Full waveform inversion by model extension using a model-space multi-scale approach

Guillaume Barnier and Ettore Biondi

ABSTRACT

We successfully show that full waveform inversion by model extension (FWIME) can mitigate the cycle-skipping issue inherent to conventional full waveform inversion (FWI). We develop and apply a model-space multi-scale workflow using spline interpolation on a 2D synthetic dataset containing only reflected data with no energy below 4 Hz and where conventional FWI fails to recover the optimal solution when starting from an inaccurate initial velocity model. We simultaneously invert all available frequency content in the data without adopting a data-space multi-scale approach. By gradually refining the node spacing of our model representation on the spline grid, we control the wavenumber content of the updates during the inversion process. We first apply FWIME to recover an accurate velocity model, and we use it as an initial guess for a conventional FWI workflow. The final inverted model has not cycle-skipped and is very close to the optimal solution.

INTRODUCTION

FWI has the potential of inverting all model scales and providing high resolution subsurface images but it is greatly hampered by its sensitivity to the accuracy of the initial model, commonly known as cycle-skipping (Virieux and Operto, 2009). Biondi and Almomin (2014) use an extended model approach to mitigate this issue, but the method, based on a nested-scheme algorithm, heavily relies on the user to tune the many inversion parameters employed in the optimization process.

Following their work, we recently proposed a modified approach, also relying on an extended Born modeling operator, and showed its potential by successfully inverting cycle-skipped transmission data (Barnier et al., 2018). The main differences in our method is that we do not separate the model parameter into a background (low-wavenumber) and a strictly high-wavenumber perturbation. Moreover, during our inversion process, we aim at completely removing all the energy present in the extended model perturbation and we attach no physical meaning to it. Finally, the use of the variable projection method to compute the adequate extended model perturbation allows us to better control the phase alignment of the data residuals (Golub and Pereyra, 1973; Rickett, 2013; Huang and Symes, 2015).

As discussed in Barnier et al. (2019), we think the success of waveform inversion at converging towards a useful solution (in an efficient manner) requires an accurate updating of the low-wavenumber components at early stages of the workflow. Biondi and Almomin (2012, 2013) successfully used scale-separation and wavelength-continuation to improve convergence, but their approach remains quite user intensive.

In this report, we aim at making this process more user independent by incorporating the model-space multi-scale method developed in Barnier et al. (2019) to our FWIME workflow. The entire data bandwidth is simultaneously inverted and the wavenumber content of the updates is controlled by the refinement rate of the spline grid on which the model is parametrized. The benefit of this technique is that all events in the data (including reflections) are used to update the low-wavenumber updates at early stages.

We begin by reviewing the FWIME theorical framework and explain how the spline parametrization is incorporated into our workflow. We then illustrate the advantages of this feature for FWIME on a simple numerical example. Next, we successfully apply our method on a subset of the Marmousi model (Martin et al., 2006), where acoustic pressure (reflection) data with no energy outside of the 4-15 Hz frequency bandwidth are generated, and a v(z) model is used as initial guess. We show that conventional FWI fails to converge to the optimal solution while FWIME manages to recover an accurate model. We then use the FWIME inverted model as an initial guess for conventional FWI, resutling in a high-resolution, accurate solution. Finally, we discuss the current challenges and future opportunities of FWIME, and discuss our strategy moving forward on how to apply it to 3D field data.

FWIME THEORY

In our new FWIME formulation, we propose to minimize the following objective function defined as

$$\Phi_\epsilon(\mathbf{m}) = \frac{1}{2}\left\|\mathbf{T}_d\left(\mathbf{f}(\mathbf{Sm}) + \tilde{\mathbf{B}}(\mathbf{Sm})\tilde{\mathbf{p}}^{\epsilon}_{opt}(\mathbf{Sm}) - \mathbf{d}^{obs}\right)\right\|_2^2 + \frac{\epsilon^2}{2}\left\|\mathbf{D}\tilde{\mathbf{p}}^{\epsilon}_{opt}(\mathbf{Sm})\right\|_2^2, \quad (1)$$

where $\mathbf{m}$ is the velocity model defined on a (coarse) spline grid, $\mathbf{S}$ is the spline interpolation operator defined in Barnier et al. (2019), $\mathbf{f}$ is the wave-equation operator (acoustic, isotropic), and $\tilde{\mathbf{B}}$ denotes the extended Born modeling operator. Possible extensions include time-lags, subsurface offsets, or shot records (Biondi and Almomin, 2014; Huang and Symes, 2015). In this report all extended operations are performed with the time-lag extension and are denoted by the $\sim$ symbol. $\mathbf{d}^{obs}$ represents the observed data, and $\mathbf{T}_d$ is a self-adjoint data tapering operator (typically employed to mute all refracted energy) whose coefficients are fixed during the entire inversion workflow. The (linear) operator $\mathbf{D}$ is an invertible version of the differential semblance optimization (DSO) operator that enhances the non-physical extended energy of extended images (Symes and Kern, 1994). ϵ is the trade-off parameter between the

data-fitting and the regularization terms. The subscript ϵ in Φ_ϵ indicates that ϵ is a fixed parameter throughout the inversion. $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ is an extended perturbation, defined as the minimizer of the following quadratic objective function: $\Phi_{\epsilon,\mathbf{m}}$,

$$\Phi_{\epsilon,\mathbf{m}}(\tilde{\mathbf{p}}) = \frac{1}{2}\left\|\mathbf{T}_d\left(\tilde{\mathbf{B}}(\mathbf{Sm})\tilde{\mathbf{p}} - \left(\mathbf{d}^{obs} - \mathbf{f}(\mathbf{Sm})\right)\right)\right\|_2^2 + \frac{\epsilon^2}{2}\left\|\mathbf{D}\tilde{\mathbf{p}}\right\|_2^2. \tag{2}$$

Assuming the Hessian matrix of $\Phi_{\epsilon,\mathbf{m}}$ (equation 2) is positive definite, $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ is given by

$$\tilde{\mathbf{p}}^{\epsilon}_{opt}(\mathbf{Sm}) = \left[\tilde{\mathbf{B}}^*(\mathbf{Sm})\ \mathbf{T}_d^2\ \tilde{\mathbf{B}}(\mathbf{Sm}) + \epsilon^2\mathbf{D}^*\mathbf{D}\right]^{-1}\tilde{\mathbf{B}}^*(\mathbf{Sm})\mathbf{T}_d^2\left(\mathbf{d}^{obs} - \mathbf{f}(\mathbf{Sm})\right), \tag{3}$$

where $*$ denotes adjoint operators. The minimization of equation 2 is performed using a linear conjugate-gradient algorithm, and is referred to as the variable projection step in FWIME. The data residual component (first term) on the right side of equation 2 is a modified FWI objective function where an additional term $\tilde{\mathbf{B}}(\mathbf{Sm})\tilde{\mathbf{p}}^{\epsilon}_{opt}(\mathbf{Sm})$ is used to ensure the phase alignment between modeled and observed data. During the optimization process we gradually reduce the contribution of this additional term by adding a regularization term on the right side of equation 1 (which gradually forces the L_2-norm of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ to vanish). It is important to notice that unlike the method proposed in Biondi and Almomin (2014), $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ may contain all wavenumber components at any stage of the inversion process, and no explicit scale mixing is applied at any step. Hence, $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ is not defined on the spline grid, but rather on the (finer) finite-difference grid.

