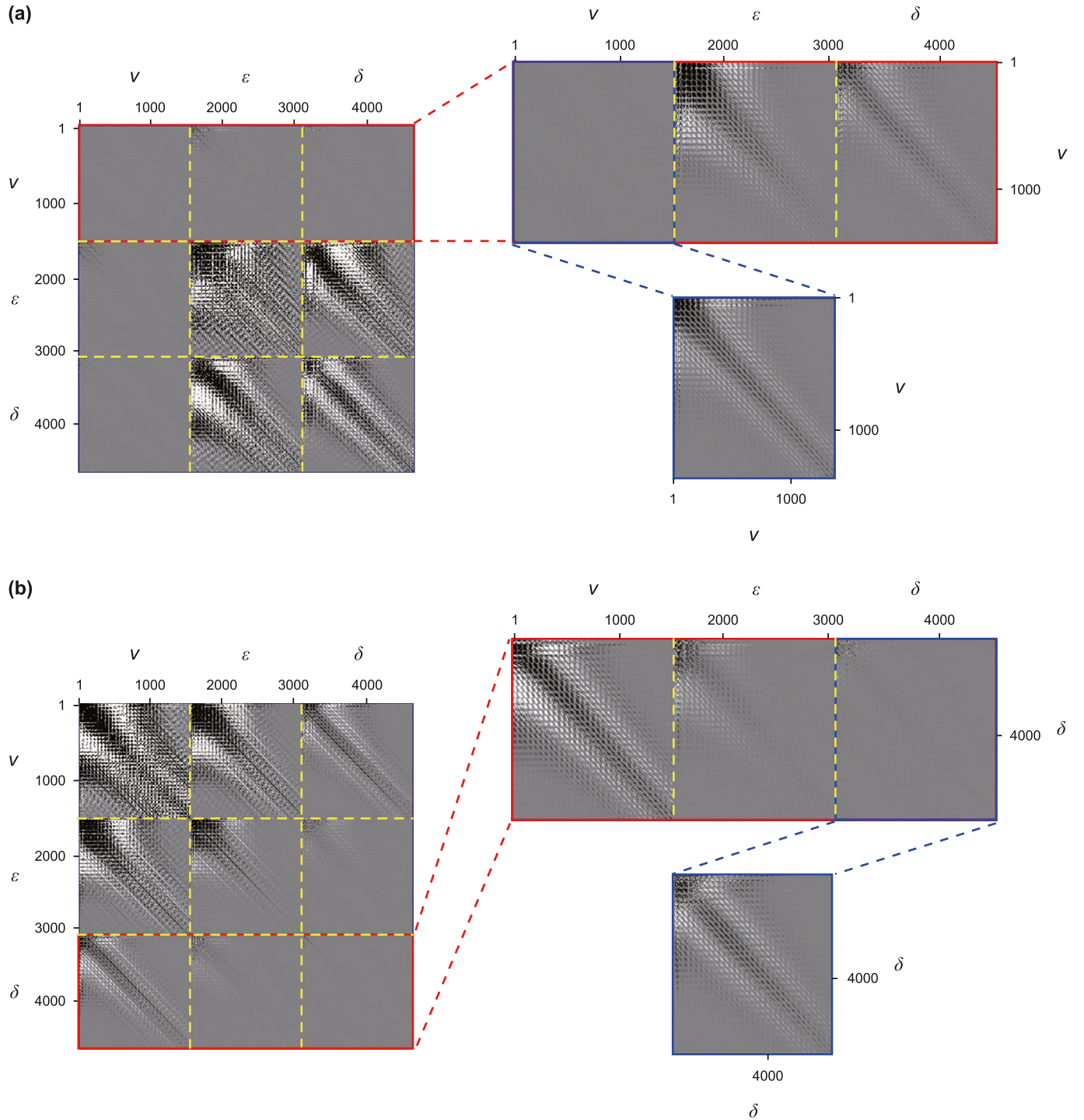


Fig. 10. Plots of the approximate Hessian (a) before and (b) after the velocity is adimensionalized for the VTI model with uncorrelated v , ε and δ structures.

$$\bar{\mathbf{B}}_m = \begin{bmatrix} \bar{B}_{m_1} & & \\ & \bar{B}_{m_2} & \\ & & \ddots \\ & & & \bar{B}_{m_N} \end{bmatrix}, \quad (59)$$

$$\bar{B}_{m_i} = \int_0^T \left(\frac{\partial \mathbf{A}}{\partial m} \mathbf{P} \right)^2 dt, m \in \{v, \varepsilon, \delta\}, i = 1, 2, \dots, N$$

s_m , $m \in \{v, \varepsilon, \delta\}$ is the scaling factor associated with the

sensitivity of seismic data to different model parameters (Yang et al., 2018; Yong et al., 2022). The Eisenstat and Walker stopping criterion is applied to determine the number of the conjugate gradient iterations (Eisenstat and Walker, 1996; Métivier et al., 2013)

$$\|\mathbf{B}_k \Delta \mathbf{m}_k + \mathbf{grad}_k\|_2 \leq \eta_k \|\mathbf{grad}_k\|_2 \quad (60)$$

where $\|\cdot\|_2$ denotes the L_2 norm, η is the forcing term and can be calculated by

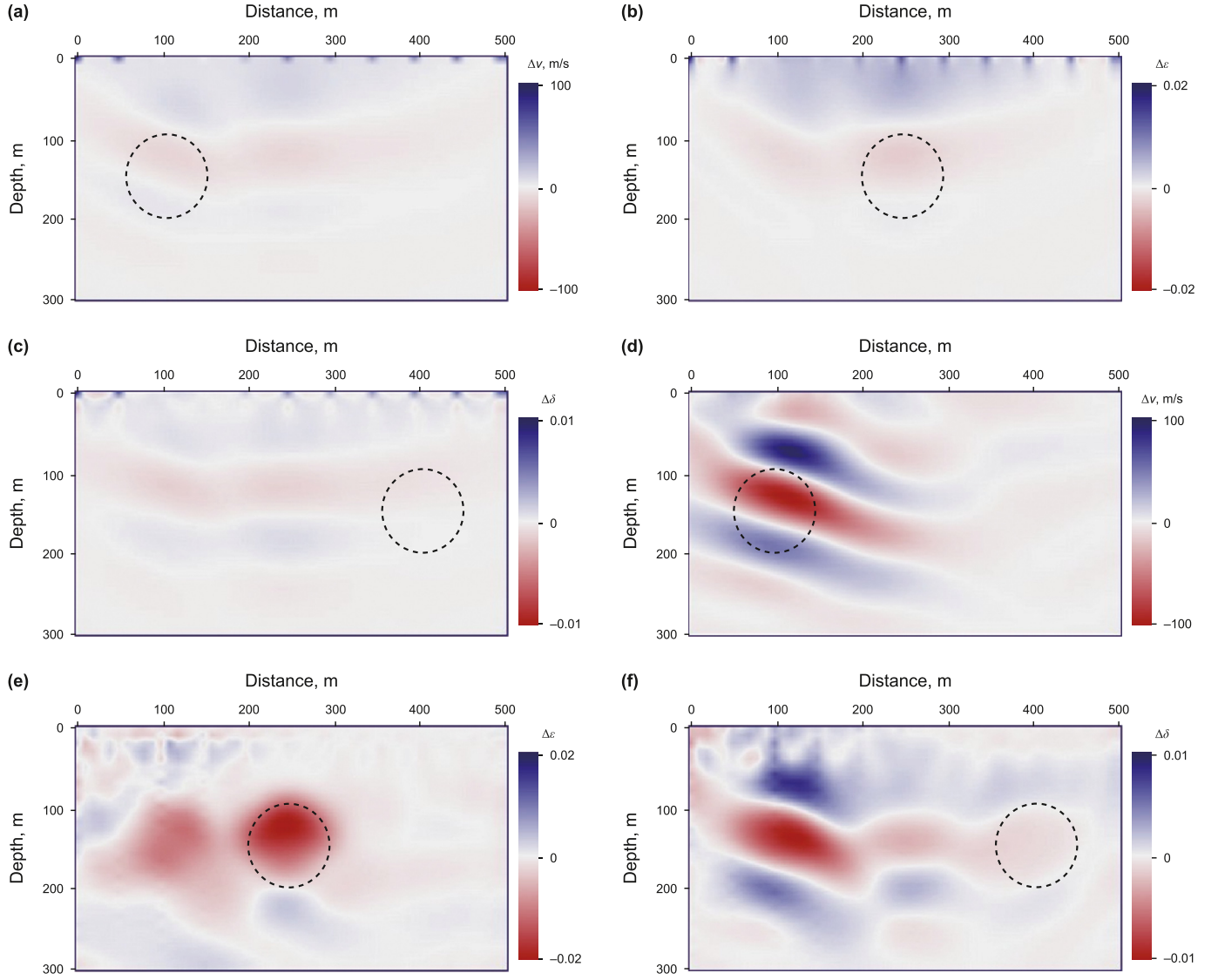


Fig. 11. FWI gradients (a)–(c) before and (d)–(f) after the inverse approximate Hessian preconditioning for the VTI model with uncorrelated v , ϵ and δ structures. Gradients with respect to (a) and (d) v , (b) and (e) ϵ and (c) and (f) δ .

$$\eta_k = \frac{\|\mathbf{grad}_k - \mathbf{grad}_{k-1} - \alpha_{k-1} \mathbf{B}_{k-1} \Delta \mathbf{m}_{k-1}\|_2}{\|\mathbf{grad}_{k-1}\|_2} \quad (61)$$

where α represents the step length for the outer iterations and can be estimated by the parabolic fitting method. We have

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha_k \Delta \mathbf{m}_k \quad (62)$$

Two additional safeguards are used to complement the Eisenstat and Walker stopping criterion:

$$\text{If } \eta_{k-1}^{(1+\sqrt{5})/2} > 0.1, \text{ then } \eta_k = \max\left\{\eta_k, \eta_{k-1}^{(1+\sqrt{5})/2}\right\} \quad (63)$$

$$\text{If } \eta_k > 1, \text{ then } \eta_k = 0.9 \quad (64)$$

We conduct the TGN method in a nested way: a linear inner loop (Eq. (57)) and a nonlinear outer loop (Eq. (62)). The formulation of TGN FWI of pure qP-waves in VTI media is summarized in Algorithm 1.

