Conditioning the full-waveform inversion gradient to welcome anisotropy

Tariq Alkhalifah¹

ABSTRACT

Multiparameter full-waveform inversion (FWI) suffers from complex nonlinearity in the objective function, compounded by the eventual trade-off between the model parameters. A hierarchical approach based on frequency and arrival time data decimation to maneuver the complex nonlinearity associated with this problem usually falls short in anisotropic media. In place of data decimation, I use a model gradient filter approach to access the parts of the gradient more suitable to combat the potential nonlinearity and parameter trade-off. The filter is based on representing the gradient in the time-lag normalized domain, in which small scattering-angles of the gradient update are initially muted out. The model update hierarchical filtering strategy include applying varying degrees of filtering to the different anisotropic parameter updates, a feature not easily accessible to simple data decimation. Using FWI and reflection-based FWI, when the modeled data are obtained with the single-scattering theory, allows access to additional low model wavenumber components. Combining such access to wavenumbers with scattering-angle filters applied to the individual parameter gradients allows for multiple strategies to avoid complex FWI nonlinearity as well as the parameter trade-off.

INTRODUCTION

Full-waveform inversion (FWI) usually requires a hierarchical strategy of data decimation over the frequency and arrival time to avoid the inherent complex nonlinearity associated with the inversion problem for conventional surface seismic frequencies (Bunks et al., 1995; Pratt et al., 1996; Virieux and Operto, 2009). These strategies are needed to help us build the long-wavelength components of the velocity model (responsible for wave propagation) prior to defining the scattering high-wavenumber components of the

velocity model. Alternatively, such a strategy can be implemented effectively by filtering and conditioning the gradient to control the wavenumbers used for updating the velocity model (Albertin et al., 2013; Almomin and Biondi, 2013; Tang et al., 2013). The recent move from the classic implementation by data decimation to those based on manipulating the model update should remind us of our transformation in imaging analysis from flattening surface offsets to focusing on subsurface offsets. Despite the value of data decimation and selection, the real objective of this process is far more apparent in the model domain, and specifically, at the gradient level (Sirgue and Pratt, 2004). For anisotropic media, model domain conditioning becomes even more effective because it can be applied to the model parameter gradients individually, a feature not accessible in the data domain.

Multiparameter inversion (Burrige et al., 1998; Plessix and Cao, 2011; Prieux et al., 2011; Operto et al., 2013) requires an appropriate choice of parameters to represent the model. A parameterization with fewer parameters representing the model tends to reduce the null space associated with FWI. Finding a minimum set of parameters that can explain the data can lead to a better inversion. Alkhalifah and Plessix (2014) analyze the radial dependency (radiation pattern) of the anisotropic parameter perturbation in acoustic transversely isotropic media with vertical symmetry axes (VTI). They advocate using certain combinations of parameters for various FWI strategies, including those that start with a model obtained from MVA, and those obtained from inverting diving wave energy. Choosing the right inversion setup for resolving anisotropy can make the difference between obtaining interpretable high-resolution results and results that may not make a lot of sense (too smooth or too biased). Representing the VTI model using the normal moveout (NMO) velocity v_n and the anisotropy parameters δ and η offers the proper perturbation radiation pattern for an inversion that includes reflections and diving waves. Because δ mildly influences the geometric aspects of the recorded wavefield, it can serve as a secondary parameter to fit the amplitude to compensate for the shortcomings of the acoustic model in representing the true elastic earth (a role that density ρ , plays in isotropic media (Plessix et al., 2013). For an

Manuscript received by the Editor 18 August 2014; revised manuscript received 22 November 2014; published online 23 April 2015.

¹King Abdullah University of Science and Technology (KAUST), Physical Sciences and Engineering Division, Thuwal, Saudi Arabia. E-mail: tariq.alkhalifah@kaust.edu.sa.

© 2015 Society of Exploration Geophysicists. All rights reserved.

inversion with a hierarchical implementation in which diving waves are used first, a VTI model represented by the horizontal velocity v_h and the anisotropy parameters η ($v_h = v_n \sqrt{1 + 2\eta}$) and ϵ offers a practical set necessary to reduce the trade-off and provide reasonable resolution. In this case, ϵ plays the role of amplitude fitting because it mildly affects the kinematics in the recorded wavefield (or in other words, the horizontal projection of the propagator).

Because one of the challenges in FWI is to have reasonably accurate reflections in the modeled data, the migrated image is used as a source for such reflections (Clement et al., 2001). This approach, referred to as *reflection FWI* (RFWI), however, produces only smooth updates because producing a reflection is not an objective of the approach. Such smooth updates are valuable in building the long-wavelength components of the velocity model. This is useful for multiparameter inversion to reduce the null space expected in the data.