Equation 1 is minimized using a gradient-based descent method and its gradient is given by

$$\nabla\Phi_\epsilon(\mathbf{m}) = \mathbf{S}^*\mathbf{M}\Big[\mathbf{B}^*(\mathbf{Sm}) + \mathbf{T}^*(\mathbf{Sm})\Big]\mathbf{T}_d^2\left(\mathbf{f}(\mathbf{Sm}) + \tilde{\mathbf{B}}(\mathbf{Sm})\tilde{\mathbf{p}}^{\epsilon}_{opt}(\mathbf{Sm}) - \mathbf{d}^{obs}\right), \tag{4}$$

where $\mathbf{B}^*$ is the adjoint of the non-extended Born modeling operator and $\mathbf{T}^*$ is the adjoint of the tomographic operator (Sava and Biondi, 2004; Biondi and Almomin, 2014). The first component of the FWIME gradient in equation 4 (the output of the application of $\mathbf{B}^*$ to the FWIME data residual) will be referred to as the "Born gradient" of FWIME. It is computed with the same operator as the one used for conventional FWI gradient but the input (i.e., the data residual) is different. In addition, the second component (stemming from the application of $\mathbf{T}^*$) updates regions of the model wavenumber spectrum missed by the first one. We refer to it as the "tomographic gradient" of FWIME. $\mathbf{M}$ is a masking operator that may be used to restrict the gradient from updating certain regions of the model (e.g., the water layer).

The effect of the spline parametrization on the gradient can be understood from equation 4. $\mathbf{S}^*$ (the adjoint of the spline interpolation operator) projects the conventional FWIME gradient onto a (usually coarser) spline grid, thereby enforcing

smoothness on the model updates (Barnier et al., 2019). Indeed, the level of smoothness can be adjusted by refining the spline grid over iterations in order to gradually include higher wavenumber updates into our inverted model. In the next two sections, we illustrate these properties with numerical examples that provide more insight on their benefit and impact on the FWIME workflow.

INCORPORATING THE MODEL-SPACE MULTI-SCALE APPROACH IN FWIME

In order to illustrate the benefits of the spline parametrization as a way to control the wavenumber content of the FWIME updates, we guide the reader through the first step of our workflow for a simple numerical example where solely reflection data are inverted. We show the potential of our model-space multi-scale approach to recover the low-wavenumber components of the velocity model, even in the absence of refracted events.

We design a 10 km-wide and 2.5 km-deep velocity model with a shallow water layer and a v(z) velocity gradient (Figures 1 and 2). We embed a strong horizontal reflector at a depth of 1.7 km. We place 100 sources every 100 m, and 500 receivers every 20 m. All acquisition devices are are placed at a depth of 20 m. Our initial velocity model is also laterally invariant, and its depth profile is shown in Figure 2b (red curve). We generate noise-free pressure data with a two-way acoustic modeling operator with a source containing energy only in the 9-20 Hz frequency range. For this numerical example, we target our inversion on reflected energy only. To that effect, we apply a data taper on all shot records to mute all events occurring at offsets greater than 2.5 km. Figures 3a and b show the recorded data and the tapered data difference (computed from the initial model) for a shot located at $x = 5$ km.

For the FWIME spline parametrization, we place 24 spline nodes in the vertical direction with a spacing of 500 m in the absorbing boundary, and 200 m spacing from 100 m above the water-bottom interface down to a depth of 3 km. In the horizontal direction, we place 20 nodes with a spacing of 500 m in the absorbing boundaries, and 1 km in the area of interest.

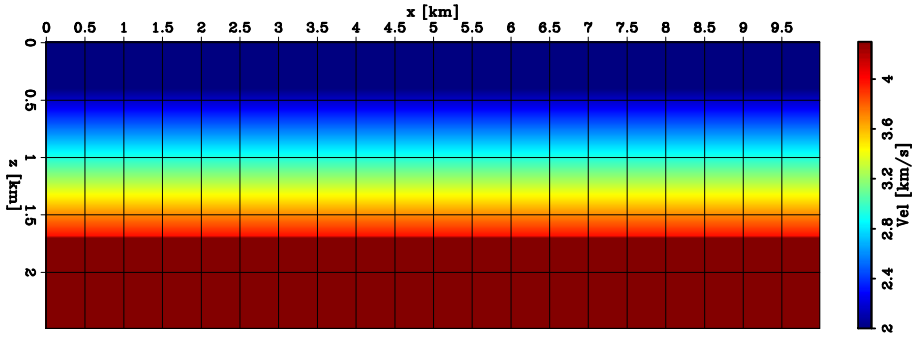


Figure 1: 2D panel of the true velocity model. [**ER**]

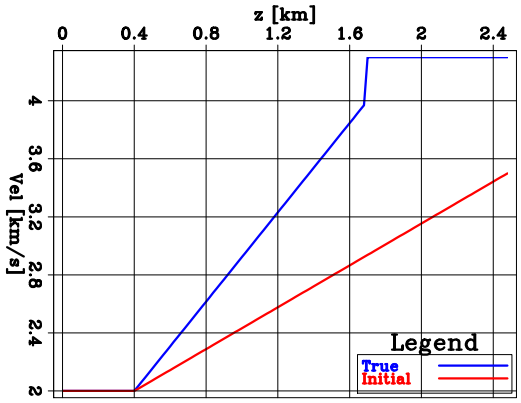
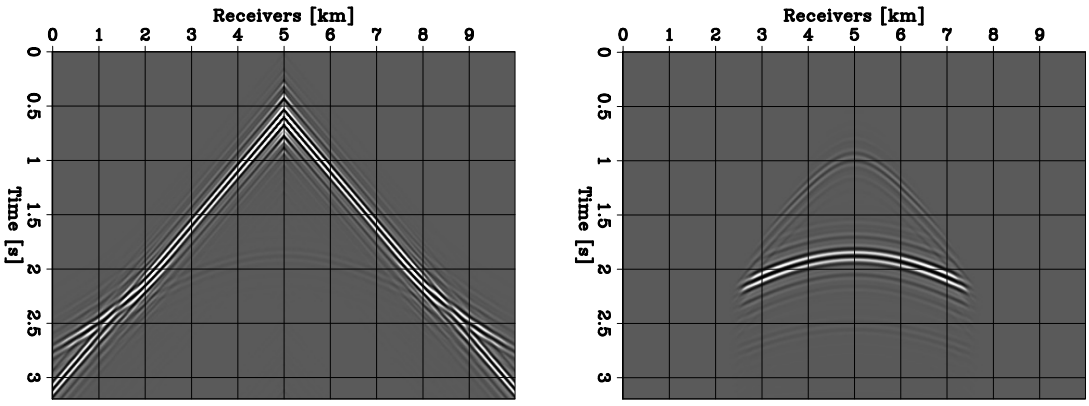


