Geophysical Journal International

Geophys. J. Int. (2015) **200,** 363–373
GJI Seismology

doi: 10.1093/gji/ggu379

Scattering-angle based filtering of the waveform inversion gradients

Tariq Alkhalifah

Department of Physical Science and Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900*, Saudi Arabia.*
E-mail: tariq.alkhalifah@kaust.edu.sa

Accepted 2014 September 26. Received 2014 September 25; in original form 2014 April 5

SUMMARY

Full waveform inversion (FWI) requires a hierarchical approach to maneuver the complex non-linearity associated with the problem of velocity update. In anisotropic media, the non-linearity becomes far more complex with the potential trade-off between the multiparameter description of the model. A gradient filter helps us in accessing the parts of the gradient that are suitable to combat the potential non-linearity and parameter trade-off. The filter is based on representing the gradient in the time-lag normalized domain, in which the low scattering angle of the gradient update is initially muted out in the FWI implementation, in what we may refer to as a scattering angle continuation process. The result is a low wavelength update dominated by the transmission part of the update gradient. In this case, even 10 Hz data can produce vertically near-zero wavenumber updates suitable for a background correction of the model. Relaxing the filtering at a later stage in the FWI implementation allows for smaller scattering angles to contribute higher-resolution information to the model. The benefits of the extended domain based filtering of the gradient is not only it's ability in providing low wavenumber gradients guided by the scattering angle, but also in its potential to provide gradients free of unphysical energy that may correspond to unrealistic scattering angles.

Key words: Inverse theory; Seismic tomography; Acoustic properties.

1 INTRODUCTION

The hierarchical approach of full waveform inversion (FWI) of starting with the low-frequency (Bunks *et al.* 1995; Pratt *et al.* 1996; Virieux & Operto 2009) and the early-arrival part of the data (Jang *et al.* 2009) is slowly giving way to more emphases on filtering and conditioning the gradient (Albertin *et al.* 2013; Tang *et al.* 2013). This is especially true as we are trying to mix migration velocity analysis (MVA) and FWI (Almomin & Biondi 2012; Fleury & Perrone 2012; Ma *et al.* 2012; Xu *et al.* 2012; Wang *et al.* 2013). This is a natural progression that 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 & Pratt 2004). With MVA integrated into FWI, we rarely damp the later arrivals as they are needed for the MVA part, but the model domain provides the opportunity to select or shape the proper updates. For anisotropic media, model domain conditioning becomes even more important as we are able to apply different filters to different parameters update, a feature not readily available in the data domain. It also goes beyond model regularization as it is extracted from the data.

However, a viable process to filter the gradient has always been either expensive or inadequate. A simple wavenumber filter could mute essential background update information like those coming from diving waves. A directional wavenumber filtering suitable for diving waves is totally orthogonal to what an MVA gradient would require. A velocity-guided Gaussian filter (Ravaut *et al.* 2004) will isolate the low wavenumbers, which is helpful, but, again, without a clear physical understanding of what was filtered within the FWI context. In fact, the diving wave directional filter very much coincides with gradients that promote reflections (perturbations), which is something we try to avoid in FWI at the early stages. In his paper, Mora (1989) clearly pointed out the role of reflections and transmissions in updating the high and low wavenumbers of a model, respectively. However, in the Born approximation, which is used to develop the gradient, such a distinction is not as clear. I will show here that the real distinction is in the scattering angle alone, and specifically within diving waves along (only) the ray path, which actually allow zero wavenumber updates that can be used for updating the background model.

Wu & Toksz (1987) and Mora (1989) in two classic papers describe the behaviour of the gradients (model updates) in the wavenumber domain by analysing the diffraction tomography behaviour at the heart of the update process (Cohen & Bleistein 1977; Panning *et al.* 2009). Such work paved the way to Sirgue & Pratt (2004) among other results, as they used such insights to confirm the hierarchical approach in dealing with data for FWI. Only recently was the concept of filtering the gradient recognized as a valid alternative. If the gradient includes updating more than one parameter,

the gradient conditioning and filtering will be more useful in the model domain.

In this paper, I extend the axis of the model update with a normalized time lag component capable of resolving the scattering angle, efficiently. This additional axis is used to properly filter the gradient to admit usable background updates for FWI from reflections and diving waves, as well as reflection (Chavent & Plessix 1999) full waveform inversion (RFWI). It utilizes the fact that a proper background update is not necessarily given by the low frequencies or low wavenumbers, but more accurately by large scattering angles, where the update actually follows the rays.

2 THE MODEL's UPDATE CHARACTERISTICS

The model update is at the heart of full waveform inversion. First, to understand the action of the update and its resolution, let us revisit the model update wavenumber, which will provide us some insight into the relation between transmissions and reflections.

2.1 The model's wavenumber update

The Born approximation provides us with the first-order sensitivity of the data to the model parameters, and thus, provides us with the gradient with respect to the model and the corresponding wavelength content. In isotropic media, the update wavelength at a model point, as illustrated by diffraction tomography, is governed by the dip of a potential reflector and the scattering angle. Specifically, the local model wavenumber vector with respect to a potential scatterer in our model (Miller *et al.* 1987; Jin *et al.* 1992; Thierry *et al.* 1999),

$$\mathbf{k_m} = \mathbf{k_s} + \mathbf{k_r} = \frac{\omega}{v}\cos\frac{\theta}{2}\mathbf{n}, \tag{1}$$

which depends on, among other things, the angular frequency, ω, with a direction guided by a unit vector, $\mathbf{n}$, normal to a potential reflector. Here, $\mathbf{k}_s$ and $\mathbf{k}_r$ are the source and receiver (or state and adjoint state) wavefield wavenumbers, respectively, at the model point, with θ as the scattering angle between the source and receiver wavefields (or any two wavefields involved in the creation of the gradient) and v is the velocity.

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. Also, lower frequencies induce long wavelength updates, but they are not the only source of long wavelength information. Clearly, large offsets, inducing large scattering angles, reduce the wavenumber of the model update. Hence, our inversion schemes prefer lower frequencies and large offsets (Pratt *et al.* 1996; Virieux & Operto 2009). Since the model updates for direct, diving and reflection waves are handled using the same Born assertions, they all conform to this wavenumber formula (1). For refractions or direct arrivals, where $\theta = \pi$, the wavenumber is zero. However, the confusing part to many is that the common banana sensitivity kernel (between a source and receiver) that we have grown accustomed to seeing (Woodward 1992) disputes this assertion. We will show below that direct and diving waves actually admit zero wavenumber velocity updates only along the ray (scattering angle equal to π) part of the banana kernel.

2.2 The myth of the transmitted and reflected waves in FWI

We have always divided waves, especially for FWI, into direct and diving waves (or transmitted waves in general) and reflected ones, though they are handled by the same Born-approximated update kernel. They, of course, provide different information, but in physics they are connected by the same wave equation that imposes the continuity of its solutions, even for a discontinuous Earth. A reflection is simply a mirror image of a potential direct wave with a virtual source placed below the reflector. In the limit of the scattering angle approaching 180° for an inhomogeneous medium, the reflected arrival becomes a diving one and, in that case, no contrast is needed as the wavenumber corresponding to this phenomenon equals zero. Waves with a wavenumber component in the direction normal to a potential reflector that is smaller in value than the maximum model wavenumber of that reflector will induce a reflection by that wavenumber. For a sharp discontinuity, the maximum wavenumber will include most wave components hitting that sharp discontinuity. However, in a smoothly increasing velocity model with depth, the wavenumber component normal to the wave path is zero (along the ray), and at some point that zero wavenumber will be smaller than the maximum wavenumber in the model, which will turn the ray or even reflect it, depending on the nature of the increase. In this case, that component is in the depth direction. For scattering angles slightly lower than 180°, the model wavenumber is small, indicating that even a smooth velocity increase with a wavenumber higher than that given by the wave can induce scattering. In fact, the model update wavenumber can be thought of as the scattering wavenumber, and for diving waves along the ray, it is zero.