In this paper, I extend the axis of the model gradient with a normalized time lag component capable of resolving the scattering-angle, efficiently. This additional axis is used to properly filter the gradient by keeping the usable background updates for FWI from reflections and diving waves, as well as RFWI. It uses the fact that a proper background update is not necessarily given by low frequencies or low wavenumbers, but more accurately by large scattering angles, where the update actually follows the rays.

THE ANISOTROPIC RADIATION PATTERNS

The radiation pattern, as opposed to the gradient wavenumber, provides us with the size of the influence, and specifically its directional dependence (sensitivity to the model point). For acoustic VTI media, Alkhalifah and Plessix (2014) derive such patterns for two sets of anisotropic parameter combinations that they deem to be the most practical. Those are v_n , η , and δ in the case in which we have the opportunity to resolve v_n first, using for example, migration velocity analysis (MVA), and v_h , η , and ϵ in the case in which we have the opportunity to invert for diving waves first. Both combinations use the fact that the kinematics of the wavefield as they are measured on the surface are dependent on v_n or v_h and η , with a disconnect to the weakly resolvable vertical velocity given by δ or ϵ , used to fit the amplitude.

Using the asymptotic Green's function (without multipathing) expressed in terms of the angular frequency, ω , $G(\mathbf{x}, \mathbf{y}, \omega) = A(\mathbf{x}, \mathbf{y}) \exp(i\mathbf{k} \cdot \mathbf{x})$; with $\mathbf{k}$ being the wavenumber vector ($\mathbf{k} = \omega \mathbf{p}$, where $\mathbf{p}$ is the slowness vector), $\mathbf{x}$ being the space location vector, i is equal to $\sqrt{-1}$, and A is the geometric amplitude between location $\mathbf{x}$ and a potential source or receiver location $\mathbf{y}$. We can then write the single-scattered wavefield as

$$p_1(\mathbf{x}_s, \mathbf{x}_r, \omega) = -\omega^2 s(\omega) \int d\mathbf{x} B(\mathbf{x}_s, \mathbf{x}, \mathbf{x}_r, \omega) \mathbf{a}_1(\mathbf{x}) \cdot \mathbf{r}_1(\mathbf{x}), \quad (1)$$

where s is the source function, ω is the angular frequency, v_0 is the background velocity, and

$$B(\mathbf{x}_s, \mathbf{x}, \mathbf{x}_r, \omega) = \frac{G(\mathbf{x}_s, \mathbf{x}, \omega) G(\mathbf{x}_r, \mathbf{x}, \omega)}{v_0^2(\mathbf{x}) \rho(\mathbf{x})}, \quad (2)$$

and

$$\mathbf{r}_1 = \begin{pmatrix} r_{v_n} \\ r_\eta \\ r_\delta \\ r_\rho \end{pmatrix}; \quad \mathbf{a}_1 = \begin{pmatrix} 2 \\ 2n_{sh}^2 n_{rh}^2 \\ -(n_{sz}^2 + n_{rz}^2) \\ 1 + \mathbf{n}_s \cdot \mathbf{n}_r \end{pmatrix}. \quad (3)$$

The vector $\mathbf{r}_1$ includes the perturbations of the individual parameters, v_n , η , δ , and ρ , from top to bottom. Thus, the coefficients of $\mathbf{a}_1$ define the radiation patterns of each parameter for the given parameterization, v_n , η , δ , ρ (Aki and Richards, 1980).

The unit vectors $\mathbf{n}_s$ and $\mathbf{n}_r$ with the source incident angle θ_s and the reflector dip angle ϕ are given by

$$\mathbf{n}_s = \begin{pmatrix} n_{sh} \\ n_{sz} \end{pmatrix} = \begin{pmatrix} \sin(\theta_s) \\ \cos(\theta_s) \end{pmatrix}; \quad \mathbf{n}_r = \begin{pmatrix} n_{rh} \\ n_{rz} \end{pmatrix} \\ = \begin{pmatrix} -\sin(\theta_s + 2\phi) \\ \cos(\theta_s + 2\phi) \end{pmatrix}. \quad (4)$$

Because $v_h = v_n \sqrt{1 + 2\eta}$ and $1 + 2\delta = (1 + 2\epsilon/1 + 2\eta)$, we have the following relations between perturbations:

$$r_{v_h} = r_{v_n} - r_\eta; \quad r_\delta = r_\epsilon - r_\eta. \quad (5)$$

Thus, the radiation patterns of the parameterization (v_h , η , ϵ , ρ) are (to obtain the radiation patterns, we use the relation $n_h^2 + n_z^2 = 1$) as follows:

$$\mathbf{r}_2 = \begin{pmatrix} r_{v_h} \\ r_\eta \\ r_\epsilon \\ r_\rho \end{pmatrix}; \quad \mathbf{a}_2 = \begin{pmatrix} 2 \\ -n_{sz}^2 n_{rh}^2 - n_{rz}^2 n_{sh}^2 \\ -(n_{sz}^2 + n_{rz}^2) \\ 1 + \mathbf{n}_s \cdot \mathbf{n}_r \end{pmatrix}. \quad (6)$$