Algorithm 1. Truncated Gauss-Newton FWI of pure qP-waves in VTI media.

Input: initial model $\mathbf{m}_0$, observed data $\mathbf{P}_{\text{obs}}$, tolerance error e , maximum iteration numbers for the inner and outer loops k_i and k_o ;

Output: $\min_{\mathbf{m}} E(\mathbf{m})$;

While $E(\mathbf{m}) > e$ and the iteration number for the outer loop is less than k_o **do**

 compute the gradients $\mathbf{grad}$ by Eqs. (14), (18) and (19);

 compute the preconditioner $\mathbf{Q}$ by Eq. (58);

 set $\Delta\mathbf{m} = \mathbf{0}$, $\eta = 0.9$, $\mathbf{g} = \mathbf{grad}$, $\mathbf{p} = \mathbf{Qg}$, and $\mathbf{x} = -\mathbf{p}$;

While $\|\mathbf{B}\Delta\mathbf{m} + \mathbf{grad}\|_2 > \eta\|\mathbf{grad}\|_2$ and the iteration number for the inner loop is less than k_i **do**

 compute the Gauss-Newton Hessian-vector products $(\mathbf{Bx})$ by Eqs. (14), (35), (39), (49) and (55);

 compute the curvature by $\beta_1 = \mathbf{x}^T \mathbf{Bx}$;

if $\beta_1 < 0$ **then**

 stop the inner iterations;

else

$\beta_2 = \mathbf{g}^T \mathbf{p}$;

$\Delta\mathbf{m} = \Delta\mathbf{m} + \frac{\beta_2}{\beta_1} \mathbf{x}$;

$\mathbf{g} = \mathbf{g} + \frac{\beta_2}{\beta_1} \mathbf{Bx}$;

$\mathbf{p} = \mathbf{Qg}$;

$\mathbf{x} = -\mathbf{p} + \frac{\mathbf{g}^T \mathbf{p}}{\beta_2} \mathbf{x}$;

end if

end while

 compute the step length α by the parabolic fitting method;

 update the model by $\mathbf{m} = \mathbf{m} + \alpha\Delta\mathbf{m}$;

 update η with the Eisenstat and Walker forcing-term formula (Eqs. (61), (63) and (64));

end while

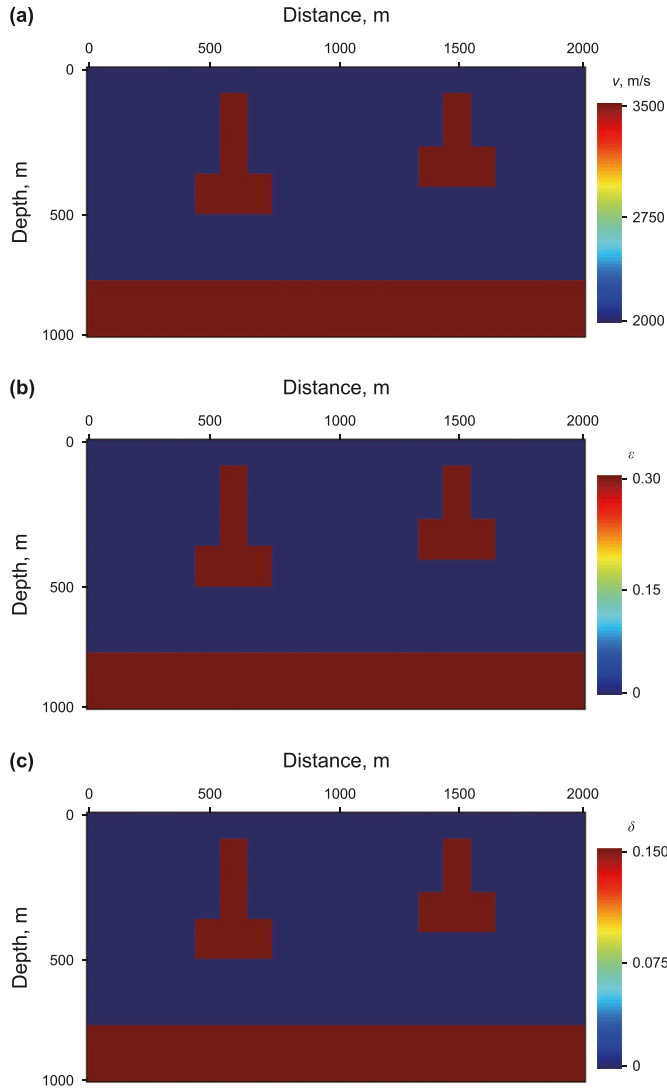


Fig. 12. A simple VTI model with correlated v , ε and δ structures. (a) v , (b) ε and (c) δ .

2.3. Interparameter crosstalk

We use the Gauss-Newton term $\mathbf{B}$ of the Hessian matrix to analyze the parameter crosstalk for small-scale models. The approximate Hessian operator is constructed by repeatedly sampling the rows or columns of the Hessian-vector products ($\mathbf{B}\mathbf{x}$) with an input vector as a Dirac delta function. The Hessian operator is ill-conditioned and the singular value decomposition (SVD) is utilized to compute an approximate inverse Hessian. To ensure the stability, the contribution associated with a number of small singular values is neglected. Fig. 6 depicts a model with correlated v , ε and δ structures. The grid dimensions are 51×31 , the grid spacing is 10 m, the time step is 1.0 ms, and the recording time is 0.5 s. A 10 Hz Ricker wavelet is used as a source. Eleven sources and 51 receivers are evenly distributed on the surface with the interval of 50 m and

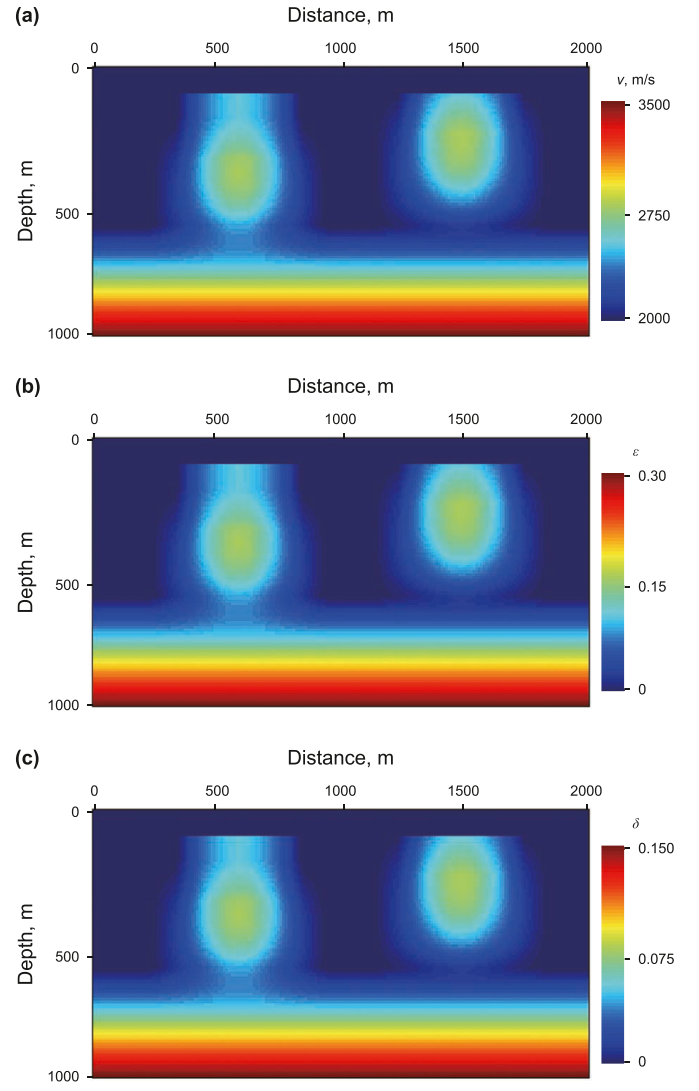


Fig. 13. The initial (a) v , (b) ε and (c) δ models for the simple VTI model with correlated v , ε and δ structures.

10 m, respectively. The initial model is a homogeneous model without a parameter anomaly. To fairly analyze the influence of model parameters on seismic response, the Hessian is built when the velocity is adimensionalized (Operto et al., 2013). Fig. 7 displays the Gauss-Newton Hessian operator, which is composed of nine blocks. The diagonal blocks show the amplitude effect and the spatial resolution with respect to model parameters of the same nature, whereas the off-diagonal blocks describe the tradeoff between the parameters of different nature. It is observed that FWI of pure qP-waves suffers severely from the parameter crosstalk. Compared to the anisotropy parameters, the velocity is more sensitive to the data residuals and can be reliably inverted for. The functional gradients for the first iteration, illustrated in Fig. 8, contain plenty of acquisition artifacts and blurring effects caused by geometry spread, incomplete illumination and band-limited data.

