Which parameterization for acoustic vertical transverse isotropic full waveform inversion? - Part 1: sensitivity and trade-off analysis

Yaser Gholami *, Romain Brossier†, Stéphane Operto‡, Alessandra Ribodetti§ and Jean Virieux¶

*Géoa, Université Nice-Sophia Antipolis, CNRS, IRD, Observatoire de la Côte d’Azur, Sophia-Antipolis, France. E-mail: gholami@geoazur.unice.fr
†ISTerre, Université Joseph Fourier Grenoble, Observatoire des Sciences de l’Univers, France. E-mail: romain.brossier@obs.ujf-grenoble.fr.
‡Géoa, Université Nice-Sophia Antipolis, CNRS, IRD, Observatoire de la Côte d’Azur, Sophia-Antipolis, France. E-mail: operto@geoazur.unice.fr
§Géoa, Université Nice-Sophia Antipolis, CNRS, IRD, Observatoire de la Côte d’Azur, Villefranche-sur-mer, France. E-mail: ribodetti@geoazur.unice.fr
¶ISTerre, Université Joseph Fourier, Grenoble, France. E-mail: Jean.Virieux@ujf-grenoble.fr.

(January 23, 2013)

Running head: VTI Full Waveform Inversion: part 1

ABSTRACT

In most geological environments, accounting for anisotropy is necessary to perform acoustic full waveform inversion (FWI) of wide-azimuth and wide-aperture seismic data because of the potential dependence of wavespeeds on the direction of the wave propagation. In the
framework of multiparameter FWI, the subsurface parameterization controls the influence of the different parameter classes on the modeled seismic data as a function of the scattering angle, and hence the resolution with which the parameters can be reconstructed and the potential trade-off between different parameters. We propose a numerical procedure based on computation of the scattering patterns of the different parameters to assess the sensitivity of the seismic data to different parameterizations of vertical transverse isotropic media in the acoustic approximation. Among the different categories we have tested, a mono-parametric FWI is proposed for imaging one wavespeed with a broad wavenumber content, keeping the Thomsen parameters fixed, which have a small influence on the data relative to the wavespeed. This raises the question of the initial information required in the background models of the Thomsen parameters to perform reliable monoparameter FWI. Alternatively, inverting simultaneously the horizontal and vertical wavespeeds introduces limited trade-off effects, as these wavespeeds have significant influence on the data for distinct ranges of scattering angles, while the influence of the Thomsen parameter $\delta$ remains weak. With such parameterization, the short-to-intermediate wavelengths of the vertical velocity are updated from the short-to-intermediate scattering angles, while the long-to-intermediate wavelengths of the horizontal velocity are updated from the wide-to-intermediate scattering angles. We conclude that the choice of the subsurface parameterization can be driven by the acquisition geometry, which controls the scattering-angle coverage and hence the resolving power of FWI, and by the accuracy of the available initial FWI models.
INTRODUCTION

The potential of full waveform inversion (FWI) for high-resolution imaging of complex media from low-frequency, wide-aperture/ wide-azimuth data has become apparent over the last decade (e.g., Ravaut et al., 2004; Sirgue et al., 2010; Plessix et al., 2012). FWI seeks to exploit the full information content of the seismic wavefield recorded over a broad range of incidence angles, to build subsurface models with a broad wavenumber content (Virieux and Operto, 2009). Diving waves, pre-critical and super-critical reflections, and diffractions potentially carry information of the subsurface at different resolution powers. FWI of wide-aperture and wide-azimuth data raises the issue, however, of the footprint of anisotropy in the imaging, because of the potential variation in the wavespeed with respect to the direction of propagation (e.g., Barnes et al., 2008; Lee et al., 2010; Plessix and Cao, 2011; Prieux et al., 2011). The need to better account for anisotropy in acoustic FWI requires to identify a suitable parameterization of the subsurface, and to define the number of parameter classes within the chosen parameterization, which can be reliably reconstructed by FWI. Parameterization should be understood as the definition of a set of independent parameters, that fully describe the subsurface properties. For these objectives, we need to assess the influence of each parameter class on the data as a function of the scattering angle, which will in turn provide insights into the trade-off between the parameters and the resolution with which they can be reconstructed. This is the central aim of this study.

Sensitivity and trade-off analysis of multi-parameter waveform inversion can be performed in the framework of inverse scattering theory. The governing idea relies on the
analysis of the scattering (or radiation) patterns of the secondary scattering sources located at the heterogeneity position. These scattering sources are the sources of the so-called partial derivative wavefields (e.g. Pratt et al., 1998), whose values at the receiver positions provide the coefficients of the sensitivity matrix in the framework of linearized waveform inversion based on the Born approximation. Such analysis were presented for elastic media in Wu and Aki (1985a), Wu and Aki (1985b), Tarantola (1986), and Forgues and Lambaré (1997) and for attenuating media in Sato (1984) and Ribodetti and Virieux (1996). Scattering patterns in elastic anisotropic transverse isotropic media were developed in Eaton and Stewart (1994), Burridge et al. (1998), Bostock (2003), Foss et al. (2005), and Calvet et al. (2006), who showed that the scattering pattern depends both on the angle between the incident-wave propagation direction and the symmetry axis, and on the angle between the scattered-wave polarization and the symmetry axis. More qualitative analysis of the sensitivity of the waveform inversion to the parameterization have also been proposed from the numerical computation of the finite-frequency sensitivity kernels of the waveform inversion (Zhou and Greenhalgh, 2009; Sieminski et al., 2009; Zhou and Greenhalgh, 2011).

Other approaches rely on eigenvalue decomposition of the Hessian operator. Kiyashchenko et al. (2004) carried out a sensitivity analysis of amplitude-preserving migration to the anisotropic parameters for simple layered examples, to determine which combinations of parameters control the amplitude-versus-offset of the reflections. They have concluded that amplitude-preserving migration is mainly governed by the normal moveout (NMO) velocity and the anellipticity parameter $\eta$ as already proposed by Alkhalifah and Tsvankin (1995). Plessix and Cao (2011) performed a parameterization analysis of the vertical transverse isotropic (VTI) FWI in the acoustic approximation. The eigenvector decomposition of the Hessian shows that the traveltimes of the diving waves are predominantly sensitive to the
horizontal velocity, while the reflection waves at short offsets are predominantly sensitive to
the vertical velocity, that is consistent with the horizontal and vertical directions of propa-
gation of diving waves and short-spread reflections, respectively. Moreover, they show from
a theoretical viewpoint that the long-wavelength variations of the $\delta$ parameter cannot be
retrieved from surface seismic data because of the intrinsic ambiguity between $\delta$ and depth.

The present study is the first of a two-part series that deals with a sensitivity and trade-
off analysis of acoustic FWI for VTI media. Although the aim of the present study is similar
to that presented by Plessix and Cao (2011), our sensitivity analysis relies on different
tools, mainly the analysis of the scattering patterns for several VTI parameterizations. One
benefit of this analysis is to provide a compact representation of the sensitivity of the data
to the parameter classes as a function of the scattering angles, and hence of the acquisition
geometry. Indeed, large scattering angles are associated with diving waves and super-
critical reflections recorded by long-offset data, while small scattering angles are associated
with short-spread reflections recorded by narrow-azimuth acquisition systems. Moreover,
as the scattering angles are closely related to the wavenumbers injected in the subsurface
models during FWI, the scattering patterns give clear insights on the bandpass filtering
effects applied to the subsurface models by the parameterization, or in other words, on
the resolution with which subsurface models associated with the different parameter classes
can be reconstructed. The analysis of the radiation patterns is validated against canonical
synthetic tests of FWI. The present study is complemented with a realistic synthetic example
and a real ocean-bottom-cable case study from the Valhall field, which are presented in a
companion report (Gholami et al., 2012), referred to as Paper II in the following.

In the first part of this study, we review the multiparameter FWI algorithm, which we
use in this report and in Paper II (Gholami et al., 2012). In the appendixes, we review
our derivation of the gradient of the misfit function for the VTI acoustic wave equation. In
the second part of this study, we numerically compute the scattering patterns for different
VTI acoustic parameterizations. Then, this analysis is validated against a grid analysis
of the misfit function. The last part of this study presents application of multiparameter
FWI on simple synthetic models. We discuss the consistency of these results against the
theoretical analysis of the scattering patterns, in terms of the spatial resolution and the
trade-off between parameters.

MULTIPARAMETER FULL WAVEFORM INVERSION

We review some key theoretical and algorithmic aspects of multiparameter FWI.

The forward problem

A VTI acoustic wave equation can be derived from the two-dimensional P-SV velocity-
stress equation in VTI media by setting to zero the shear-wave velocity on the symmetry
axis (Duveneck et al. (2008) and Appendix A), that gives:

\[-\iota \omega v_x = b \left( \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right),\]

\[-\iota \omega v_z = b \left( \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right),\]

\[-\iota \omega \sigma_{xx} = c_{11} \frac{\partial v_x}{\partial x} + c_{13} \frac{\partial v_z}{\partial z},\]

\[-\iota \omega \sigma_{zz} = c_{13} \frac{\partial v_x}{\partial x} + c_{33} \frac{\partial v_z}{\partial z},\] (1)

where \( \iota = \sqrt{-1} \) is the purely imaginary term, \( \omega \) denotes the angular frequency, \( \mathbf{x} = (x, z) \) the
spatial Cartesian coordinates, \( (v_x(\omega, \mathbf{x}), v_z(\omega, \mathbf{x})) \) and \( (\sigma_{xx}(\omega, \mathbf{x}), \sigma_{zz}(\omega, \mathbf{x}), \sigma_{xz}(\omega, \mathbf{x})) \) the
particle velocities and stresses, respectively. In the framework of the acoustic approximation
of VTI media, we shall consider the wavefield \( \mathbf{p} = \frac{1}{2} (\sigma_{xx} + \sigma_{zz}) \) as the pressure wavefield
recorded by the hydrophone component (Appendix A). The subsurface is described by
buoyancy \( b \) (the inverse of density), and the elastic moduli, \( c_{11}, c_{33}, \) and \( c_{13} \). We model the
wave propagation in two-dimensional visco-acoustic VTI media with a frequency-domain
finite-element discontinuous Galerkin method (Brossier et al., 2008, 2010b; Brossier, 2011).
More details on the implementation of the acoustic VTI wave equation and its interfacing
with the inversion code is provided in Appendixes A and B.

The inverse problem

We seek to minimize a misfit function \( C \) given by:

\[
C(m) = \frac{1}{2} \Delta d^\dagger W_{d} \Delta d + \frac{1}{2} \sum_{i=1}^{N_p} \lambda_i \left( m_i - m_{\text{prior},i} \right)^\dagger W_{m_i} \left( m_i - m_{\text{prior},i} \right),
\]  

(2)

where the data residual vector, \( \Delta d = d_{\text{cal}}(m) - d_{\text{obs}} \), is the difference between the modeled
data \( d_{\text{cal}}(m) \) and the recorded data \( d_{\text{obs}} \). The conjugate transpose is denoted by the sign
\( \dagger \). In the present study, we consider only pressure wavefields recorded by the hydrophone
component. The multi-parameter subsurface model is denoted by \( m = (m_1, ..., m_{N_p}) \),
where \( N_p \) denotes the number of parameter classes to be updated during the FWI. As we
set the density to be constant and equal to 1 and we do not consider attenuation, the
acoustic VTI medium is described by three classes of parameter, for example the elastic
moduli \( c_{11}, c_{33}, c_{13} \) in equation 1. Therefore \( N_p \leq 3 \). In the present study, we minimize the
misfit function, equation 2, with respect to normalized model parameters that are scaled
by their mean value in the initial model, such that each class of parameter has the same
order of magnitude: for example, \( m_1 = c_{11}/c_{11,0} \), \( m_2 = c_{33}/c_{33,0} \), \( m_3 = c_{13}/c_{13,0} \), where \( c_{11,0} \),
\( c_{33,0} \), and \( c_{13,0} \) denote the mean values of the subsurface parameters. As discussed later, our
motivation behind this normalization is to steer the inversion towards the reconstruction of
the parameters that have a dominant influence on seismic data.