Figure 2: 1D profiles of the true model (blue curve), and the initial model (red curve). [**ER**]



(a) (b)

Figure 3: Shot gather for a source located at $x = 5$ km. (a) Observed data with no tapering applied. (b) Tapered data difference between observed and predicted data using the initial velocity model. We use a data taper to mute all energy present beyond 2.5 km offset. For clarity purposes, (a) and (b) are displayed on different scales. [**ER**]

We first conducted two conventional (i.e., without spline) data-space multi-scale FWI, (1) using all available data (including refracted energy with both diving and head waves shown in Figure 3a), and (2) using only reflected energy with the same data tapering as the one applied in Figure 3b. Though not shown in this report, both converged to very inaccurate local minima.

We perform the first variable projection step of our FWIME workflow (minimization of equation 2). We use a total 101 points for the extended axis at 8 ms sampling, allowing time-lags to range from -0.4 s to 0.4 s. Figure 4 shows a time-lag common image gather (CIG) at $x = 5$ km of the inverted optimal extended perturbation $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ after 100 iterations of linear conjugate gradient (Biondi et al., 2019), and with $\epsilon = 1.0 \times 10^{-6}$. The strong visible event in Figure 4 corresponds to the reflection from the sharp horizontal interface present in the true model (also visible in the data space in Figure 3b). As expected, the position of its maximum energy is shifted away from the zero time-lag axis and is located at a time-lag of approximately $\tau = -0.2$ s, which indicates that the velocity used to "migrate" this event was lower than the true velocity. This also illustrates how the optimal extended perturbation $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ provides crucial kinematic information used when computing the tomographic gradient of FWIME, which is key to mitigate the cycle-skipping issue. Moreover, for this numerical example, we tested various values of ϵ and observed that they had very limited impact on the results. However, this behavior is not generally true.

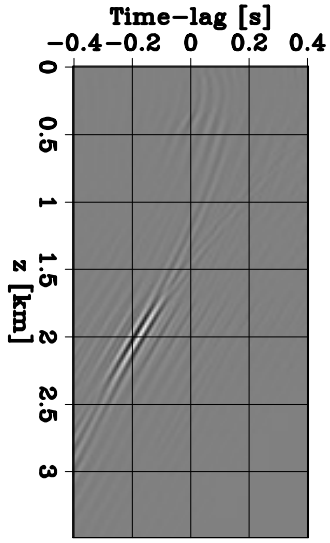


Figure 4: Time-lag CIG of the optimal extended perturbation $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ extracted at $x = 5$ km. $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ was computed by minimizing equation 2 with 100 iterations of linear conjugate gradient, and $\epsilon = 1.0 \times 10^{-6}$. [**ER**]

The FWIME search direction is then computed using equation 4. Figures 5a-c show the Born, tomographic, and the total (their sum) search directions on the finite-difference grid, before projection onto the (coarse) spline grid (before the application of $\mathbf{S}^*$). As expected, the Born search direction (kinematically equivalent to the

conventional FWI first search direction) contains a strong high-wavenumber update (the sharp reflector located at an approximate depth of 1.5 km in Figure 5a), but incorrectly positioned due to the large kinematic error present in the initial velocity model. On the other hand, the (lower-wavenumber) tomographic search direction in Figure 5b seems very promising but its amplitude is dominated by the Born search direction. However, the projection onto the coarse spline grid enables us to automatically recover its features. Figure 5d shows the FWIME total search direction on the spline grid, and Figure 5e shows the re-projection of the FWIME search direction onto the finite-difference grid (Figure 5e is computed by applying operator $\mathbf{S}$ to the panel in Figure 5d).

By comparing Figure 5c with Figure 5e, we can clearly see the potential of using such a spline parametrization. The projection onto the coarse grid allowed us to automatically recover and extract the coherent (and crucial) low-wavenumber components of the gradient (contained solely in the tomographic search direction shown in Figure 5b). This was done without having to explicitly and manually apply any filter nor scale mixing to the FWIME gradient. Indeed, for a more complex subsurface (as shown in the next section), this spline disposition may not be adequate, and may require some refinement at some stage of the inversion process. Therefore, as described in Barnier et al. (2019), we propose to adopt a model-space multi-scale approach. The inversion begins with a very coarse spline grid, and is automatically refined when convergence is reached (for that parametrization), thereby reducing the need for user input.

APPLICATION OF FWIME ON REFLECTION DATA FROM THE MARMOUSI MODEL

We test our FWIME workflow described in the previous sections to invert reflection data generated from a subset of the Marmousi model (Martin et al., 2006), as shown in Figure 8a. We model noise-free pressure data with a two-way acoustic isotropic wave-equation operator, and a finite-difference grid spacing of 20 m in both directions. For this numerical experiment, we use seismic sources containing energy limited to the 4-15 Hz band (Figure 10). We place 100 sources every 100 m, and 500 receivers every 20 m. Sources and receivers are located at a depth of 20 m.

For all our tests, we use a laterally invariant initial model shown in Figure 8b and Figure 9 (red curve), which is far from the true solution. We first apply conventional data-space multi-scale FWI using all available data (including refracted energy) and show that it fails to retrieve the optimal solution. We then apply FWIME to invert a subset of the data (only reflected events), and we are able to retrieve an accurate model. We then use the output of FWIME as a starting guess for conventional FWI. The final inverted model is close to the optimal solution.