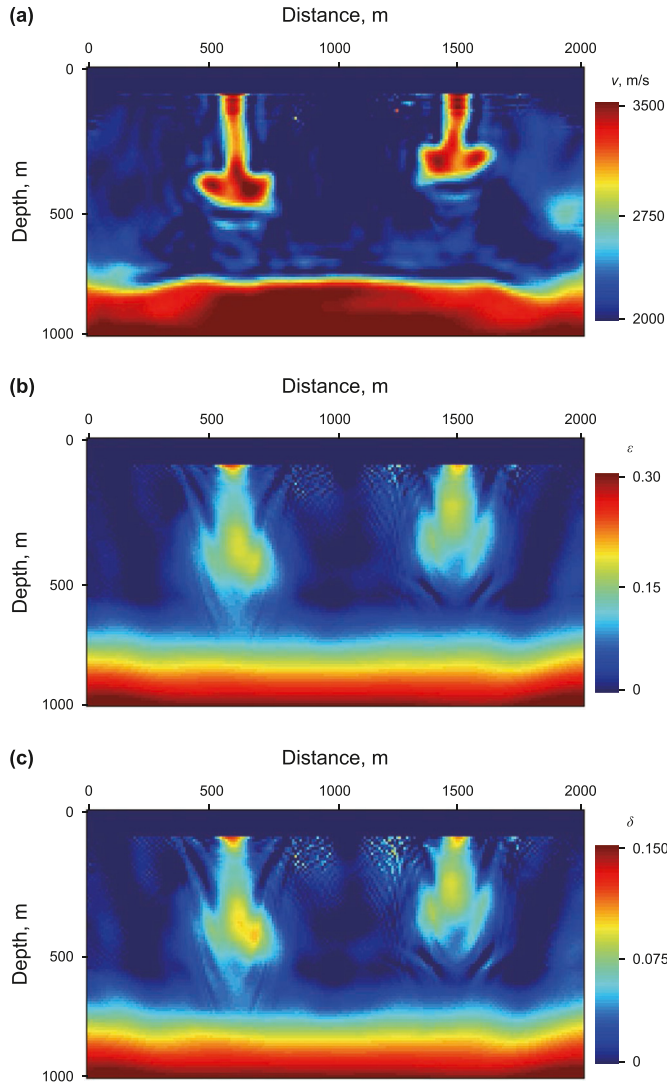


Fig. 14. The inverted (a) v , (b) ε and (c) δ models by the pseudoacoustic FWI for the simple VTI model with correlated v , ε and δ structures.

By multiplying the inverse approximate Hessian to the gradients, the acquisition artifacts and blurring effects are greatly suppressed. The gradient with respect to ε is much smoother than that with respect to v and δ , suggesting that the long-wavelength component of the ε model can be retrieved. A model with uncorrelated v , ε and δ structures, shown in Fig. 9, is used to further evaluate the parameter tradeoff effect. Fig. 10 displays the Gauss-Newton term of the Hessian operator. Note that the energy in the diagonal blocks decreases from shallow to deep and smears over a wide area around the diagonal, degrading the accuracy and resolution of the inversion. The nonzero elements in the off-diagonal blocks impact the parameter crosstalk. Fig. 11 depicts the gradients with respect to model parameters for the first iteration before and after the inverse

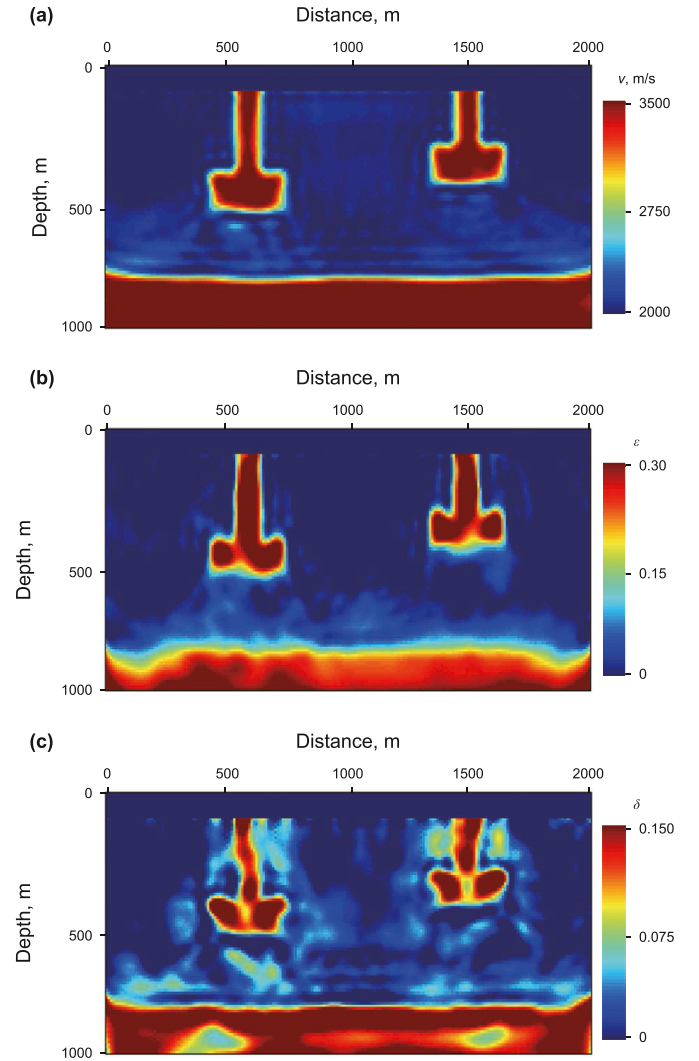


Fig. 15. The inverted (a) v , (b) ε and (c) δ models by the pure acoustic FWI for the simple VTI model with correlated v , ε and δ structures.

approximate Hessian preconditioning. It is observed that the velocity is a dominant parameter and the gradients with respect to anisotropy parameters are contaminated by the crosstalk produced by the velocity anomaly. The application of the inverse approximate Hessian can focus and deblur the gradients and reduce the ambiguity produced by the parameter crosstalk. The tradeoff effect cannot be totally eliminated because of the truncation of the SVD operator for small singular values in the computation of the inverse Hessian operator.

2.4. Computational cost analysis

We implement TGN FWI of pure qP-waves in VTI media in the time domain. The boundary-value source wavefield reconstruction

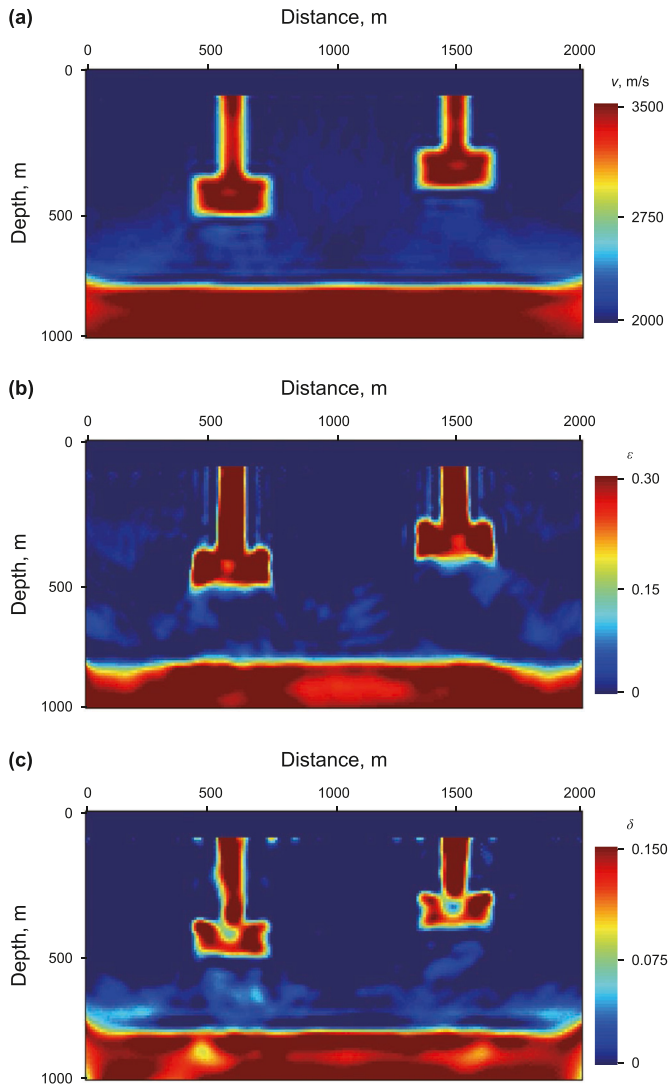


Fig. 16. The inverted (a) v , (b) ϵ and (c) δ models by the pure acoustic TGN FWI for the simple VTI model with correlated v , ϵ and δ structures.

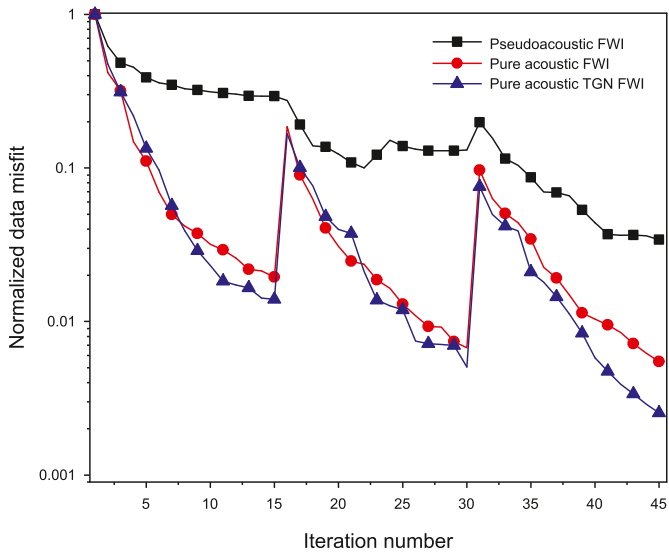


Fig. 17. Convergence curves of different FWI methods for the simple VTI model with correlated v , ϵ and δ structures.