Considering the parameter combination of v_n , η , and δ , Figure 1a shows a vector plot with vector components given by the r_η (horizontal) and r_δ (vertical) (η and δ radiation patterns, respectively) as a function of the scattering angle and reflector dip. The shaded regions represent the potential locations of reflections (bottom), diving waves (middle top), and RFWI (top corners) information. The size of the arrow (amplitude of the vector) is the size of the sensitivity with respect to perturbations in η and δ , and the direction of the arrow reveals the dependency distribution. For an arrow to point vertically implies that the sensitivity is due to a δ perturbation. An arrow pointing horizontally implies an η dependency. For an arrow to point diagonally implies a trade-off between the two parameters. On the other hand, the radiation pattern for the NMO velocity is invariant of direction, as given in equation 3, and thus it has a trade-off with η or δ across the potential point illumination angles. As expected, the η resolution is strongest with diving waves and very large dips. However, in all cases, it inherits a trade-off with v_n . To invert for these model parameters, it would be wise to fix δ , which does not affect the kinematics, and focus initially on v_n and η .

On the other hand, for the parameter set v_h , η , and ϵ , Figure 1b shows that the dependency on r_η (the vector's horizontal component) and r_ϵ (the vector's vertical component) is weak in certain areas and specifically for diving waves, which makes the resolution of v_h possible, regardless of the other parameters. However, we also note that even if we resolve v_h , we will need to set ϵ initially (and

hopefully close to its accurate value) constant to resolve η because the trade-off between the two parameters is apparent everywhere. We can also invert for ϵ first using reflections and then finally invert for η by accessing the larger angles or larger dips of reflection.

Instead of decimating the data to give the appropriate optimal gradient contribution for certain parameters, we can organize the model update by filtering the components we need at certain stages. In the next section, we consider an approach for doing so.

FILTERING THE GRADIENT

As an alternative to the conventional hierarchical data selection approach used initially to isolate the long-wavelength updates to the model (Sirgue and Pratt, 2004), we can filter the gradient to provide such updates in spite of the data (Albertin et al., 2013; Almomin and Biondi, 2013; Tang et al., 2013). This recent development is based on the separation of scales in the update by applying the proper wavenumber filtering to the model update or matching the directional components of the source and receiver wavefields and, thus, decomposing the update into reflection and transmission components. In both cases, dipping reflections, in which the low-wavenumber components are present in a direction not easily isolatable from the transmission, can cause problems. In fact, diving waves with limited source and receiver coverage may have high-wavenumber components (at high frequencies) in certain directions, in spite of their value in updating the background, and specifically, the path through which the wave traveled.

To improve the action of the gradient and specifically to allow it to focus on the appropriate long-wavelength components, we can use an approach devised by Khalil et al. (2013) to filter out such components for a cleaner RTM image free of low-frequency artifacts, to do exactly the opposite and enhance these components. We consider source and receiver (computed adjointly) wavefields u_{s_i} and u_{r_i} , respectively, in the frequency domain, for a shot location indexed with i . For RFWI, the gradient also requires the image perturbed source and receiver wavefields δu_{s_i} and δu_{r_i} , respectively, again in the frequency domain (Xu et al., 2012; Wang et al., 2013). In line with Khalil et al. (2013), we introduce a slightly modified time lag (velocity scaled ζ) to our conventional gradient (Alkhali-fah, 2014), and as a result,

$$R_1(\mathbf{x}, \zeta) = \sum_i u_{s_i}(\mathbf{x}) u_{r_i}(\mathbf{x}) e^{-4i\omega \tau_0(\mathbf{x}) \zeta}, \quad (7)$$

$$R_2(\mathbf{x}, \zeta) = \sum_i (u_{s_i}(\mathbf{x}) \delta u_{r_i}(\mathbf{x}) + u_{r_i}(\mathbf{x}) \delta u_{s_i}(\mathbf{x})) e^{-4i\omega \tau_0(\mathbf{x}) \zeta}, \quad (8)$$

are the extended gradients for FWI and RFWI, respectively, and $\zeta = \frac{\tau}{2} v_0(\mathbf{x})$, where τ is the classic time lag (Sava and Fomel, 2006). An inherent feature of this modified time-lag (in distance units) representation is that the relationship of the scattering angle to the wavenumber of the gradient is free of an explicit velocity (space) dependency. In fact, the scattering angle θ , is then given by the following formula:

$$\cos^2 \frac{\theta}{2} = \frac{|\mathbf{k}_m|^2}{k_\zeta^2}, \quad (9)$$

where $\mathbf{k}_m$ is the model wavenumber vector and k_ζ is the wavenumber (Fourier transform) corresponding to ζ . A 4D Fourier transform of $R(\mathbf{x}, \zeta)$ (3D in 2D), will allow us to map $\hat{R}(\mathbf{k}, k_\zeta)$ to its angle gather equivalence $\hat{R}(\mathbf{k}, \theta)$ using equation 9, where $\hat{R}$ is the space Fourier transform of R . In our case, we use equation 9 to filter out the gradient energy corresponding to small θ (reflections), starting for example with $\theta < 170^\circ$, and just sum the rest over k_ζ (the zero ζ imaging condition). As a result, there is no need to map to the angle gathers. Of course, we will have to inverse Fourier transform the gradient back to space to apply the gradient in space. The mute region to eliminate reflections in $\hat{R}(\mathbf{k}, k_\zeta)$ is interestingly given by low conventional wavenumbers and high lag wavenumbers. Obviously, we also mute all energy lying in regions of $\hat{R}(\mathbf{k}, k_\zeta)$, where $\cos \frac{\theta}{2} > 1$.

Relating equation 9 to the model update wavenumber demonstrates that $k_\zeta = \frac{\omega}{v}$, which allows us to directly control the frequency content scaled by the velocity in the update. In other words, k_ζ controls the update wavenumber scale. Thus, the model update wavenumber can be written as

$$\mathbf{k}_m = k_\zeta \cos \frac{\theta}{2} \mathbf{n}, \quad (10)$$

where $\mathbf{n}$ is a unit vector pointing in the direction of $\mathbf{k}_m$. This allows for full control of the model update wavelength regardless of frequencies.

WAVENUMBER DISTRIBUTION FOR ANISOTROPIC MEDIA

Because a dip is typically treated as an unknown, and the Born linearized update is based on scattering theory (that is, from a point in the model), the analysis here focuses on the dependency (radiation pattern) of the anisotropic parameters on the scattering angle and frequency.

Before we start filtering the anisotropic parameter gradients, let us gain some understanding of the wavenumber distribution of the sensitivity of the data to these parameters and learn more about the

Figure 1. Vector plots depicting the sensitivity to η (horizontal component of the vector) and to either (a) δ or (b) ϵ (vertical component of the vector) for the combinations led by v_n and v_h , respectively. The effective angular range for the scattering angle θ is 0° – 180° , and for dip it is ϕ , -90° – 90° .

trade-off between the parameters in the domain we intend to filter in, specifically k_ζ . From equation 9, we note that the scattering angle depends on the magnitude of $\mathbf{k}_m$ and k_ζ . Because $\mathbf{k}_m$ is given by the gradient function represented in the wavenumber domain, we investigate the behavior of the anisotropic parameter sensitivity with respect to the reflector dip (ϕ) and k_ζ . Figure 2a shows such dependency for the parameter combination given by v_n , η , and δ . Like Figure 1a, Figure 2a is a vector plot displaying the sensitivity magnitudes for η (horizontal direction) and δ (vertical direction). The three highlighted regions correspond to our data, with the left block depicting the locations of conventional reflections (with reasonable maximum offset, 6 km), the bottom-right representing the location of diving waves, with their close-to-180° scattering angles, and the top-right corresponds to RFWI, where the effective dip (given by the normal to $\mathbf{n}$ in equation 10) of the kernel is more vertical for horizontal reflectors. For vertical reflectors, the region can extend to zero dip because that part acts like diving waves. However, for FWI, the region is reflection dependent. For the anisotropic combination led by v_n , the sensitivity is largest for δ for reflections in conventional FWI for small dips (vectors pointing vertically on the left side). The parameter η takes over for reflections corresponding

to large dips (vectors pointing horizontally on the left side) or large offsets. For diving waves, the data are sensitive mainly to η , and for RFWI, δ has the upper hand, unless the dips are large or the offset is extremely large, which induces smaller effective dip angles in the gradient, and η starts to have more influence. For the v_h , η , and ϵ combination, the story changes (Figure 2b). For reflections in conventional FWI, the data are more sensitive to ϵ at small dips and η starts to have influence at moderate dips, although both parameters have a very mild influence at large dips. They also have very little influence on the diving wave part, which can be used solely to invert for v_h . For RFWI, there is a trade-off between η and ϵ in favor of ϵ for small offsets or small dips, and η for larger offsets or dips.

A cut-off filter controlled by the scattering angle, in which we mute a region of k_ζ below a certain value, will exclude the classic FWI reflection contribution (left side of Figure 2a and 2b). Actually, in practical implementations, we should initially spare a small band at the rightmost part of k_ζ . For the combination led by v_n , we seemingly would get η from the diving waves and δ from RFWI, both with trade-off potential with v_n , which has an angular independent radiation pattern. For the combination led by v_h , we have little sensitivity to η and ϵ for diving waves, but we have reasonable sensitivity to both parameters in RFWI.

Figure 2. Vector plots synonymous with those shown in Figure 1a and 1b depicting the sensitivity to η (horizontal component of the vector) and to either (a) δ or (b) ϵ (vertical component of the vector) for the combinations led by v_n and v_h , respectively. The highlighted regions correspond to FWI reflections (left), to RFWI (upper right), and FWI diving waves (lower left).

