A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study

Tariq Alkhalifah¹ and René-Édouard Plessix²

ABSTRACT

In multiparameter full-waveform inversion (FWI) and specifically one describing the anisotropic behavior of the medium, it is essential that we have an understanding of the parameter resolution possibilities and limits. Because the imaging kernel is at the heart of the inversion engine (the model update), we drew our development and choice of parameters from what we have experienced in imaging seismic data in anisotropic media. In representing the most common (first-order influence and gravity induced) acoustic anisotropy, specifically, a transversely isotropic medium with a vertical symmetry direction (VTI), with the P-wave normal moveout velocity, anisotropy parameters δ, and η, we obtained a perturbation radiation pattern that has limited trade-off between the parameters. Because δ is weakly resolvable from the kinematics of P-wave propagation, we can use it to play the role that density plays in improving the data fit for an imperfect physical model that ignores the elastic nature of the earth. An FWI scheme that starts from diving waves would benefit from representing the acoustic VTI model with the P-wave horizontal velocity, η, and ε. In this representation, the diving waves will help us first resolve the horizontal velocity and then reflections, if the nonlinearity is properly handled, could help us resolve η, and ε could help improve the amplitude fit (instead of the density). The model update wavenumber for acoustic anisotropic FWI is very similar to that for the isotropic case, which is mainly dependent on the scattering angle and frequency.

INTRODUCTION

In the slew of possible full-waveform inversion (FWI) implementations on surface seismic 1C recording over an anisotropic subsurface, we present a possible recipe for what we deem to be a practical implementation. A practical implementation in multiparameter inversion accounts for resolvability, resolution limits, trade-off, and common sense. We get the common sense from the vast literature on the anisotropic parameter dependency paradox in the linearized form of the inversion given by the conventional imaging process.

At the heart of FWI is the velocity model update process. In the classical implementation (Lially, 1983; Tarantola, 1984a, 1984b), the update is based on the gradient method represented so eloquently by the Born approximation (Cohen and Bleistein, 1977; Panning et al., 2009). Considering the high dimensionality of our inverted model, using the physics of the data dependency on the model parameters to update the model is a logical and intuitive choice. The convergence of FWI depends on the specific features of this gradient-based update including the radiation pattern and the involved update wavelengths (Sirgue and Pratt, 2004). Despite the constraints on the model exercised by Sirgue and Pratt (2004) (i.e., horizontal reflector), a more general formulation is readily available from the illumination principles we encounter in imaging.

We base our analysis on the pseudoacoustic wave equations. This is currently the most common approximation made in real-size FWI applications (see, e.g., Plessix et al., 2013; Vigh et al., 2013; Wang et al., 2013). We can question the validity of this assumption to invert real elastic data. Here, we consider only monocomponent (mainly pressure) data sets and assume that the most energetic events are the P-waves. This is a classic assumption with transmitted waves. We further assume that the elastic effects on the amplitudes of the reflected waves can be approximately included in the short-wavelength variations of one of the acoustic parameters we invert for. As we shall see, depending on the parameters, it can
be density, $\delta$, or $\varepsilon$. The implicit assumption is that the contrasts are weak and we neglect the wave conversions. Because our goal is to define a parameterization for the inverse problem, despite its limitations, this framework is interesting because we generally assume that, for several geologic settings, the earth is locally smooth almost everywhere, except at specific interfaces.

To analyze the dependency of FWI on the anisotropy parameters, it is imperative that we understand the angular influence of the perturbed parameters and, specifically, the perturbation radiation pattern including the trade-off between the parameters (Wu and Aki, 1985; Virieux and Operto, 2009), some of which have been addressed in the past within the contest of certain anisotropy parameter combinations and specific considerations in the model and data (Červený, 2001; Plessix and Cao, 2011; Gholami et al., 2013a, 2013b). The right combination of anisotropy parameters is essential to reduce the inherent ambiguity embedded in the formulation when we apply a hierarchical approach by selecting parts of the data at the beginning of the inversion. Because large offsets are essential to recover the intermediate wavelengths of the model, they are equally important in recovering the anisotropy parameters (Alkhalifah, 1997). Thus, the classic trade-off between inhomogeneity and anisotropy translates to a trade-off between model wavelength-based inhomogeneity and anisotropy. Some constraints are necessary to resolve this trade-off, some of which are as simple as forcing the anisotropy parameters to have only long-wavelength components.

The Born approximation provides us with the first-order sensitivity of the data to the parameters, and thus, it provides us with the gradient of the model update and the corresponding wavelength of that update. In isotropic media, the update wavelength at a model point is governed by the dip of the reflector and the scattering angle. Specifically, the wavenumber $k_m$ is equal to $k \cos \theta$, with $\theta$ being the scattering angle; where $k_m = k_s + k_r$ is the midpath (model) wavenumber vector; and $k_s$ and $k_r$ are the source and receiver wavefield wavenumbers, respectively, at the model point. For isotropic media, the same equation holds, but the relation between $k_m$ and the scattering angle is more complicated as discussed in Appendix A and shown in Figure 1.

