Conditioning the full waveform inversion gradient to welcome anisotropy

Tariq Alkhalifah, KAUST

SUMMARY

Multi-parameter full waveform inversion (FWI) suffers from the complex nonlinearity in the objective function, compounded by the eventual tradeoff 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 the small scattering angles of the gradient update is initially muted out. A model update hierarchical filtering strategy includes applying varying degree of filtering to the different parameter updates. A feature not easily accessible to simple data decimation. Using both FWI and reflection based FWI (RFWI), two strategies to combat the tradeoff between anisotropic parameters are outlined.

INTRODUCTION

Multi parameter inversion (Burrige et al., 1998; Plessix and Cao, 2011; Prioux et al., 2011) requires an appropriate choice of parameters to represent the model. The choice of parametrization will reduce the null space of the inversion only when we reduce the number of parameters we invert for. Alkhalifah and Plessix (2014) analytically analyzed the radial dependency (radiation pattern) of the anisotropic parameter perturbation in an acoustic transversely isotropic (VTI) medium. They advocated using certain parameter combinations for various FWI strategies, including those that start with a model obtained from MVA or reflection tomography, and those dependent first on diving waves. Choosing the right inversion set up 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).

An acoustic VTI medium can be described using the NMO velocity, v_n , and two dimensionless parameters, δ (which relates the v_n to the vertical velocity), and ε (which relates the vertical velocity to the horizontal one, v_h). For surface seismic data, imaging can benefit from using η (which relates the v_n to v_h , avoiding a dependence on the vertical velocity) in place of either δ or ε . Alkhalifah and Plessix (2014) showed that representing the VTI model using the NMO velocity (v_n), δ , and η (two dimensionless anisotropic parameters) offers the proper perturbation radiation pattern for an inver-

sion that includes reflections and diving waves. Since δ mildly effects the geometrical 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 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 , η , and ε offers a practical set necessary to reduce the tradeoff and provide reasonable resolution. In this case, ε plays the role of the amplitude fitting as it mildly effects the kinematics in the recorded wavefield.

Alkhalifah (2014) proposes a filter based on the scattering angle between the wavefields involved in the construction of the gradient to enhance the background components of the update. This filter can be even more useful with a multi parameter setup of an anisotropic inversion. So in this abstract, I analyze the data sensitivity to the anisotropic parameters, and specifically considering the combinations suggested by Alkhalifah and Plessix (2014). The analysis involves the action of this new filter on the various parameter updates for classic FWI and reflection based FWI (RFWI) (Plessix et al., 1995; Xu et al., 2012; Wang et al., 2013).

THE ANISOTROPIC RADIATION PATTERNS

The radiation pattern, as opposed to the gradient wavenumber, provides the size of the influence, and specifically it's directional dependence. For acoustic VTI media, Alkhalifah and Plessix (2014) derive such patterns for mainly two sets of anisotropic parameter combinations they deem to be the most practical. Using the asymptotic Green's function (without multi-pathing),

$$G(\mathbf{x}, \mathbf{y}, \omega) = A(\mathbf{x}, \mathbf{y}) \exp(i\mathbf{k} \cdot \mathbf{x}), \quad (1)$$

with $\mathbf{k}$ is the wavenumber given by $\mathbf{k}_m = \cos \frac{\theta}{2} \frac{\omega}{v_0} \mathbf{n}$, where v_0 is the background velocity, and A is the geometrical amplitude. we can then write (Alkhalifah and Plessix, 2014),