GRADIENTS UNDER FILTERING IN ANISOTROPIC MEDIA

One of the virtues of filtering in the model domain is that it can be applied to the model parameter updates separately, which is not easily done by decimating the data over frequency or offset because these model parameters share the same data. For certain model parameters, the update can be exposed to excessive filtering, and we can loosen the filtering on other parameters to allow for higher resolution contributions to the update. In the examples below, we consider the effect of filtering on classic FWI and RFWI gradients for perturbations in η and δ or ϵ . For the analysis of the scattering-angle filter on v_n and v_h (both with angular invariant radiation patterns), I refer you to Alkhalifah (2014), which is devoted to isotropic media. We first consider the model parameterized by v_n , δ , and η , a combination more suitable as discussed earlier for an inversion that relies first on RFWI. Thus, we show the RFWI gradients first.

The v_{NMO} combination

Figure 3 shows the gradient in the η perturbation for a monochromatic wavefield of 10 Hz for the case of RFWI. Obviously, the components of the gradient corresponding to wavefields traveling nearly horizontal (i.e., a vertical reflector) have more energy, like the sides. On the other hand, the middle part of the gradient is dominated by having one of the wavefields (source or receiver or model perturbation) travel vertically, and thus the amplitude is small. Figure 4 shows the ζ extended version. The vertical section along ζ from the low-energy middle part shows energy only at the perturbed model point. Figure 5 shows the gradients after filtering with various low-cut scattering angles. Despite the low-amplitude contribution for the η perturbation, the filtering managed to isolate such contributions

Figure 3. The RFWI model update response (sensitivity kernel) for the η perturbation for a monochromatic wavefield of 10 Hz and a source located at 1 km and a receiver at 4 km, both at a 0.1-km depth, and a model point at 2.5 km laterally and a 2.0-km depth. The white dots correspond to the locations of the source, receiver, and model point perturbation.

along the wavepath. As a result, we obtain “rabbit ears” for certain cut-off angles, granted that the magnitude is small, compared, for example, with the δ perturbation as we will see later. This is inherent in the low sensitivity to perturbations in η for such a setup. Larger offsets, providing rabbit ears with a wider opening, will admit more sensitivity to η , with similar behavior as a response to the filtering.

In the case of δ perturbation, the gradient for RFWI for a single frequency of 10 Hz is shown in Figure 6. The response shows more energy in parts of the gradient corresponding to near-vertical wavefield propagation, including those along the rabbit ears. Thus, the ζ extension shows energy in the vertical section from the middle crossing the perturbed model point (Figure 7). The application of scattering-angle filtering shown in Figure 8 shows a similar response to that for the η perturbations (Figure 5), but with the opposite sign, and more importantly, a factor of magnitude higher. Thus, δ with this setup is better resolved than η . The opposite behavior happens with classical FWI. However because δ hardly affects the kinematics of the modeled or recorded data, the scaling of the kernel provided by the residuals will be extremely small. Thus, its

effective influence is very small. In other words, in RFWI, the δ influence on the propagator will cancel out between the migration and demigration processes.

Figure 6. The RFWI model update response (sensitivity kernel) for the δ perturbation for a monochromatic wavefield of 10 Hz and a source located at 1 km and a receiver at 4 km, both at 0.1 km depth, and a model point at 2.5 km laterally and 2.0-km depth. The white dots correspond to the locations of the source, receiver, and model point perturbation.

Figure 4. The model update (sensitivity kernel) of Figure 3 with a ζ extension to allow for scattering-angle identification.

Figure 5. The model update (sensitivity kernel) for RFWI of that in Figure 3 after applying the low-cut scattering-angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 170°, and (f) 160°. The white dots correspond to the locations of the source, receiver, and model point perturbation.

After a proper, hopefully good enough for FWI, anisotropy model is attained from RFWI, we usually use that model as an initial model for classic FWI. The key requirement here is that the initial model can produce reflections at the available low frequencies that are a half-cycle away from the observed data. However, because we are filtering the updates for FWI, we can relax such requirements because our focus will still include updating the background model. The move from RFWI to FWI is mainly to address anisotropy. RFWI is used here to establish a good velocity model, and FWI is hopefully ready to give us a good η result. The δ as discussed previously will come from the higher resolution amplitudes. Thus, to understand the effect of filtering on the updates of FWI, we consider again an isotropic constant-velocity background medium. Figure 9 shows the classic FWI sensitivity kernel for the 10-Hz monochromatic wavefields. It is similar to the velocity one, but it has low energy for parts of the gradient corresponding to waves traveling vertically, such as reflections, especially those with small scattering angles. Figure 10 shows the extended ζ gradient. Figure 11 shows the gradient response to scattering-angle filtering. Unlike for RFWI,

the η influence is large and the filter emphasizes that contribution along the ray. A strong low-cut scattering filtering will distribute the direct ray contribution over a large region centered between the source and receiver.