A proper scheme for an appropriate FWI is to update the low-wavenumber components necessary to take us to the global minimum region (basin of attraction), prior to updating the high-wavenumber components. Of course, the model update wavenumber $k_m$ depends on, among other things, the angular frequency $\omega$; $k = v n$, where $v$ is the velocity and $n$ is the unit vector normal to the reflector. Thus, lower frequencies induce long wavelength updates, but it is not the only source of long-wavelength information. Clearly, data from large offsets inducing large scattering angles reduce the wavenumber of the model update. Hence, our inversion schemes prefer low frequencies and large offsets (Pratt et al., 1996; Virieux and Operto, 2009). However, the model wavelength formula extracted from the Born approximation does not reveal the whole story. For refractions or direct arrivals, where $\theta = \pi$, the model wavenumber is zero, and thus, resolution is embedded in the nonlinear terms. Despite these basic and intuitive insights, we will also look at the second-order term of the Born series to discuss parameter trade-off. We will assume only a local perturbation. This means that we won’t comment on the fact that the Hessian of the FWI misfit function is not diagonally dominant when correlations between the incident and the back-propagated fields happen along the full wavefield as in the so-called tomography mode of FWI with transmitted waves. Our discussion aims at understanding the parameterization at a given subsurface point.

In this paper, we analyze the amplitude and phase of the Born perturbation by decomposing the gradient update into its plane-wave components. The amplitude is represented by the radiation pattern, whereas the phase provides the model update wavenumber first-order information. In anisotropic media, the trade-off between parameters for a model point is embedded in the radiation patterns and the off-diagonal elements of the Hessian for a model point derived from the second term of the Born series. We analyze such a trade-off to pick what we deem to be the optimal parameter combination for a practical FWI in anisotropic media. In the next section, we recall the role of the anisotropy parameters in a vertical transversely isotropic medium and argue for a potentially viable parameterization. Then, we derive analytical expressions for the VTI radiation patterns and discuss the relevance of the different parameterizations. Next, we relate the radiation patterns to the gradient and Hessian of the FWI misfit function.

**THE CHOICE OF PARAMETERS**

A successful multiparameter inversion (Burrige et al., 1998; Plessix and Cao, 2011; Priex et al., 2011) requires an appropriate choice of parameters to represent the model. The choice of the parametrization will reduce the null space of the inversion only when we reduce the number of parameters we invert for. Finding a minimum set of parameters that can explain the data can lead to a better inversion. Moreover, because of the nonlinearity, an improper choice can also render the inversion results to be suboptimal. Imposing physicsless conditions, such as regularization, can replace an otherwise high-resolution model with a smooth (poorly fitting the data) model of the earth. The combination of our surface seismic measurements and the tendency of the symmetry axis to be near vertical have resulted in well-established trade-offs in the classic elastic coefficients representation of the anisotropic influence on our data. These trade-offs are systematic for certain model conditions, such as vertically varying media, but they also exist in more complex media. The data are sensitive to the traveltimes and the corresponding (traveltimes driven) geometrical amplitudes, both dependent on the wavefront, that generally provide background.

![Figure 1. A schematic plot of the source and receiver rays as they connect at a model point and the angles involved in an anisotropic model; n is normal to the reflector.](Image)
information. Because the imaging of the discontinuities requires background information, our built-up knowledge of the wavefront behavior as it interacts with the earth’s surface, providing our data, should be exercised here (Plessix and Cao, 2011; Alkhalifah, 2013).

The first-order variation of the wavefront along the seismic horizontal recording surface in VTI media is controlled by the NMO velocity \( v_n \) and the horizontal velocity \( v_h \). The NMO velocity \( v_n \) is the effective extension of the isotropic velocity for waves traveling near vertically, and the horizontal velocity is the effective extension for waves traveling horizontally. For short-to-intermediate incident angles, the NMO velocity has the same influence on the recorded wavefield (described by the horizontal wavenumber and frequency) as the isotropic velocity. In addition to the NMO velocity, the VTI acoustic-wave equation depends on the vertical velocity or, alternatively, \( \delta \) (\( v_n = v_h \sqrt{1 + 2\delta} \)). We can choose to invert for the velocities or for anisotropic variations (Plessix and Cao, 2011; Prieux et al., 2013). We use \( \delta \) instead of the vertical velocity to describe the transversely isotropic model. Similarly, we choose the parameter \( \eta \) as opposed to the horizontal velocity \( v_h \) (\( v_n = v_h \sqrt{1 + 2\eta} \)) in our parameterization. Thus, for acoustic VTI media, we consider the combination \( v_n, \delta, \eta, \) and possibly density, in our study as a practical combination. In Appendix B, we briefly illustrate the case of when we work with the NMO and horizontal velocities.