As transmission (i.e. diving) waves admit near zero wavenumber updates according to the model wavenumber update formula (1), and depending on the frequency (and specifically with high frequencies), the wavenumber can increase within the first Fresnel zone reasonably fast away from the ray path, yielding reasonable resolution to the model update. Defining the border between transmissions and reflections, based on the first Fresnel zone, is both frequency and experiment dependent, and more importantly it ignores the natural continuity between transmissions and reflections as they adhere to the same model update formula. Though beyond the first zero crossing away from the ray may admit the wrong background update direction (in the reflection zone), they are needed to admit proper weighting for the updates over the whole model. We will look into a more natural way to develop smooth model updates for the early part of FWI, and specifically to maneuver the classic local minima of the FWI objective function.

3 GRADIENTS FOR CLASSICAL AND REFLECTION-BASED FWI

The classic objective function for FWI is given by (Tarantola 1987)

$$E_1(m) = \sum_i \left|d_{o_i} - d_{s_i}(m)\right|^2, \tag{2}$$

where the i corresponds to the source index, d_o is the observed data, $d_s = u_s(x, y, z = 0, t)$ is the modelled synthetic data and $m(\mathbf{x})$ is the velocity model described in space using the vector $\mathbf{x} = \{x, y, z\}$, along the Cartesian coordinates, with t as time. From this point on, and for simplicity of derivation, I will represent wavefields and data in the frequency domain. In this case, the modelled wavefield, u_s, satisfies the wave equation

$$L(m)u_{s_i} = f\delta(\mathbf{x} - \mathbf{x}_{s_i}), \tag{3}$$

for a particular source location, $\mathbf{x}_{s_i}$, and f is the source function given in the frequency domain with

$$L(m) = L^t(m) = \nabla^2 + \omega^2 m(\mathbf{x}), \tag{4}$$

as L is the Helmholtz operator, which in this case is self adjoint, ω is the angular frequency and ∇ is the Laplacian operator. The gradient corresponding to such an objective function is given by

$$R_1(\mathbf{x}) = \omega^2 \sum_i u_{s_i}(\mathbf{x}) u_{r1_i}(\mathbf{x}) \tag{5}$$

with

$$L(m)u_{r1_i} = \sum_j \delta(\mathbf{x} - \mathbf{x}_{r_j})(d_{o_{ij}} - d_{s_{ij}})^*, \tag{6}$$

where $\mathbf{x}_{r_j}$ are the locations of the receivers for a particular receiver, j.

Since one of the biggest challenges in FWI is to have reasonably accurate reflections in the modelled data, the migrated image is used as a source for such reflections (Clément *et al.* 2001). This approach, however, produces only smooth updates since producing a reflection is not an objective of the approach. Nevertheless, such inverted models can be used as a starting model for classic FWI. This new objective functional is given by:

$$E_2(m, I) = \sum_i \left| d_{o_i} - d_{m_i}(m, I) \right|^2, \tag{7}$$

where d_m is the modelled data from the image, $I(\mathbf{x})$, obtained by classic (hopefully true amplitude) migration is

$$I(\mathbf{x}) = \sum_i u_{s_i}(\mathbf{x}) u_{a_i}(\mathbf{x}), \tag{8}$$

where the receiver wavefield, u_{a_i}, satisfies

$$L^t(m)u_{a_i} = \sum_j \delta(\mathbf{x} - \mathbf{x}_{r_j}) d^*_{o_{ij}}, \tag{9}$$

and thus,

$$L(m)\delta u_{s_i} = I(\mathbf{x}) u_{s_i}(\mathbf{x}), \tag{10}$$

and $d_{m_i} = \delta u_{s_i}(\mathbf{x}_{r_i})$. The gradient corresponding to this objective functional is given by

$$R_2(\mathbf{x}) = \omega^2 \sum_i \left[u_{s_i}(\mathbf{x}) \delta u_{r_i}(\mathbf{x}) + u_{r2_i}(\mathbf{x}) \delta u_{s_i}(\mathbf{x}) \right], \tag{11}$$

where

$$L^t(m)u_{r2_i} = \sum_j \delta(\mathbf{x} - \mathbf{x}_{r_j})(d_{o_{ij}} - d_{m_{ij}})^*, \tag{12}$$

$$L^t(m)\delta u_{r_i} = I(\mathbf{x}) u_{r_i}(\mathbf{x}). \tag{13}$$

For cross-correlation objective functions (Choi & Alkhalifah 2012; Xu *et al.* 2012), we obtain similar gradient formulations, but the residual is replaced by the observed data, a feature that will mitigate the role of the source and specifically reduce the role of amplitudes to a scaling factor.

4 FILTERING THE GRADIENT

As an alternative to the conventional hierarchal data selection approach used initially to isolate the long wavelength updates to the model (Sirgue & Pratt 2004), we can filter the gradient to provide such updates despite the data (Albertin *et al.* 2013; Almomin & Biondi 2013; Tang *et al.* 2013). This recent development is based on the separation of scales in the updates 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 updates to reflection and transmission parts. In both cases, reflections corresponding to dipping reflectors, 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, despite their value in updating the background, and specifically the path through which the wave travels.

To improve the action of the gradient and specifically to allow it to focus on the appropriate long wavelength components of the gradient, we can utilize 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. This approach is based on a partial decomposition of the image to it's angle gather components using a slightly modified approach to that originally suggested by Sava & Fomel (2006). In line with Khalil *et al.* (2013), we introduce a slightly modified time lag (velocity scaled, ζ) to our conventional gradient (Alkhalifah 2014),

$$R_1(\mathbf{x}, \zeta) = \sum_i u_{s_i}(\mathbf{x}) u_{r_i}(\mathbf{x}) \mathrm{e}^{-4\mathrm{i}\omega \frac{\zeta}{v(\mathbf{x})}}, \tag{14}$$

$$R_2(\mathbf{x}, \zeta) = \sum_i \left[u_{s_i}(\mathbf{x}) \delta u_{r_i}(\mathbf{x}) + u_{r_i}(\mathbf{x}) \delta u_{s_i}(\mathbf{x}) \right] \mathrm{e}^{-4\mathrm{i}\omega \frac{\zeta}{v(\mathbf{x})}}, \tag{15}$$

where v is the velocity, and $\zeta = \frac{\tau}{2} v(\mathbf{x})$. An inherent feature of this modified time-lag (distance units) representation is that the relationship between the scattering angle and the wavenumber of the gradient is free of a velocity (space) dependency. In fact, the scattering angle, θ, is then given by the following relation:

$$\cos^2 \frac{\theta}{2} = \frac{|\mathbf{k}|^2}{k_\zeta^2}, \tag{16}$$

where $\mathbf{k} = \mathbf{k}_\mathrm{m}$ is the space wavenumber vector of the gradient and k_ζ is the wavenumber (Fourier transform) corresponding to ζ. A 4-D Fourier transform of $R(\mathbf{x}, \zeta)$ (3-D in 2-D), will allow us to map $\hat{R}(\mathbf{k}, k_\zeta)$ to its angle gather equivalence, $\hat{R}(\mathbf{k}, \theta)$ using eq. (16). In our case, we use eq. (16) to filter out the gradient energy corresponding to small θ (reflections), starting probably by muting energy 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 angle gathers. Of course, we will have to inverse Fourier transform the image 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 space wavenumbers and high ζ wavenumbers. Obviously, we also mute all energy laying in regions of $\hat{R}(\mathbf{k}, k_\zeta)$, where $\cos \frac{\theta}{2} > 1$, which is unphysical for the considered model (i.e. acoustic).

Relating eq. (16) 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 wavenumber of the model updates can be written:

$$\mathbf{k}_m = k_\zeta \cos \frac{\theta}{2} \mathbf{n}. \tag{17}$$

This allows for a full control of the model update wavelengths regardless of frequencies.