Table 1
The RMS error between the real and initial/inverted models for the simple VTI model.

Parameter	Initial model, m/s	Pseudoacoustic FWI, m/s	Pure acoustic FWI, m/s	Pure acoustic TGN FWI, m/s
v	387.4603	269.9959	216.2441	179.3746
ϵ	0.0775	0.0725	0.0654	0.0591
δ	0.0387	0.0349	0.0308	0.0269

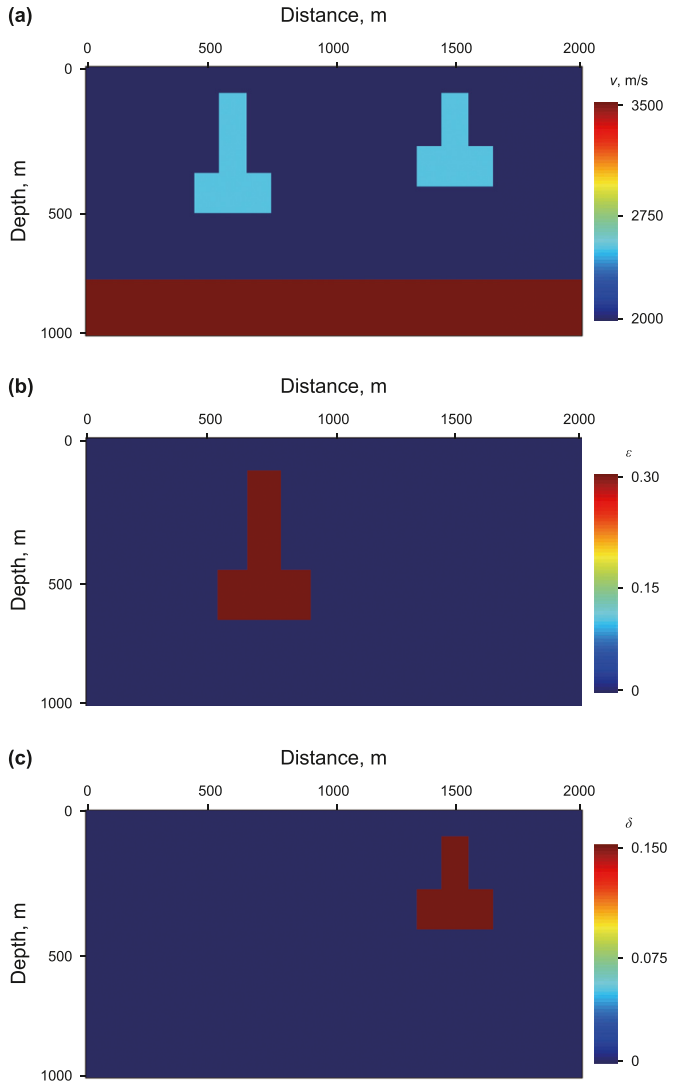


Fig. 18. A simple VTI model with uncorrelated v , ϵ and δ structures. (a) v , (b) ϵ and (c) δ .

method is used to reduce the storage requirements (Dussaud et al., 2008). Five forward computations, the propagation of the source and adjoint wavefield (Eq. (14) and (18)), the reconstruction of the source wavefield and two additional wave propagation to estimate the iteration step, are needed in each outer iteration. Four forward computations, the propagation of the source wavefield (Eq. (14)), the propagation of the perturbed source and adjoint wavefield (Eqs. (39) and (49)) and the reconstruction of the source wavefield, are

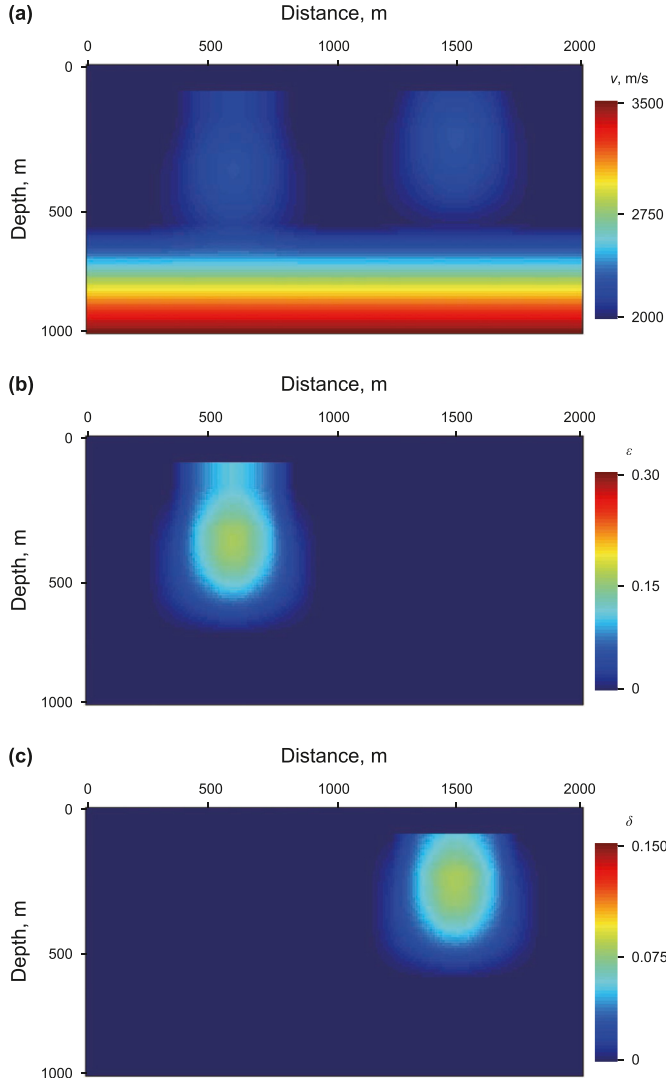


Fig. 19. The initial (a) v , (b) ϵ and (c) δ models for the simple VTI model with uncorrelated v , ϵ and δ structures.

needed in each inner iteration. The computational cost of the TGN method in terms of wave propagation simulation is given by

$$C = (4k_i + 5)k_o \quad (65)$$

where k_i and k_o are the maximum number of iterations for the inner and outer loops, respectively. In contrast, the L-BFGS method require $5k_o$ forward computations. The TGN method is expensive because of the inter loop involved. A suitable pre-conditioner (Eq. (58)) and the Eisenstat and Walker stopping criterion (Eqs. (61), (63) and (64)) are adopted to accelerate the convergence. Meanwhile, the inner iterations are automatically terminated when a negative curvature is detected. We set $k_i = 5$ in the following TGN FWI examples to avoid prohibitive computational cost. The efficiency can be further improved by the source encoding or shot or model subsampling strategy (Castellanos et al., 2015; Matharu and Sacchi, 2019; Wang et al., 2022).

3. Examples

We use synthetic and field data to test the proposed method. The pseudoacoustic FWI, the pure acoustic FWI, and the pure

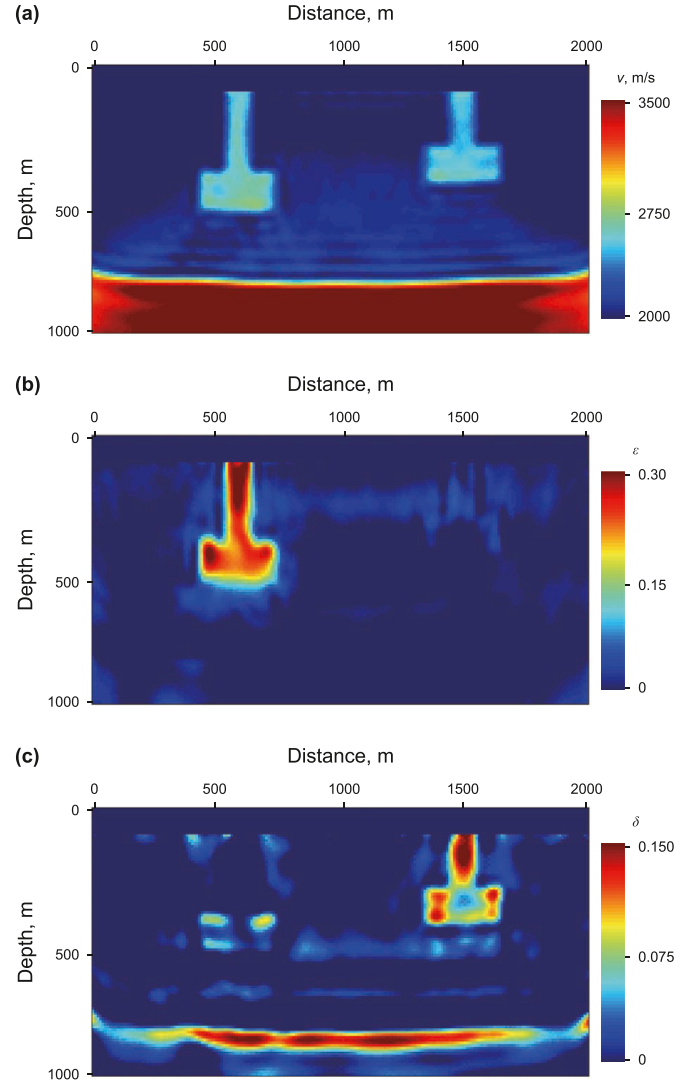


Fig. 20. The inverted (a) v , (b) ϵ and (c) δ models by the pure acoustic FWI for the simple VTI model with uncorrelated v , ϵ and δ structures.

acoustic TGN FWI are compared. We apply the pseudoacoustic solver in Chu et al. (2011) and impose a constraint condition

$$\eta = \frac{\epsilon - \delta}{1 + 2\delta} \geq 0 \quad (66)$$

to guarantee the stability of the pseudoacoustic FWI. The L-BFGS method (Nocedal and Wright, 1999) is adopted to update velocity and anisotropy parameter models in the pseudoacoustic and pure acoustic FWI.