The dispersion relation for a constant-density VTI medium is given by the following formula (Alkhalifah et al., 2001):

\[
v_n^2 k_z^2 (2v_n^2 \eta(k_x^2 + k_y^2) - \omega^2) - v_n^2 (2\eta + 1) \omega^2 (k_x^2 + k_y^2) + \omega^4 = 0,
\]

where \( k_z = \frac{k}{\sqrt{2\delta}} \) (Plessix and Cao, 2011). Because \( k_z \) is not something we measure in our surface seismic data, this substitution is natural and holds mainly for the \( r(z) \) medium. We can stretch the depth axis later to correct for the proper \( \delta \). (Alkhalifah et al., 2001) show that surface seismic P-wave data are very much independent of \( \delta \) in complex media, granted that \( \delta \) does not vary laterally. This can be realized from equation 1 because the scaling of the vertical wavenumber is, in this case, laterally consistent.

Because two parameters (\( v_n \) and \( \eta \)) describe the kinematic and geometrical behavior of the wave propagation in VTI media as it interacts with the horizontal free surface, we are left with two parameters (\( \delta \) and density) to handle the behavior at a discontinuity. Density is generally used in the acoustic waveform inversion to compensate for the inadequate amplitude fitting due to ignoring the elastic nature of the earth. This role can alternatively be assigned to \( \delta \), or to \( \epsilon \) depending on our parameter choice. Because we are using the NMO velocity to describe the medium, \( \delta \) (or \( \epsilon \) when we are using the horizontal velocity) appears in the NMO equation of the reflection coefficient expansion as a function of angle (Ruger, 1997; Plessix and Bork, 2000). Of course, \( \delta \) in this case (in elastic media) will not properly place reflections at the right depth because that task is reserved for other sources of information, such as well logs. Only if the medium is close to acoustic, then the inverted \( \delta \) may serve in this capacity. This observation means that we may use constant-density or fixed-density wave equations. This reduces the number of parameters to invert for. We could similarly fix \( \delta \) or \( \epsilon \) and invert for density.

In the next sections, we try to refine this analysis based on first- and second-order terms of the Born series.

### Using the NMO Velocity

We can gain insights into the behavior of FWI by applying plane-wave-component analysis, although it assumes a constant or smooth background. This process, based on the Born approximation, exposes the linear dependency of the data on the parameters, and it specifically defines the update resolution and angular dependency (radiation pattern). Starting from the VTI acoustic-wave equations presented by Dvurechek et al. (2008), Plessix and Cao (2011) show that they can be written in terms of a pressure field \( p \) and its anisotropic perturbation \( q \) (see also Zhou et al., 2006):

\[
\begin{align*}
-\frac{1}{v_n^2} \omega^2 p - \partial_z \left( \frac{1}{n} \partial_z (p + q) \right) - \partial_z \left( \frac{1}{n} \partial_z (p + q) \right) \\
- \frac{1}{v_n^2} \omega^2 q = \frac{1}{n} \left( \partial_z \left( \frac{1}{n} \partial_z (p + q) \right) + \partial_z \left( \frac{1}{n} \partial_z (p + q) \right) \right) = 0,
\end{align*}
\]

where \( s \) is the source term, \( \omega \) is the angular frequency, and \( \rho \) is the density.

To derive the Born approximation, we define the perturbations \( v_n, \rho, \eta, \) and \( \delta \) by

\[
\begin{align*}
v_n &= v_n(1 + r_n); & \rho &= \rho_0(1 + r_p); \\
\eta &= \eta_0 + r_\eta; & \delta &= \delta_0 + r_\delta.
\end{align*}
\]

The subscript 0 denotes the earth parameters in the background medium. Assuming weak anisotropy, we linearize around an isotropic smooth background (i.e., \( \eta_0 = 0 \) and \( \delta_0 = 0 \)) and neglect the spatial derivatives of the background earth parameters. This allows us to derive an expression for the perturbed pressure field. The background pressures \( p_0 \) and \( q_0 \) satisfy

\[
\begin{align*}
-\frac{1}{v_0^2} \omega^2 p_0 - \frac{1}{\rho_0} \left( \partial_{xx} p_0 + \partial_{yy} p_0 + \partial_{zz} p_0 \right) = s; \\
q_0 = 0.
\end{align*}
\]

The scattered fields \( p_1 \) and \( q_1 \) satisfy the Born equations. Because \( q_0 = 0 \), the field \( q_1 \) only depends on \( p_0 \). The perturbed \( q_1 \) field is obtained directly from the unperturbed field \( p_0 \):

\[
-\frac{1}{v_0^2} \omega^2 q_1 = 2r_\eta \left( \partial_{xx} p_0 + \partial_{yy} p_0 \right).
\]