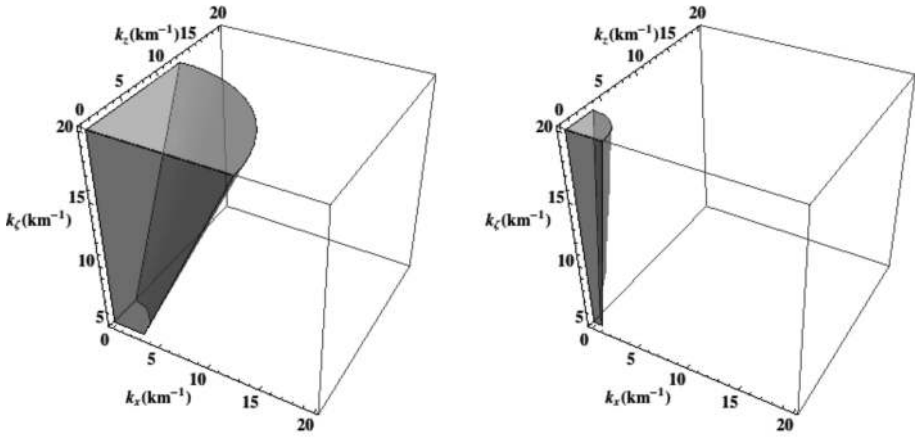

Figure 1. A plot showing in grey the regions in the model update (gradient) that are spared the muting to allow a minimum scattering angle of (a) 100° and (b) 160°. The highlighted regions correspond to reflections (bottom), to RFWI (upper corners) and diving waves (upper centre).

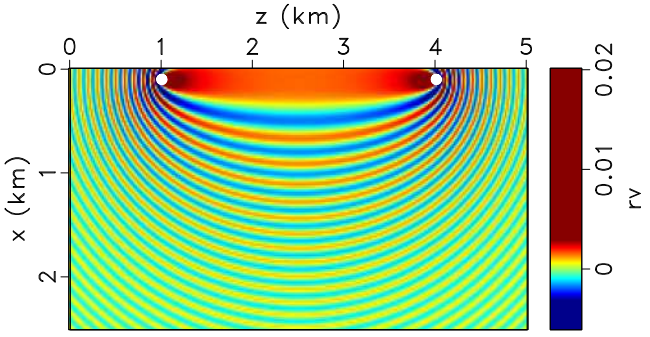

Figure 2. The model update response for a monochromatic wavefield at 10 Hz and a source located at 1 km and a receiver at 4 km, both at 0.1 km depth. The scale bar reflects the relative update that needs to be scaled by the residual and a normalized Hessian. The white dots correspond to the locations of the source and receiver.

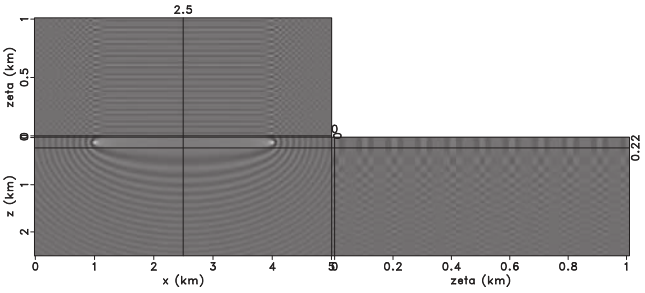

Figure 3. The model update (sensitivity kernel) of Fig. 2 with a ζ extension to allow for scattering angle identification.

For minimum scattering angles of 100° and 160°, the grey area in Figs 1(a) and (b), respectively, shows the areas of the gradient represented in the space and ζ wavenumber domain that will be spared the muting. Below, we mute scattering angles lower than 178°, which will provide a focus on the ray part of the gradient that admits actually almost zero wavenumber updates.

5 GRADIENTS UNDER FILTERING

Since the update (gradient) is best understood by analysing its impact on the model with a single frequency, as well as with a single source, receiver, and, in the case of using reflections, a single model point perturbation, we focus here on analysing the effect of filtering the update on such sensitivity kernels. We consider monochromatic wavefields initially in a homogeneous background velocity model, with velocity equal to 2 km s^{-1}, and a source and receiver at 0.1 km depth, with an offset of 3 km. Fig. 2 shows the 10 Hz monochromatic gradient function. Clearly displayed is the classic sensitivity kernel we have grown accustomed to and attributable to Woodward (1992). The velocity gradient scale is meant here for comparison, as the actual value depends on the residual and is scaled by the Hessian, or simply by an appropriate line search. In Fig. 3, the ζ extended version of the gradient is displayed. In the prescribed implementation, we next perform a 3-D Fourier transformation to this extended gradient, and then filter out the low scattering angle energy using eq. (16). Figs 4(a)–(f) display the gradient after filtering for various low-cut scattering angles. For low-cut at high scattering angles, we obtain a low wavelength model update, with most of the update focused between the source and receiver. As we include lower scattering angles, the gradient includes high wavenumber information. Since we cut most of the angles, the amplitude of the update is smaller. This is the gradient for a single 10 Hz frequency, corresponding to a single source and receiver. As we include more sources and receivers, the gradient will develop into a more complete representation of the required update. More importantly, the energy associated with reflections, the dominant source of non-linearity, is mitigated at high scattering angle cuts. A simple wavenumber filtering of the gradient in Fig. 2 will not yield such updates, as some of the reflection energy includes zero wavelength components, specifically along the reflector. This response to a single frequency and source and receiver position reveals the essence of the capability of this filter. Note that a zero scattering angle cut update provides the gradient of our classical FWI update shown in Fig. 2.

If we use reflections to update the background medium as suggested by Xu *et al.* (2012), the update for a perturbation in the model at 2.5 km depth between the source and receiver is shown in Fig. 5. Embedded in the update is the classic rabbit ear response masked slightly by the monochromatic nature of the wavefields involved. Fig. 6 includes the ζ extended gradient. Figs 7(a)–(f) show the response to a low-cut scattering angle filter for various angles. We obtain a similar effect to that for the classic FWI gradient. However, we also obtain the rabbit ears at scattering angle cuts of 170–176° even for this monochromatic operator. In fact, with a single

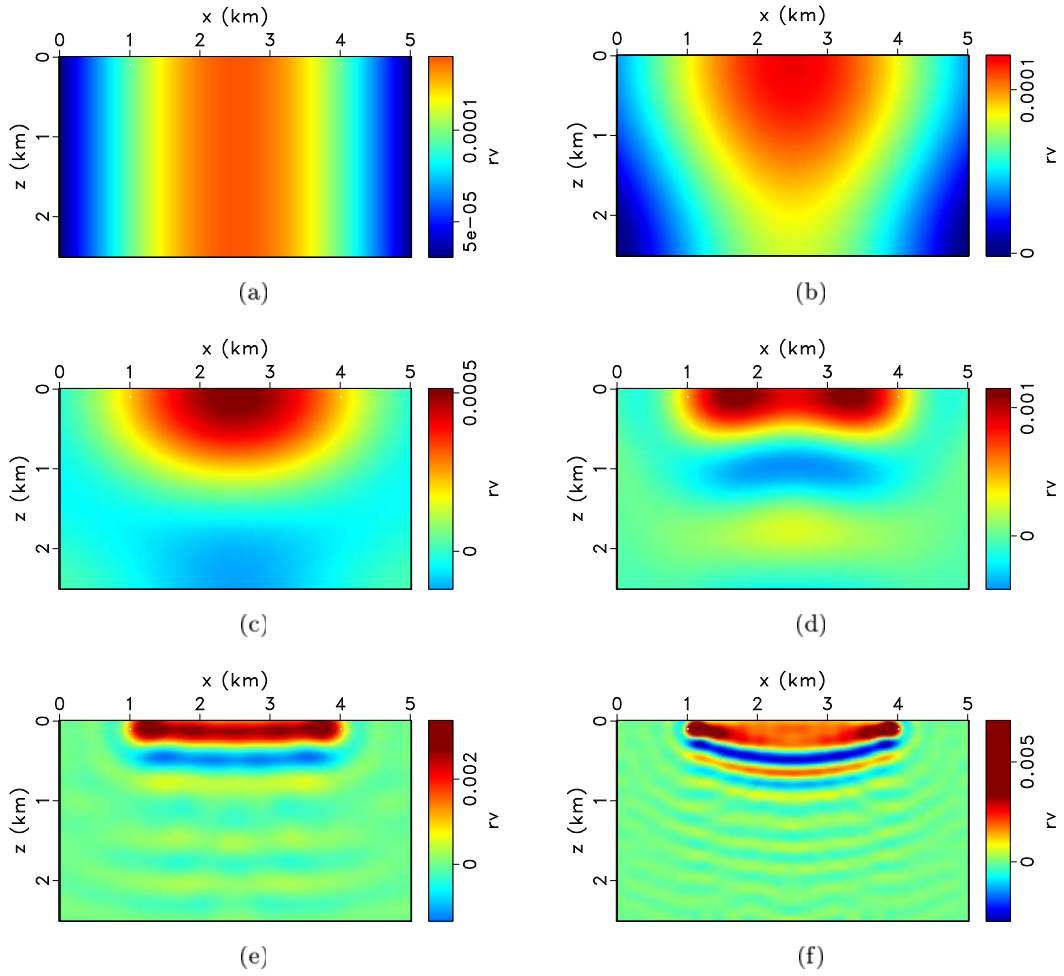

Figure 4. The model update (sensitivity kernel) after applying the low-cut scattering angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 160° and (f) 140°. The white dots correspond to the locations of the source and receiver.