3.1. A simple model example

Fig. 12 illustrates a simple VTI model with correlated v , ϵ and δ structures. The time step is 1.0 ms, the recording time is 2.0 s, the grid spacing is 10 m, and the grid dimensions are 201×101 . A Ricker wavelet with the peak frequency of 10 Hz is used as a source. Twenty-one sources and 201 receivers are evenly distributed on the surface. The initial model, shown in Fig. 13, is a Gaussian smoothed version of the real model. For each method, the inversion is performed at three scales and the cut-off frequencies for coarse, medium and fine scales are 5, 10 and 20 Hz, respectively. The number

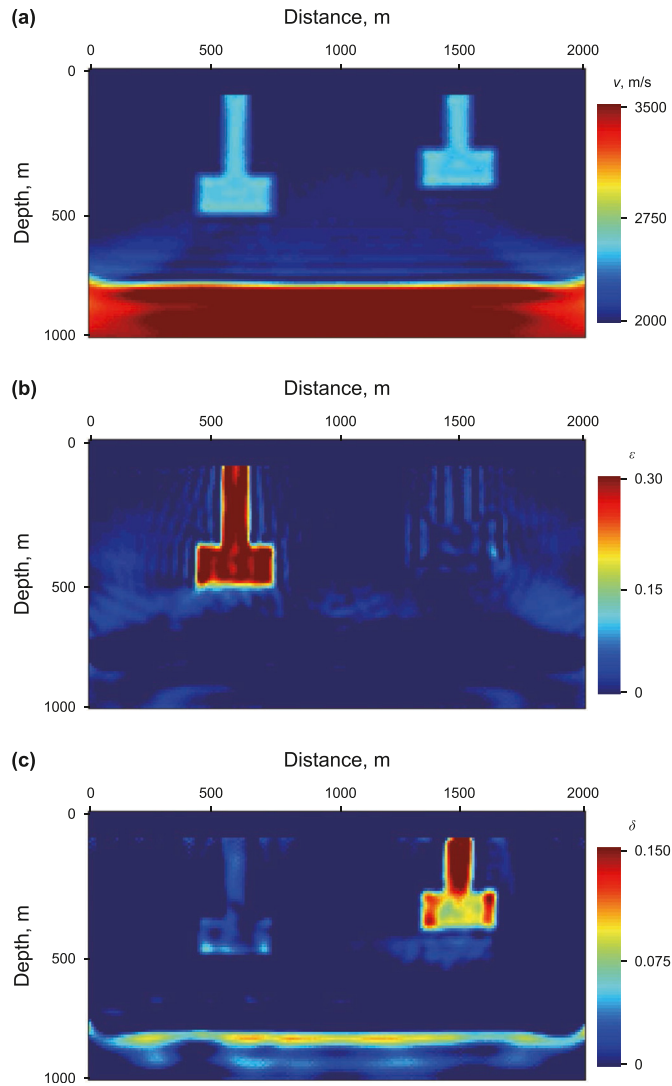


Fig. 21. The inverted (a) v , (b) ϵ and (c) δ models by the pure acoustic TGN FWI for the simple VTI model with uncorrelated v , ϵ and δ structures.

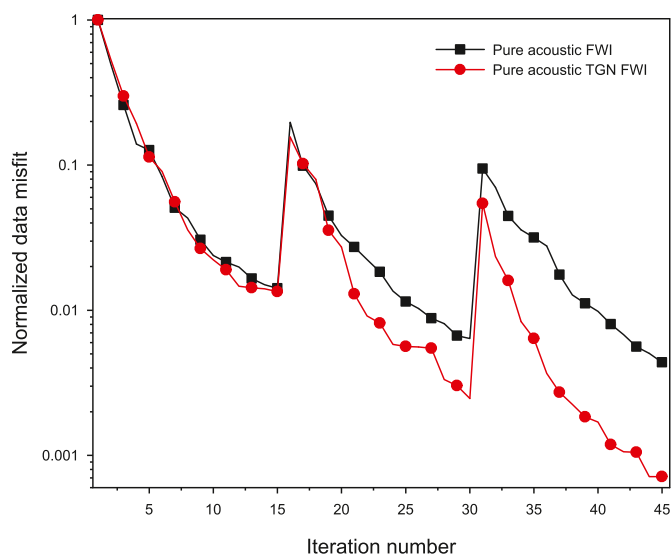


Fig. 22. Convergence curves of different FWI methods for the simple VTI model with uncorrelated v , ϵ and δ structures.

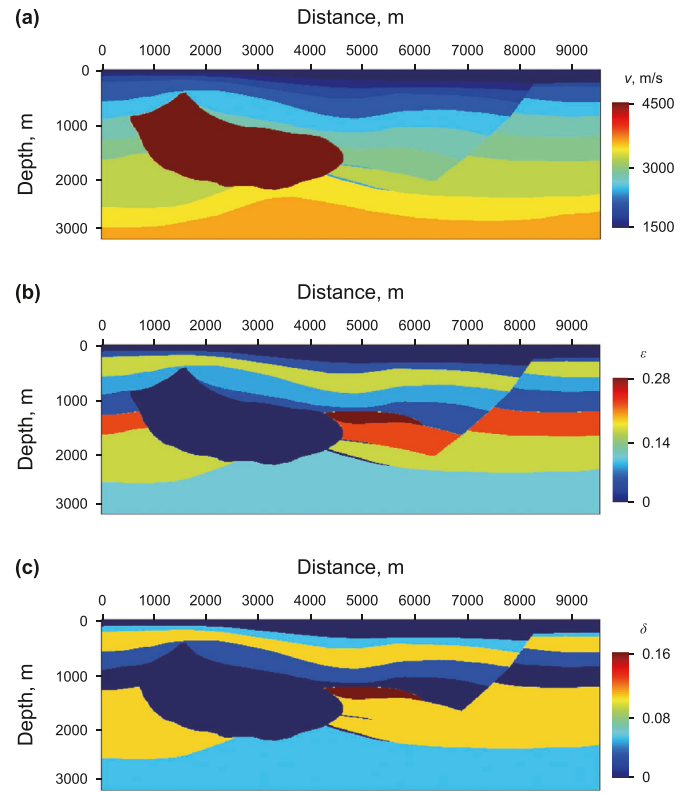


Fig. 23. The modified Hess VTI model. (a) v , (b) ϵ and (c) δ .

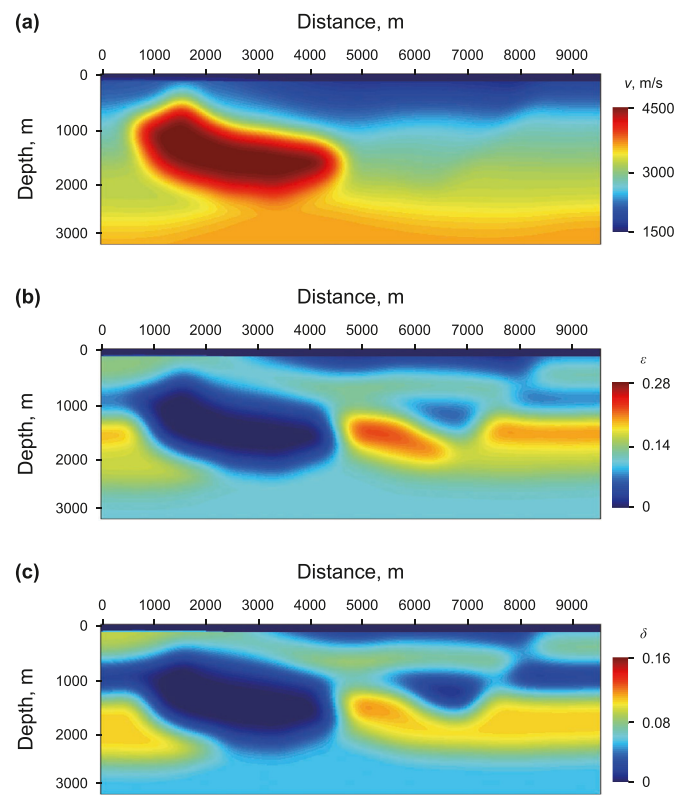


Fig. 24. The initial (a) v , (b) ϵ and (c) δ models for the modified Hess VTI model.