The operators $W_d$ and $W_{m_i}$ are data-space and model-space weighting operators, respectively. As we consider noise-free data and idealized acquisition in the present study, we set $W_d = I_d$, where $I_d$ denotes the identity matrix. Therefore, no weighting is applied to the data residuals. In the present study, the operators $W_{m_i}$ are roughness operators, which are introduced to steer the inversion towards smooth models. The inverse of the $W_{m_i}$ are smoothing operators, which are given by:

$$W^{-1}_{m_i}(z, x, z', x') = \sigma_i^2(z, x) \exp\left(\frac{-|x - x'|}{\tau_x}\right) \exp\left(\frac{-|z - z'|}{\tau_z}\right),$$

(3)

where $\tau_x$ and $\tau_z$ are horizontal and vertical correlation lengths. As we normalize the subsurface parameters by their mean values, we do not introduce standard deviations in the operators $W_{m_i}$: $\sigma_i^2(z, x) = 1$. The scalar hyper-parameters $\lambda_i$ control the respective weight of the data-space and model-spaced misfit functions in the equation 2.

Local minimization of the misfit function gives the Newton descent direction at iteration $k$:

$$p^{(k)} = -\left[\frac{\partial^2 C(m^{(k)})}{\partial m^2}\right]^{-1} \frac{\partial C(m^{(k)})}{\partial m},$$

(4)

along which the model update is searched:

$$m^{(k+1)} = m^{(k)} + \gamma^{(k)} p^{(k)}.$$

(5)

The first and second derivatives of the misfit function on the right-hand side of the equation 4 are the gradient and the Hessian of the misfit function. The step length $\gamma$, equation 5, is estimated by line search through the parabolic fitting of the misfit function.
The descent direction as a function of the sensitivity matrix $J$ is given by:

$$
p^{(k)} = \Re \left( \tilde{W}_m^{-1} (J^{(k)})^T W_d J^{(k)} + \tilde{W}_m^{-1} \left( \frac{\partial (J^{(k)})^T}{\partial m} \right) \left( (\Delta d^{(k)})^* \ldots (\Delta d^{(k)})^* \right) + \Lambda \right)^{-1} \Re \left( \tilde{W}_m^{-1} (J^{(k)})^T W_d (\Delta d^{(k)})^* + \Lambda (m^{(k)} - m_{\text{prior}}) \right), \quad (6)
$$

where $\Lambda$ is a block diagonal damping matrix:

$$
\Lambda = \begin{pmatrix}
\lambda_1 I_M & \cdots & 0 \\
0 & \ddots & \vdots \\
0 & \cdots & \lambda_{N_p} I_M
\end{pmatrix},
$$

and $I_M$ is the identity matrix of dimension $M$, where $M$ denotes the number of model parameters per class. In equation 6, $\Re$, $T$, and $*$ denote the real part of a complex number, the transpose of a matrix, and the complex conjugate, respectively. The matrix $\tilde{W}_m^{-1}$ is an $N_p \times N_p$ block diagonal matrix, where each block is formed by the $W_m^{-1}$ matrices.

We use the quasi-Newton L-BFGS optimization algorithm to recursively compute an approximate solution of the equation 6 from a few gradients and a few solution vectors from previous iterations (Nocedal, 1980; Nocedal and Wright, 1999). As an initial guess of the inverse of the Hessian, we use a diagonal approximation of the approximate Hessian (the linear term) damped by the $\Lambda$ matrix,

$$
H_0 = \left( \tilde{W}_m^{-1} \text{diag} \left\{ J^{(k)} \right\} W_d J^{(k)} \right) + \Lambda.
$$

The gradient of the misfit function in equation 4 is computed with the adjoint-state method (Plessix, 2006), and is given by:

$$
\frac{\partial C(m^{(k)})}{\partial m_{ij}} = \sum_\omega \sum_s \sum_r \Re \left\{ \tilde{W}_m^{-1} W_m v^T \left( \frac{\partial B}{\partial m_{ij}} \right)^T \beta_1 \right\}, \quad (9)
$$

where $m_{ij}$ denotes the $j^{th}$ model parameter of parameter class $i$ ($j \in [1; M]$), $v = (v_x, v_z)$ is the incident particle-velocity wavefield, $\beta_1$ is the corresponding backpropagated adjoint
wavefield, and \( B \) is the forward modeling operator for particle velocities (Appendix A). The signature of the parameter class on the gradient expression is given by the radiation pattern term, \( \left( \frac{\partial B}{\partial m_{ij}} \right) \). The detailed expression of the gradient of the misfit function is derived in Appendix B.

The so-called partial derivative wavefield from the \( j^{th} \) parameter of the class \( i \) satisfies the wave equation:

\[
B(\omega, m(x)) \frac{\partial v}{\partial m_{ij}} = -\frac{\partial B(\omega, m(x))}{\partial m_{ij}} v,
\]

where the right-hand side is refer to as the so-called virtual source (Pratt et al., 1998). The scattering or radiation pattern of this virtual source, \( \left( \frac{\partial B}{\partial m_{ij}} \right) \), gives some insight into the sensitivity of the data to the parameter \( m_{ij} \) as a function of the scattering angle.

**Resolution power of full waveform inversion**

We explicitly introduce the summation over the frequency, source and receiver in the expression of the gradient of the misfit function, equation 9, to remember the factors that impact on the resolution power of FWI. Among others, Wu and Toksöz (1987), and Sirgue and Pratt (2004) showed that in the framework of diffraction tomography, the gradient of the misfit function can be recast as a truncated Fourier series, where the arguments of the basis function are the wavenumber vectors injected in the subsurface model. These wavenumber vectors \( k \) are related to the scattering angle \( \theta \), the local wavespeed \( c \), and the angular frequency \( \omega \) by

\[
k = \frac{2\omega}{c} \cos(\theta/2) n,
\]

where \( n \) is a unit vector in the direction of the vector formed by the sum of the slowness vectors associated with the rays that connect the source and the receiver to the diffractor,
and the scattering angle $\theta$ is the angle formed by the two slowness vectors (Thierry et al., 1999, their figure 1). The truncation of the Fourier series results on the one hand from the limited bandwidth of the source and on the other hand from the limited range of scattering angles, as shown by equation 11. The local range of scattering angles in the subsurface is controlled by the acquisition geometry and the structural complexity of the target. This truncation is the first factor that limits the resolution with which the subsurface is imaged. The second limiting factor is related to the radiation pattern ($\partial B/\partial m$) of the virtual source of the partial derivative wavefields, equation 10. The amplitude variations of the partial derivative wavefields with the scattering angles $\theta$, which results from the directivity of the virtual source, act as a weighting of the wavenumber spectrum of the gradient, equation 9. The radiation patterns of model parameters in multiparameter FWI are generally not isotropic, and therefore they impact significantly on the resolution of the multiparameter imaging, as well as on the trade-off between parameters. Indeed, the weighting of the wavenumber spectrum of the subsurface model can be partly corrected by the deconvolution (i.e., whitening) action of the Hessian in the framework of Newton or quasi-Newton optimizations.

**SENSITIVITY AND TRADE-OFF ANALYSIS OF VTI FWI**

A VTI medium in the acoustic approximation ($c_{44}=0$) can be parametrized by three kinematic parameters, without considering the density and attenuation. Among the possible parameterization, we propose to investigate here three kinds (Table 1): the first one, which is referred to as type 1 parameterization, involves one wavespeed and two dimensionless Thomsen parameters; the second one, referred to as type 2 parameterization, involves two wavespeeds and one Thomsen parameter and the last one, referred to as type 3 parame-
terization, involves the elastic coefficients $c_{11}$, $c_{33}$, and $c_{13}$. For velocity we consider the
vertical velocity $V_{P0}$, the horizontal velocity $V_h$, and the NMO velocity $V_{NMO}$, while we
shall consider the Thomsen parameters $\delta$ and $\epsilon$, as well as a combination of both, which
corresponds to the $\eta$ parameter.

The relationships between these parameters in the acoustic approximation is given by:

$$c_{33} = \rho V_{P0}^2, \quad c_{11} = \rho V_h^2, \quad c_{13} = \rho V_{P0}^2 (\delta + 1).$$

$$V_{NMO} = V_{P0} \sqrt{1 + 2 \delta}, \quad V_h = V_{P0} \sqrt{1 + 2 \epsilon} = V_{NMO} \sqrt{1 + 2 \eta}, \quad \eta = \frac{(\epsilon - \delta)}{1 + 2 \delta}. \quad (12)$$

In practice, the radiation pattern terms, $(\partial B/\partial m)$, in the gradient of the misfit function,
equation 9, are computed for the $(c_{11}, c_{33}, c_{13})$ parameterization in our FWI code, and the
radiation pattern terms for any other parameterizations are inferred easily by applying the
chain rule on partial derivatives.

10 Radiation patterns of virtual sources

In this section, we estimate numerically the radiation patterns for the above-mentioned
parameterizations of VTI media. Numerically, to achieve this goal, we first compute the
partial derivative of the pressure wavefield with respect to each parameter classes in a finite-
difference sense. From these partial derivatives, we compute, for each parameter class $i$,
the wavefield perturbations $\Delta p_{ij}$ resulting from a point model perturbation $\Delta m_i \delta (x - x_j)$
located in the middle of the finite-difference grid:

$$\Delta p_{ij} \approx \left( \frac{\partial p}{\partial m_{ij}} \right) \Delta m_i \delta (x - x_j). \quad (13)$$

where $\delta$ denotes the Dirac delta function. We define the model perturbation $\Delta m_i$ as a
percentage of the value of the parameter in the background model: $\Delta m_{ij} = perc \times m_{0i}$. The
amplitude variations of the wavefield perturbations $\Delta p_{ij}$ around the diffractor point $j$ give some insight into the sensitivity of the data to the parameter as a function of the scattering angle $\theta$, and hence will be considered as the radiation pattern of the virtual sources in this study. We use perturbation wavefields instead of partial derivative wavefields to assess the real (i.e., physical) influence of the parameters on the data, because the former are independent to the range of values of the model parameters. We use the same value of $perc$ whatever the parameter class $i$ is for a fair comparison between the influence of the different parameter classes on the data. The absolute value of $perc$ (set to 1) is not meaningful, because equation 13 is linear and we are interested in the influence of the parameters on the data in a relative sense. For this value of $perc$, the wavefield perturbations in equation 13 are equivalent to partial derivative wavefields with respect to normalized parameters $\tilde{m}_i$: $\partial p/\partial \tilde{m}_i = (\partial p/\partial m_i) \cdot m_0_i = \Delta p_i$, where $\tilde{m}_i = m_i/m_0_i$. The normalization consists of scaling each parameter class by $m_0_i$, such that all of the parameter classes have the same range of values. This normalization is also applied to the Thomsen parameters, although they are dimensionless.