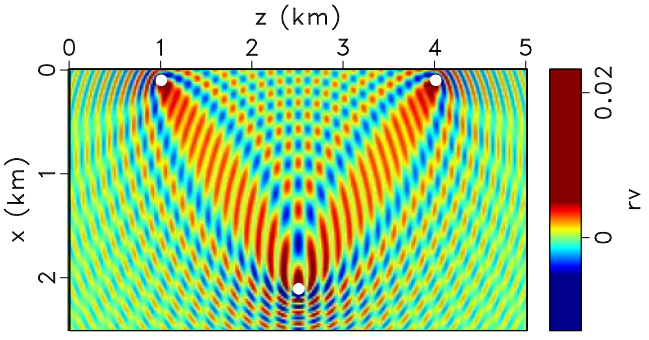

Figure 5. The model update response for a monochromatic wavefield at 10 Hz and a source located at 1 km and a receiver at 4 km, both at 0.1 km depth, and a model point at 2.5 km laterally and 2.5 km depth. The white dots correspond to the locations of the source, receiver and model point perturbation.

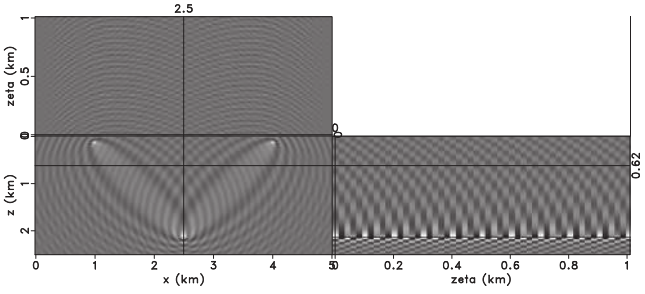

Figure 6. The model update (sensitivity kernel) of Fig. 5 with a ζ extension to allow for scattering angle identification.

relatively high frequency (in the FWI world), we obtain an operator equivalent to that obtained using RFWI for a band of frequencies (Xu *et al.* 2012).

If the reflector is horizontal, instead of assuming a point scatterer as used by Xu *et al.* (2012), the update for a perturbation in the reflector at 2.5 km depth is shown in Fig. 8. Embedded in the update is the actual rabbit ear response masked slightly by the monochromatic nature of the wavefields involved. Figs 9(a)–(f) show the response to a low-cut scattering angle filter for various angles. A focusing of the energy for the update near the reflector is a result of the contribution of the reflector as a group of scatterers.

6 FOR INHOMOGENEOUS MEDIA

For inhomogeneous media, and for efficiency purposes only, I use Born theory (Beydoun & Mendes 1989; Moser 2012) for rays to develop the sensitivity kernels. Equivalent results are expected from wave extrapolation based gradients as the implementation is independent of the wavefield solution approach. We repeat the setup for the classic FWI and the reflection one first for a vertical linear

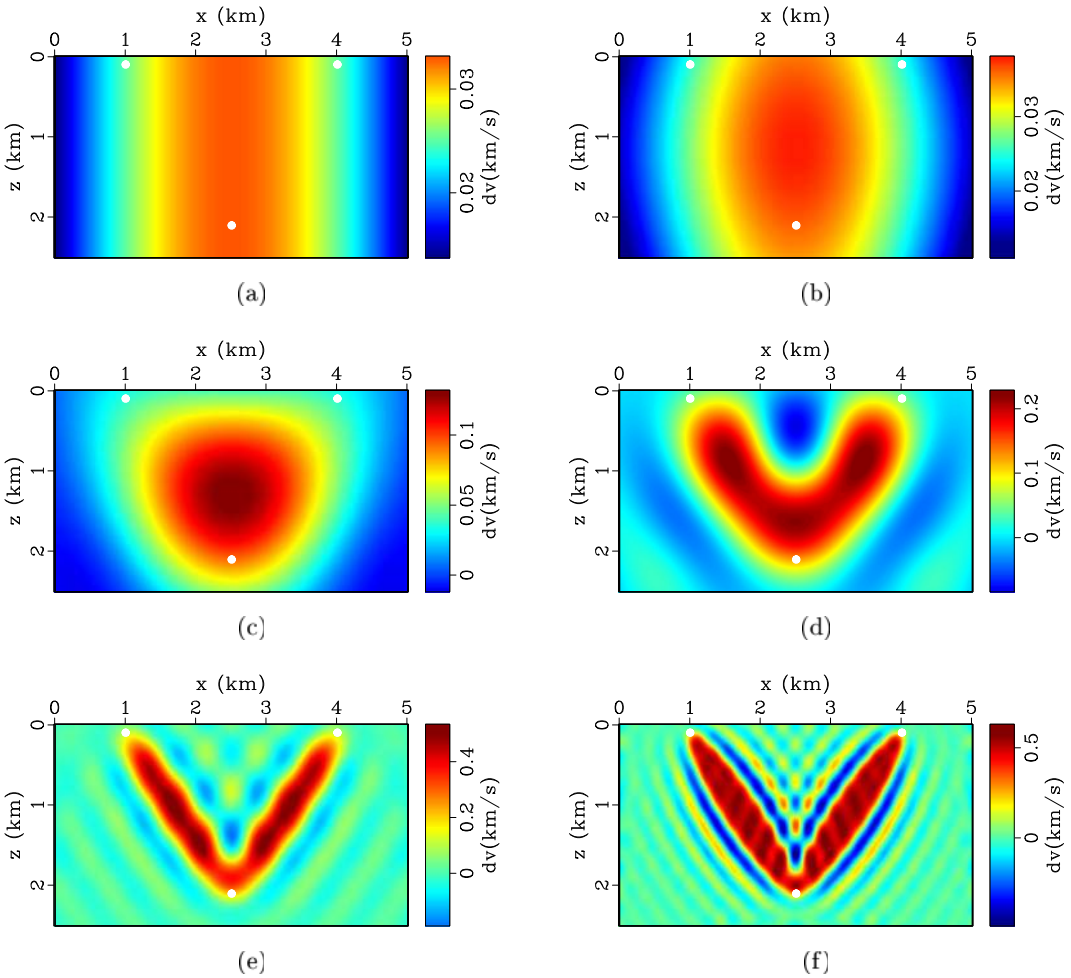

Figure 7. The model update (sensitivity kernel) for the reflection-based inversion after applying the low cut scattering angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 160° and (f) 140°. The white dots correspond to the locations of the source, receiver and model point perturbation.

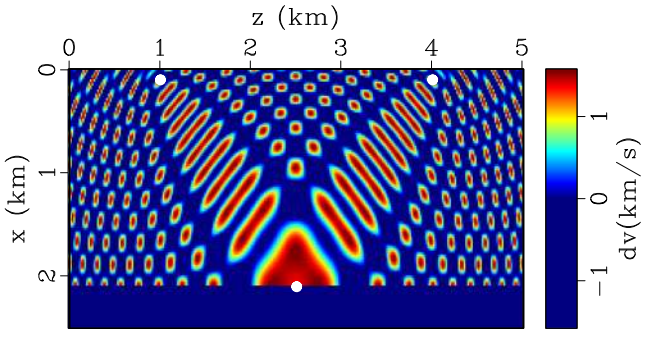

Figure 8. The model update response for a monochromatic wavefield at 10 Hz and a source located at 1 km and a receiver at 4 km, both at 0.1 km depth, and a model point at 2.5 km laterally and 2.5 km depth.

velocity variation of $v(z) = 1500 + z$ m s^{-1}. Fig. 10 shows the classic FWI monochromatic sensitivity kernel for this $v(z)$, with its diving wave characteristic (semi circle) central ray path. The kernel also includes the response of scattering in the full model providing the common imaging isochrons for this source and receiver pair. Figs 11(a)–(f) display the gradient after filtering for various low-cut scattering angles. Again, for high scattering angles, we obtain a low wavelength model update, with most of the update focused between the source and receiver. As we include lower scattering angles, the gradient includes high wavenumber information and follows the ray path of the diving waves. Reflections, especially those corresponding to horizontal reflectors (where scattering angles are the largest) start to appear. We slowly obtain the classical update, like that given by Fig. 2, as we allow the small scattering angles to pass.