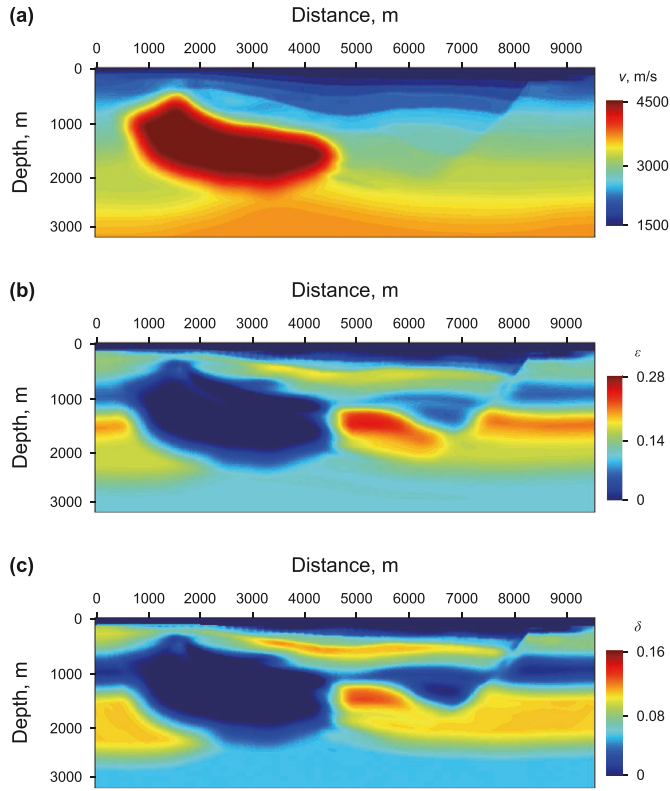


Fig. 25. The inverted (a) v , (b) ϵ and (c) δ models by the pseudoacoustic FWI for the modified Hess VTI model.

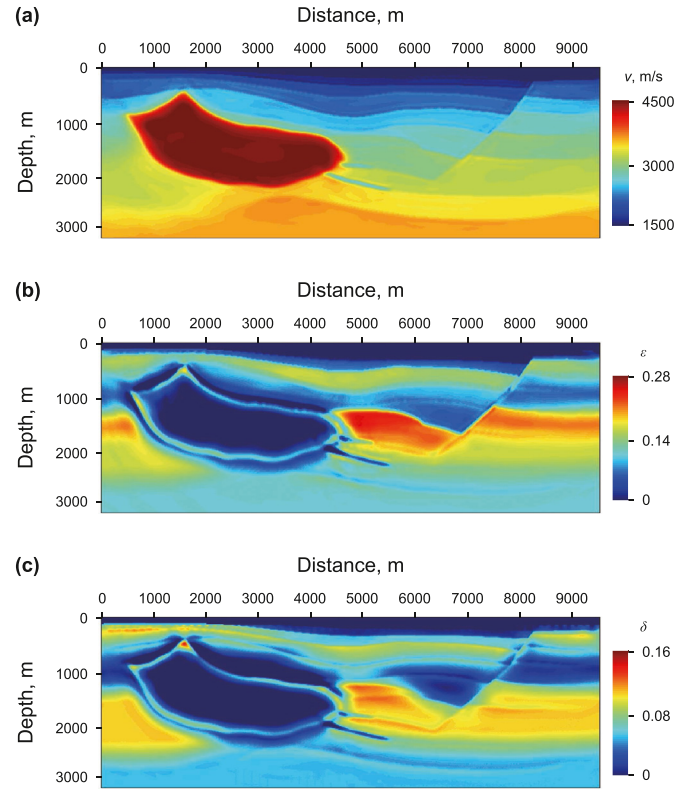


Fig. 27. The inverted (a) v , (b) ϵ and (c) δ models by the pure acoustic TGN FWI method for the modified Hess VTI model.

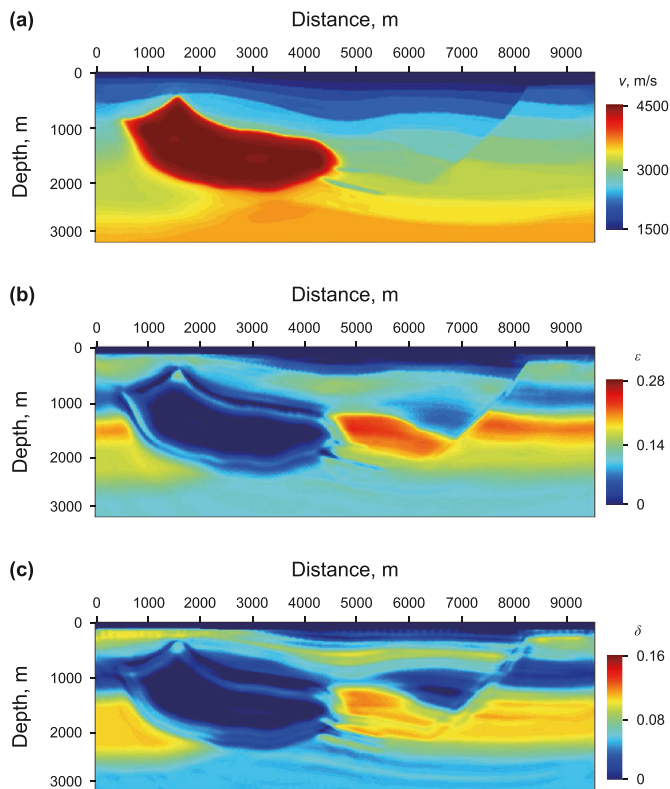


Fig. 26. The inverted (a) v , (b) ϵ and (c) δ models by the pure acoustic FWI for the modified Hess VTI model.

of iterations is equal to 15 at each scale. Figs. 14–16 show the inversion results of different methods. The spurious S-waves, caused by the pseudoacoustic solver, inevitably increase the ill-posedness of the inversion, and the constraint condition (Eq. (66)) improves the stability at the expense of inversion accuracy. As a result, the pseudoacoustic FWI obtains the velocity model with modest accuracy and ϵ and δ models with low accuracy. In contrast, the pure acoustic FWI, free of the S-wave artifacts and instability, recovers plausible velocity and anisotropy parameter models. Owing to the inversion Hessian operator, the pure acoustic TGN FWI further reduces the artifacts and increases the accuracy of ϵ and δ models. Fig. 17 shows the convergence curves of different methods, further proving the effectiveness of the pure acoustic approximation and the TGN method in anisotropic FWI. To quantitatively evaluate the inversion results, Table 1 displays the root mean square (RMS) error of different methods. The error of the pure acoustic TGN FWI method is much smaller than the pseudoacoustic and pure acoustic FWI methods.

We then apply a VTI model with uncorrelated v , ϵ and δ structures to analyze the parameter tradeoff. The modeling and inversions parameters are the same as those used in the correlated v , ϵ and δ model example. The real and initial models are shown in Figs. 18 and 19, respectively. Figs. 20 and 21 illustrate the inversion results for different methods. It is observed that the pure acoustic FWI can obtain plausible velocity model, whereas the inverted anisotropy parameter models are far from the true model. The crosstalk between v and δ is serious. The pure acoustic TGN FWI greatly improves the focusing, resolution and accuracy of ϵ and δ anomalies and reduces the amplitude of the false interface produced by the crosstalk between v and δ . Fig. 22 depicts the convergence curves of different methods. The pure acoustic TGN FWI exhibits a faster convergence rate than the pure acoustic FWI.

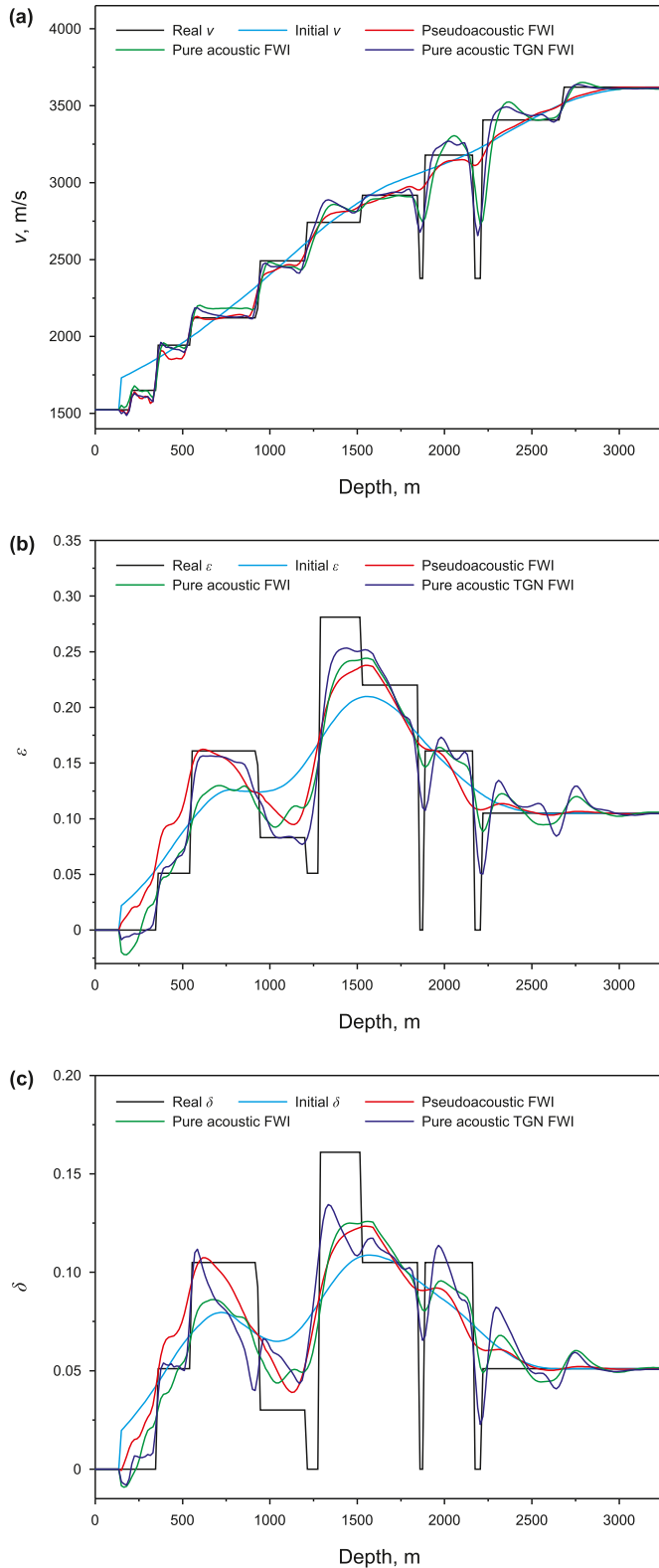


Fig. 28. The inverted (a) v , (b) ϵ and (c) δ at the distance of 4860 m by different FWI methods for the modified Hess VTI model.