To compute the partial derivative wavefields in a finite-difference sense, we consider a homogeneous VTI background model defined by $V_{P0} = 4$ km/s, $\delta = 0.05$, $\epsilon = 0.10$, and small parameter perturbations: $\Delta V_{P0} = 0.2$ km/s, $\Delta \delta = 0.05$, $\Delta \epsilon = 0.1$. The modeled frequency is 20 Hz. The radiation patterns for three incidence angles with respect to the vertical symmetry axis are shown in Figures 1 and 2 for the $(V_{P0}, \delta, \epsilon)$ and $(V_{P0}, \delta, V_h)$ parameterizations, respectively. We first show how the radiation patterns can vary with the incidence angle (Calvet et al., 2006) (Figures 1 and 2, top to bottom panels). For example, the sensitivity of the data to $\delta$ and $\epsilon$ is quite small in the $(V_{P0}, \delta, \epsilon)$ parameterization when the source is located on the symmetry axis as shown by the numerical noise in Figure
and the sensitivity improves progressively as the incidence angle increases. The same comment applies to \( \delta \) and \( V_h \) for the \( (V_{P0}, \delta, V_h) \) parameterization (Figure 2b,c). Secondly, we show how the radiation pattern of one parameter class can vary depending on the other parameter classes involved in the subsurface parameterization. For example, while the radiation pattern of the vertical wavespeed does not change with the incidence angle in the \( (V_{P0}, \delta, \varepsilon) \) parameterization and spans the full range of scattering angles (Figures 1a,d,g), it shows some notches as the incidence angle increases when the \( (V_{P0}, \delta, V_h) \) parameterization is used (Figure 2g).

The variation of the radiation patterns with the incidence angle makes their interpretation quite difficult. To facilitate the analysis of the sensitivity of the data to the parameterization, we plot the radiation patterns for several parameterizations in a more compact form in Figures 3 and 4 to facilitate the analysis. We consider the case of the specular reflection angle when the incidence angle \( \varphi \) is half of the scattering angle \( \theta \) and when the incidence angle is defined with respect to the vertical and horizontal axes. The first case (Figure 3) is representative of reflections from a horizontal reflector recorded at the surface, while the second case is representative of reflections from a vertical reflector recorded by a vertical source-receiver array, such as vertical seismic profiling (VSP). We observe that, because the direction of propagation is rotated by 90°, the radiation patterns in Figure 4 are the mirror image of the radiation patterns in Figure 3 with respect to the horizontal axis \( \theta = 90° \) and \( \theta = 270° \). Therefore, a parameter that shows some influence on the data for a scattering angle \( \theta \) in Figure 3 will show the same influence on the data for a scattering angle of \( 180° - \theta \) in Figure 4.
Type 1 parameterizations: one wavespeed + two Thomsen parameters

The specular radiation patterns for the \( (V_P, \delta, \varepsilon), (V_h, \delta, \varepsilon) \) and \( (V_{NMO}, \delta, \eta) \) parameterizations are shown in Figures 3b-d and 4b-d. The first conclusion is that the radiation patterns of the wavespeeds are similar for all three parameterizations, and they show a significant influence of the wavespeed on the data for all of scattering angles: reconstruction of the wavespeed with a broad wavenumber content is therefore expected. As the wavespeed parameter has an influence on the data over the full range of scattering angles, the trade-off between the wavespeed and the two other parameters of the parameterization are unavoidable. A second major conclusion is that the amplitudes of the radiation patterns for the Thomsen parameters are much smaller than those for the wavespeed.

When combined with either the vertical or horizontal wavespeeds, the Thomsen parameter \( \delta \) has a small influence on the data at intermediate aperture angles, and has no influence on the vertical and horizontal wave paths (solid black lines in Figures 3b,c and 4b,c). The radiation pattern of \( \delta \) is the same, when the incidence angle is defined with respect to the vertical and horizontal axis (in other words, the radiation pattern of \( \delta \) is symmetric with respect to the horizontal axis \( \theta = 90^\circ \) and \( \theta = 270^\circ \)). This implies that the image of a horizontal \( \delta \) reflector built from a surface acquisition will be the same as that of a vertical reflector built from a vertical acquisition. The Thomsen parameter \( \varepsilon \) has an influence on the data for large scattering angles, with a maximum of sensitivity for the pure transmission regime \( (\theta=180^\circ, \varphi=90^\circ) \) when a surface acquisition is considered and when the vertical wavespeed is used in the parameterization (dashed line in Figure 3b). For the vertical acquisition, \( \varepsilon \) has a maximum influence on the data at small scattering angles \( (\theta=0^\circ, \varphi=0^\circ) \) (dashed line in Figure 4b). In both cases, the maximum influence on the data is shown
for the horizontal wave paths, which is consistent with the close relationship between $\epsilon$ and the horizontal wave speed. When $\epsilon$ is combined with the horizontal wavespeed rather than with the vertical wavespeed, the opposite scenario is shown: the Thomsen parameter $\epsilon$ has influence mainly on wave paths propagating near vertically (dashed line in Figures 3c and 4c).

The different radiation patterns of $\epsilon$ in the $(V_{P0}, \delta, \epsilon)$ and $(V_h, \delta, \epsilon)$ parameterizations, as well as the weaker influence of $\epsilon$ on the data relative to the wavespeeds, raise the following comment. If one parameter has a small influence on the data and if some prior information provides benefits for this parameter coming from well or traveltime tomographic methods, only the parameter with the major influence on the data during FWI might be updated keeping the background models of the parameters with the minor influence on the data fixed. Generally, this prior information corresponds to the large wavelengths of the parameter built by traveltime tomography. Under these circumstances, it would be preferable to use a parameterization for which the parameter kept fixed during FWI has an influence on the data for large scattering angles because these large scattering angles govern the reconstruction of the large wavelengths of the medium (equation 11). In this case, the influence of the fixed parameter on the data should be predicted sufficiently accurately by the large-wavelength background model. According to this reasoning, the $(V_{P0}, \delta, \epsilon)$ parameterization should be chosen for surface acquisition because $\epsilon$ scatters energy at wide scattering angles.

The $(V_{NMO}, \delta, \eta)$ parameterization for surface acquisition shows an uncoupled influence of the parameters $\delta$ and $\eta$ at short and wide scattering angles, respectively (Figure 3d). Both influences have the same order of magnitude although they remain small relative to that of the NMO velocity and, consequently, a significant trade-off between the NMO velocity and
the two anisotropic parameters $\delta$ and $\eta$ is expected with this parameterization.

Type 2 parameterizations: two wavespeeds + one Thomsen parameter

The radiation pattern for the $(V_{P0}, V_{h}, \delta)$ and $(V_{N,MO}, V_{h}, \delta)$ parameterizations are shown in Figure 3e,f and 4e,f. The radiation pattern of the horizontal wavespeed has a significant influence on the horizontal wave paths, and has the same order of magnitude as the radiation pattern of $V_{P0}$ for vertical wavepaths. Compared to the $(V_{P0}, \delta, \epsilon)$ parameterization, the vertical wave speed has no influence on horizontal wavepaths. Therefore, a limited trade-off between the vertical and horizontal wavespeeds can be expected with the $(V_{P0}, V_{h}, \delta)$ parameterization, as these two parameters have radiation patterns that do not overlap significantly. The counterpart is that the range of scattering angles spanned by the radiation pattern of the vertical velocity is narrower than with the previous parameterization. In consequence, a more limited bandwidth reconstruction of the vertical velocity is expected with the $(V_{P0}, V_{h}, \delta)$ parameterization. For surface acquisitions, long wavelengths of $V_{h}$ should be built from the wide scattering angles (i.e., diving waves and super-critical reflections), while the short wavelengths of $V_{P0}$ should be built from the short scattering angles (i.e., short-spread reflections). The radiation pattern of $\delta$ shows the same trend as for the previous parameterization, with an even smaller influence on the data (Figure 3e).

It is unlikely that the $\delta$ parameter can be reconstructed by FWI of noisy data, because the influence of this parameter on the data will be dominated by the noise.

It is worth noting that the $\delta$ parameter has a stronger influence on the data when it is combined with the NMO velocity rather than with the vertical velocity, with both the $(V_{N,MO}, \delta, \eta)$ or $(V_{N,MO}, \delta, V_{h})$ parameterizations (Figure 3d,f). This highlights the trade-off
between $V_{NMO}$ and $\delta$ in equation 12. A significant trade-off between the NMO velocity and $\delta$ is expected with both parameterizations. If $\delta$ is kept fixed during FWI, a parameterization that minimizes the influence of $\delta$ should be favored, which directs us towards a parameterization involving the vertical velocity rather than the NMO velocity.

Type 3 parameterization: elastic moduli

The radiation patterns for the $(c_{33}, c_{13}, c_{11})$ parameterization show a similar trend to the radiation patterns of the previous parameterization (Figures 3g and 4g), which is consistent with the close relationship between the elastic moduli and the wave speeds (equation 12). The two parameters $c_{11}$ and $c_{33}$, which are related to the horizontal and vertical wave speeds, have radiation patterns with dominant amplitudes relative to the $c_{13}$ parameter. The $c_{33}$ parameter, which is related to the vertical velocity, has a maximum influence on the vertical wave paths, while the $c_{11}$ parameter, which is related to the horizontal velocity, has a maximum influence on the horizontal wave paths. Similar to $\delta$, the $c_{13}$ parameter has a maximum imprint in the data at intermediate scattering angles, which is consistent with the relationship between these parameters in (equation 12). As for the vertical and horizontal waves speeds, the radiation patterns of $c_{11}$ and $c_{33}$ have amplitudes of the same order of magnitude for distinct ranges of incident angles. Therefore, a significant influence of these parameters on the data is expected, as well as a limited trade-off between them. The amplitude of the radiation patterns of the $c_{13}$ parameter is around three-times less than that of the $c_{11}$ and $c_{33}$ parameters, and shows a greater influence on the data than $\delta$ in the two previous parameterizations.
Grid analysis of the misfit function

Now, we assess the convexity of the misfit function as a function of the parameterization. We consider an anisotropic model corresponding to an inclusion in a homogeneous elliptic background model (Figure 5a). An elliptic background model is used to prevent excitation of unwanted shear waves at the source position, which would impact on the values of the misfit function (Grechka et al., 2004). The model space is parametrized with three model parameters describing the anisotropic properties of the homogeneous inclusion, the geometry of which is assumed to be known. We analyze the variations of the misfit function for the three kinds of parameterizations described in the previous section when the values of the three parameters in the inclusion deviate from the true ones. The vertical velocity and the Thomsen’s parameters $\delta$ and $\epsilon$ are (3 km/s, 0.05, 0.05) in the background and (3.3 km/s, 0.1, 0.2) in the inclusion. Nine frequencies between 4.8 Hz and 19.5 Hz are used for the computation of the misfit function. The radius of the inclusion is 300 m. The sources and receivers surround the inclusion to provide a complete seismic illumination of the target in terms of incidence ($\phi$) and scattering ($\theta$) angles (Figure 5a). The maximum deviations from the true parameters in the inclusion are $\pm 0.5$ km/s, $\pm 0.1$ and $\pm 0.2$ for $V_{P0}$, $\delta$ and $\epsilon$, respectively. These values are representative of the model perturbations that are expected to be found in realistic geological targets by FWI. For a fair comparison of the different parameterizations, the misfit function is sampled over the same model space, whatever the parameterization is: the model space is defined by the minimum and maximum values of the parameters for a reference parameterization (for example, $c_{11}$, $c_{13}$, $c_{33}$) and the minimum and maximum values for the other parameterizations are inferred from the the ones of the reference parameterization using the relationships between parameters, equation 12.
The contours of the misfit function for the \((V_{P0}, \delta, \epsilon)\) parameterization are shown in Figure 6a,b with their corresponding one-dimensional profiles in Figure 7a,b. The misfit function in the \((V_{P0}, \epsilon)_{\delta=0.1}\) plane shows that \(V_{P0}\) has a greater influence on the data than \(\epsilon\), as the contours of the misfit function become tighter and tilt towards the \(\epsilon\)-axis (Figure 6a). The misfit function in the \((V_{P0}, \delta)_{\epsilon=0.2}\) plane shows that the influence of \(\delta\) is negligible, as the contours of the misfit function are almost parallel to the \(\delta\)-axis (Figure 6b). The presence of local minimum are shown in the profiles of the misfit function plotted as a function of \(\delta\) for the true values of \(V_{P0}\) and \(\epsilon\) (Figure 7b, lower panel). The dominant influence of \(V_{P0}\) relative to \(\epsilon\) and more significantly to \(\delta\) is consistent with the radiation pattern analysis presented in the previous section (Figure 3b).