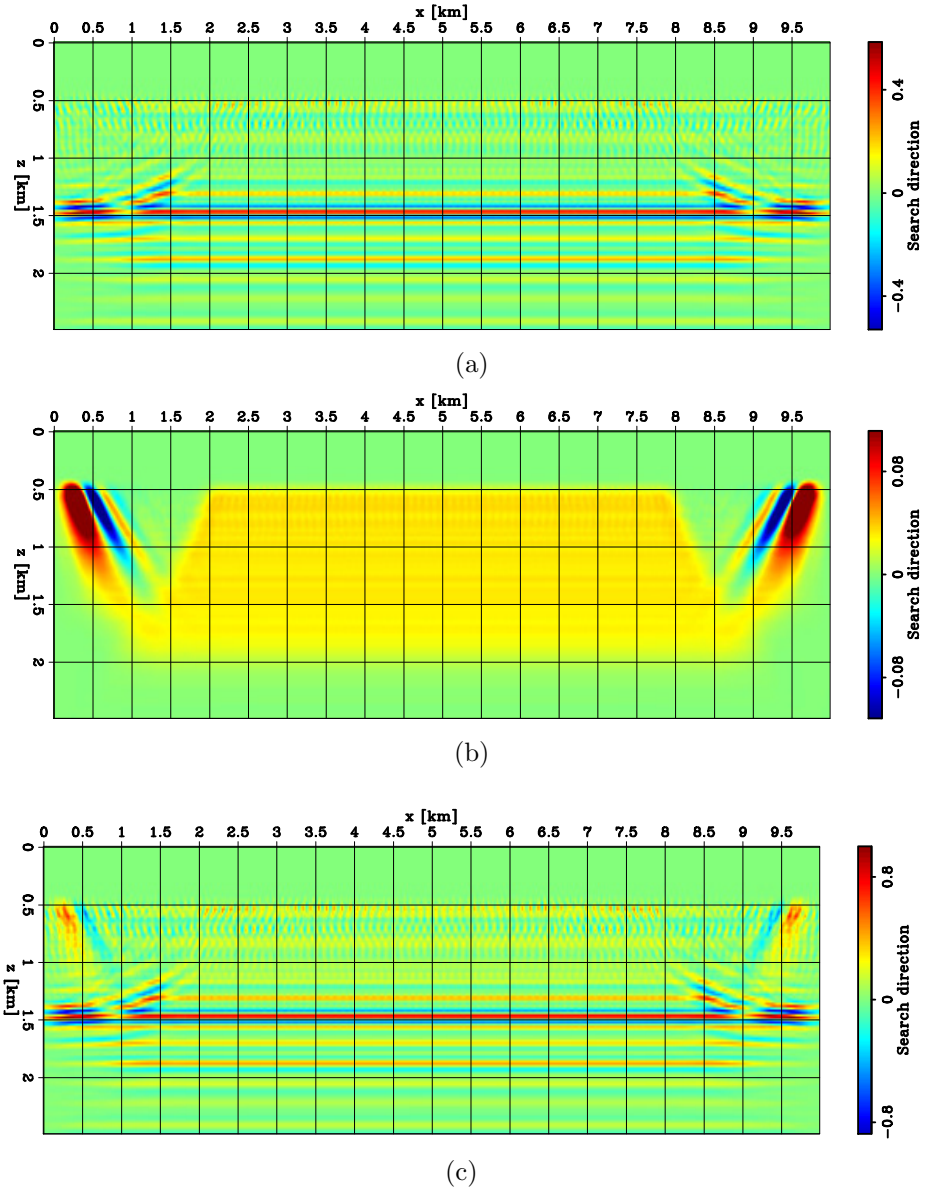


Figure 5: FWIME first search directions. (a) Normalized Born search direction. (b) Normalized tomographic search direction. (a) and (b) are normalized by the same value. (c) Normalized total FWIME search direction before projection on the spline grid (sum of (a) and (b)). [**ER**]

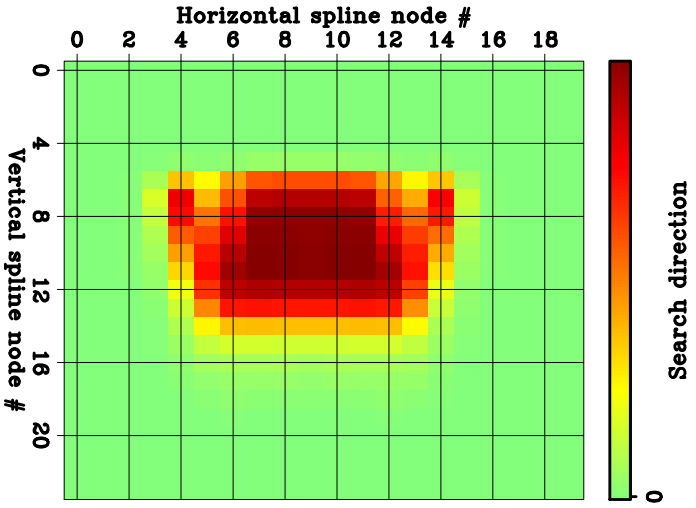


Figure 6: FWIME first search direction displayed on the spline grid. [**ER**]

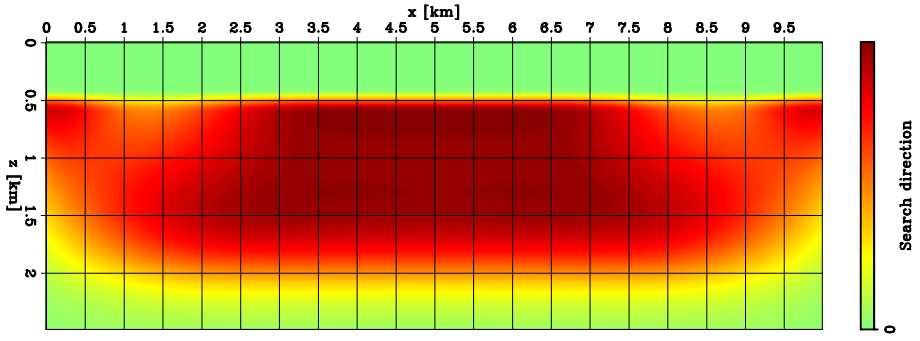


Figure 7: Projection of the first search direction (Figure 6) on the finite-difference propagation grid. [**ER**]

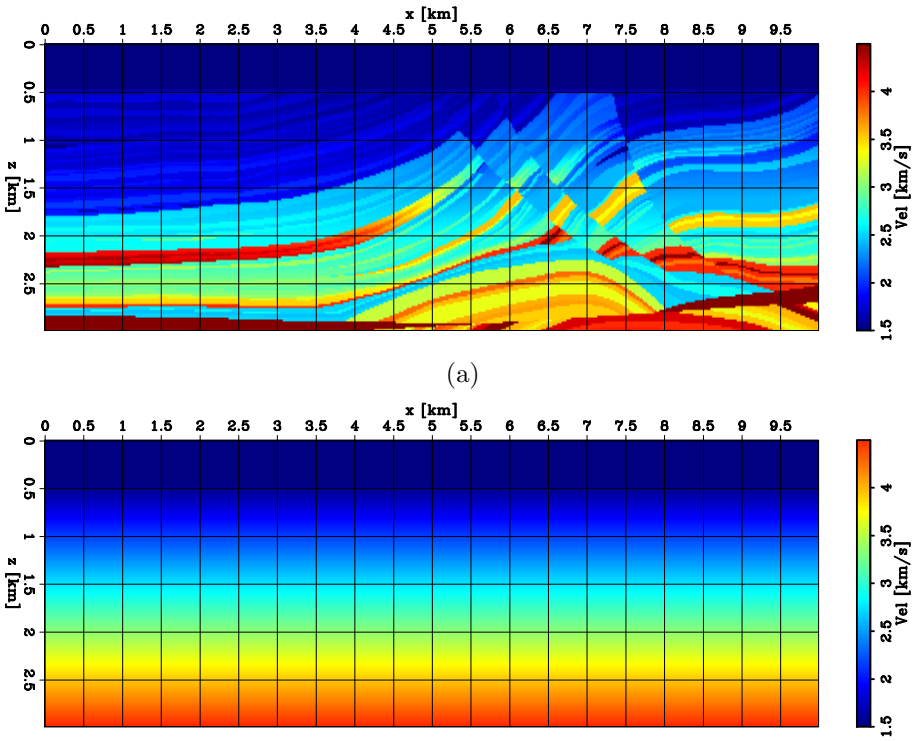


Figure 8: (a) True velocity model (subset of the Marmousi model). (b) Initial velocity model. [**ER**]