Figure 9. The model update response (sensitivity kernel) for the η perturbation for a monochromatic wavefield of 10 Hz and a source located at 1 km and a receiver at 4 km, both at a 0.1-km depth, and a model point at 2.5 km laterally and 2.0 km depth. The white dots correspond to the locations of the source and receiver.

Figure 7. The model update (sensitivity kernel) of Figure 6 with a ζ extension to allow for scattering-angle identification.

Figure 8. The model update (sensitivity kernel) for RFWI of that in Figure 6 after applying the low-cut scattering-angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 170°, and (f) 160°. The white dots correspond to the locations of the source, receiver, and model point perturbation.

On the other hand, the δ gradient (Figure 12) has little energy between the source and receiver because that energy corresponds to horizontally traveling waves. The ζ extension (Figure 13) also emphasizes the lack of energy in the ζ axis from the center. Because the δ has almost no contribution especially along the line between the sources and receivers, gradient filtering produces mixed results. Because there is no energy in the δ perturbation, some of the cut-off angles produce low-energy results void of direction. At very high cut scattering angles (Figure 14) such as 179.5° and 179.7° , we obtain extremely long-wavelength energy, but in the right direction update. Otherwise, the response has a high wavenumber, which is possibly not helpful in avoiding the local minima. Thus, for δ , we can keep the high-cut filter until a proper background is calculated for v_n and η .

In all cases because the NMO velocity (the third parameter in this combination) has a stationary radiation pattern with angle, there will be a trade-off with both parameters (Alkhalifah and Plessix, 2014). This means that in this setup, if we set δ and η to some fixed value, our inversion for v_n will be dependent on the accuracy of that value. In the parameterization, led by the horizontal velocity, we have a region in the radiation pattern that is exclusively reserved for v_h .

The v_h combination

In this combination, as suggested by Alkhalifah and Plessix (2014), the model is described by v_h , η , and ϵ . Because Figure 2b shows that the horizontal velocity has practically no trade-off with

Figure 12. The model update response (sensitivity kernel) for the δ perturbation for a monochromatic wavefield of 10 Hz and a source located at 1 km and a receiver at 4 km, both at 0.1-km depth and a model point at 2.5 km laterally and 2.0-km depth. The white dots correspond to the locations of the source and receiver.

Figure 10. The model update (sensitivity kernel) of Figure 9 with a ζ extension to allow for scattering-angle identification.

Figure 11. The model update (sensitivity kernel) for the classic inversion of that in Figure 9 after applying the low-cut scattering-angle filter muting angles less than (a) 179.4° , (b) 179° , (c) 178° , (d) 176° , (e) 170° , and (f) 160° . The white dots correspond to the locations of the source and receiver.

the other two parameters for diving waves, we start the inversion and our analysis with classic FWI.

Figure 15 shows the classic FWI sensitivity kernel for the 10-Hz monochromatic wavefields for this combination. Unlike in the previous model representation, the data are less sensitive now to η because v_h mainly describes the data for horizontal propagation. Figure 16 shows the extended ζ gradient. Figure 17 shows the gradient response to scattering-angle filtering. Unlike for the v_n combination, the η influence is small and the filtering indicates the lack of contribution along the ray. In fact, because ϵ has a response similar to δ , shown in the previous subsection, η and ϵ mildly affect the data for direct waves. Such data can be used to invert exclusively for v_h , and we can keep both η and ϵ at a high low-cut scattering angle, which admits only smooth updates. On the other hand, the ϵ gradient is exactly similar to the δ one from the previous section because they have the same radiation pattern for the different combinations.

As soon as a good v_h model is extracted from direct and diving waves, we can start using RFWI to get a good background model for η . Figure 18 shows the gradient in η perturbation for a monochromatic wavefield of 10 Hz for the case of RFWI. Now, unlike in

the v_n combination, we have energy in the center. Figure 19 shows the extended ζ version. Figure 20 shows the gradient after filtering with various low-cut scattering angles. The filtering managed to lower the wavenumber of the gradient and localize it to the potential wavepaths. As a result, we obtain rabbit ears for certain cut-off

Figure 15. The model update response (sensitivity kernel) for the η perturbation for a monochromatic wavefield of 10 Hz and a source located at 1 km and a receiver at 4 km, both at depth 0.1 km, and a model point at 2.5 km laterally and 2.0 km in depth. The white dots correspond to the locations of the source and receiver.

Figure 13. The model update (sensitivity kernel) of Figure 12 with a ζ extension to allow for scattering-angle identification.

Figure 14. The model update (sensitivity kernel) for the classic inversion of that in Figure 12 after applying the low-cut scattering-angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 170°, and (f) 160°. The white dots correspond to the locations of the source and receiver.

angles. As soon as a viable smooth η is obtained, we return back to FWI to resolve the high-wavenumber components of the velocity model by relaxing the scattering-angle filter. For the v_h parameterization, η has more influence at these angles than it had in the v_n parameterization.