The misfit function for the \((V_{P0}, V_h, \delta)\) parameterization is shown in Figure 6c,d with their corresponding one-dimensional profiles in Figure 7c,d. The misfit function in the \((V_{P0}, V_h)_{\delta=0.1}\) plane (Figure 6c) shows that the influences of \(V_{P0}\) and \(V_h\) are of the same order of magnitude for the full acquisition illumination, as the contours of the misfit function are almost circles in the \((V_{P0}, V_h)\) plane. This is consistent with the radiation patterns of \(V_{P0}\) and \(V_h\) shown in Figure 3e. Indeed, if a surface reflection acquisition would have been considered (source and receivers on the top side), the contours of the misfit function would have shown a dominant influence of \(V_{P0}\), while a vertical cross-hole acquisition would have shown a dominant influence of \(V_h\). The misfit function in the \((V_h, \delta)_{V_{P0}=3300\text{m/s}}\) plane (Figure 6d) shows that the influence of \(\delta\) is negligible, as the contours of the misfit function are parallel to the \(\delta\)-axis.

The contours of the misfit function for the parameterization \((c_{11}, c_{13}, c_{33})\) are shown in Figure 6e,f with their one-dimensional profile in Figure 7e,f. The misfit function in the \((c_{11}, c_{33})_{c_{13}=23.958\epsilon+9P_n}\) plane defined by the true value of \(c_{13}\) (Figure 6e) confirms that \(c_{11}\)
and $c_{33}$ have equal influence on the data, as the contours are almost circular. The profiles of the misfit function with respect to $c_{11}$ and $c_{33}$ (Figure 7e) confirm the equal sensitivity of the misfit function with respect to these two parameters. The elastic coefficient $c_{13}$ has a similar influence on the data as $c_{33}$ in the vicinity of the minimum of the misfit function (Figure 6f). However, a lack of sensitivity of the misfit function can be noted when both $c_{13}$ and $c_{33}$ increase (Figures 6f and top panel of 7f). The same trend is shown for the ($V_{P0}$, $\delta$, $\epsilon$) and ($V_{P0}$, $V_h$, $\delta$) parameterizations, where the sensitivity of the misfit function decreases as $\delta$ increases (Figure 7b,d).

**Synthetic examples of full waveform inversion: inclusion model**

**Experimental setup**

In this section, we validate the conclusions of the radiation pattern and of the grid analysis against synthetic examples of FWI, where we seek to image an inclusion in the middle of a homogeneous background model (Figure 5b) using an idealized acquisition geometry surrounding the target. The complete illumination of the target allows us to remove the influence of the acquisition geometry in the FWI results, and hence to focus on the filtering effects induced by the model parameterization. The medium properties are the same as the grid analysis: the vertical velocity and the Thomsen’s parameters $\delta$ and $\epsilon$ are equal to 3 km/s, 0.05 and 0.05, respectively, in the background and equal to 3.3 km/s, 0.1 and 0.2 in the inclusion. Compared to the grid analysis, the radius of the inclusion is decreased to 100 m to remain in the domain of validity of the Born approximation and to prevent cycle-skipping artifacts. Nine frequencies between 4.8 Hz and 19.5 Hz are inverted simultaneously with a total of 15 iterations. For all of the inversions shown below, the misfit function is
decreased by two orders of magnitude during the iterations.

We consider the same parameter scaling as the radiation pattern analysis. Therefore, the misfit function is minimized with respect to parameters that are normalized by their mean value in the background model. This setting allows us to cross-validate the conclusions of the radiation pattern analysis with these FWI tests.

For the regularization, we use $m_{\text{prior}} = m_k$ in equation 6. In this case, the regularization reduces to the smoothing of the gradient and of the Hessian with the operator $\hat{W}_m^{-1}$. For the smoothing of the model perturbations (equation 3) we use horizontal and vertical correlation lengths $\tau_x$ and $\tau_z$ of 12 m for each parameter class. We use the same value $\lambda$ for the damping coefficients $\lambda_i$ in equation 7:

$$\Lambda = \lambda \mathbf{I}_{N_p \times M},$$

with

$$\lambda = 0.001 \cdot \text{Max} \left[ \text{diag} \left\{ \mathbf{J}^\dagger \mathbf{J} \right\} \right].$$

The damping coefficient $\Lambda$ is added to the diagonal approximate Hessian in equation 8 which is used as an initial guess for the L-BFGS optimization. A factor of 0.001 in equation 15 is similar to the one used in previous FWI application (Ravaut et al., 2004). The approximate Hessian is a $N_p \times N_p$ block matrix, where each diagonal block represents the zero-lag correlation between the partial derivative of the wavefields with respect to model parameters belonging to the same class (Figure 8). If a diagonal block of the approximate Hessian has much smaller coefficients than the selected damping, it is unlikely that the inversion will succeed in correctly scaling the model perturbations of the corresponding parameter class, because the high damping term will cancel out the scaling action of the Hessian. We have shown that the normalization of the model parameters by their mean
value will steer the inversion towards the parameters with the strongest influence on the
data at the expense of the parameters that have the weakest. Therefore, we expect the FWI
results to reflect the conclusion of the radiation pattern analysis in terms of sensitivity of
seismic data to anisotropic parameters.

As an example, we show in Figure 8 the Hessian computed for the normalized parameters
in the \((V_P^0, \delta, \epsilon)\) parameterization. The diagonal block associated with \(V_P^0\) has coefficients
several order of magnitudes higher than those associated with \(\delta\) and \(\epsilon\) (Table 2). This is
consistent with the radiation patterns of \(V_P^0\), \(\delta\), and \(\epsilon\) in Figures 3b and 4b. In this case,
the damping parameter \(\lambda\) is equal to 0.1% of the maximum coefficient of the \(V_P^0 \times V_P^0\)
block, and it is unlikely that the true amplitudes of \(\delta\) and \(\epsilon\) can be reconstructed. The
maximum coefficients of each block of the approximate Hessian are provided in Table 2 for
the 3 parameterizations investigated here.

In the following we present three categories of FWI tests for different subsurface param-
eterization.

The first category (referred to as mono-parameter test) performs three independent
mono-parameter FWI where only one parameter class is updated during inversion. For
each of the three monoparameter inversions, we compute the data to be inverted in the
true homogeneous background models for the two parameter classes that are not updated
during the inversion and in the true model with the inclusion for the updated parameter
(Figure 5b). The initial models of the three parameter classes are the true homogeneous
background models. For each monoparameter FWI, pressure data are inverted and the
resulting data residuals contain only the imprint of the parameter to be updated. For
these mono-parameter tests, the FWI results are only impacted by the bandpass filtering
effects induced by the model parameterization, and no trade-off effects are investigated. The damping term in equation 15 is scaled to the maximum coefficient of the mono-parameter Hessian.

The second category (referred to as "multi-parameter test 1") consists of one multi-parameter FWI where the three parameter classes are jointly updated during inversion. We compute the data to be inverted in the true models that contain the inclusion for each parameter class. The initial models are the true homogeneous background models. Pressure data are inverted to jointly update the three parameter classes. The damping term is scaled to the maximum coefficient of the multi-parameter Hessian (equation 15). Comparison between the results of this multi-parameter test and those of the mono-parameter test provides insights on the trade-off between the different parameter classes and on the relative influence of the parameter classes on the inversion.

As the data residuals of multi-parameter test 1 contain the influence of the three classes of parameters, an interpretation of the trade-off effects might be difficult. This difficulty prompts us to design a second multi-parameter FWI test (referred to as multi-parameter test 2). The only difference to multi-parameter test 1 is that the inclusion is set in the true model of one parameter class only, while the true models of the two remaining parameter classes are the true homogeneous background models. In this case, the data residuals contain the influence of one parameter class only. These residuals are inverted to jointly update the three parameter classes. The leakage of the model perturbations over the different parameter classes highlight the trade-off between parameters.

For all these tests, the extraction of vertical and horizontal profiles across the center of the inclusion allows us to determine how the resolution of the reconstruction is influenced
by the radiation pattern of the model parameters as a function of the incidence angle.

*Type 1 parameterizations: one wavespeed + two Thomsen parameters*

The results of the mono-parameter test for the \((V_P, \delta, \epsilon)\) parameterization are shown in Figure 9. Consistent with the almost isotropic radiation pattern of the vertical wavespeed in the \((V_P, \delta, \epsilon)\) parameterization (Figures 3b and 4b), the \(V_P\) parameter is reconstructed with a good resolution along the vertical and horizontal directions (Figure 9a-c). The inversion of \(\delta\) reconstructs the true peak-to-peak amplitude of the model perturbations (Figure 9d-f). However, the horizontal and vertical profiles show a deficit of small and high wavenumbers, which is consistent with the influence of \(\delta\) at intermediate angles (Figures 3b and 4b). The footprint of the radiation pattern of \(\delta\) is clearly visible in the FWI model (Figure 9d). The horizontal and vertical profiles of \(\delta\) are strictly identical, which is consistent with the symmetry of the radiation patterns with respect to the axis \(\theta = 90^\circ\) and \(\theta = 270^\circ\) as seen in Figures 3 and 4. The reconstruction of \(\epsilon\) shows a more complex anisotropic reconstruction with correct peak-to-peak amplitudes of the model perturbations (Figure 9g-i). The vertical profile shows a deficit of small and high wavenumbers consistent with the lack of influence of \(\epsilon\) on vertical normal-incidence reflection and vertical transmitted wavepaths, respectively. The complex shape of the horizontal profile might result from a deficit of intermediate wavenumbers. Indeed, the large wavelengths of the horizontal profile of the \(\epsilon\) perturbation are reconstructed from transmitted horizontal wavepaths, while the short wavelengths are reconstructed from short-spread reflections that propagate sub-horizontally. A lack of sensitivity of the wide-spread reflections to \(\epsilon\) might explain the lack of intermediate wavenumbers in the reconstruction of the horizontal profile of \(\epsilon\) (Figure 1f-i).
The results of the multi-parameter test 1 for the \((V_{P0}, \delta, \epsilon)\) parameterization are shown in Figure 10. Now, the amplitudes of the wavespeed perturbations are now overestimated (Figure 10a-c) while the perturbations of the Thomsen’s parameter perturbations are strongly underestimated (Figure 10d-i). These amplitude effects result from the trade-off between parameters and the dominant influence of \(V_{P0}\) on the inversion. The reconstruction of the wavespeed differs significantly in the vertical and horizontal profiles (Figure 10b-c) although they should be nearly identical (Figure 9b-c). The vertical profile shows a deficit of small wavenumbers in the reconstructed \(V_{P0}\) model (a high-pass filtering of the inclusion profile would give a similar shape to the reconstructed model) while the horizontal profile shows a more complex shape, suggesting a deficit of intermediate wavenumbers in the spectrum of the reconstructed model.