By replacing \( q_1 \) by its values, the field \( p_1 \) satisfies the (isotropic) wave equation with a source term depending on \( p_0 \) and the VTI earth parameter perturbations:

\[
\begin{align*}
-\frac{1}{v_0^2} \omega^2 p_1 - \frac{1}{\rho_0} \left( \partial_{xx} p_1 + \partial_{yy} p_1 + \partial_{zz} p_1 \right) &= -\omega^2 \left( \frac{1}{v_0^2} r_\rho \right) (r_p + 2r_v) p_0 \\
- \frac{1}{\rho_0} \left( \partial_z r_\rho \partial_z p_0 + \partial_z r_\rho \partial_z p_0 + \partial_z r_\rho \partial_z p_0 \right) &= - \frac{1}{\rho_0} \left( r_\rho \partial_z p_0 \right) + \frac{2 v_0^2}{\rho_0 \omega^2} \left( \partial_{xx} r_\rho \partial_{xx} p_0 + \partial_{yy} r_\rho \partial_{yy} p_0 + \partial_{zz} r_\rho \partial_{zz} p_0 \right) \\
- \frac{1}{\rho_0} \left( \partial_{zz} r_\rho \partial_{zz} p_0 \right) &= - \frac{1}{\rho_0} \left( r_\rho \partial_{zz} p_0 \right) + \partial_{zz} r_\rho p_0.
\end{align*}
\]
We now define the Green’s function $G$ by
\[
\left[ -\frac{1}{v_0^2(x)\rho_0(x)}\omega^2 - \frac{1}{\rho_0}(\partial_{xx} + \partial_{yy} + \partial_{zz}) \right] G(x, x', \omega) = \delta(x - x'),
\]
and we express $p_0$ and $p_1$ with the Green’s functions, with $\Delta_h = \partial_{xx} + \partial_{yy}$ and $\Delta_v = \partial_{zz}$:
\[
\begin{align*}
p_0(x, x, \omega) &= s(\omega)G(x, x, \omega); \\
p_1(x, x, \omega) &= -s(\omega)\int d\omega^2 \frac{1}{v^0_0} \left( r^v_0(x) + 2r^v_0(x)G(x, x, \omega)G(x, x, \omega) \right) \\
&\quad + s(\omega)\int d\omega^2 r^v_0(\nabla G(x, x, \omega), \nabla G(x, x, \omega)) \\
&\quad - s(\omega)\int d\omega^2 r^v_0(\Delta_0 G(x, x, \omega) \Delta_0 G(x, x, \omega)) \\
&\quad - s(\omega)\int d\omega^2 r^v_0(\Delta_0 G(x, x, \omega) \Delta_0 G(x, x, \omega)) + \Delta_h G(x, x, \omega)G(x, x, \omega).
\end{align*}
\]
(8)

Here, we have used the fact that the Green’s function is reciprocal; i.e., $G(x, x', \omega) = G(x', x, \omega)$. Using the asymptotic Green’s function (without multipathing), $G(x, y, \omega) = A(x, y) \exp \{i \frac{2}{\rho_0} \mathbf{p} \cdot \mathbf{x}\}$; with $\mathbf{p}$ being the unitary vector related the wavenumber vector by $\mathbf{k} = \frac{2}{\rho_0} \mathbf{p}$. Here, we make the assumption of a locally smooth medium, which is an assumption we also use in extended images. We recognize that this may raise a question within the context of the FWI high-resolution objective. However, based on the Born linearized implementation, which adheres to the same constraints, the analysis is valid. We can then write
\[
p_1(x, x, \omega) = -\omega^2 s(\omega)\int d\mathbf{A}(\mathbf{x}, \mathbf{x}, \mathbf{x}, \omega) a_1(\mathbf{x}) \cdot \mathbf{r}_1(\mathbf{x})
\]
(9)

with
\[
A(\mathbf{x}, \mathbf{x}, \mathbf{x}, \omega) = \frac{G(x, x, \omega)G(x, x, \omega)}{v^0_0(x)\rho_0(x)}.
\]
(10)

and
\[
\mathbf{r}_1 = \begin{pmatrix} r^v_0 \\ r^\eta \\ r^\delta \\ r^p \end{pmatrix}, \quad a_1 = \begin{pmatrix} 2 \\ 2p^2_s sin^2 p^r_{zh} \\ -p^r_{z^2} - p^r_{z^2} \\ 1 + p^r_{z^2} \end{pmatrix}.
\]
(11)

The coefficients of $a_1$ define the radiation patterns (Aki and Richards, 1980) of each parameter for the given parameterization, $v^0_0$, $\eta$, $\delta$, $\rho$. We investigate other parameter combinations later, and thus, we use the subscript 1 for this radiation pattern.