Next, we investigate the Marmousi response. I use the eikonal solver, for simplicity, to provide me with the traveltimes required for the ray-Born representation. Although we will miss the later arrivals, the point here is to investigate the action of the filter. The Marmousi model is shown in Fig. 12 along with a traveltime contour map courtesy of the eikonal solver for a source at 2 km location and 0.1 km depth. For the Marmousi model, we consider monochromatic wavefields with a frequency of 7 Hz, for a change. Fig. 13 shows the corresponding sensitivity kernel from classic FWI. It bares all the complexity expected from such a complicated model. The ζ extended version is shown in Fig. 14, which we will use to filter out energy below certain scattering angles. Figs 15(a)–(f) show the response to a low-cut scattering angle filter for various angles. Note that the update energy is focused more toward the (low velocity) source part of the kernel. Some of this discrepancy will be corrected by the geometrical spreading (given by the diagonal elements of the Hessian), not applied here. However, there is also a natural bias of update energy to low velocities due to the larger amplitudes and shorter wavelengths of the involved waves in such areas.

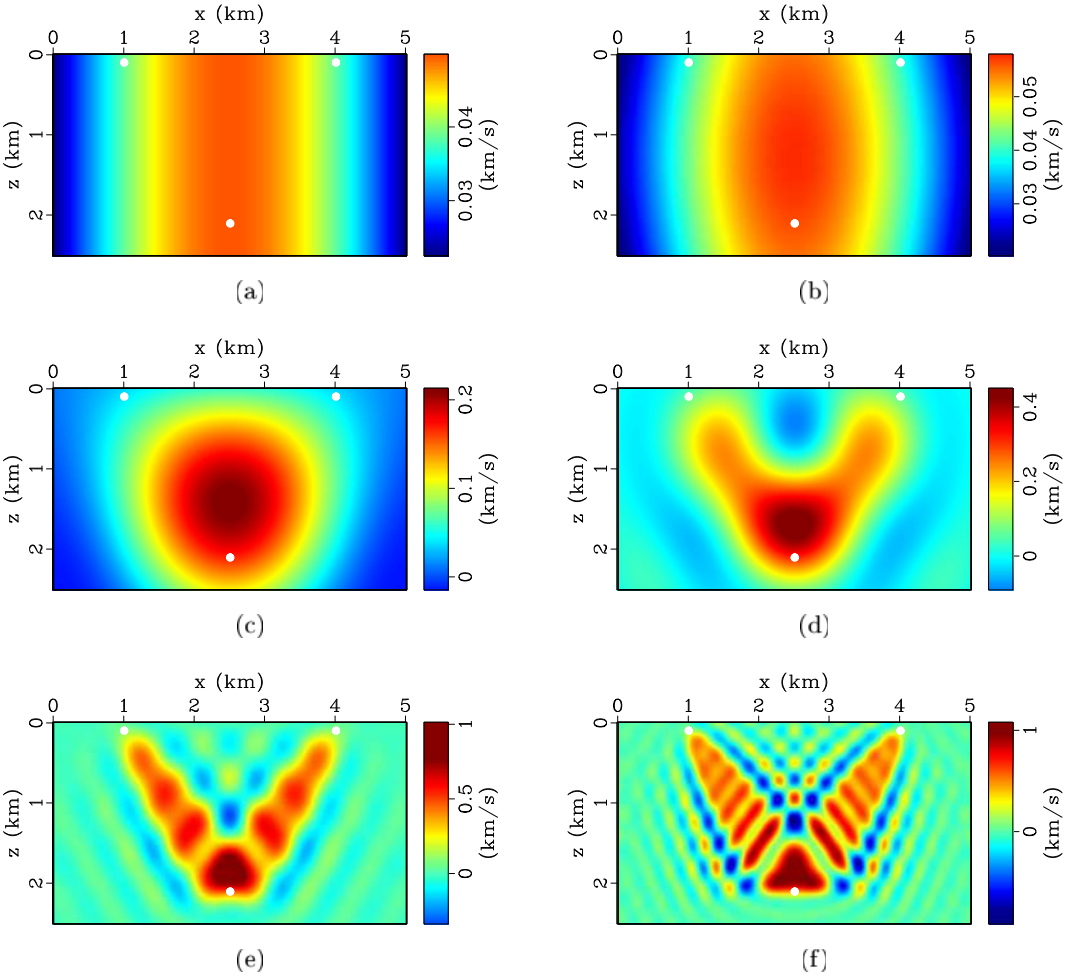

Figure 9. The model update (sensitivity kernel) for the reflection-based inversion after applying the low cut scattering angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 160° and (f) 140°.

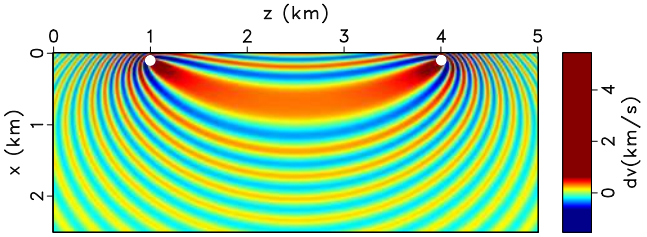

Figure 10. The model update response for a monochromatic wavefield at 10 Hz and a source located at 1 km and a receiver at 4 km, both at 0.1 km depth. The scale bar reflects the relative update that needs to be scaled by the residual and a normalized Hessian. The white dots correspond to the locations of the source and receiver.

For the reflection FWI, I place the perturbation at 2.5 km depth, located within the presumed reservoir location as shown in Fig. 16. Fig. 17 shows the corresponding sensitivity kernel for the background model corresponding to this model perturbation. Figs 19(a)–(f) display the result of the application of the low-cut scattering angle filter. The bias seen in the classical FWI is mildly mitigated by the model point perturbation influence and location near the receiver. The high-cut scattering angle resulting in the gradient energy is concentrated almost in the middle between the source, receiver, and the model perturbation point. Again, as we lower the low-cut scattering angle, we obtain higher resolution information.

7 THE APPLICATION

Despite that the scattering angle filtering, applied throughout the paper, was for a single source and receiver. It's application in practice will include all available (or used) source and receivers. Thus, the filtering maybe applied directly on the full gradient. The resultant filtered gradient may also be used to obtain an approximate Hessian along the lines of the Gauss–Newton implementation. Fig. 20(a) shows the gradient obtained from the Marmousi model considering an initial homogenous background model and sources and receivers covering most of the surface. The gradient corresponds to solving the Helmholtz wave equation for a frequency of 5 Hz. The gradient clearly includes some of the update features we need and expect for the Marmousi model, but it also includes some of the high wavenumber information we want to avoid at this stage since the background velocity model is not accurate.