These trade-off effects are more clearly shown by the multi-parameter test 2 where the inclusion is present in the true \(\epsilon\) model only (Figure 11). Most of the model perturbations are contained within the vertical velocity model (Figure 11a-c). This highlights the dominant influence of \(V_{P0}\) for the chosen parameterization and scaling, and the trade-off between the vertical velocity and \(\epsilon\).

\textit{Type 2 parameterizations: two wavespeeds + one Thomsen parameter}

The results of the mono-parameter test for the \((V_{P0}, \delta, V_h)\) parameterization are shown in Figure 12. The vertical wavespeed is well reconstructed along both the horizontal and vertical directions (Figure 12a-c). The narrower radiation pattern of the \(V_{P0}\) parameter in \((V_{P0}, \delta, V_h)\) parameterization relative to the \((V_{P0}, \delta, \epsilon)\) parameterization yields to some Gibbs effects. The \(V_h\) model is close to the \(\epsilon\) model obtained by mono-parameter FWI.
in the \((V_{P0}, \delta, \epsilon)\) parameterization (compare Figures 12g-i and 9g-i) because the radiation
patterns of \(V_h\) in the \((V_{P0}, \delta, V_h)\) parameterization and of \(\epsilon\) in the \((V_{P0}, \delta, \epsilon)\) parameterization
are similar (Figures 3b, e and 4b, e). The \(\delta\) model obtained by mono-parameter FWI in
the \((V_{P0}, \delta, V_h)\) parameterization is identical to the one obtained by mono-parameter FWI in
the \((V_{P0}, \delta, \epsilon)\) parameterization (compare Figures 12d-f and 9d-f), as their radiation patterns
have strictly identical shape.

Results of the multi-parameter test 1 for the \((V_{P0}, \delta, V_h)\) parameterization are shown in
Figure 13. We show differences between the vertical and horizontal profiles of \(V_{P0}\) obtained
by multiparameter and monoparameter FWI (Figure 13a-c) suggesting trade-off effects.
The horizontal profile of \(V_{P0}\) shows overestimated peak-to-peak amplitude of the model
perturbations due to artificial negative perturbations. The difference between the trade-off
effects in the vertical and horizontal profiles are caused by the non-symmetrical radiation
patterns of \(V_{P0}\) and of \(V_h\) with respect to the axis \(\theta = 90^\circ\) and \(\theta = 270^\circ\). The model
perturbations of \(V_h\) have similar amplitudes for the mono and the multi-parameter tests 1
(compare Figures 12g-i and 13g-i). This differs from the model perturbations of \(\epsilon\) which
shows stronger amplitudes in the mono-parameter test than in the multi-parameter test 1
(compare Figures 9g-i and 10g-i). Therefore, \(V_h\) has a stronger influence in the data relative
to \(\epsilon\), as highlighted by the amplitude of their respective radiation patterns (Figures 3b, e
and 4b, e).

The results of the multi-parameter test 2 is shown in Figure 14 where the inclusion is
present in the true \(V_{P0}\) model. The results show a leakage of the \(V_{P0}\) model perturbations
in the \(V_h\) and \(\delta\) models.

The results of the multi-parameter test 1 for the \((V_{NMO}, \delta, V_h)\) parameterization are
shown in Figure 15. The reconstruction of $V_{NMO}$ and $V_h$ are close to those of $V_P^0$ and $V_h$
obtained with the $(V_P^0, \delta, V_h)$ parameterization. However, unlike the reconstruction of $V_P^0$
(Figure 12b), the amplitudes in the vertical profile of $V_{NMO}$ (Figure 15b) are underesti-
mated. Moreover, the perturbations of $\delta$ have the wrong polarity. These artifacts result
probably from the significant trade-off between $V_{NMO}$ and $\delta$ as well as the stronger influence
of $\delta$ in the $(V_{NMO}, \delta, V_h)$ parameterization relative to the $(V_P^0, \delta, V_h)$ one (Figures 3f and
4f, gray and black lines).

Type 3 parameterization: elastic moduli

The results of the multi-parameter test1 for the $(c_{11},c_{33},c_{13})$ parameterization are shown
in Figure 16. We first note that the vertical and horizontal profiles of the $c_{13}$ model are
not identical, as they should be according to the symmetric radiation pattern of $c_{13}$ in
Figure 3g. However, the amplitudes of the $c_{13}$ perturbations are better reconstructed than
those of $\delta$ for the first two parameterizations. The wavenumber contents of both $c_{11}$ and
$c_{33}$ are similar to that of $V_h$ from the $(V_P^0, \delta, V_h)$ parameterization because their radiation
patterns have identical elliptical shapes (Figures 3e,g and 4e,g). The radiation pattern of
$V_P^0$ in the $(V_P^0, \delta, V_h)$ parameterization spans a broader scattering-angle range than $c_{33}$ in
the $(c_{11},c_{33},c_{13})$ parameterization. This explains why the vertical profile of $c_{33}$ is not re-
constructed as well as $V_P^0$ in the $(V_P^0, \delta, V_h)$ parameterization. The multiparameter FWI
reconstructions based upon the $(c_{11},c_{33},c_{13})$ parameterization are clearly hampered by sig-
ificant trade-off artifacts resulting from the significant influence of $c_{13}$ on the horizontal
and vertical wavepaths. This is shown, for example, by the differences between the verti-
cal and horizontal profiles of $c_{13}$, which should be identical according the symmetry of its
radiation pattern in Figures 3g and 4g.
DISCUSSION

In the present study, we have identified two main categories of parameterization for acoustic VTI media. For the type 1 parameterization, the wavespeed parameter has a dominant influence on the data, and has a radiation pattern that spans over the full range of scattering angles. The two other Thomsen parameters, have weaker influence on the data over a narrower range of scattering angles. As significant trade-off was shown between the wavespeed and the two Thomsen parameters, one reliable strategy might be to keep the Thomsen parameters fixed during FWI and only update the dominant parameter, provided that sufficiently accurate starting models of the Thomsen parameters are available. If the strategy is relevant, we would tend to favor a parameterization that involves the vertical velocity rather than the NMO velocity, because $\delta$ has a stronger influence on the data when combined with the NMO velocity. In this case, a more accurate initial model of the $\delta$ parameter would be required to guarantee a reliable update of the velocity model during mono-parameter FWI. We would also recommend a parameterization that involves the vertical velocity rather than the horizontal velocity if a smooth background model of $\epsilon$ is available. When combined with the vertical velocity in the parameterization, the smooth background model of $\epsilon$ should allow sufficiently accurate prediction of the influence of $\epsilon$ on the wide-aperture components of the data.

Alternatively, a parameterization that involves two wavespeeds and one Thomsen parameter can be used. This kind of parameterization yields to limited trade-off between the vertical and horizontal wavespeeds at intermediate scattering angles, while $\delta$ has a very minor influence on the data. In this case, a joint update of the two wavespeeds should be possible with different resolutions. For surface acquisition, the intermediate and short
wavelengths of the vertical velocity are reconstructed from the short and intermediate scattering angles, while the intermediate and long wavelengths of the horizontal velocity are reconstructed from the intermediate and wide scattering angles.

The choice between the type 1 and type 2 parameterizations might be driven by the accuracy of the available initial models and the geometry of the acquisition. For narrow-azimuth reflection acquisitions, the type 1 parameterization might be chosen to update the dominant parameter $V_{P0}$, as $\epsilon$ is expected to have a weak influence. For wide-aperture surface acquisitions or cross-hole acquisitions, the type 1 parameterization might be also chosen, if the background model of $\epsilon$ describes the large wavelengths of the subsurface sufficiently accurately. If this condition is not satisfied, trade-off effects between $V_{P0}$ and $\epsilon$ might hamper the reconstruction of the vertical velocity model, in particular if $\epsilon$ is kept fixed during the inversion (residuals resulting from inaccuracies of the $\epsilon$ background model will be interpreted with vertical velocity perturbations).

If the large-scale variations of the horizontal wavespeeds need to be updated to match the traveltimes of wide-aperture data (diving waves and super-critical reflections) and if the initial models of $V_{P0}$ and $\delta$ are sufficiently accurate to predict the traveltimes of short-spread reflections, the type 2 parameterization may be favored because $V_h$ has a stronger influence on the data than $\epsilon$, and hence will be updated more easily with a more limited trade-off with $V_{P0}$.

For the type 2 parameterization, we recommend avoiding a parameterization that involves the NMO velocity combined with the horizontal velocity, because the influence of $\delta$ on the data is higher thus requiring a more accurate background model for $\delta$ to allow the reconstruction of both NMO and horizontal velocities. We would also tend to favor the
\((V_{P0}, \delta, V_h)\) parameterization as opposed to the \((c_{11}, c_{33}, c_{13})\) one for similar reasons.

The present study has not investigated the tuning of the inversion algorithm to optimize the FWI results. We rather focus on the sensitivity of seismic data to different VTI parameterization. Parameter classes can be arbitrarily scaled such that some of the parameters have a dominant numerical weight in the Hessian, whatever their physical influence on the data. In this case, the inversion may be steered towards the reconstruction of these parameters. The challenge consists in finding an optimal trade-off between the need to improve the conditioning of the Hessian by giving a similar weight to all of the parameters in the Hessian, and the risk to get unstable results by giving too much weight to secondary parameters whose influence in the data is dominated by data noise.