DISCUSSION

From the sensitivity kernel maps, we saw that for both suggested three-parameter combinations v_n and v_h have angular independent radiation patterns with an influence on the propagator (background) and scattering parts of the wavefields. For reflections that are reasonably deep, specifically those we care about in surface seismic data, η mainly affects the anisotropic component of the wavefield propagator (the background Green's functions used to map the scatterers), but has a mild influence on the scattering part (near zero scattering potential). Thus, no high-resolution information of η , those responsible for reflections, at depth is present in our surface recorded data. In both cases, δ and ϵ have an opposite influence compared with η for reflections. They hardly affect the propagator, but they have reasonable scattering potential. Thus, they can be used

to help match the amplitude for a lack an accurate physical representation of reflections for velocity. In this case, they do not have a real physical meaning, but they provide a degree of freedom that

Figure 18. The RFWI model update response (sensitivity kernel) for the η perturbation for a monochromatic wavefield of 10 Hz and a source located at 1 km and a receiver at 4 km, both at 0.1-km depth, and a model point at 2.5 km laterally and 2.0-km depth. This η perturbation is based on the model parameterization given by v_h , η , and ϵ . The white dots correspond to the locations of the source, receiver, and model point perturbation.

Figure 16. The model update (sensitivity kernel) of Figure 15 with a ζ extension to allow for scattering-angle identification.

Figure 17. The model update (sensitivity kernel) for the classic inversion of that in Figure 15 after applying the low-cut scattering-angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 170°, and (f) 160°. The white dots correspond to the locations of the source and receiver.

will help us when we impose an acoustic assumption on an elastic medium.

We can summarize the approaches described above into two main strategies to address anisotropy in FWI. The first one is for the case in which we have large offsets and credible diving waves. As discussed earlier, in this case, I recommend using the combination v_h , η , and ϵ . Figure 21a shows the strategy for filtering, for this case, using RFWI and FWI. However, if large offsets or credible diving waves are not available, we have to resort to the second strategy outlined in Figure 21b, which tends to have more trade-off between the parameters. In this case, we invert for v_n , η , and δ . In both cases, we expect to end up with a high-resolution velocity model, a smooth η (an inherent limitation of our surface acquisition) (R. Djebbi, personal communication, 2014), and a high-resolution erroneous ϵ or δ . In every step, the scattering-angle filtering starts with high-angle low cuts down to a low angle. In both strategies, large offsets are desired to obtain smoother updates using the scattering angle filter,

as well as to resolve anisotropy. These strategies are also meant for data dominated by near-horizontal reflections.

Thus, the anisotropic models we obtain mainly play the role of enabling us to extract a better high-resolution velocity model, hopefully free of the acoustic and isotropic limiting assumptions. The η model, which tends to be smooth, has a physical meaning because it represents the anelliptic nature of the transversely isotropic assumption.

The analysis presented in this paper is with respect to the FWI and RFWI sensitivity kernels, which are independent of the data residuals. The data residuals could be based on amplitude or phase misfits. The sensitivity analysis here reflects the high- and low-resolution parameter dependency regardless of the misfit used in FWI or RFWI. Also, the analysis in this paper is meant to provide insights into the VTI multiparameter inversion summarized in the “Discussion” and “Conclusions” sections. These insights include parameter trade-off and suggestions on proper representation,

Figure 19. The model update (sensitivity kernel) of Figure 18 with a ζ extension to allow for scattering-angle identification.

Figure 20. The model update (sensitivity kernel) for RFWI of that in Figure 18 after applying the low-cut scattering-angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 170°, and (f) 160°. The white dots correspond to the locations of the source, receiver, and model point perturbation.

Figure 21. Work flows for applying scattering-angle filters to FWI and RFWI (a) for the case of usable diving waves in the data, and thus inverting for a combination given by v_n , η , and ϵ and (b) for the case in which horizontal reflectors are dominant in which diving waves are poorly recorded, and thus inverting for a combination given by v_n , η , and δ .

hierarchical implementations, as well as the reaction of these parameter updates to scattering-angle filters. Based on the Born approximation and the resulting update kernel, I expect that the imprint of these analytical observations on the inversion process will also depend on the implementation.

CONCLUSIONS

Despite the many features that filtering the gradient brings to the table in the isotropic case, the benefits for anisotropic media are even larger. It allows us to apply different filtering for different parameters. We can apply a specific filter for FWI and another for RFWI, with specific features corresponding to the parameters of perturbation. For an acoustic-anisotropic medium, in which we invert for three anisotropic parameters, using FWI and RFWI allows for six different filtering strategies. These strategies depend on the data content including the amount of diving waves present in the data. Such strategies can be formulated from the analysis provided here. Although velocity affects the propagator and scattering part of the wavefield for surface seismic data, η mainly affects the propagator, and δ or ϵ affects the scattering. They will both enable us to get a higher resolution velocity model.