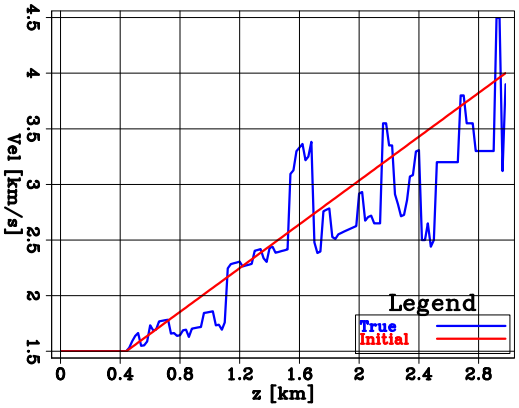


Figure 9: Depth velocity profiles extracted at $x = 5$ km of the true model (blue curve), and the initial model (red curve). [**ER**]

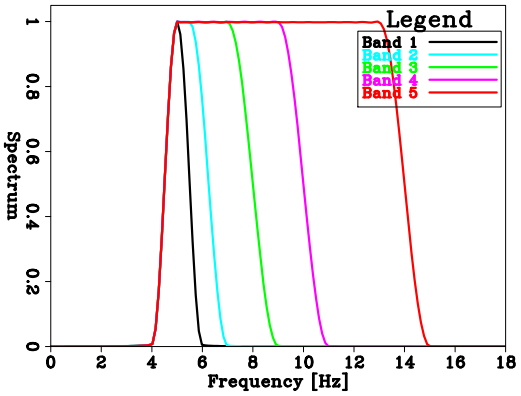
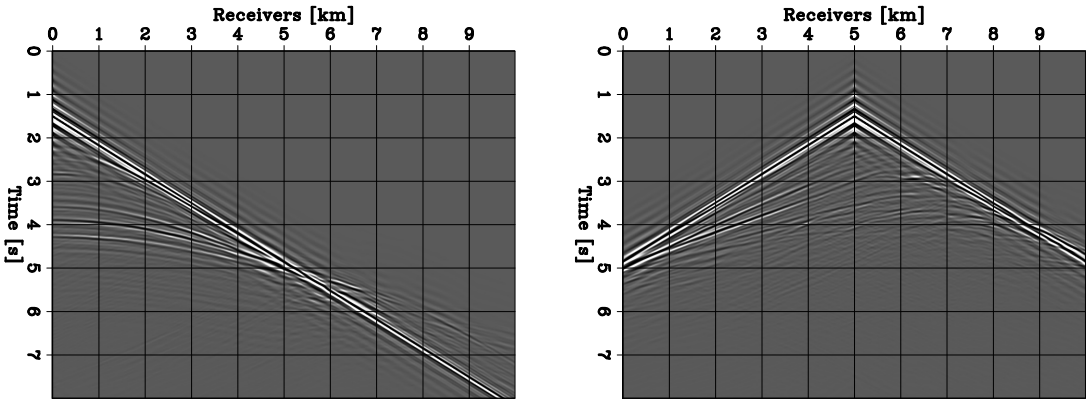


Figure 10: Spectra of the various seismic sources used in the Marmousi numerical experiment. [**ER**]

Conventional FWI

We apply a conventional data-space multi-scale FWI using five frequency bands (Figure 10) and the full dataset (including some refracted energy). Figures 8a and b show two shot gathers generated with sources located at $x = 0$ km and $x = 5$ km, respectively (using the full 4-15 Hz frequency band). For each bandwidth, we conduct 100 iterations of BFGS using the library implemented by Biondi et al. (2019). Figure 12 shows the inverted model after the last bandwidth. Conventional multi-scale FWI is not able to retrieve the correct solution, especially for depths greater than 1 km. Even with the presence of some refracted energy, the inaccuracy of the initial model as well as the lack of low-frequency energy in the data pushes the inversion to converge to a local minimum.



(a) (b)

Figure 11: Shot gathers generated with seismic sources containing energy between 4-15 Hz (red curve in Figure 10), positioned at $x = 0$ km and $x = 5$ km, respectively. [**ER**]

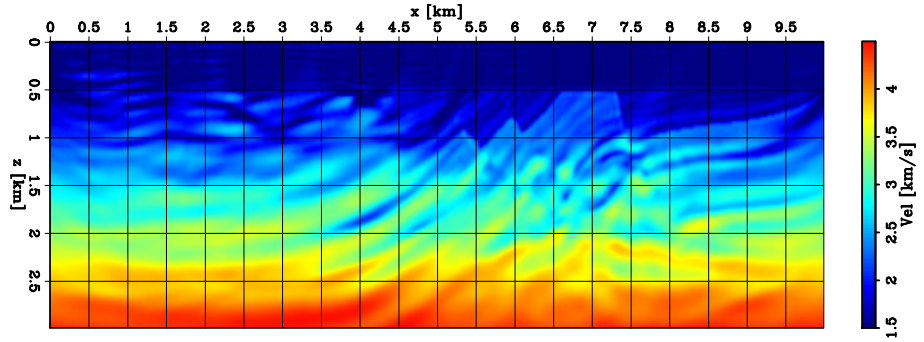


Figure 12: Final inverted velocity model using a conventional data-space multi-scale approach with five frequency bands shown in Figure 10. For each frequency band, the full dataset was inverted and no muting was applied. [**ER**]