To further analyze the parameterization, we express $p_s$ and $p_r$ with the source incident angle $\theta$ and the reflector dip angle $\alpha$ as

Assuming a mainly horizontal earth, in Figure 2 we plot the radiation patterns for a horizontal reflector. A perturbation in $\delta$ is driven by the vertical component of the wavenumber for the source and receiver wavefields. As expected, the NMO velocity has an angle-invariant radiation pattern, similar to the isotropic case, whereas $\eta$ is mainly associated with the horizontal component of the wavenumber. From this simple observation, we can directly conclude that, in the parameterization, $(v^0_0, \eta, \delta, \rho)$, $\delta$ is not resolvable from the geometrical behavior of the wavefield in the background medium. Even if we ignore the perturbation in $\rho$, there is a clear trade-off between $\delta$, $v^0_0$, and the depth of the reflector. We must use the reflection signature to extract the $\delta$ influence. On the other hand, the radiation pattern of the $\eta$ perturbation clearly shows the importance of large scattering angles or large offsets (in other words, large horizontal wavenumbers) to resolve $\eta$. However, in multi-parameter inversions, there is always the issue of trade-off. At large angles, a trade-off between $v^0_0$ and $\eta$ exists, but not at short angles. This may suggest a hierarchical approach by first retrieving $v^0_0$ from reflection velocity analysis and then improving the intermediate wavelengths with an FWI approach that focuses more on the diving waves. This is the approach in which the initial guess is obtained by reflection tomography. For a horizontal reflector, the trade-off between $\delta$ and density cannot be resolved because they have the same radiation pattern. We can therefore choose to fix one and reduce the number of parameters.

**USING THE HORIZONTAL VELOCITY**

We know that the kinematics (i.e., geometrical) behavior of the wave propagation in a smooth VTI medium depends mainly on two parameters: $v^0_0$ and $\eta$. To update the long-to-intermediate wavelengths of the earth parameters with surface seismic data, FWI needs to start with diving waves. In this context, the waves predominantly depend on the horizontal velocity $v^0_0$. Therefore, we consider replacing $v^0_0$ with $v^0_h$ as the key velocity parameter (the natural extension of the isotropic velocity for waves propagating horizontally or as the horizontal wavenumber is relatively large). We can use, in addition to $v^0_h$, two of the following three parameters: $\delta$, $\eta$, $\epsilon$. The main obvious feature of using $\eta$ and $\epsilon$ is that both parameters are directly tied to the horizontal velocity, and thus, their radiation patterns are influenced by that. In the following, we will see the benefit of this feature. In Appendix B, we also briefly present parameterizations that involve $v^0_0$ and $v^0_h$ together.

For any parameterization, we can write the scattered field as follows:
\[
p_1(x, x, \omega) = -\omega^2 s(\omega)\int d\mathbf{A}(\mathbf{x}, \mathbf{x}, \mathbf{x}, \omega) a(\mathbf{x}) \cdot \mathbf{r}(\mathbf{x})
\]
(13)

with $a$ and $r$ having components that depend on the choice of parameterization. Thus, we obtain the scattered field with different
parameterizations by using just the linear relation between the perturbations. Because \( v_h = v_n \sqrt{1 + 2 \eta} \) and \( 1 + 2 \delta = \frac{1 + 2 \varepsilon}{1 + 2 \varepsilon} \), we have the following relations between perturbations:

\[
\begin{align*}
\rho_{v_h} &= \rho_{v_n} - \rho_\eta; \\
\delta_{v_h} &= \delta_v - \delta_\eta. \\
\end{align*}
\]  

Thus, the radiation patterns of the parameterization \((v_h, \eta, \varepsilon, \rho)\) are

\[
\mathbf{r}_2 = \begin{pmatrix} r_{v_h} \\ r_\eta \\ r_\varepsilon \\ r_\rho \end{pmatrix}; \quad \mathbf{a}_2 = \begin{pmatrix} 2 \\ -p_{2h}^2 - p_{2h}^2 - p_{2c}^2 \\ -p_{2c}^2 \\ 1 + \mathbf{p}_x \cdot \mathbf{p}_r \end{pmatrix}
\]

(to obtain the radiation patterns, we use the relation \( p_{2h}^2 + p_{2c}^2 = 1 \)).

In Figure 3, we show the radiation patterns for a horizontal reflector. In this case, the trade-off between \( v_h \) and \( \eta \) is maximum at 45°. For horizontally traveling waves, the horizontal velocity is the lone controller of wave propagation. The trade-off between \( v_h \) and \( \varepsilon \) is largest in the vertical direction, and only if we knew the depth of the reflector can we resolve \( \varepsilon \) from the kinematics. This setup is most relevant for an inversion scheme that starts with diving waves, in which we invert first for the horizontal velocity, and then we slowly include data that can resolve \( \eta \), specifically, the reflection kinematics (phases). Finally, with a reasonable \( v_h \) and \( \eta \), we can use the amplitude part of the reflections to resolve \( \varepsilon \) or the density because we cannot resolve those two parameters independently. This is a recipe that may be valid in cases where we have the luxury of accessing data in such a sequence.

Let us now discuss the parameterization \((v_h, \eta, \delta, \rho)\). The radiation patterns are given by

\[
\mathbf{r}_3 = \begin{pmatrix} r_{v_h} \\ r_\eta \\ r_\delta \\ r_\rho \end{pmatrix}; \quad \mathbf{a}_3 = \begin{pmatrix} 2 \\ 2p_{sh}^2 p_{rh}^2 - 2(p_{2c}^2 + p_{2c}^2) \\ 1 + \mathbf{p}_s \cdot \mathbf{p}_r \end{pmatrix}
\]

For a horizontal reflector, we display the radiation patterns in Figure 4. A model representation using \( \delta \) instead of \( \varepsilon \) moves the trade-off for a horizontal reflector between \( v_h \) and \( \eta \) to the vertical direction, which is the same region where we have the trade-off between \( v_h \) and \( \rho \) resulting in inherent ambiguity in resolving \( \eta \). We cannot use the hierarchical approach discussed previously. We may, however, use a simultaneous inversion and risk increasing the null space.