ACKNOWLEDGMENTS

I thank R. Plessix and R. Djebbi for their fruitful comments and suggestions. I thank the associate editor A. Baumstein, as well as D. Vigh and two anonymous reviewers for their valuable and critical suggestions and comments. I thank KAUST for its support.

REFERENCES

- Aki, K., and P. G. Richards, 1980, Quantitative seismology: Theory and methods: W. H. Freeman and Company.
- Albertin, U., G. Shan, and J. Washbourne, 2013, Gradient orthogonalization in adjoint scattering-series inversion: 83rd Annual International Meeting, SEG, Expanded Abstracts, 1058–1062.
- Alkhalifah, T., 2014, Scattering angle base filtering of the FWI gradients: 76th Annual International Conference and Exhibition, EAGE, Extended Abstracts, E106 10.
- Alkhalifah, T., and R. Plessix, 2014, A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study: *Geophysics*, **79**, no. 3, R91–R101, doi: [10.1190/geo2013-0366.1](https://doi.org/10.1190/geo2013-0366.1).
- Almomin, A., and B. Biondi, 2013, Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations: 83rd Annual International Meeting, SEG, Expanded Abstracts, 1048–1052.
- Bunks, C., F. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: *Geophysics*, **60**, 1457–1473, doi: [10.1190/1.1443880](https://doi.org/10.1190/1.1443880).
- Burridge, R., M. V. de Hoop, D. Miller, and C. Spencer, 1998, Multiparameter inversion in anisotropic elastic media: *Geophysical Journal International*, **134**, 757–777.
- Clement, F., G. Chavent, and S. Gomez, 2001, Migration based traveltime waveform inversion of 2-D simple structures: A synthetic example: *Geophysics*, **66**, 845–860, doi: [10.1190/1.1444974](https://doi.org/10.1190/1.1444974).
- Khalil, A., J. Sun, Y. Zhang, and G. Poole, 2013, RTM noise attenuation and image enhancement using time-shift gathers: 83rd Annual International Meeting, SEG, Expanded Abstracts, 3789–3793.
- Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: *The Leading Edge*, **32**, 1040–1054, doi: [10.1190/le32091040.1](https://doi.org/10.1190/le32091040.1).
- Plessix, R.-É., and Q. Cao, 2011, A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium: *Geophysical Journal International*, **185**, 539–556, doi: [10.1111/j.1365-246X.2011.04957.x](https://doi.org/10.1111/j.1365-246X.2011.04957.x).
- Plessix, R., P. Milcik, H. Rynja, A. Stopin, K. Matson, and S. Abri, 2013, Multiparameter full-waveform inversion: Marine and land examples: *The Leading Edge*, **32**, 1030–1038, doi: [10.1190/le32091030.1](https://doi.org/10.1190/le32091030.1).
- Pratt, R. G., Z.-M. Song, P. Williamson, and M. Warner, 1996, Two-dimensional velocity models from wide-angle seismic data by wavefield

- inversion: *Geophysical Journal International*, **124**, 323–340, doi: [10.1111/j.1365-246X.1996.tb07023.x](https://doi.org/10.1111/j.1365-246X.1996.tb07023.x).
- Prieux, V., R. Brossier, Y. Gholami, S. Operto, J. Virieux, O. I. Barkved, and J. H. Kommedal, 2011, On the footprint of anisotropy on isotropic full waveform inversion: The Valhall case study: *Geophysical Journal International*, **187**, 1495–1515, doi: [10.1111/j.1365-246X.2011.05209.x](https://doi.org/10.1111/j.1365-246X.2011.05209.x).
- Sava, P. C., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: *Geophysics*, **71**, no. 6, S209–S217, doi: [10.1190/1.2338824](https://doi.org/10.1190/1.2338824).
- Sirgue, L., and R. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: *Geophysics*, **69**, 231–248, doi: [10.1190/1.1649391](https://doi.org/10.1190/1.1649391).
- Tang, Y., S. Lee, A. Baumstein, and D. Hinkley, 2013, Tomographically enhanced full wavefield inversion: 83rd Annual International Meeting, SEG, Expanded Abstracts, 1037–1041.
- Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC1–WCC26, doi: [10.1190/1.3238367](https://doi.org/10.1190/1.3238367).
- Wang, S., F. Chen, H. Zhang, and Y. Shen, 2013, Reflection-based full waveform inversion (RFWI) in the frequency domain: 83rd Annual International Meeting, SEG, Expanded Abstracts, 877–881.
- Xu, S., D. Wang, F. Chen, G. Lambare, and Y. Zhang, 2012, Inversion on reflected seismic wave: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: [10.1190/segam2012-1473.1](https://doi.org/10.1190/segam2012-1473.1).