FWIME

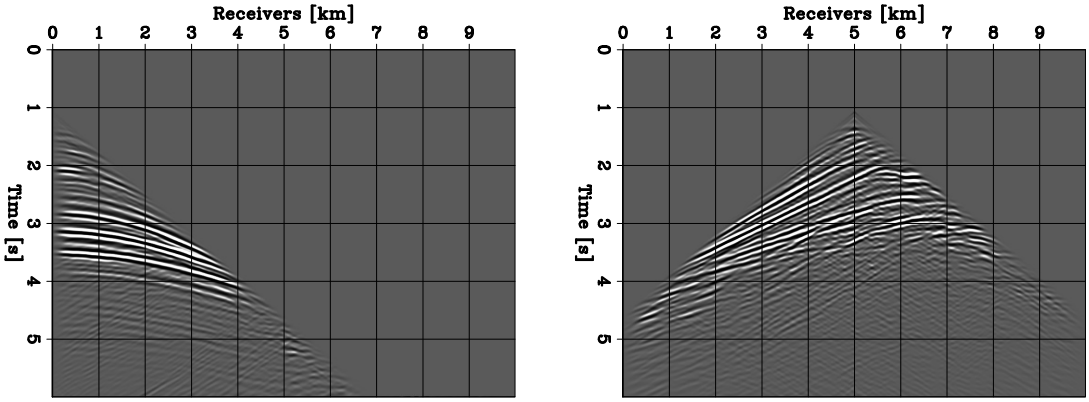
We apply the model-space multi-scale FWIME workflow described in the previous section. We use a subset of the data generated with the 4-15 Hz frequency band (red curve in Figure 10) by applying a taper to mute all events occurring at offsets greater than 6 km. After tapering, the data mostly consists of reflected events. Figures 13a and b show two representative shot gathers of the data used for the inversion. We assume no prior knowledge of the subsurface geometry, and we follow our prescribed workflow by choosing a very coarse spline representation (200 m node spacing in the vertical direction, 1 km spacing in the horizontal direction, and 500 m in the absorbing boundaries). For the extended perturbation $\tilde{\mathbf{p}}^{\epsilon}_{opt}$, we use 101 points on the extended axis, with time-lags ranging from $\tau = -0.4$ s to $\tau = 0.4$ s. Each variable projection step is performed with 100 iterations of linear conjugate gradient.

Initial search direction with spline parametrization

Figures 14a shows the first FWIME search direction $\mathbf{s}^f_0$ on the (fine) finite-difference propagation grid, before projection onto the spline grid. It is given by the following expression:

$$\mathbf{s}^f_0 = -\mathbf{M}_w \Big[\mathbf{B}^*(\mathbf{Sm_0}) + \mathbf{T}^*(\mathbf{Sm_0})\Big] \mathbf{T}_d^2 \left(\mathbf{f}(\mathbf{Sm_0}) + \tilde{\mathbf{B}}(\mathbf{Sm_0})\tilde{\mathbf{p}}^{\epsilon}_{opt}(\mathbf{Sm_0}) - \mathbf{d}^{obs}\right), \qquad (5)$$

where $\mathbf{m}_0$ is the initial model expressed on the spline grid, and $\mathbf{M}_w$ is a masking operator that mutes the updates in the water layer. We can clearly see the presence of high-wavenumber features in the vicinity of the water-bottom interface. As shown in Figure 14b, these features disappear when applying operator $\mathbf{SS}^*$ to $\mathbf{s}^f_0$ (i.e., projecting the search direction onto the spline grid, and back). Similarly, Figure 14c shows the



(a) (b)

Figure 13: data used for the FWIME inversion workflow, generated with seismic sources containing energy between 4-15 Hz (red curve in Figure 10), and located at (a) $x = 0$ km and (b) $x = 5$ km. A taper was applied to the modeled data to mute the majority of the refracted energy. [**ER**]

"ideal" search direction (i.e., the difference between the true and the initial models on the fine grid), and Figure 14d is obtained by applying operator $\mathbf{SS}^*$ to the panel in Figure 14c. The first FWIME search direction (Figure 14b) possesses promising features that seem to capture the long wavelength components of the ideal update, at least in the shallow part of the model, down to a depth of 2 km.

Inversion results

We conduct our FWIME workflow for the first spline grid parametrization. Figure 15 displays the normalized convergence curves for the data fitting term, the model regularization term, and the total FWIME objective function. For this first spline parametrization, the FWIME scheme converged after four iterations of nonlinear conjugate gradient. At this point, the spline grid needs to be refined to continue the inversion process.

To better understand the behavior of FWIME sequence applied with the first spline parametrization, we show the inverted model at the fourth iteration, projected on the finite-difference grid (Figure 16). We can observe some significant updates in the left part of the model (located between $x = 0$ km and $x = 5$ km), which seem to be guiding the inversion towards a promising solution. Figures 17a and b show two time-lag CIGs extracted (at $x = 3.0$ km) from the initial optimal perturbation $\tilde{\mathbf{p}}^{\epsilon}_{opt}(\mathbf{Sm_0})$, and from the one computed after four iterations, $\tilde{\mathbf{p}}^{\epsilon}_{opt}(\mathbf{Sm_4})$, respectively. We can clearly see that the strong event initially located at an approximate depth of $z = 1.8$ km, and at a time-lag value of $\tau = 0.2$ s (Figure 17a) has been shifted closer to the zero time-lag axis (Figure 17b). This indicates an improvement of the accuracy of the overburden velocity in that region of the model.

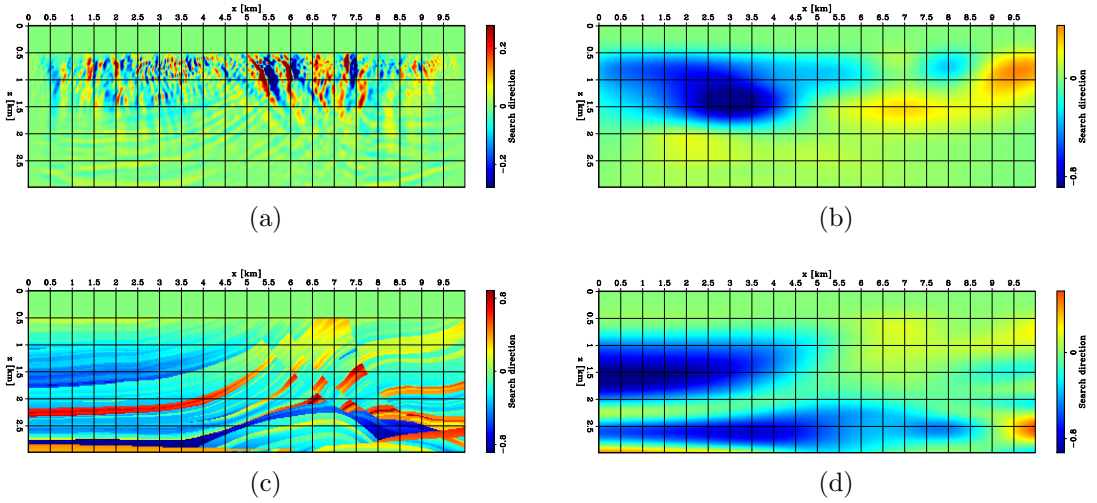


Figure 14: FWIME normalized search directions. (a) FWIME first search direction before spline interpolation. (b) FWIME first search direction after spline interpolation. (c) Ideal search direction on the fine grid. (d) Ideal search direction after spline interpolation. [**CR**]