**GRADIENT AND HESSIAN OF THE FWI MISFIT FUNCTION**

In FWI, we minimize the least-squares misfit \( E \) between the observed data \( d \) and the synthetic data:

\[
E = \frac{1}{2} \int d\mathbf{x}, d\mathbf{x}, d\omega \| p(\mathbf{x}_r, \omega) - d(\mathbf{x}_r, \omega) \|^2.
\]

Because of the computational cost, we rely on a local optimization. The gradient and the Hessian directly dependent on the radiation patterns.

Using equation 13 for the scattered field, the gradient \( g \) of \( E \) is given by

Figure 2. Radiation patterns for the parameterization \((v_h, \eta, \delta, \rho)\) and a horizontal reflector. The radiation patterns here and throughout are unitless.
\[ g(x) = -R e \left[ \omega^2 s^*(\omega) \int d\omega A^\dagger(x, x_s, \omega) a(x) \right] \times \left( p(x_s, x_r, \omega) - d(x_s, x_r, \omega) \right) \]

with the superscript * denoting the complex conjugate. At subsurface point \( x \), the relative strength of the gradient with respect to the different quantities of the parameterization is given by the radiation patterns. The displays of the radiation patterns provide us with the strength of the influence of the different parameters, which depends

Figure 3. Radiation patterns for the parameterization \((v_h, \eta, \varepsilon, \rho)\) and a horizontal reflector.

Figure 4. Radiation patterns for the parameterization \((v_h, \eta, \delta, \rho)\) and a horizontal reflector.
also on the scattering angle or the aperture of the data. In Figure 3, we clearly see that $r_D$ has a much smaller influence than $r_\eta$. Although this is not an issue in a hierarchical and sequential inversion, it should be taken into account during a simultaneous inversion. In the parameterization $(v_h, \eta, \epsilon, \rho)$, the maximum values for the radiation pattern of $v_h$ is two, whereas for $\eta$ is 0.5. Thus, we may invert for $r'_\eta = 4r_\eta$ instead of $r_\eta$ to balance the influence. In fact, a gradient preconditioned by the pseudoinverse of the Hessian gives a much better descent direction than the gradient alone. Under the Born approximation, the Hessian is given by

$$H(x, y) = \mathcal{R}_e \left[ \int dx, dx, d\omega B(x, x, y, \omega) a(x, x, x) a^T(x, x, y) \right],$$

(19)

with $B$ being a factor independent of the parameterization,

$$B(x, x, y, x, \omega) = \omega^4 |s(\omega)|^2 A^T(x, x, x, \omega) A(x, y, x, x, \omega).$$

(20)

The Hessian, in addition to its scaling properties, tells us about the trade-off between the different parameters in the inversion. The Hessian of the FWI misfit function is not always diagonal dominant. For instance, in the so-called tomography model with transmitted/diving waves, there are large correlations along the wavepaths connecting the sources and the receivers, especially for low frequencies. This means that we may have a trade-off between earth parameters located at different subsurface points. In this paper, we don’t discuss this aspect of the inverse problem. We are just interested in the trade-off between parameters at the same subsurface point. This means that we only consider a local perturbation, as in the diffraction tomography analysis, or in other words, for low frequencies, we are assuming large grid points.

We can analyze the trade-off between the parameters by studying the matrix

$$W(x, x) = \int dx, dx, a(x, x, x) a^T(x, x, x).$$

(21)

For the parameterization $(v_h, \eta, \delta)$, we have

$$W_1 = \int \int dx, dx, 
\begin{pmatrix}
4p_{2h}^2p_{2h}^2 & 4p_{2h}^2p_{2h}^4 & -2(p_{2h}^2 + p_{2h}^2) \\
-2(p_{2h}^2 + p_{2h}^2) & 2(p_{2h}^2 + p_{2h}^2) & (p_{2h}^2 + p_{2h}^2)^2 \\
-2(p_{2h}^2 + p_{2h}^2) & -2(p_{2h}^2 + p_{2h}^2) & (p_{2h}^2 + p_{2h}^2)^2
\end{pmatrix},$$

(22)

and for the parameterization $(v_h, \eta, \epsilon)$, we have

$$W_2 = \int \int dx, dx, 
\begin{pmatrix}
4 & -2p_{2h}^2p_{2h}^2 & -2(p_{2h}^2 + p_{2h}^2) \\
-2p_{2h}^2p_{2h}^2 & (p_{2h}^2 + p_{2h}^2)^2 & (p_{2h}^2 + p_{2h}^2)(p_{2h}^2 + p_{2h}^2) \\
-2(p_{2h}^2 + p_{2h}^2) & (p_{2h}^2 + p_{2h}^2)(p_{2h}^2 + p_{2h}^2) & (p_{2h}^2 + p_{2h}^2)^2
\end{pmatrix}. $$

(23)

In $W_1$ and $W_2$, we did not include the density because its influence for horizontal reflectors is well represented by $\delta$ or $\epsilon$.