**CONCLUSION**

Updating multiple parameters with FWI is a difficult challenge. This is because the ill-posed nature of the inversion increases when the number of degrees of freedom in the model space increases. In this study, we propose a pragmatical approach to analyze the influence of different subsurface parameterizations on wide-aperture data, along with some criteria to choose a suitable parameterization for either mono-parameter and multi-parameter FWI in VTI media under the acoustic approximation. Our approach relies on the numerical analysis of the radiation patterns of different parameter classes. It gives some insights into the influence of the parameters on the data as a function of the scattering angle. This analysis is validated with synthetic examples that are performed in a simple subsurface model with an ideal illumination to highlight the bandpass filtering and trade-off effects the parameterization induce. One strategy consists of choosing a parameterization that combines one parameter, updated with FWI, with a dominant influence on the data over a broad range
of scattering angles, with two secondary parameters (kept fixed during FWI) with much smaller influences on the data. A parameterization that combines the vertical wavespeed and the Thomsen parameters $\delta$ and $\epsilon$ would be the most suitable for this purpose. This mono-parameter strategy should be suitable for narrow-azimuth reflection data for which the horizontal wavespeed has a weak influence on the data. It might be suitable for wide-azimuth/wide-aperture surface data or cross-hole data provided that the background model for $\epsilon$ describes the large wavelengths of the medium sufficiently accurately. If this condition is not satisfied, trade-off effects between the vertical velocity and $\epsilon$ might hamper the reconstruction of the vertical velocity. Alternatively, we have considered parameterizations that combine two wavespeeds and the Thomsen parameter $\delta$. Because the two wavespeeds have significant influence on the data for different ranges of scattering angles, a limited-trade-off between the two wavespeeds is expected. In this case, the simultaneous update of the two wavespeeds should be possible yielding a model with two different resolutions: the large-to-intermediate wavelengths of the horizontal wavespeed can be updated from the wide-aperture data while the short-to-intermediate wavelengths of the vertical wavespeed can be updated from the short-spread reflection wavefields. This parameterization should be suitable for wide-azimuth/wide-aperture surface data assuming that the initial vertical velocity model is sufficiently-accurate to match the traveltimes of the short-spread reflections. This condition should be satisfied as the inversion will be insensitive to the large-scale variations of the vertical velocity and hence will fail to update them. Although the theoretical analysis of the radiation pattern is quite general, we have steered the multi-parameter inversion, towards the parameters that have the strongest physical influence on seismic data. The design of suitable regularization and parameter scaling strategies, as well as the incorporation of prior constraints in the inversion to reconstruct parameters that have a
weak influence in the data will be the aim of future studies.

ACKNOWLEDGMENTS

This study was funded by the SEISCOPE consortium http://seiscope.oca.eu, sponsored by BP, CGGVeritas, ENI, EXXON-MOBIL, PETROBRAS, SAUDI ARAMCO, SHELL, STATOIL and TOTAL. The linear systems were solved with the MUMPS package, available on http://graal.ens-lyon.fr/MUMPS/index.html. The mesh generation was performed with the help of TRIANGLE, available on http://www.cs.cmu.edu/~quake/triangle.html. This study was granted access to the high-performance computing facilities of the SIGAMM (Observatoire de la Côte d’Azur) and to the HPC resources of [CINES/IDRIS] under the allocation 2010- [project gao2280] made by GENCI (Grand Equipement National de Calcul Intensif). We gratefully acknowledge both of these Facilities and the support of their staff. We would like to thank the Associate Editor for the time dedicated to the improvement of our manuscript and for his constructive comments. We thank René-Edouard Plessix (Shell), Andrew Ratcliffe (CGGVeritas) and two anonymous reviewers for their constructive comments, which significantly improved the manuscript.

APPENDIX A

ACOUSTIC VTI MODELING

In this appendix we review the two VTI acoustic forward modeling operators used for wave propagation modeling and gradient computation with the adjoint state method. These two forward modeling operators are different for reasons explained below and in Appendix B.

We start from the two-dimensional P-SV velocity-stress equation in VTI media where
the stiffness coefficient $c_{44}$ is set to zero (equation 1):

$$
-i\omega v_x = b \left( \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right),
$$

$$
-i\omega v_z = b \left( \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right),
$$

$$
-i\omega \sigma_{xx} = c_{11} \frac{\partial v_x}{\partial x} + c_{13} \frac{\partial v_z}{\partial z},
$$

$$
-i\omega \sigma_{zz} = c_{13} \frac{\partial v_x}{\partial x} + c_{33} \frac{\partial v_z}{\partial z}.
$$

where $i = \sqrt{-1}$. A change of variables can be applied to the stress vector to explicitly introduce the pressure wavefield $p$ (Brossier et al., 2008; Brossier, 2011). The new stress components are given by $(p, q) = (\frac{\sigma_{xx} + \sigma_{zz}}{2}, \frac{\sigma_{xx} - \sigma_{zz}}{2})$, which can be interpreted as the mean and deviatoric components.

After the change of variables, the first-order hyperbolic P-SV system is given by:

$$
-i\omega v_x = b \left( \frac{\partial (p + q)}{\partial x} + \frac{\partial r}{\partial z} \right) + b f_x \delta(x - x_s),
$$

$$
-i\omega v_z = b \left( \frac{\partial r}{\partial x} + \frac{\partial (p - q)}{\partial z} \right) + b f_z \delta(x - x_s),
$$

$$
-i\omega p = \frac{c_{11} + c_{13}}{2} \frac{\partial v_x}{\partial x} + \frac{c_{13} + c_{33}}{2} \frac{\partial v_z}{\partial z} - i\omega p_0 \delta(x - x_s),
$$

$$
-i\omega q = \frac{c_{11} - c_{13}}{2} \frac{\partial v_x}{\partial x} + \frac{c_{13} - c_{33}}{2} \frac{\partial v_z}{\partial z}.
$$

In equation A-2, we introduce two-dimensional point sources (line sources in three dimensions): vertical and horizontal forces, $f_x \delta(x - x_s)$ and $f_z \delta(x - x_s)$, respectively, or explosive source, $p_0 \delta(x - x_s)$, where the Dirac delta function and the source position are denoted by $\delta$ and $x_s$, respectively. This approach based on the P-SV elastodynamic system is close to that developed by Duveneck et al. (2008). The difference is that we use the first-order system to perform seismic modeling, while Duveneck et al. (2008) eliminate the particle-velocity wavefields from the first-order system to build a second-order wave equation for normal stresses, following a parsimonious approach. Second, we use a change of variables
on the normal stresses, which gives a direct access to the pressure.

This system of linear equations, equation A-2, can be recast in matrix form (Marfurt, 1984) as:

\[ \mathbf{A}(\mathbf{m}(\mathbf{x}), \omega)\mathbf{w}(\mathbf{x}, \omega) = \mathbf{s}_A(\mathbf{x}, \omega), \]

(A-3)

where \( \mathbf{A} \) is referred to as the velocity-stress impedance matrix. The vectors \( \mathbf{w} = (v_x, v_z, p, q) \) and \( \mathbf{s}_A = (bf_x \delta(x-x_s), bf_{xz} \delta(x-x_s), \omega p_0 \delta(x-x_s), 0) \) are the monochromatic velocity-stress wavefields and source, respectively. We discretize the velocity-stress wave equation A-2 with a nodal formulation of the discontinuous Galerkin method, based on Lagrange polynomials of order 0, 1, or 2 (referred to as P0, P1, and P2, respectively) to perform seismic modeling in VTI acoustic media (Brossier et al., 2010a; Brossier, 2011).

A second-order wave equation for particle velocities, useful for FWI implementation, can be inferred from the velocity-stress wave equation A-2 by eliminating the stress wavefields \( p \) and \( q \) in the first and second rows as follows:

\[
\omega^2 \rho v_x + \frac{\partial}{\partial x} \left( c_{11} \frac{\partial v_x}{\partial x} + c_{13} \frac{\partial v_z}{\partial z} \right) = \omega p_0 \frac{\partial \delta(x-x_s)}{\partial x} + \omega f_x \delta(x-x_s),
\]

\[
\omega^2 \rho v_z + \frac{\partial}{\partial x} \left( c_{13} \frac{\partial v_x}{\partial x} + c_{33} \frac{\partial v_z}{\partial z} \right) = \omega p_0 \frac{\partial \delta(x-x_s)}{\partial z} + \omega f_z \delta(x-x_s). \quad (A-4)
\]

This system of linear equations can be recast in matrix form as:

\[ \mathbf{B}(\mathbf{m}(\mathbf{x}), \omega)\mathbf{v}(\mathbf{x}, \omega) = \mathbf{s}_B(\mathbf{x}, \omega), \]

(A-5)

where the particle-velocity vector and the source are denoted respectively by \( \mathbf{v} = (v_x, v_z) \) and \( \mathbf{s}_B = (\omega p_0 \frac{\partial \delta(x-x_s)}{\partial x} + \omega f_x \delta(x-x_s), \omega p_0 \frac{\partial \delta(x-x_s)}{\partial z} + \omega f_z \delta(x-x_s)) \). The second-order system in equation A-5 is self-adjoint, making this formulation suitable for computing the gradient of the misfit function with the adjoint-state method (see Appendix B). We discretize the forward-problem operator \( \mathbf{B} \) with the finite-volume scheme of Brossier et al. (2008), which is equivalent to the P0 discontinuous Galerkin scheme.
The first and second-order hyperbolic systems in equations A-3 and A-5 give the same solutions for the particle velocities as long as the correct relationships between the sources \( s_A \) and \( s_B \) are as follows:

\[
\begin{align*}
\mathbf{s}_A &= (bf_x \delta(x - x_s), bf_z \delta(x - x_s), 0, 0) \quad \rightarrow \quad \mathbf{s}_B = (i\omega f_x \delta(x - x_s), i\omega f_z \delta(x - x_s)) \quad \text{for a force.} \\
\mathbf{s}_A &= (0, 0, \nu \rho \partial \delta(x - x_s) \partial x, 0) \quad \rightarrow \quad \mathbf{s}_B = (i\omega \rho \partial \delta(x - x_s) \partial x, i\omega \rho \partial \delta(x - x_s) \partial z) \quad \text{for an explosion.} 
\end{align*}
\]

(A-6)

Indeed, the explosive source in the velocity-stress system is transformed into vertical and horizontal dipoles in the second-order wave equation.

**APPENDIX B**

**COMPUTING THE GRADIENT OF THE MISFIT FUNCTION WITH THE ADJOINT-STATE METHOD**

In this appendix we review the derivation of the gradient of the misfit function with the adjoint-state method based on the two forward modeling operators \( A \) and \( B \), which are introduced in Appendix A. Our motivation behind the use of these two modeling operators is many-fold. On one hand, first-order differential operators in matrix \( A \) are easier to discretize with the discontinuous Galerkin method for seismic modeling. On the other hand, the second-order forward problem operator \( B \) allows us to derive the expression of the gradient of the misfit function from the self-adjoint operator. Moreover, it allows us to save memory during the computation of the gradient as stress wavefields do not need to be kept in memory after seismic modeling (Brossier, 2011). In addition, when pressure data are inverted, the velocity-stress forward problem operator \( A \) allows us to introduce pressure residuals into the source term of the adjoint equation in Appendix A easily. The self-adjointness of the forward modeling operator (and hence its symmetry) is not very useful in frequency-
domain modeling because one can multiply the forward modeling operator or its transpose to a vector to compute the adjoint-state variable (Plessix and Cao, 2011). However, for time-domain modeling, the self-adjointness of the forward modeling operator allows one to use the same numerical scheme to compute the state and adjoint-state wavefields, and hence can greatly facilitate the FWI implementation (Castellanos et al., 2011).