Figs 20(b)–(d) show the result of applying the scattering angle filtering to cut low scattering angle energy under 178°, 176° and 174°, respectively. In all cases the new gradient includes only low wavenumber updates, the kind we would like to have at the early stages of FWI. The updates at lower scattering angle low cut includes some of the key features of the Marmousi model. The difference

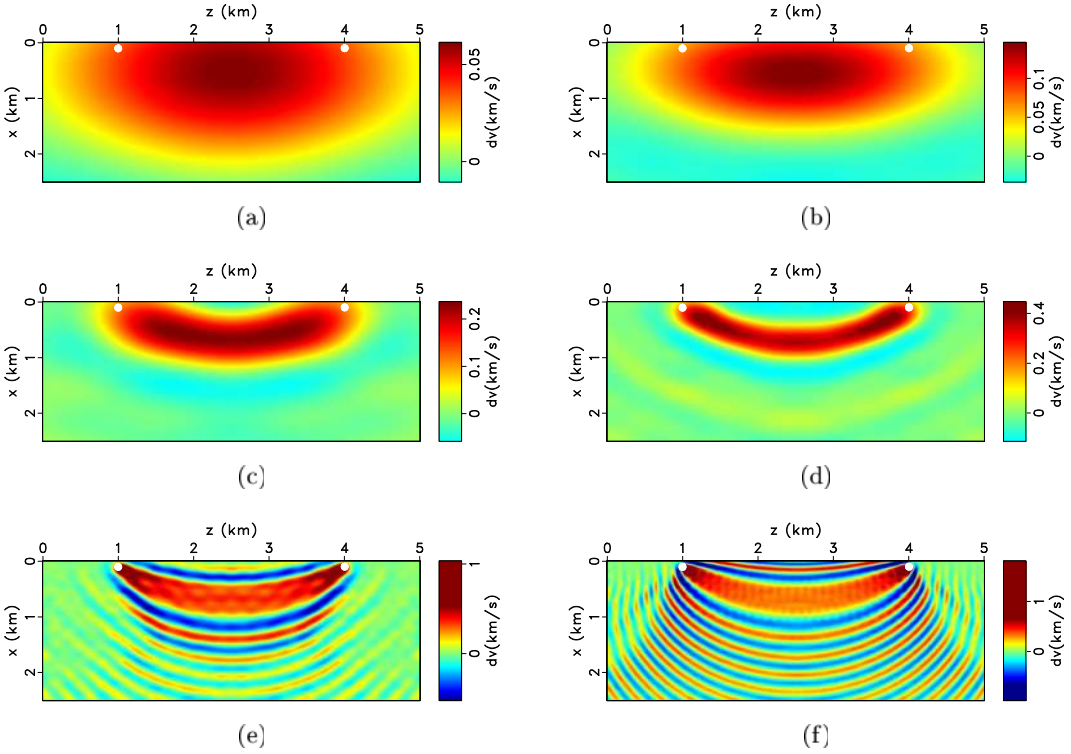

Figure 11. The model update (sensitivity kernel) after applying the low-cut scattering angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 160° and (f) 140°. The white dots correspond to the locations of the source and receiver.

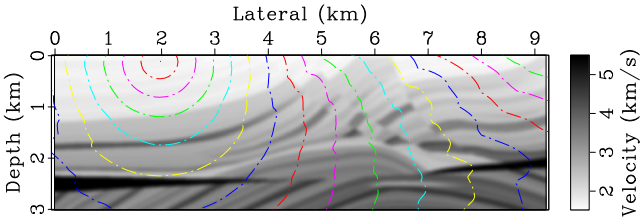

Figure 12. The Marmousi model filtered with a 20 sample box window in both directions to its original sampling at 4 m. The contours correspond to the traveltime solution of the eikonal equation for a source located near the surface, given by the white dot.

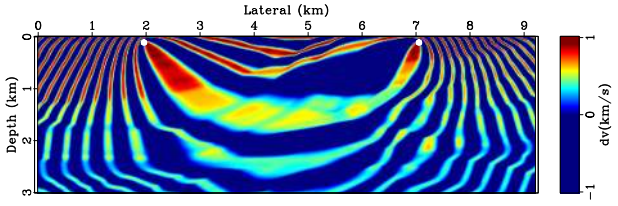

Figure 13. The model update response for a monochromatic wavefield at 7 Hz and a source located at 2 km and a receiver at 7 km, both at 0.1 km depth. The scale bar reflects the relative update that needs to be scaled by the residual and a normalized Hessian. The white dots correspond to the locations of the source and receiver.

between updates conditioned by the scattering angle and those obtained from a simple filter is that the updates here have a physical meaning tied with one component of our diffraction tomography model wavenumber formula.

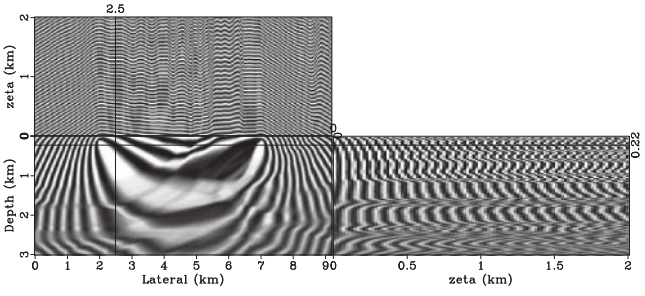

Figure 14. The model update (sensitivity kernel) of Fig. 2 with a ζ extension to allow for scattering angle identification.

8 DISCUSSION

The new filtering method described here takes the ray information and smears it over the model domain. This is what low frequency tends to do over the Fresnel zone, but no low frequency was required here. The physical meaning of these filters is highlighted by isolating the energy that provides the spared scattering angles, and for a very narrow band around the transmission angle of 180°, this corresponds to reducing the wavefield from the source and receiver to plane waves. As the narrow band widens slightly, the plane waves start to have a width controlled by the distance of the model point from the source and receiver. In other words, some point source characteristics come into play. We actually have a phenomenon similar to what we encounter with beams for even a monochromatic wavefield. Low frequencies present in the data are still important as the dispersion phenomena guaranties that the wavefield behaviour at low frequencies is different and cannot be predicted from high frequencies, and contains its unique information. The objective of a scattering angle filter strategy is to avoid falling into local minima at the prescribed frequency.

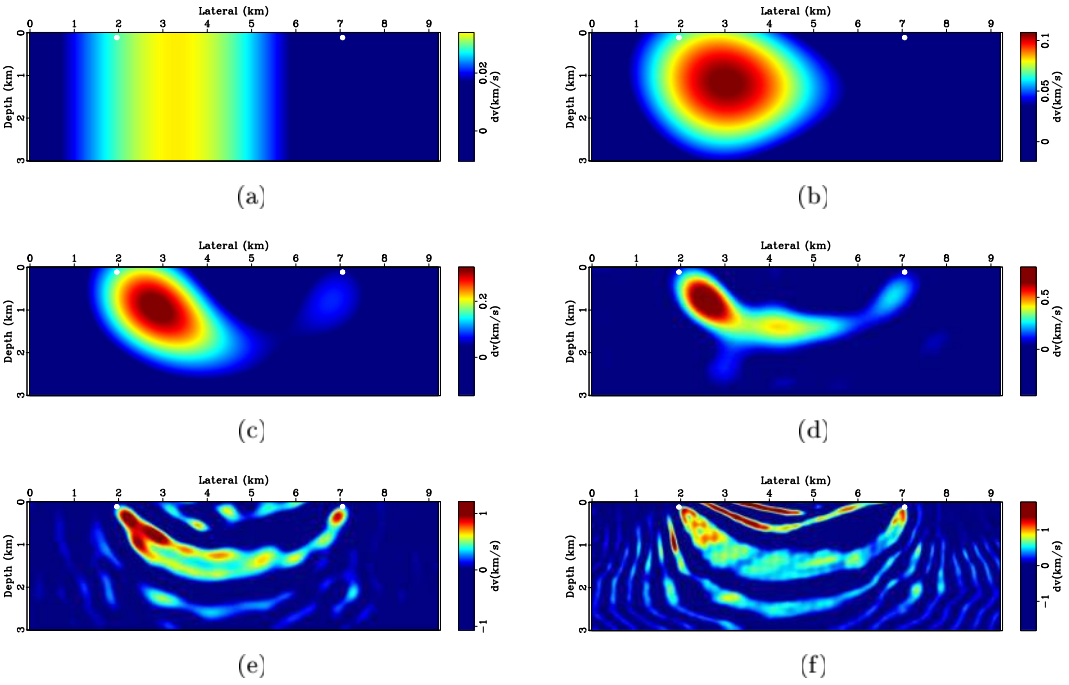

Figure 15. The model update (sensitivity kernel) after applying the low-cut scattering angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 160° and (f) 140°. The white dots correspond to the locations of the source and receiver.