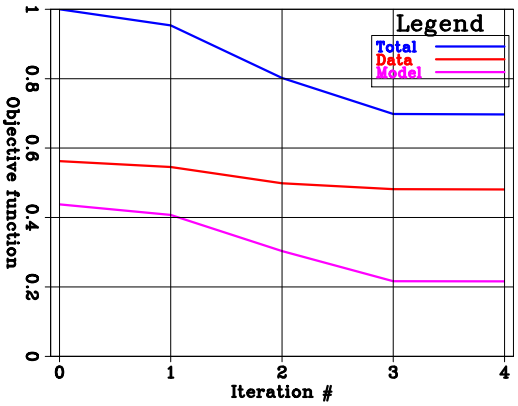


Figure 15: Normalized convergence curves for the FWIME scheme using the first spline parametrization. [**CR**]

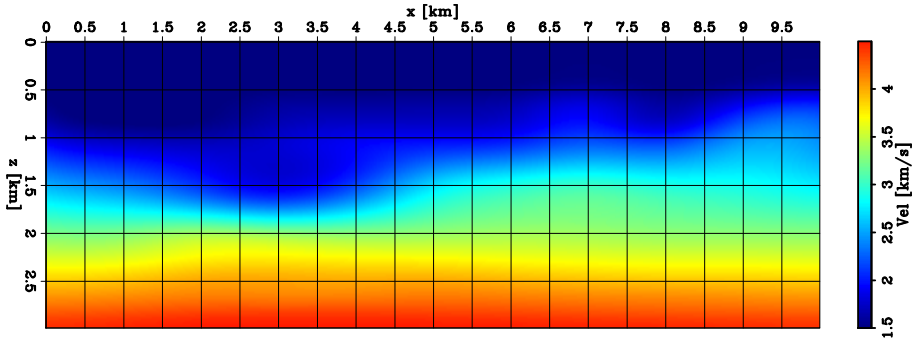
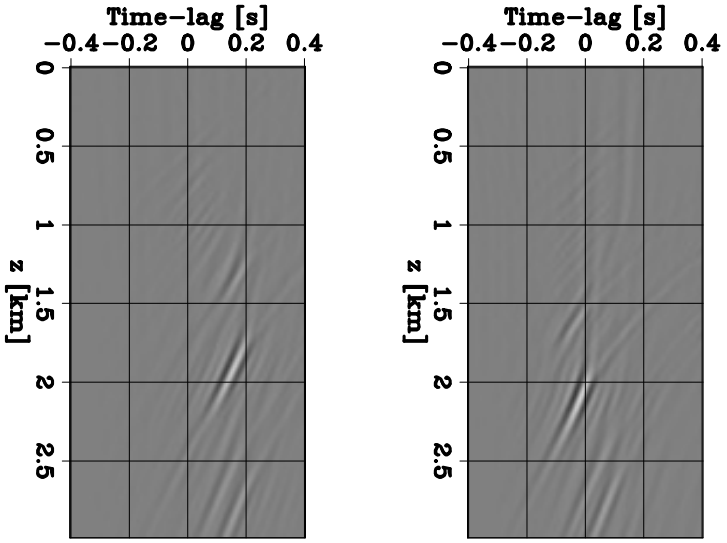


Figure 16: Inverted model after five iterations of the FWIME scheme using the first spline parametrization. [**CR**]



(a) (b)

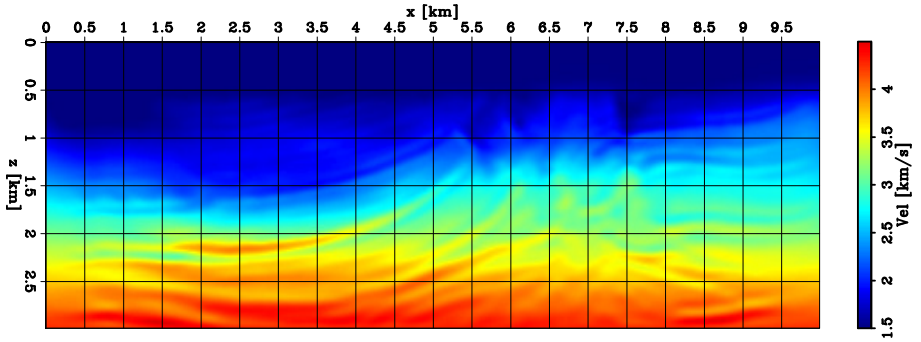
Figure 17: Time-lags CIGs extracted at $x = 3.0$ km from (a) the initial optimal extended perturbation, and (b) the optimal perturbation computed after five iterations. [**CR**]

We continue our FWIME workflow and conduct a total of four inversions (for four different spline parametrizations, each time using the output of the inversion as the initial model for the following one), and display the final inverted model in Figure 18a. For each refinement step, we iterate the FWIME scheme until convergence.

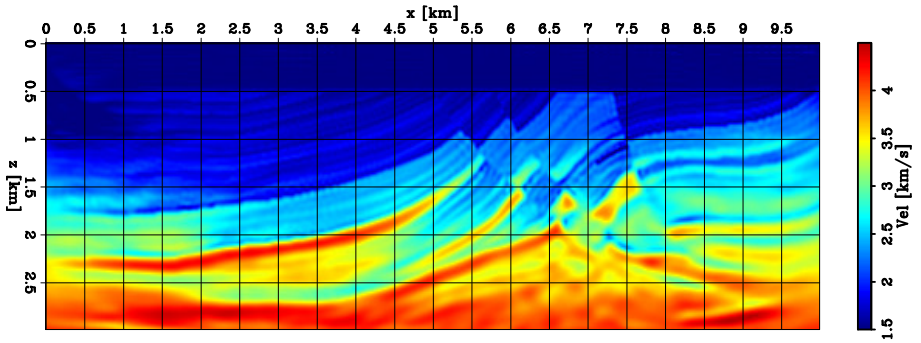
The accuracy of the recovered model is quite good, even though some regions still require improvement, especially in the shallow part of the left side of the model, as well as in the deepest region (for depths grater than 2.5 km). Due to time constraints, we decided to terminate the inversion process after a total 30 iterations of the FWIME workflow (including for the four grid parametrization steps), even though the inversion had not converged yet. Finally, we use the last inverted model as an initial guess for a conventional FWI. The final model is shown in Figure 18b, and is close to the true solution.