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

with s as the source function, ω is the angular frequency, and v_0 is the background velocity. The function

$$A(\mathbf{x}_s, \mathbf{x}, \mathbf{x}_r, \omega) = \frac{G(\mathbf{x}_s, \mathbf{x}, \omega) G(\mathbf{x}_r, \mathbf{x}, \omega)}{v_0^2(\mathbf{x}) \rho(\mathbf{x})}, \quad (3)$$

with a phase driven by the model update wavenumber $\mathbf{k}_m = \cos \theta \frac{\omega}{v_0} \mathbf{n}$, where $\mathbf{n}$ is a unit vector pointing normal to the dip of a potential reflector. The coefficients of $\mathbf{a}_1$ define the radiation patterns of each parameter for the

given parameterization, v_n , η , δ , for perturbations $\mathbf{r}$, and are given by

$$\mathbf{r}_1 = \begin{pmatrix} r_{v_n} \\ r_\eta \\ r_\delta \end{pmatrix}; \quad \mathbf{a}_1 = \begin{pmatrix} 2 \\ 2n_{sh}^2 n_{rh}^2 \\ -(n_{sz}^2 + n_{rz}^2) \end{pmatrix}. \quad (4)$$

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 $\mathbf{n}_s = (n_{sh}, n_{sz}) = (\sin \theta_s, \cos \theta_s)$, and $\mathbf{n}_r = (n_{rh}, n_{rz}) = (-\sin(\theta_s + 2\alpha), \cos(\theta_s + 2\alpha))$.

Since $v_h = v_n \sqrt{1 + 2\eta}$ and $1 + 2\delta = \frac{1+2\varepsilon}{1+2\eta}$, we have the following relations between perturbations: $r_{v_n} = r_{v_h} - r_\eta$, and $r_\delta = r_\varepsilon - r_\eta$. Thus, the radiation patterns of the parameterization (v_h , η , ε) are

$$\mathbf{r}_2 = \begin{pmatrix} r_{v_h} \\ r_\eta \\ r_\varepsilon \end{pmatrix}; \quad \mathbf{a}_2 = \begin{pmatrix} 2 \\ -n_{sz}^2 n_{rh}^2 - n_{rz}^2 n_{sh}^2 \\ -(n_{sz}^2 + n_{rz}^2) \end{pmatrix}. \quad (5)$$

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 lead by v_{nmo} and v_h , respectively.

Considering the parameter combination of v_n , η , and δ , Figure 1a shows a vector plot with vector components given by the coefficients of r_η (horizontal) and r_δ (vertical) (η and δ radiation patterns, respectively) as a function of scattering angle and potential reflector dip (normal to the model update wavenumber). The shaded regions represent the potential locations of reflection (bottom), diving waves (middle top), RFWI (sides top) gradients for conventional surface seismic data. The size of the arrow (amplitude of the vector) is the size of the sensitivity with respect to perturbations in η and δ , while the direction reveals the dependency distribution. For an arrow to point vertical implies that the sensitivity is due to a δ perturbation. An arrow pointing horizontally implies an η dependency. For an arrow pointing diagonally implies a trade off between the two parameters. On the other hand, the radiation pattern for the NMO velocity is direction invariant, and thus, has a tradeoff with η or δ throughout the potential model point illumination angles. As expected, the η resolution is strongest with diving waves and very large dips. However, in all

cases, it inherents a tradeoff with v_n . To invert for this model parameters, we should first fix δ , which does not effect 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 both r_η (vector horizontal component) and r_ε (vector vertical component) is weak in certain areas and specifically for diving waves, which makes the resolution of v_h possible, without fixing any parameter. However, we also note that even if we resolve v_h , we will need to set initially ε (and hopefully close to it's accurate value) fixed to resolve η as the tradeoff between the two parameters is apparent everywhere.