To compute the gradient of the misfit function with the adjoint-state method, we introduce a Lagrangian function \( \mathcal{L} \)

\[
\mathcal{L}(\mathbf{d}_{\text{cal}}, \mathbf{v}, \mathbf{m}, \beta_1, \beta_2) = \frac{1}{2} < \mathbf{d}_{\text{cal}} - \mathbf{d}_{\text{obs}} | \mathbf{d}_{\text{cal}} - \mathbf{d}_{\text{obs}} >_D + \Re < \beta_1 | \mathbf{Bv} - \mathbf{s}_B >_V \\
+ \Re < \beta_2 | \mathbf{d}_{\text{cal}} - \mathbf{RCv} >_D.
\]

where the complex-valued subspaces \( D \) and \( V \) span over the receiver positions and the full computational domain, respectively (in equation B-1, we omit the regularization terms for reasons of compactness). The inner product between \( \mathbf{x} \) and \( \mathbf{y} \) is denoted by \( < \mathbf{x} | \mathbf{y} >_D = \int_D \mathbf{x}^\ast \mathbf{y} \). The restriction operator \( \mathbf{R} \) samples the pressure wavefield at the receiver positions. The Lagrangian function \( \mathcal{L} \) in equation B-1, is equivalent to the misfit function \( \mathcal{C} \) subject to the constraint that the state equations are satisfied. The state variables are the particle-velocity wavefields \( \mathbf{v} \) and the pressure data \( \mathbf{d}_{\text{cal}} \) at the receiver positions. The Lagrange multipliers \( \beta_1 \) and \( \beta_2 \) are the adjoint state variables. The two state equations allow us to introduce the self-adjoint forward modeling operator \( \mathbf{B} \) for particle velocities and the relationships between the modeled particle velocities and the pressure data \( \mathbf{d}_{\text{cal}} \), through the operator \( \mathbf{C} \). The operator \( \mathbf{C} \) corresponds to the third row of equation A-2:

\[
\mathbf{C} = \begin{bmatrix} \frac{i}{\omega} c_{11} + c_{13} & \frac{i}{\omega} c_{13} + c_{33} \\
2 & 2 \frac{\partial}{\partial z} \\
\end{bmatrix}.
\]

At the saddle points of \( \mathcal{L} \) with respect to the state and adjoint-state variables we have

\[
\frac{\partial \mathcal{C}(\mathbf{m}(k))}{\partial \mathbf{m}} = \frac{\partial \mathcal{L}}{\partial \mathbf{m}} = < \beta_1 | \frac{\partial \mathbf{B}}{\partial \mathbf{m}} \mathbf{v} > + < \beta_2 | \mathbf{R} \frac{\partial \mathbf{C}}{\partial \mathbf{m}} \mathbf{v} >.
\]
If we assume that the medium properties at the receiver positions are known, the term $< \beta_2 | R \frac{\partial C}{\partial m} v >$ vanishes. The adjoint-state variables $\beta_1$ and $\beta_2$ satisfy the adjoint-state equations $\partial C / \partial v = 0$ and $\partial C / \partial d_{cal} = 0$:

$$B^T \beta_1^* = C^\dagger R^T \beta_2^*, \quad \beta_2 = \Delta d.$$

(B-4)

Exploiting the symmetry of matrix $B$ and injecting the expression of $\beta_2$ into the right-hand-side of the first state equation yields:

$$B \beta_1^* = C^\dagger R^T \Delta d^*.$$

(B-5)

Equation B-5 indicates that the adjoint wavefield $\beta_1$ can be computed by back-propagation of the pressure residuals after conversion of these into particle velocities for interfacing with the second-order wave equation. The adjoint of $C$ allows us to perform the pressure-to-particle-velocity conversion, and is derived from the first and second rows of equation A-2:

$$C^\dagger = \frac{i b}{\omega} \left[ \frac{\partial}{\partial x} \quad \frac{\partial}{\partial z} \right]^T,$$

(B-6)

where we have assumed that the wavefield $q$ is zero at the receiver positions. Indeed, the aim of the operator $C^\dagger$ in equation B-5 is to represent the residual pressure source by two dipoles through the gradient operator. Because building $B$ is difficult with the discontinuous Galerkin method, we infer the state variables $v$ and the adjoint-state variables $\beta_1$ from the numerical solutions of the first-order wave equation in equation A-3. For this, the correct relationships between the sources in equation A-6, should be used to guarantee that the two wave equations give the same solutions. This leads to:

$$v = PR_v w \quad \text{and} \quad \beta_1 = PR_v \beta_A,$$

(B-7)

where $w$ and $\beta_A$ are solutions of the first-order wave equation:

$$A w = s_A, \quad A \beta_A^* = \frac{b}{i \omega} R_p^T R^T \Delta d^*.$$

(B-8)
$\mathbf{R}_v$ and $\mathbf{R}_p$ are the restriction operators of the velocity-stress vector to the particle velocities and to the pressure, respectively. The operator $\mathbf{P}$ projects the solutions of the discontinuous Galerkin method computed at the nodes of the P0/P1/P2 mesh onto the barycenter of the triangular elements, for consistency with the finite-volume discretization of the forward-modeling operator $\mathbf{B}$ (Brossier, 2011). Notice that the pressure residuals on the right-hand side of equation B-8 need to be scaled by the factor $\frac{b}{\omega}$ to derive the correct expression of the adjoint wavefield $\beta_1$. Inserting the expression of $\mathbf{v}$ and $\beta_1$ in equations B-7 and B-8 into the expression of the gradient in equation B-3 yields:

$$
\frac{\partial C(m^{(k)})}{\partial m} = \Re \left\{ (\mathbf{P} \mathbf{R}_v \left( \mathbf{A}^{-1} s_A \right))^T \left( \frac{\partial \mathbf{B}}{\partial m} \right)^T \mathbf{P} \mathbf{R}_v \left( \mathbf{A}^{-1} \frac{b}{\omega} \mathbf{R}_p^T \mathbf{R}^T (\mathbf{R} \mathbf{R}_p \mathbf{A}^{-1} s_A - \mathbf{d}_{\text{obs}})^* \right) \right\}.
$$

Equation B-9 is the detailed expression of equation 9.
REFERENCES


SEG Technical Program Expanded Abstracts, 30, 2793–2798.


Wu, R.-S. and K. Aki, 1985a, Elastic wave scattering by a random medium and the small scale inhomogeneities in the lithosphere: Journal Geophysical Research, 90, 10237–10261.


Zhou, B. and S. A. Greenhalgh, 2009, On the computation of the Fréchet derivatives for

Type 1 parameterization | Type 2 parameterization | Type 3 parameterization
---|---|---
\((V_{P0},\delta,\epsilon),(V_{NMO},\delta,\epsilon),(V_{NMO},\delta,\eta)\) | \((V_{P0},\delta,V_{h}),(V_{NMO},\delta,V_{h})\) | \(c_{11}, c_{33}, c_{13}\)

Table 1: Parameterization of VTI media in the acoustic approximation \((c_{44} = 0)\) investigated in the present study.

| \((V_{P0},\delta,\epsilon)\) | \(5e+25 (V_{P0})\) | \(1e+22 (\delta)\) | \(2.5e+22 (\epsilon)\) |
| \((V_{P0},\delta,V_{h})\) | \(8e+24 (V_{P0})\) | \(2.5e+21 (\delta)\) | \(3e+24 (V_{h})\) |
| \((c_{33},c_{13},c_{11})\) | \(1.8e+24 (c_{33})\) | \(4.5e+24 (c_{13})\) | \(3e+24 (c_{11})\) |

Table 2: Maximum value of the diagonal coefficients of the undamped Hessian for the three parameterizations \((V_{P0},\delta,\epsilon), (V_{P0},\delta,V_{h})\) and \((c_{33},c_{13},c_{11})\).
FIGURE CAPTIONS
Figure 1: Radiation patterns for parameterization $(V_{P0}, \delta, \epsilon)$ for incidence angles $0^\circ$ (a-c), $45^\circ$ (d-f) and $90^\circ$ (g-i). These radiation patterns are shown for the $V_{P0}$ (a,d,g), $\delta$ (b,e,h) and $\epsilon$ (c,f,i) model perturbations, respectively. The incidence angle $\varphi$ is defined with respect to the vertical symmetry axis, and this defines the source position in the finite-difference modeling, as denoted by the white circle. The diffractor point associated with the model perturbation of each parameter class is located in the center of the finite-difference grid. The source-receiver scattering angle $\theta$ is labeled around the dashed white circle for each incidence angle.

Figure 2: As for Figure 1, for the $(V_{P0}, \delta, V_h)$ parameterization. The radiation patterns shown are of $V_{P0}$ (a, d, g), $\delta$ (b, e, h) and $V_h$ (c, f, i). Note how the radiation pattern of $V_{P0}$ differs from that in Figure 1 because the other parameter classes in this parameterization have changed.
Figure 3: Radiation patterns for a specular reflection on a horizontal reflector. (a) Procedure to extract the value of the partial derivative wavefields (Figures 1 and 2) at a receiver position (black square) for the specular reflection on a horizontal reflector (thick black segment in (a)), given the source position (black circle in (a)). The incidence angle $\varphi$ and the reflection angle $\theta$ are defined with respect to the vertical axis. This extraction is repeated for a source moving on the dashed circle around the diffractor point, located in the middle of the panel. The assemblage of the values of the partial derivative wavefields at the specular positions for each shot position is shown in (b-g) for different parameterizations. (b-g) Radiation patterns of the $(V_{P0}, \delta, \epsilon)$ (b), $(V_h, \delta, \epsilon)$ (c), $(V_{NMO}, \delta, \eta)$ (d), $(V_{P0}, \delta, V_h)$ (e), $(V_{NMO}, \delta, V_h)$ (f) and $(c_{33}, c_{13}, c_{11})$ (g) parameterizations with respect to the reflection angle $\theta$ as defined in (a). Note that the radiation patterns of the Thomsen parameters are magnified by a factor of 10 in b, c, d, e and f.

Figure 4: As for Figure 3, except that specular reflections from a vertical interface (thick black segment in (a)) are considered. The incidence and aperture angles are defined with respect to the horizontal axis. Notice that the radiation patterns shown here are the mirror images of those in Figure 3 (see text for details).
Figure 5: Inclusion model and acquisition geometry for the grid analysis of the misfit function (a), and for VTI FWI (b). Both sources and receivers surround the inclusion providing complete seismic illumination. The radius of the inclusion is higher in (a) to increase the sensitivity of the misfit function to the inclusion.
Figure 6: (a-b) Grid analysis of the misfit function in the \((V_{P0}, \delta, \epsilon)\) parameterization. In the cross-section of the misfit function in (a), the \((V_{P0}, \epsilon)\) plane is defined by the true value of \(\delta\) (\(\delta = 0.1\)), and the \((V_{P0}, \delta)\) plane is defined by the true value of \(\epsilon\) (\(\epsilon = 0.2\)). The dashed lines show the directions of the maximum (black) and minimum (gray) sensitivities, respectively. (c-d) As for (a-b) for the \((V_{P0}, V_h)\) (c) and \((\delta, V_h)\) (d) planes for the \((V_{P0}, \delta, V_h)\) parameterization. (e-f) As for (a-b) for the \((c_{11}, c_{33})\) (e) and \((c_{13}, c_{33})\) (f) planes for the \((c_{11}, c_{33}, c_{13})\) parameterization.