A nonzero off-diagonal (with a significant value in practice) indicates that there is a trade-off between the parameters. For instance, we see in $W_1$ that the trade-off between $v_h$ and $\eta$ is maximum at 90°, i.e., for the transmitted waves. In $W_2$, the trade-off between $v_h$ and $\eta$ is maximum at 45°. Thus, we retrieve the conclusions made during the study of the radiation patterns.

We can also apply a singular value decomposition of $W$ to analyze the direction cosine for each parameter. Using the parameterization $(r_{\eta}, 4r_{\eta}, r_{\rho})$, we display the eigenvalues and eigenvectors in Figure 5. We clearly see that the first eigenvector is mainly along $v_h$, the second eigenvector is mainly along $\eta$ with the incident angles between 45° and 90°. This confirms the previous analysis on the radiation patterns. We recall that this Hessian analysis is based on a single diffractor. This is not the complete analysis of the Hessian of the FWI misfit function.

**DISCUSSION**

Multiparameter FWI from surface seismic data is not trivial. The potential null space induced by the additional degrees of freedom for parameters not resolvable from surface seismic data, in addition to the inherent trade-off between these parameters, can deem the inversion a failure. A proper understanding of the parameter dependency in anisotropic media is vital to an effective and successful setup of the inversion problem. It is well known that the Born approximation has a limited range of validity, especially for transmitted waves. However, when we assume just a local perturbation, we can gain considerable insights into the trade-off between different parameters at a given subsurface point. Though the radiation patterns, gradient, and Hessian are products of the first two terms of a Taylor series expansion of the dependency of the data on the inverted parameters, and thus they do not provide us with the full (nonlinear) story, they give us valuable insights into the inversion process. These insights are especially informative for low frequencies (specifically around the basin of attraction of the global minimum) because the Taylor’s series approximation, even with only two terms, is accurate.

Despite the analysis and the choice of parameters, trade-off is inherent in the anisotropic multiparameter inversion and enhanced by the limitations in acquisition with limited frequency bandwidths and aperture. Some constraints are necessary to resolve this trade-off, some of which are as simple as forcing the anisotropy parameters to have only long-wavelength components. This is not a bad constraint because the anisotropy parameters represent ratios of velocities, and thus, their smooth representation still defines a complex velocity model, which is another reason why working with the dimensionless anisotropy parameters, instead of velocities, can help FWI.

The parameters $\delta$ or $\epsilon$ (when using NMO velocity or the horizontal velocity) can induce major null space in the kinematic (more important) component of FWI. Their influence appears mainly in the amplitude component of the recorded wavefield, and especially from the reflection behavior. In the acoustic isotropic wavefield, density has a similar influence, and thus, we tend to use it to fit the amplitude. Therefore, as suggested earlier, $\delta$ (or $\epsilon$) can play a similar role. If the medium is near acoustic, this $\delta$ might be accurate; otherwise, it will serve only to help us fit the amplitude. The analysis we pursued in this paper assumed
a vertical transversely isotropic medium and (more or less) horizontal discontinuities. To illustrate that the choice of parameters may depend on the geologic context, in Figure 6, we display the radiation patterns for the parameterization \( (v_h, \eta, \varepsilon, \rho) \) for a reflector dipping at 45°. We notice that the trade-offs have changed. For instance, the density and \( \varepsilon \) now have different zones of influence. This, however, does not mitigate the fact that we are able to mainly resolve only two parameters from the kinematics of

Figure 5. Eigenvalues (a) and eigenvectors (b-d) of the matrix \( W \) (equation 23) with the parameterization \( (r_v, 4r_{\eta}, r_{\varepsilon}) \) and a horizontal reflector. In the eigenvalues, the black line represents the first eigenvalue, the dotted-dashed line represents the second one, and the dotted line represents the third one. In the eigenvector plots, the solid line corresponds to the direction cosine associated to \( r_v \), the dotted-dashed line corresponds to the one associated to \( r_{\eta} \), and the dotted line corresponds to the one associated to \( r_{\varepsilon} \).

Figure 6. Radiation patterns for the parameterization \( (v_h, \eta, \varepsilon, \rho) \) and a reflector dipping at 45°.
FWI with monocomponent (pressure) data. These two parameters, thanks to our surface seismic acquisition, are the $v_n$ or $v_h$, and $\eta$.

The analysis here applies mainly for FWI with a data hierarchical approach in which the parameters are inverted predominantly in a sequential manner. When simultaneously inverting several parameters, this analysis indicates how we could build an approximated block diagonal Hessian, where each block would depend on the subsurface point and an estimation of the angle range and reflector dips. Although this would be just an approximation, it can serve as a preconditioning of the inversion and may be combined with a truncated Newton approach (Métiévier et al., 2013).