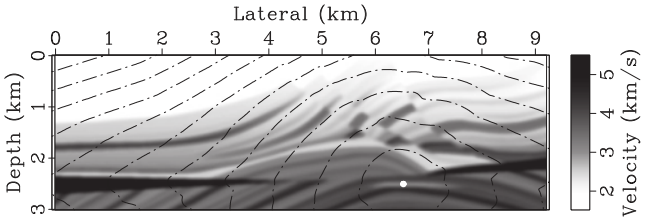

Figure 16. The contours now correspond to the traveltime solution of the eikonal equation for a source located at the reservoir area, given by the white dot.

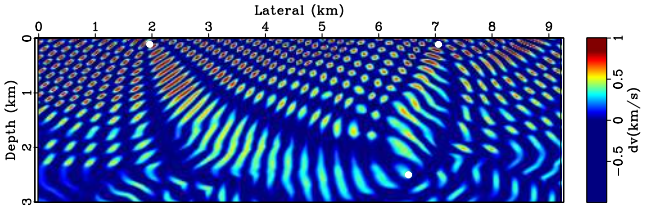

Figure 17. The model update response for a monochromatic wavefield at 10 Hz and a source located at 1 km and a receiver at 4 km, both at 0.1 km depth, and a model point at 2.5 km laterally and 2.5 km depth. The white dots correspond to the locations of the source, receiver and model point perturbation.

There are a lot of weighting aspects that were ignored in the development of the update. Specifically, those extracted from the full Hessian. The diagonal part of the Hessian provides the appropriate geometrical spreading correction to the sensitivity kernel. Nevertheless, the contribution of the Hessian or its approximation will not alter the performance of the filtering. The filtering will still highlight the energy corresponding to the angles that we pass resulting in more localized energy along the ray path regions. What would differ is the distribution of energy within the filtered gradient. If an approximate Hessian is developed from the filtered Gradient using for example Quasi–Newton methods, the action of the Hessian

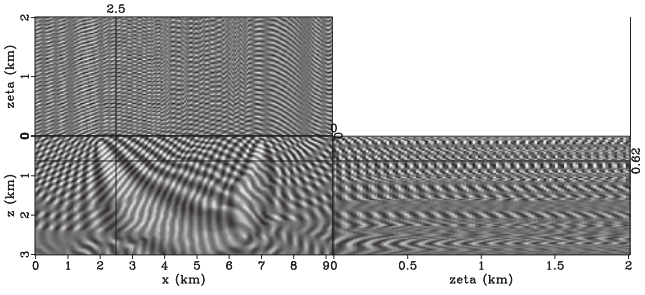

Figure 18. The model update (sensitivity kernel) of Fig. 17 with a ζ extension to allow for scattering angle identification.

will be consistent with the gradient (i.e. the information within the Hessian will correspond to the spared scattering angles).

Despite the many features that filtering the gradient brings to the table in the isotropic case, the benefits for anisotropic media are even greater. It allows us to apply different scattering-angle filtering on 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. Scattering angles for multiparameter inversion provides the key information to mitigate trade-off (Alkhalifah & Plessix 2014), and thus, its control will help make the inversion better posed.

For global Earth tomography, we are capable of producing the banana shaped kernels from monochromatic wavefields. Though the effective span of a monochromatic wavefield is controlled by the frequency, constrained by the first Fresnel zone, the filters here span beyond that limit to take on a regularization role as well, in FWI. Figs 11(a)–(f) are examples of the kind of sensitivity kernels that we can extract from the scattering angle filter. Since such filters are applied to the gradient constructed for all sources and receivers, the cost of this filtering depends mainly on the extension axis and

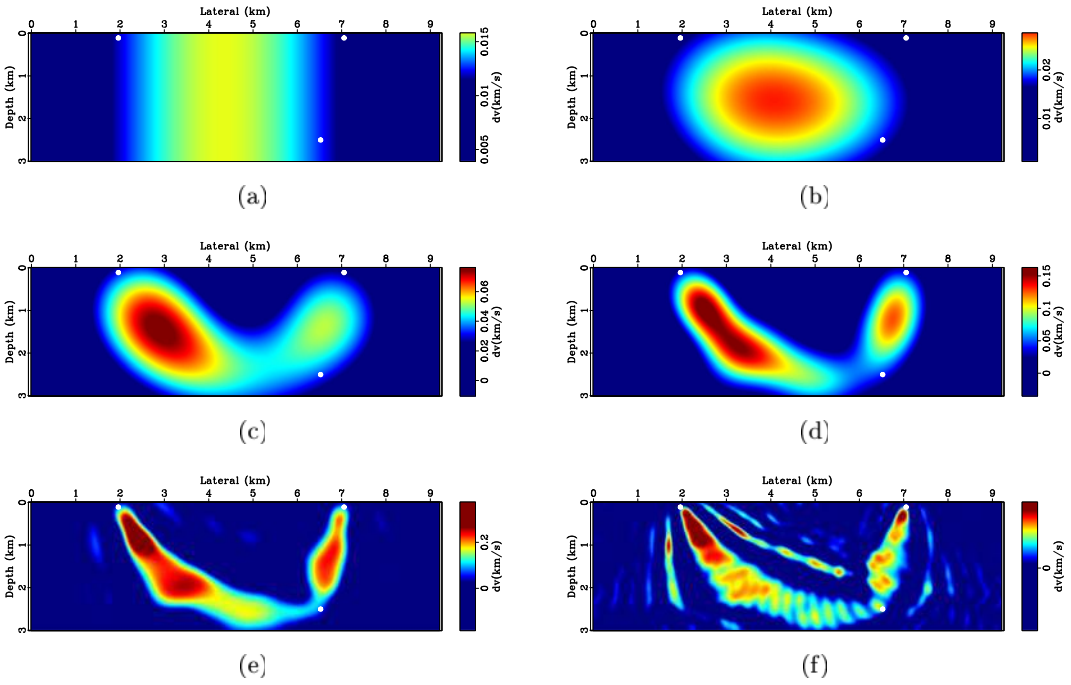

Figure 19. The model update (sensitivity kernel) for the reflection-based inversion after applying the low-cut scattering angle filter muting angles below (a) 179.4°, (b) 179°, (c) 178°, (d) 176°, (e) 160° and (f) 140°. The white dots correspond to the locations of the source, receiver and model point perturbation.

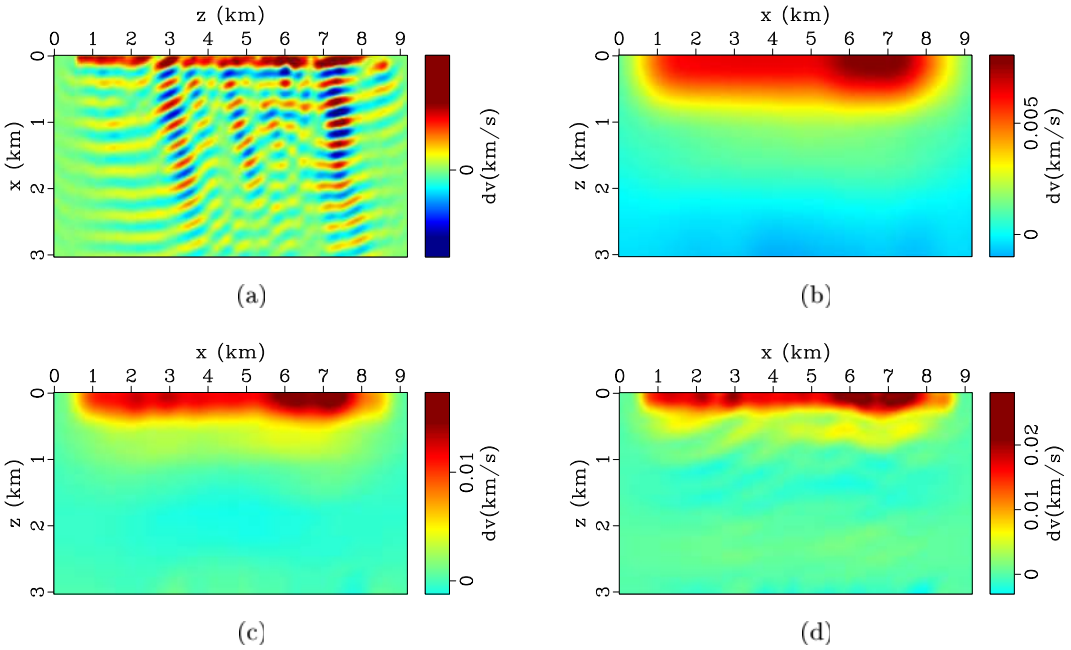

Figure 20. (a) The FWI gradient corresponding to a background homogenous model with a velocity of 3 km s^{-1} at 5 Hz frequency. The gradient after scattering angle filtering with a low cut under (b) 178°, (c) 176° and (d) 174°.

it's sampling. The sampling of the extension axis is governed by the maximum k_ζ we plan to keep, which may be small for high scattering angles. Of course, the range of the ζ extension controls the sampling of k_ζ. In all cases, we need only limited number of ζ samples, far less than what is commonly used to sample the time or frequency axis needed to construct the wavefields. As a result, the cost of the scattering angle filter pales in comparison to the cost of building the gradient. Zuberi & Alkhalifah (2014) even formulate a mapping approach to the angle gather formulation that allows us to avoid the extended domain.