We use the number of forward computations to evaluate the efficiency. The number of forward computations for the pure acoustic FWI and the pure acoustic TGN FWI is 225 and 795, respectively.

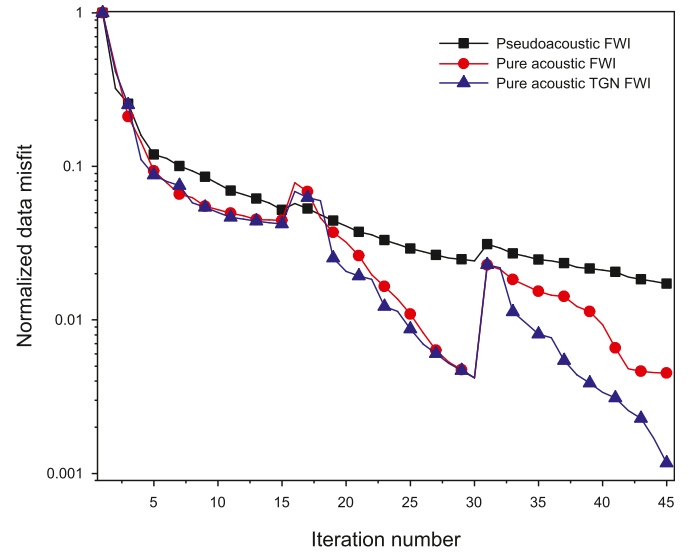


Fig. 29. Convergence curves of different FWI methods for the modified Hess VTI model.

3.2. Hess VTI model example

The second example is for the resampled Hess VTI model, displayed in Fig. 23. The time step is 1.5 ms, the recording time is 3.0 s, the grid spacing is 15 m, and the computational domain is divided into 630×218 grids. Sixty-three sources and 630 receivers are evenly deployed on the surface with the interval of 150 m and 15 m, respectively. The source is represented by a Ricker wavelet with the peak frequency of 10.0 Hz. Fig. 24 shows the initial model, which is a smoothed version of the real model. We apply a multi-scale inversion (the cut-off frequency of 5, 10 and 20 Hz, respectively) and perform 15 iterations at each scale. Figs. 25–27 illustrate the inversion results of different methods. The results of the pure acoustic TGN FWI are more accurate and focused than those of the pseudoacoustic and pure acoustic FWI, especially for the anisotropy parameter models. To see details, we depict the results at the distance of 4860 m by different FWI methods in Fig. 28. Because of the interparameter crosstalk and the weak sensitivity of anisotropy parameters, the pseudoacoustic FWI and pure acoustic FWI fail to produce reasonable ϵ and δ structures. By using the inverse Hessian to refocus the functional gradients and minimize the parameter tradeoff, the pure acoustic TGN FWI drastically increases the inversion accuracy of the ϵ model and marginally improves the velocity and δ models. Note that the anisotropy parameters are insensitive to the seismic response and can be inverted for only when the velocity is accurate. However, the pure acoustic TGN FWI still obtains reliable velocity and ϵ structures. The convergence curves and the RMS error of different methods, shown in Fig. 29 and Table 2, respectively, revealing that the pure acoustic TGN FWI is superior to the other two methods based on the first-order optimization algorithm. From the point of view of efficiency, the pure acoustic FWI and the pure acoustic TGN FWI require 225 and 816 forward computations, respectively.

4. Conclusions

We have developed an acoustic anisotropic FWI method based on a simplified pure qP-wave equation and implemented TGN FWI of pure qP-waves in VTI media. We derive the formulation of Hessian-vector products based on the Lagrange multiplier method and introduce an efficient matrix-free Gauss-Newton method. The

Table 2
The RMS error between the real and initial/inverted models for the Hess VTI model.

Parameter	Initial model, m/s	Pseudoacoustic FWI, m/s	Pure acoustic FWI, m/s	Pure acoustic TGN FWI, m/s
ν	195.2420	172.6188	136.7062	104.4177
ϵ	0.0302	0.0281	0.0273	0.0219
δ	0.0202	0.0189	0.0182	0.0145

parameter tradeoff is analyzed by the approximate Hessian operator and the proposed method is tested on synthetic data. Numerical results show that the pure acoustic approximation is preferable to the pseudoacoustic approximation in anisotropic FWI. Moreover, the pure acoustic TGN FWI helps to mitigate the parameter crosstalk and improve the inversion accuracy of anisotropy parameters.

The pure acoustic TGN FWI is computationally intensive compared to the methods based on the first-order optimization algorithms. We only perform TGN FWI of pure qP-waves in VTI media for the 2D case and the method for the 3D case is expensive and unaffordable even with current high-performance computers. Further improvements in computational efficiency are needed before large-scale applications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Zhi-Ming Ren: Methodology. **Lei Wang:** Writing – original draft. **Qian-Zong Bao:** Writing – review & editing.

Acknowledgments

We would like to express our sincere appreciation to the editor and two anonymous reviewers for their valuable comments and suggestion. This research is supported by the National Natural Science Foundation of China (grant No. 42174156), the Young Science and Technology Star Project of Shaanxi Province (grant No. 2023KJXX-021), and the Fundamental Research Funds for the Central Universities, CHD (grant Nos. 300102263401 and 300102264204).

References

Alkhalifah, T., 2000. An acoustic wave equation for anisotropic media. *Geophysics* 65 (4), 1239–1250. <https://doi.org/10.1190/1.1444815>.
Alkhalifah, T., Plessix, R.-E., 2014. A recipe for practical full-waveform inversion in anisotropic media: an analytical parameter resolution study. *Geophysics* 79 (3), R91–R101. <https://doi.org/10.1190/geo2013-0366.1>.
Bunks, C., Saleck, F.M., Zaleski, S., et al., 1995. Multiscale seismic waveform inversion. *Geophysics* 60 (5), 1457–1473. <https://doi.org/10.1190/1.1443880>.
Castellanos, C., Métivier, L., Operto, S., et al., 2015. Fast full waveform inversion with source encoding and second-order optimization methods. *Geophys. J. Int.* 200 (2), 720–744. <https://doi.org/10.3997/2214-4609.20130834>.
Chu, C., Macy, B.K., Anno, P.D., 2011. Approximation of pure acoustic seismic wave propagation in TTI media. *Geophysics* 76 (5), WB97–WB107. <https://doi.org/10.1190/geo2011-0092.1>.
Dussaud, E., Symes, W., Williamson, P., et al., 2008. Computational strategies for reverse-time migration. In: 78th Ann. Int. Meet. SEG, Expand. Abstr., pp. 2267–2271. <https://doi.org/10.1190/1.3059336>.
Duveneck, E., Milcik, P., Bakker, P.M., et al., 2008. Acoustic VTI wave equations and their application for anisotropic reverse-time migration. In: 78th Ann. Int. Meet. SEG, Expand. Abstr., pp. 2186–2189. <https://doi.org/10.1190/1.3059320>.
Eisenstat, S.C., Walker, H.F., 1996. Choosing the forcing terms in an inexact Newton method. *SIAM J. Sci. Comput.* 17, 16–32. <https://doi.org/10.1137/0917003>.
Fichtner, A., Trampert, J., 2011. Hessian kernels of seismic data functionals based upon adjoint techniques. *Geophys. J. Int.* 185 (2), 775–798. <https://doi.org/10.1111/j.1365-246X.2011.04966.x>.