CONCLUSIONS

Choosing the right inversion setup for resolving anisotropy can make the difference between interpretable high-resolution results and results that may not make a lot of sense (too smooth or too biased). Analyzing the trade-off by assessing the first- and second-order terms of the Born series reveals the proper parameterization for the inversion implementation. Representing the VTI model using the NMO velocity, $\delta$, and $\eta$ offers the proper perturbation radiation pattern for an inversion that includes reflections and diving waves. Because $\delta$ mildly influences the geometrical aspects of the recorded waveform, it can serve as a secondary parameter to fit the amplitude to compensate for the shortcomings of the acoustic model in representing the true earth (a role that density plays in isotropic media). For an inversion with a hierarchical implementation in which diving waves are used first in the inversion, a VTI model represented by $v_h$, $\eta$, and $\epsilon$ offers a practical set necessary to reduce the trade-off and provide reasonable resolution. In this case, $\epsilon$ serves the role of the amplitude fitting as it mildly affects the kinematics in the recorded waveform. The model wavenumber resolution is generally similar to what we experience in isotropic media, but the relation between the scattering angle and the model update wavenumber vector points in the direction normal to the reflector dip, which is the same formula governing the isotropic case. This is an approximation based on a plane-wave representation of the source and receiver wavefields in which the model is homogeneous with respect to the dominant wavelength around the point of interest. In anisotropic media, however, the relation between $k_m$ and the scattering angle is different from that in the isotropic case. Only for horizontal reflectors is the relation between the scattering angle and the wavenumbers identical, with the velocity replaced by the phase velocity. Specifically, $k_m = -\frac{\omega}{v_p} \cos \left(\frac{\theta_s}{2}\right)$, where $v_p$ is the phase velocity, and $\theta_s$ is the scattering angle.

For the special case of a dip-oriented transversely isotropic medium (Alkhalifah and Sava, 2010), in which the symmetry axis is set to be normal to the reflector dip, the relation is also simple, and it is given by

$$k_m = \frac{\omega}{v_p} \cos \left(\frac{\theta_s}{2}\right),$$

where $\mathbf{n}$ is normal to the reflector.

For general TI in 2D, the relation is far more complicated (Sava and Alkhalifah, 2013) and cannot be written in an explicit closed form with respect to the scattering angle. Nevertheless, the model update wavenumber vector points in the direction normal to the reflector dip as follows:

$$k_m = k_\eta + k_\epsilon = A(\theta_s) \mathbf{n},$$

where $A$ is some complicated (possibly not explicit) function of the scattering angle, velocity, and frequency.

APPENDIX B

OTHER PARAMETERIZATIONS

For completeness, we look at parameterizations using $v_n$ and $v_h$ together. Of course, in this case, $\eta$ is not needed because it is described by these two velocities. The radiation patterns for the parameterization $(v_h, v_n, \delta, \rho)$ are

$$r_4 = \begin{pmatrix} r_{v_h} \\ r_{v_n} \\ r_\delta \\ r_\rho \end{pmatrix}, \quad a_4 = \begin{pmatrix} 2p_{v_h}^2p_{v_n}^2 \\ 2 - 2p_{v_h}^2p_{v_n}^2 \\ -(p_{\delta z}^2 + p_{v_h}^2) \\ 1 + p_\rho \cdot \mathbf{p}_r \end{pmatrix},$$

and for the parameterization $(v_h, v_n, \epsilon, \rho)$

$$r_5 = \begin{pmatrix} r_{v_h} \\ r_{v_n} \\ r_\epsilon \\ r_\rho \end{pmatrix}, \quad a_5 = \begin{pmatrix} 2 - p_{\delta z}^2p_{v_h}^2 - p_{v_h}^2p_{v_n}^2 \\ p_{}\delta z p_{v_h} + p_{v_h}^2p_{\delta z} \\ -(p_{\delta z}^2 + p_{v_h}^2) \\ 1 + p_\rho \cdot \mathbf{p}_r \end{pmatrix}.$$

For a horizontal reflector, the radiation patterns are displayed in Figures B-1 and B-2. The parameterization $(v_h, v_n, \delta, \rho)$ has a trade-off between $v_n$, $\delta$, and $\rho$ at small angles that cannot be resolved at other angles. Similar to the parameterization $(v_n, \eta, \delta, \rho)$, this parameterization is not really adequate for a hierarchical...
approach. The radiation patterns for the parameterization \((v_h, v_n, \varepsilon, \rho)\) is similar to that of \((v_h, \eta, \varepsilon, \rho)\), with \(v_h\) having angular dependence.

Figure B-1. Radiation patterns for the parameterization \((v_h, v_n, \delta, \rho)\) and a horizontal reflector.

Figure B-2. Radiation patterns for the parameterization \((v_h, v_n, \varepsilon, \rho)\) and a horizontal reflector.

REFERENCES