The scattering angle filter has additional features. By isolating scattering angle ranges we are not only, in some cases, smoothing the gradient, but also, muting out non-physical waves, that may induce unphysical scattering angles, like those with $\cos\frac{\theta}{2} > 1$. This can happen for recorded (or converted to) shear wave events under the acoustic model assumption. It may also be the case with multisource modelling and the unphysical scattering angles associated with the crosstalks.

Developing an efficient space domain equivalent to the proposed scattering angle filter is possible, but might not be efficient.

The equivalence to the abrupt low scattering angle cut off in the wavenumber domain is a convolution with an isotropic non-stationary Bessel function in the space domain in 2-D. The cost of applying such a filter will depend on the filter size and the tapering methodology. The wavenumber domain provides us with the opportunity to apply the filter in an efficient matter.

9 CONCLUSION

A scattering-angle based filtering of the full waveform inversion (or the reflected version) gradient is capable of isolating the background update information regardless of the frequency. For monochromatic wavefields, I show that cutting out energy corresponding to scattering angles below 170° yields velocity updates along the general ray path of the wavefields involved. Stronger cutting thresholds (∼179°), yield smooth updates covering most of the model, all from a single frequency of 10 Hz.

ACKNOWLEDGEMENTS

I thank KAUST for it's support, and specifically the seismic wave analysis group (SWAG) within KAUST for their support and specifically Yunseok Choi for the many fruitful discussions on the subject. I thank the associate editor, Stephane Operto and anonymous reviewers for their critical review of the paper and their helpful suggestions.

REFERENCES

Albertin, U., Shan, G. & Washbourne, J., 2013. Gradient orthogonalization in adjoint scattering-series inversion, in *Proceedings of the 76th EAGE Int. Conf. SEG Technical Program,* Expanded Abstracts 205, pp. 1058–1062.

Alkhalifah, T., 2014. Scattering angle base filtering of the fwi gradients, in *Proceedings of the 76th EAGE Int. Conf.,* Extended Abstract, EAGE, Amsterdam.

Alkhalifah, T. & Plessix, R., 2014. A recipe for practical full-waveform inversion in anisotropic media: an analytical parameter resolution study, *Geophysics,* **79,** R91–R101.

Almomin, A. & Biondi, B., 2012. Tomographic full waveform inversion: practical and computationally feasible approach, in *Proceedings of the SEG Technical Program,* Expanded Abstracts 500, pp. 1–5.

Almomin, A. & Biondi, B., 2013. Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations, in *Proceedings of the SEG Technical Program,* Expanded Abstracts 203, pp. 1048–1052.

Beydoun, W.B. & Mendes, M., 1989. Elastic ray-born l2-migration/inversion, *Geophys. J. Int.,* **97,** 151–160.

Bunks, C., Saleck, F., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics,* **60,** 1457–1473.

Chavent, G. & Plessix, R., 1999. An optimal true-amplitude least-squares prestack depth-migration operator, *Geophysics,* **64,** 508–515.

Choi, Y. & Alkhalifah, T., 2012. Application of multi-source waveform inversion to marine streamer data using the global correlation norm, *Geophys. Prospect.,* **60,** 748–758.

Clément, F., Chavent, G. & Gómez, S., 2001. Migration-based traveltime waveform inversion of 2-D simple structures: a synthetic example, *Geophysics,* **66,** 845–860.

Cohen, J. & Bleistein, N., 1977. Seismic waveform modelling in a 3-D Earth using the born approximation: potential shortcomings and a remedy, *J. Appl. Math.,* **32,** 784–799.

Fleury, C. & Perrone, F., 2012. Bi-objective optimization for the inversion of seismic reflection data: combined FWI and MVA, in *Proceedings of the SEG Technical Program,* Expanded Abstracts 548, pp. 1–6.

Jang, U., Min, D. & Shin, C., 2009. Comparison of scaling methods for waveform inversion, *Geophys. Prospect.,* **57,** 49–59.

Jin, S., Madariaga, R., Virieux, J. & Lambare, G., 1992. Two-dimensional asymptotic iterative elastic inversion, *Geophys. J. Int,* **108,** 575–588.

Khalil, A., Sun, J., Zhang, Y. & Poole, G., 2013. RTM noise attenuation and image enhancement using time-shift gathers, in *Proceedings of the SEG Technical Program,* Expanded Abstracts 733, pp. 3789–3793.

Ma, Y., Hale, D., Gong, B. & Meng, Z., 2012. Image-guided sparse-model full waveform inversion, *Geophysics,* **77,** R189–R198.

Miller, D., Oristaglio, M. & Beylkin, G., 1987. A new slant on seismic imaging: migration and integral geometry, *Geophysics,* **52,** 943–964.

Mora, P., 1989. Inversion = migration + tomography, *Geophysics,* **54,** 1575–1586.

Moser, T., 2012. Review of ray-born forward modeling for migration and diffraction analysis, *Stud. Geophys. Geod.,* **56,** 411–432.

Panning, M., Capdeville, Y. & Romanowicz, A., 2009. An inverse method for determining small variations in propagation speed, *Geophys. J. Int.,* **177,** 161–178.

Pratt, R.G., Song, Z.-M., Williamson, P. & Warner, M., 1996. Two-dimensional velocity models from wide-angle seismic data by wavefield inversion, *Geophys. J. Int.,* **124,** 323–340.

Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A. & Dell'Aversana, P., 2004. Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt, *Geophys. J. Int.,* **159,** 1032–1056.

Sava, P.C. & Fomel, S., 2006. Time-shift imaging condition in seismic migration, *Geophysics,* **71,** S209–S217.

Sirgue, L. & Pratt, R., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics,* **69,** 231–248.

Tang, Y., Lee, S., Baumstein, A. & Hinkley, D., 2013. Tomographically enhanced full wavefield inversion, in *Proceedings of the SEG Technical Program,* Expanded Abstracts 201, pp. 1037–1041.

Tarantola, A., 1987. *Inverse Problem Theory,* Elsevier.

Thierry, P., Operto, S. & Lambare, G., 1999. Fast 2-d ray+born migration/inversion in complex media, *Geophysics,* **64,** 162–181.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics,* **74,** WCC1–WCC26.

Wang, S., Chen, F., Zhang, H. & Shen, Y., 2013. Reflection-based full waveform inversion (RFWI) in the frequency domain, in *Proceedings of the SEG Technical Program,* Expanded Abstracts 171, pp. 877–881.

Woodward, M. J., 1992. Wave-equation tomography, *Geophysics,* **57,** 15–26.

Wu, R. & Toksz, M., 1987. Diffraction tomography and multisource holography applied to seismic imaging, *Geophysics,* **52,** 11–25.

Xu, S., Wang, D., Chen, F., Lambare, G. & Zhang, Y., 2012. Inversion on reflected seismic wave, in *Proceedings of the SEG Technical Program,* Expanded Abstracts 509, pp. 1–7.

Zuberi, M. & Alkhalifah, T., 2014. Frequency scaling for angle gathers, in *Proceedings of the 76th EAGE Conference and Exhibition,* 76, doi:10.3997/2214–4609.20141495.