Figure 7: Profiles across the planes of the misfit function shown in Figure 6. The profiles are shown at the true values of the two parameters that are held constant. (a) \(V_{P0}\) and \(\epsilon\) profiles across the planes of Figure 6a. (b) \(V_{P0}\) and \(\delta\) profiles across the planes of Figure 6b. (c) \(V_{P0}\) and \(V_h\) profiles across the planes of Figure 6c. (d) \(\delta\) and \(V_h\) profiles across the planes of Figure 6d. (e) \(c_{11}\) and \(c_{33}\) profiles across the planes of Figure 6e. (f) \(c_{13}\) and \(c_{33}\) profiles across the planes of Figure 6f.
Figure 8: Hessian for the \((V_{P0}, \delta, \epsilon)\) parameterization. The upper-left diagonal, middle, and lower-right diagonal blocks are associated with \(V_{P0}\), \(\epsilon\) and \(\delta\), respectively. The Hessian is computed for parameters normalized by their background value. This normalization amounts to multiply the Hessian coefficients by the square of the value of the parameter in the background model. With such normalization, the \(V_{P0}\) parameter has a dominant weight in the Hessian.
Figure 9: Mono-parameter FWI test - \((V_{P0}, \delta, \epsilon)\) parameterization. Mono-parameter FWI results for \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(\epsilon\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. The initial models are the homogeneous background model.

Figure 10: Multi-parameter FWI test 1 - \((V_{P0}, \delta, \epsilon)\) parameterization. Joint update FWI results for \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(\epsilon\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Comparison between these results and those shown in Figure 9 allows for a first assessment of the effects of the trade-off between the parameters. The initial models are the homogeneous background models.
Figure 11: Multi-parameter FWI test 2 - $(V_{P0}, \delta, \epsilon)$ parameterization. Joint update FWI results of $V_{P0}$ (a-c), $\delta$ (d-f) and $\epsilon$ (g-i). The true model is homogeneous in $V_{P0}$ and $\delta$ and contains an inclusion in $\epsilon$. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. The main perturbations are erroneously concentrated in the $V_{P0}$ model. This shows both the trade-off between $V_{P0}$ and $\epsilon$, and the dominant influence of $V_{P0}$. The initial models are the homogeneous background models.
Figure 12: Mono-parameter FWI test - \((V_{P0}, \delta, V_h)\) parameterization. Mono-parameter FWI results for \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(V_h\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. The initial models are the homogeneous background models.

Figure 13: Multi-parameter FWI test 1 - \((V_{P0}, \delta, V_h)\) parameterization. Joint update FWI results for \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(V_h\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Comparison between these results and those shown in Figure 12 allows for a first assessment of the effects of the trade-off between the parameters. Notice the significant difference between the horizontal profile of \(V_{P0}\) shown in Figures 12 and 13. The initial models are the homogeneous background models.
Figure 14: Multi-parameter FWI test 2 - \((V_{P0}, \delta, V_h)\) parameterization. Joint update FWI results for \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(V_h\) (g-i). The true model is homogeneous in \(V_h\) and \(\delta\), and contains an inclusion in \(V_{P0}\). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed one (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Notice how erroneous model perturbations are reconstructed in both the \(\delta\) and \(V_h\) models, hence highlighting the trade-off between \(V_{P0}\), \(\delta\), and \(V_h\). The initial models are the homogeneous background models.
Figure 15: Multi-parameter FWI test 1 - \((V_{NMO}, \delta, V_h)\) parameterization. Joint update FWI results for \(V_{NMO}\) (a-c), \(\delta\) (d-f) and \(V_h\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Compared to the results obtained with the \((V_{P0}, \delta, V_h)\) parameterization (Figure 13), the \(\delta\) perturbations have stronger amplitudes, at the expense of underestimated amplitudes of the \(V_{NMO}\) perturbations. The initial models are the homogeneous background models.

Figure 16: Multi-parameter FWI test 1 - \((c_{33}, c_{13}, c_{11})\) parameterization. Joint update FWI results for \(c_{33}\) (a-c), \(c_{13}\) (d-f) and \(c_{11}\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. The initial models are the homogeneous background models.
FIGURE
Figure 1: Radiation patterns for parameterization ($V_{P0}$, $\delta$, $\epsilon$) for incidence angles 0˚ (a-c), 45˚ (d-f) and 90˚ (g-i). These radiation patterns are shown for the $V_{P0}$ (a,d,g), $\delta$ (b,e,h) and $\epsilon$ (c,f,i) model perturbations, respectively. The incidence angle $\varphi$ is defined with respect to the vertical symmetry axis, and this defines the source position in the finite-difference modeling, as denoted by the white circle. The diffractor point associated with the model perturbation of each parameter class is located in the center of the finite-difference grid. The source-receiver scattering angle $\theta$ is labeled around the dashed white circle for each incidence angle.
Figure 2: As for Figure 1, for the \((V_P^0, \delta, V_h)\) parameterization. The radiation patterns shown are of \(V_P^0\) (a, d, g), \(\delta\) (b, e, h) and \(V_h\) (c, f, i). Note how the radiation pattern of \(V_P^0\) differs from that in Figure 1 because the other parameter classes in this parameterization have changed.
Figure 3: Radiation patterns for a specular reflection on a horizontal reflector. (a) Procedure to extract the value of the partial derivative wavefields (Figures 1 and 2) at a receiver position (black square) for the specular reflection on a horizontal reflector (thick black segment in (a)), given the source position (black circle in (a)). The incidence angle $\phi$ and the reflection angle $\theta$ are defined with respect to the vertical axis. This extraction is repeated for a source moving on the dashed circle around the diffractor point, located in the middle of the panel. The assemblage of the values of the partial derivative wavefields at the specular positions for each shot position is shown in (b-g) for different parameterizations. (b-g) Radiation patterns of the $(V_{P0}, \delta, \epsilon)$ (b), $(V_h, \delta, \epsilon)$ (c), $(V_{NMO}, \delta, \eta)$ (d), $(V_{P0}, \delta, V_h)$ (e), $(V_{NMO}, \delta, V_h)$ (f) and $(c_{33}, c_{13}, c_{11})$ (g) parameterizations with respect to the reflection angle $\theta$ as defined in (a). Note that the radiation patterns of the Thomsen parameters are magnified by a factor of 10 in b, c, d, e and f.
Figure 4: As for Figure 3, except that specular reflections from a vertical interface (thick black segment in (a)) are considered. The incidence and aperture angles are defined with respect to the horizontal axis. Notice that the radiation patterns shown here are the mirror images of those in Figure 3 (see text for details).
Figure 5: Inclusion model and acquisition geometry for the grid analysis of the misfit function (a), and for VTI FWI (b). Both sources and receivers surround the inclusion providing complete seismic illumination. The radius of the inclusion is higher in (a) to increase the sensitivity of the misfit function to the inclusion.
Figure 6: (a-b) Grid analysis of the misfit function in the ($V_{P0},\delta,\epsilon$) parameterization. In the cross-section of the misfit function in (a), the ($V_{P0}-\epsilon$) plane is defined by the true value of $\delta$ ($\delta = 0.1$), and the ($V_{P0}-\delta$) plane is defined by the true value of $\epsilon$ ($\epsilon = 0.2$). The dashed lines show the directions of the maximum (black) and minimum (gray) sensitivities, respectively. (c-d) As for (a-b) for the ($V_{P0},V_h$) (c) and ($\delta,V_h$) (d) planes for the ($V_{P0},\delta,V_h$) parameterization. (e-f) As for (a-b) for the ($c_{11},c_{33}$) (e) and ($c_{13},c_{33}$) (f) planes for the ($c_{11},c_{33},c_{13}$) parameterization.
Figure 7: Profiles across the planes of the misfit function shown in Figure 6. The profiles are shown at the true values of the two parameters that are held constant. (a) $V_{P0}$ and $\epsilon$ profiles across the planes of Figure 6a. (b) $V_{P0}$ and $\delta$ profiles across the planes of Figure 6b. (c) $V_{P0}$ and $V_h$ profiles across the planes of Figure 6c. (d) $\delta$ and $V_h$ profiles across the planes of Figure 6d. (e) $c_{11}$ and $c_{33}$ profiles across the planes of Figure 6e. (f) $c_{13}$ and $c_{33}$ profiles across the planes of Figure 6f.
Figure 8: Hessian for the \((V_{P0}, \delta, \epsilon)\) parameterization. The upper-left diagonal, middle, and lower-right diagonal blocks are associated with \(V_{P0}\), \(\epsilon\) and \(\delta\), respectively. The Hessian is computed for parameters normalized by their background value. This normalization amounts to multiply the Hessian coefficients by the square of the value of the parameter in the background model. With such normalization, the \(V_{P0}\) parameter has a dominant weight in the Hessian.
Figure 9: Mono-parameter FWI test - $(V_{P0}, \delta, \epsilon)$ parameterization. Mono-parameter FWI results for $V_{P0}$ (a-c), $\delta$ (d-f) and $\epsilon$ (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. The initial models are the homogeneous background model.
Figure 10: Multi-parameter FWI test 1 - \((V_{P0}, \delta, \epsilon)\) parameterization. Joint update FWI results for \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(\epsilon\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Comparison between these results and those shown in Figure 9 allows for a first assessment of the effects of the trade-off between the parameters. The initial models are the homogeneous background models.
Figure 11: Multi-parameter FWI test 2 - \((V_{P0}, \delta, \epsilon)\) parameterization. Joint update FWI results of \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(\epsilon\) (g-i). The true model is homogeneous in \(V_{P0}\) and \(\delta\) and contains an inclusion in \(\epsilon\). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. The main perturbations are erroneously concentrated in the \(V_{P0}\) model. This shows both the trade-off between \(V_{P0}\) and \(\epsilon\), and the dominant influence of \(V_{P0}\). The initial models are the homogeneous background models.
Figure 12: Mono-parameter FWI test - \((V_{P0}, \delta, V_h)\) parameterization. Mono-parameter FWI results for \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(V_h\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. The initial models are the homogeneous background models.
Figure 13: Multi-parameter FWI test 1 - \((V_{P0}, \delta, V_h)\) parameterization. Joint update FWI results for \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(V_h\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Comparison between these results and those shown in Figure 12 allows for a first assessment of the effects of the trade-off between the parameters. Notice the significant difference between the horizontal profile of \(V_{P0}\) shown in Figures 12 and 13. The initial models are the homogeneous background models.
Figure 14: Multi-parameter FWI test 2 - \((V_{P0},\delta,V_h)\) parameterization. Joint update FWI results for \(V_{P0}\) (a-c), \(\delta\) (d-f) and \(V_h\) (g-i). The true model is homogeneous in \(V_h\) and \(\delta\), and contains an inclusion in \(V_{P0}\). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed one (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Notice how erroneous model perturbations are reconstructed in both the \(\delta\) and \(V_h\) models, hence highlighting the trade-off between \(V_{P0}\), \(\delta\), and \(V_h\). The initial models are the homogeneous background models.
Figure 15: Multi-parameter FWI test 1 - \((V_{NMO}, \delta, V_h)\) parameterization. Joint update FWI results for \(V_{NMO}\) (a-c), \(\delta\) (d-f) and \(V_h\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Compared to the results obtained with the \((V_P, \delta, V_h)\) parameterization (Figure 13), the \(\delta\) perturbations have stronger amplitudes, at the expense of underestimated amplitudes of the \(V_{NMO}\) perturbations. The initial models are the homogeneous background models.
Figure 16: Multi-parameter FWI test 1 - \((c_{33}, c_{13}, c_{11})\) parameterization. Joint update FWI results for \(c_{33}\) (a-c), \(c_{13}\) (d-f) and \(c_{11}\) (g-i). (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. The initial models are the homogeneous background models.