FILTERING THE GRADIENT

As an alternative to the conventional hierarchical data selection approach used to isolate initially the long wavelength updates to the model, we can filter the gradient to provide such updates in spite of the data. Alkhalifah (2014) proposed to filter the gradient (update) using the scattering angle between the wavefields involved in computing the gradient. This is accomplished efficiently by extending the gradient, $R(\mathbf{x})$ with a normalized time-lag axis, ζ , for a single frequency gradient by multiplying it by $e^{-4i\omega \frac{\zeta}{v(\mathbf{x})}}$, where v for anisotropic media is the background phase velocity. So a four dimensional Fourier transform of $R(\mathbf{x}, \zeta)$ (three dimensional in 2D), will allow us to map the gradient $\hat{R}(\mathbf{k}_m, k_\zeta)$ to it's angle gathers equivalence, $\hat{R}(\mathbf{k}_m, \theta)$ using $\cos^2 \frac{\theta}{2} = \frac{|\mathbf{k}_m|^2}{k_\zeta^2}$, where k_ζ is the wavenumber (Fourier transform) corresponding to ζ . Here, I use the relation to filter out the gradient energy corresponding to small θ (reflections).

GRADIENTS UNDER FILTERING

One of the virtues of the filtering in the model domain, the filtering can be applied to the model parameters updates separately. Something that is not easily done by decimating the data over frequency or offset, as these model parameters share the same data.

The v_n combination

Since we are using v_n , η , and δ , we start by analyzing reflection FWI. Figure 2(a) shows the gradient in η perturbation for a monochromatic wavefield of 10Hz for the case of RFWI. Obviously, the components of the gradient corresponding to wavefields traveling nearly horizontal (i.e. vertical reflector) have more energy, like on 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. Figures 2(b) to 2(d) show the gradients after filtering with various scattering angle low cut off filters. Despite the low amplitude contribution for the η perturbation, the filtering managed to

isolate the contributions along the wave path. As a result, I obtain the rabbit ears for certain cut-off angles, however, the magnitude is small, compared for example to a δ perturbation. The δ perturbation (not shown here for brevity) admitted similar responses to those shown in Figures 2(b) to 2(d), but with a 10 times larger amplitude. This is inherent to the low sensitivity to perturbations in η for such a setup. Larger offsets, providing rabbit ears with a wider opening angles admits more sensitivity to η , with similar behavior as a response to the filtering.

Figure 2: a) The RFWI model update response (sensitivity kernel) for the η perturbation for a monochromatic wavefield of 10 Hz. The model update after applying the low cut scattering angle filter muting angles below b) 179, c) 176, and d) 160 degrees. 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. However, since we are filtering the updates for FWI, we can relax such requirements as our focus will still include the background model. So the move from RFWI to FWI is mainly to treat the potential tradeoff in the anisotropy parameters. Figure 3(a) shows the classic FWI sensitivity kernel for the 10 Hz monochromatic wavefields. It is similar to the velocity one (Alkhalifah, 2014), but it has low energy for parts of the gradient corresponding to waves traveling vertically, like reflections, especially those with small scattering angle. Figures 3(b) to 3(d) show the gradient response to scattering angle filtering. Unlike for RFWI, the η influence is large, 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. On the other hand, the δ gradient has very little sensitivity to the data at these angles. In all cases, since the NMO velocity (the third parameter in this combination) has a stationary radiation pattern with angle, it will have a tradeoff with both parameters (Alkhalifah and Plessix, 2014). In the parametrization, lead by the horizontal velocity, we have a region in the radiation

Figure 3: a) The model update response (sensitivity kernel) for the η perturbation for a monochromatic wavefield of 10 Hz. The model update after applying the low cut scattering angle filter muting angles below b) 179, c) 176, and d) 160 degrees. The white dots correspond to the locations of the source and receiver.

pattern that is exclusively reserved for v_h .

The v_h combination

With a model described by v_h , η , and ϵ , as an alternative, and since, as Figure 1b shows, the horizontal velocity has practically no tradeoff with the other two parameters for diving waves, we start the inversion with the classic FWI focussing on v_h . As soon as a good v_h model is extracted from direct and diving waves, we can start using RFWI to hopefully get a good background velocity for η . Figure 4(a) shows the gradient in η perturbation for a monochromatic wavefield of 10Hz for the case of RFWI. Now, unlike in the v_{nmo} combination, we have energy in the center. Figures 4(b) to 4(d) show the gradient after filtering with various scattering angle low cut off angles. The filtering managed to lower the wavenumber of the gradient and localize it to the potential ray paths. As a result, I obtain the rabbit ears for certain cut-off angles. For the v_h parametrization, η has more influence at these angles than it had in the v_{nmo} parametrization.