Finkelstein, B., Kastner, R., 2007. Finite difference time domain dispersion reduction schemes. *J. Comput. Phys.* 221, 422–438. <https://doi.org/10.1016/j.jcp.2006.06.016>.
Forgues, E., Lambaré, G., 1997. Parameterization study for acoustic and elastic ray + Born inversion. *J. Seismic Explor.* 6, 253–278.
Gao, L., Pan, Y., Rieder, A., et al., 2021. Multiparameter viscoelastic full-waveform inversion of shallow-seismic surface waves with a pre-conditioned truncated Newton method. *Geophys. J. Int.* 227 (3), 2044–2057. <https://doi.org/10.1093/gji/ggab311>.
Gholami, Y., Brossier, R., Operto, S., et al., 2013. Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 1: sensitivity and trade-off analysis. *Geophysics* 78 (2), R81–R105. <https://doi.org/10.1190/geo2012-0204.1>.
Hadden, S., Pratt, R.G., Smithyman, B., 2019. Anisotropic full-waveform inversion of crosshole seismic data: a vertical symmetry axis field data application. *Geophysics* 84 (1), B15–B32. <https://doi.org/10.1190/geo2017-0790.1>.
He, W., Plessix, R.-E., Singh, S., 2018. Parametrization study of the land multiparameter VTI elastic waveform inversion. *Geophys. J. Int.* 213 (3), 1660–1672. <https://doi.org/10.1093/gji/ggy099>.
Jeong, W., Lee, H., Min, D., 2012. Full waveform inversion strategy for density in the frequency domain. *Geophys. J. Int.* 188 (3), 1221–1242. <https://doi.org/10.1111/j.1365-246X.2011.05314.x>.
Kamath, N., Tsvankin, I., 2016. Elastic full-waveform inversion for VTI media: Methodology and sensitivity analysis. *Geophysics* 81 (2), C53–C68. <https://doi.org/10.1190/geo2014-0586.1>.
Li, X., Zhu, H., 2018. A finite-difference approach for solving pure quasi-P-wave equations in transversely isotropic and orthorhombic media. *Geophysics* 83 (4), C161–C172. <https://doi.org/10.1190/geo2017-0405.1>.
Liu, F., Morton, S., Jiang, S., et al., 2009. Decoupled wave equations for P and SV waves in an acoustic VTI media. In: 79th Ann. Int. Meet. SEG, Expand. Abstr., pp. 2844–2848. <https://doi.org/10.1190/1.3255440>.
Liu, Y., Wu, Z., Kang, H., et al., 2020. Use of prismatic waves in full-waveform inversion with the exact Hessian. *Geophysics* 85 (4), R325–R337. <https://doi.org/10.1190/geo2018-0625.1>.
Malinowski, M., Operto, S., Ribodetti, A., 2011. High-resolution seismic attenuation imaging from wide-aperture onshore data by viscoacoustic frequency-domain full-waveform inversion. *Geophys. J. Int.* 186 (3), 1179–1204. <https://doi.org/10.1111/j.1365-246X.2011.05098.x>.
Matharu, G., Sacchi, M., 2019. A subsampled truncated-Newton method for multiparameter full-waveform inversion. *Geophysics* 84 (3), R333–R340. <https://doi.org/10.1190/geo2018-0624.1>.
Métivier, L., Bretaudeau, F., Brossier, R., et al., 2014. Full waveform inversion and the truncated Newton method: quantitative imaging of complex subsurface structures. *Geophys. Prospect.* 62 (6), 1353–1375. <https://doi.org/10.1111/1365-2478.12136>.
Métivier, L., Brossier, R., Virieux, J., et al., 2013. Full waveform inversion and the truncated Newton method. *SIAM J. Sci. Comput.* 35, B401–B437. <https://doi.org/10.1088/1742-6596/386/1/012013>.
Mu, X., Huang, J., Mao, Q., et al., 2023. Efficient pure qP-wave simulation and reverse time migration imaging for vertical transverse isotropic (VTI) media. *J. Geophys. Eng.* 20, 712–722.
Nocedal, J., Wright, S.J., 1999. *Numerical Optimization*, second ed. Springer.
Oh, J.-W., Alkhalifah, T., 2016. Elastic orthorhombic anisotropic parameter inversion: an analysis of parameterization elastic orthorhombic anisotropic FWI. *Geophysics* 81 (6), C279–C293. <https://doi.org/10.1190/geo2015-0656.1>.
Operto, S., Gholami, Y., Prieux, V., et al., 2013. A guided tour of multi-parameter full-waveform inversion with multicomponent data: from theory to practice. *Lead. Edge* 32, 1040–1054. <https://doi.org/10.1007/s10596-019-09897-6>.
Pan, W., Innanen, K.A., Wang, Y., 2019. Parameterization analysis and field validation of VTI-elastic full-waveform inversion in a walk-away vertical seismic profile configuration. *Geophysics* 85 (3), B87–B107. <https://doi.org/10.1190/geo2019-0089.1>.
Pan, W., Innanen, K.A., Margrave, G.F., et al., 2016. Estimation of elastic constants for HTI media using Gauss-Newton and full-Newton multiparameter full waveform inversion. *Geophysics* 81 (5), R275–R291. <https://doi.org/10.1190/geo2015-0594.1>.
Plessix, R.E., Cao, Q., 2011. A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium. *Geophys. J. Int.* 185 (1), 539–556. <https://doi.org/10.1111/j.1365-246X.2011.04957.x>.
Pratt, R.G., Shin, C., Hick, G., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophys. J. Int.* 133 (2), 341–362. <https://doi.org/10.1046/j.1365-246X.1998.00498.x>.
Prieux, V., Brossier, R., Operto, S., et al., 2013. Multiparameter full waveform

- inversion of multicomponent ocean-bottom-cable data from the Valhall field—Part 1: imaging compressional wave speed, density and attenuation. *Geophys. J. Int.* 194 (3), 1640–1664. <https://doi.org/10.1093/gji/ggt177>.
- Prieux, V., Brossier, R., Gholami, Y., et al., 2011. On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study. *Geophys. J. Int.* 187 (3), 1495–1515. <https://doi.org/10.1111/j.1365-246X.2011.05209.x>.
- Ren, Z., Liu, Y., 2016. A hierarchical elastic full-waveform inversion scheme based on wavefield separation and the multistep-length approach. *Geophysics* 81 (3), R99–R123. <https://doi.org/10.1190/geo2015-0431.1>.
- Ren, Z., Bao, Q., Gu, B., 2021. Joint wave-equation traveltime inversion of diving/direct and reflected waves for P- and S-wave velocity macromodel building. *Geophysics* 86 (4), R603–R621. <https://doi.org/10.1190/geo2020-0762.1>.
- Ren, Z., Dai, X., Wang, L., 2024. Full-waveform inversion of pure quasi-P waves in tilted transversely isotropic media based on different parameterizations. *Geophysics* 89 (2), 1–17. <https://doi.org/10.1190/GEO2023-0367.1>.
- Ren, Z., Li, Z., Gu, B., 2019. Elastic reflection waveform inversion based on the decomposition of sensitivity kernels. *Geophysics* 84 (2), R235–R250. <https://doi.org/10.1190/geo2018-0220.1>.
- Rusmanugroho, H., Modrak, R., Tromp, J., 2017. Anisotropic full-waveform inversion with tilt-angle recovery. *Geophysics* 82 (3), R135–R151. <https://doi.org/10.1190/geo2016-0025.1>.
- Strang, G., 2007. *Computational Science and Engineering*. Wellesley. Cambridge Press.
- Sears, T.J., Singh, S.C., Barton, P.J., 2008. Elastic full waveform inversion of multicomponent OBC seismic data. *Geophys. Prospect.* 56 (6), 843–862. <https://doi.org/10.1111/j.1365-2478.2008.00692.x>.
- Sheen, D.-H., Tuncay, K., Baag, C.-E., et al., 2006. Time domain Gauss-Newton seismic waveform inversion in elastic media. *Geophys. J. Int.* 167 (3), 1373–1384. <https://doi.org/10.1111/j.1365-246X.2006.03162.x>.
- Singh, S., Tsvankin, I., Naein, E.Z., 2020. Full-waveform inversion with borehole constraints for elastic VTI media. *Geophysics* 85 (6), R553–R563. <https://doi.org/10.1190/geo2019-0816.1>.
- Singh, S., Tsvankin, I., Naein, E.Z., 2021. Facies-based full-waveform inversion for anisotropic media: a North Sea case study. *Geophys. Prospect.* 69 (8–9), 1650–1663. <https://doi.org/10.1111/1365-2478.13139>.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation. *Geophysics* 49 (8), 1259–1266. <https://doi.org/10.1190/1.1441754>.
- Tarantola, A., 1986. A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics* 51 (10), 1893–1903. <https://doi.org/10.1190/1.1442046>.
- Thomsen, L., 1986. Weak elastic anisotropy. *Geophysics* 51 (10), 1954–1966. <https://doi.org/10.1190/1.1442051>.
- Virieux, J., Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics. *Geophysics* 74 (6), WCC1–WCC26. <https://doi.org/10.1190/1.3238367>.
- Wang, T., Cheng, J., 2017. Elastic full waveform inversion based on mode decomposition: the approach and mechanism. *Geophys. J. Int.* 209 (2), 606–622. <https://doi.org/10.1093/gji/ggx038>.
- Wang, T., Cheng, J., Geng, J., 2022. Reflection-based traveltime and waveform inversion with second-order optimization. *Petrol. Sci.* 19, 1582–1591. <https://doi.org/10.1016/j.petsci.2022.02.003>.
- Wang, Y., Dong, L., 2015. Multi-parameter full waveform inversion for acoustic VTI media using the truncated Newton method. *Chin. J. Geophys.* 58, 2873–2885. <https://doi.org/10.6038/cjg20150821> (in Chinese).
- Yang, L., Yan, H., Liu, H., 2016. Optimal implicit staggered-grid finite-difference schemes based on the sampling approximation method for seismic modelling. *Geophys. Prospect.* 64 (3), 595–610. <https://doi.org/10.1111/1365-2478.12325>.
- Yang, P., Brossier, R., Métivier, L., et al., 2018. A time-domain preconditioned truncated Newton approach to visco-acoustic multiparameter full waveform inversion. *SIAM J. Sci. Comput.* 40, B1101–B1130. <https://doi.org/10.1137/17M1126126>.
- Yong, P., Brossier, R., Métivier, L., 2022. Parsimonious truncated Newton method for time-domain full-waveform inversion based on the Fourier-domain full-scattered-field approximation. *Geophysics* 87 (1), R123–R146. <https://doi.org/10.1190/geo2021-0164.1>.
- Zhou, W., Brossier, R., Operto, S., et al., 2015. Full waveform inversion of diving and reflected waves for velocity model building with impedance inversion based on scale separation. *Geophys. J. Int.* 202, 1535–1554. <https://doi.org/10.1093/gji/ggv228>.