FILTERING STRATEGIES

We can summarize the approaches described above into two main strategies to address anisotropy in FWI. The first one is for the case where 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 5 shows the strategy for filtering, for this case, using both RFWI and FWI. However, if large offsets, or credible diving waves, are not available we have to resort to a second strategy outlined in Figure 6, 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 are expected to end up with a high resolution velocity model, a smooth η (an inherent limitation of our surface ac-

Figure 4: a) The RFWI model update response (sensitivity kernel) for the η perturbation for a monochromatic wavefield of 10 Hz. The model update after applying the low cut scattering angle filter muting angles below b) 179, c) 176, and d) 160 degrees.

quisition (Djebbi et al., 2014)), and a high resolution erroneous ϵ or δ . In every step, the scattering angle filtering starts from high down to low angles. In both strategies large offsets are desired to obtain smoother background updates using the scattering angle filter, as well as to resolve anisotropy. These strategies are also meant for data dominated by near horizontal reflections.

CONCLUSIONS

Despite the many features that the scattering angle filtering of the gradient brings to the table in the isotropic case, the benefits for anisotropic media are even greater. It allows us to apply various filtering for different parameters for FWI and RFWI gradients. For an acoustic anisotropic medium, in which we invert for three anisotropic parameters, using FWI and RFWI allows for six different filtering strategy possibilities. These possibilities depend on the data content including the amount of diving waves present in the data. If we had large offsets with good quality early arrivals, a parametrization of the model given by v_h , η , and ϵ allows us to maneuver the FWI nonlinearity and the anisotropy parameter tradeoff. Without large offsets, but alternatively with good quality reflections, a parametrization of the model given by v_n , η , and δ seems optimal.

ACKNOWLEDGMENTS

I thank Rene Plessix for the many insights I managed to develop in FWI. I thank KAUST for it's support.

Figure 5: A work flow for applying scattering angle filters to FWI and RFWI for the case of usable diving waves in the data, and thus, inverting for a combination given by v_h , η , and ϵ .

Figure 6: A work flow for applying scattering angle filters to FWI and RFWI for the case where horizontal reflectors are dominant in which diving waves are poorly recorded, and thus, inverting for a combination given by v_n , η , and δ .

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

- Alkhalifah, T., 2014, Scattering angle base filtering of the FWI gradients: Presented at the 76th Annual International Conference and Exhibition, EAGE.
- 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.
- Burridge, R., M. V. de Hoop, D. Miller, and C. Spencer, 1998, Multiparameter inversion in anisotropic elastic media: *Geophysical Journal International*, **134**, 757–777.
- Djebbi, R., T. Alkhalifah, and P. R., 2014, Analysis of the traveltimes sensitivity kernels for the acoustic vertical transverse isotropic medium: *Geophysical Prospecting*, in press.
- Plessix, R., and Q. Cao, 2011, A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium: *Geophysical Journal International*, **185**, no. 1, 539–556, <http://dx.doi.org/10.1111/j.1365-246X.2011.04957.x>.
- Plessix, R., Y. De Roeck, and G. Chavent, 1995, 1995, Automatic and simultaneous migration velocity analysis and waveform inversion of real data using a MBTT/WKJB formulation: 65th Annual International Meeting, SEG, Expanded Abstracts, **304**, 1099–1102.
- 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**, no. 3, 1495–1515, <http://dx.doi.org/10.1111/j.1365-246X.2011.05209.x>.
- 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, doi: 10.1190/segam2013-0671.1.
- Xu, S., D. Wang, F. Chen, G. Lambare, and Y. Zhang, 2012, Inversion on reflected seismic wave: Presented at the 82nd Annual International Meeting, SEG.