Role of the optimal extended perturbation in FWIME

One of the main conceptual features in the design of FWIME (and also an important difference with the approach proposed by Biondi and Almomin (2014)) can be understood by analyzing the role of the optimal extended perturbation $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ as a function of nonlinear iterations. As mentioned in the theory section, the purpose of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ is to generate all the events existing in the observed data that were missed by our current model estimate, $\mathbf{f}(\mathbf{Sm})$. As the FWIME scheme correctly updates our model estimate, more events in the data will be captured directly by the term $\mathbf{f}(\mathbf{Sm})$. Hence,



(a)



(b)

Figure 18: Inverted velocity models from the FWIME scheme. (a) FWIME inverted model. (b) FWI inverted model using (a) as a starting model. [**CR**]

less energy will be contained in the term $\mathbf{d}^{obs} - \mathbf{f}(\mathbf{Sm})$, which acts as the "observed" data for the variable projection step (equation 2). Therefore, if the inversion converges to the optimal solution, the L_2-norm of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ should vanish completely. This clearly highlights the fact that the final $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ of our FWIME scheme should not just be a non-extended perturbation (as required at the end of each set of inner loops in Biondi and Almomin (2014)), but a zero vector.

In order to guide our inversion to such a solution, we use a modified DSO regularization operator, referred to as $\mathbf{D}$ in our main objective function (equation 1). $\mathbf{D}$ is very similar to the one proposed by Symes and Kern (1994). One common application of this operator falls in the context of wave-equation migration velocity analysis (WEMVA) (Sava and Biondi, 2004), where it is used to penalize the energy located in the non-physical regions of extended images (e.g., non-zero time-lags or non-zero subsurface offsets). In addition, Sava and Biondi (2004) show the advantages of using such an operator in order to mitigate the cycle-skipping issue in image-space waveform inversions. However, it does not penalize energy present in the physical axis of the extended images because its aim is to recover a strong and well-focused migrated image. In FWIME, we take advantage of the beneficial properties of this operator, but we modify it by adding a penalty (with a lesser weight) to also force the energy present at zero time-lags in $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ to vanish.

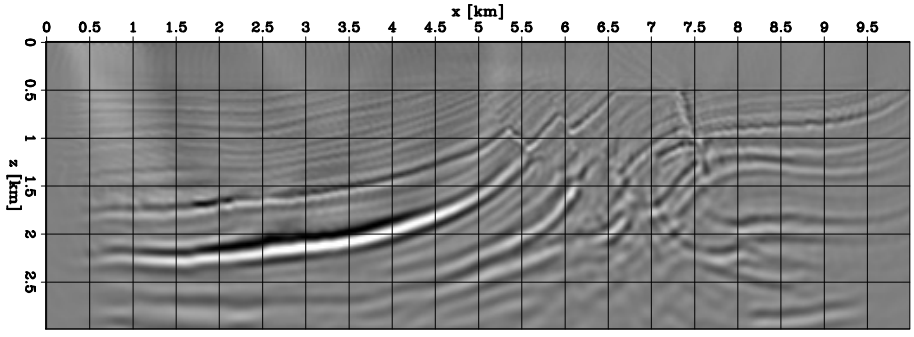
Consequently, a successful FWIME scheme should start reducing the energy along the extended axis of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ (over nonlinear iterations) by adjusting the long-wavelength components of the velocity model. This step is crucial in mitigating the cycle-skipping issue present in conventional FWI. Once the energy of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ is focused in the vicinity of the zero time-lag axis (for each CIG), it should gradually decrease until all the observed data are predicted by the inverted velocity model.

Even though our FWIME scheme has not finished converging (due to time constraints), we are still able to observe the behavior of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ described in the previous paragraph (on certain regions of the model). Figure 19a shows the zero time-lag cross-section of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ computed at the initial step of our last inversion (last spline grid), while Figure 19b shows the same cross-section of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ computed with the last FWIME inverted model (Figure 18a). Both cross-sections are plotted on the same scale, and we can clearly see that in the left region of the model, the energy has almost disappeared from Figure 19b, indicating that all events have been matched directly without the help of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$.

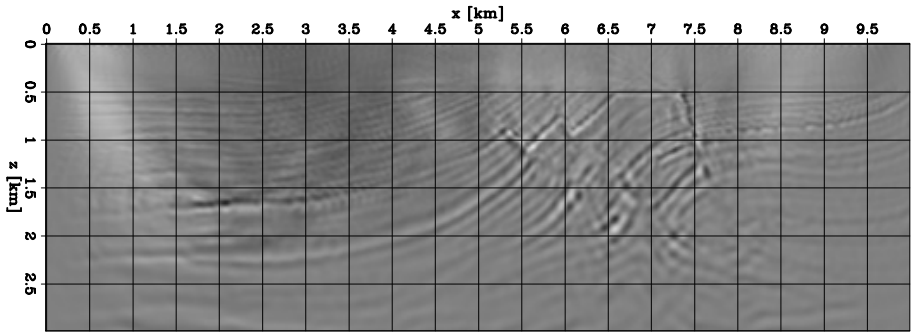
This can also be observed by analyzing the energy of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ along the extended axis throughout the inversion process. Figure 20a and b show two time-lag CIGs extracted at $x = 3.0$ km from Figures 19a and b, respectively. By comparing them to the same CIGs but at early stages of our FWIME workflow (Figures 17a and b), we can see that the energy is now much more focused (notice that the horizontal axis in Figure 20 only ranges from -0.2 s to 0.2 s), which highlights the improvement in the accuracy of the velocity model. Moreover, we observe that the majority of the energy originally present in Figure 20a has almost disappeared in Figure 20b, confirming our observations from the zero time-lag cross-sections.

However, it is clear from Figure 19a that there are still improvements to be made (probably due to the fact that we did not wait until convergence). We can still see some remaining energy in the region located between $x = 5$ km and $x = 7.5$ km, and from the water-bottom interface down to a depth of 2.5 km. The last two columns in Figure 20 show two CIGs extracted at $x = 5.5$ km and $x = 7.5$ km from both cross-sections in Figure 19. Even though we see some noticeable dimming of the deeper events (Figure 20e and f), there is still energy mapped in the shallower region. This is corroborated by our final inverted velocity model that still lacks resolution in the same region (Figure 18a). We believe that by conducting the inversion until full convergence, and possibly performing more spline grid refinement steps, we should be able to improve these features in the velocity model.

Another major difference between the two cross-sections is the presence (also located in the left part of the model) of stronger low-wavenumber energy in Figure 19b. This energy stems from the back-scattering of the source/receiver wavefields due to the high-wavenumber content of inverted velocity model. In fact, it can be seen from Figure 18a that a strong reflector has been added to the velocity model (at an approximate depth of 2.25 km).



(a)



(b)

Figure 19: Zero time-lag cross-sections of $\tilde{\mathbf{p}}^{\epsilon}_{opt}$. (a) First step of the last FWIME inversion (last spline grid parametrization). (b) Last step of the last FWIME inversion (computed using the velocity model shown in Figure 18a). [**CR**]

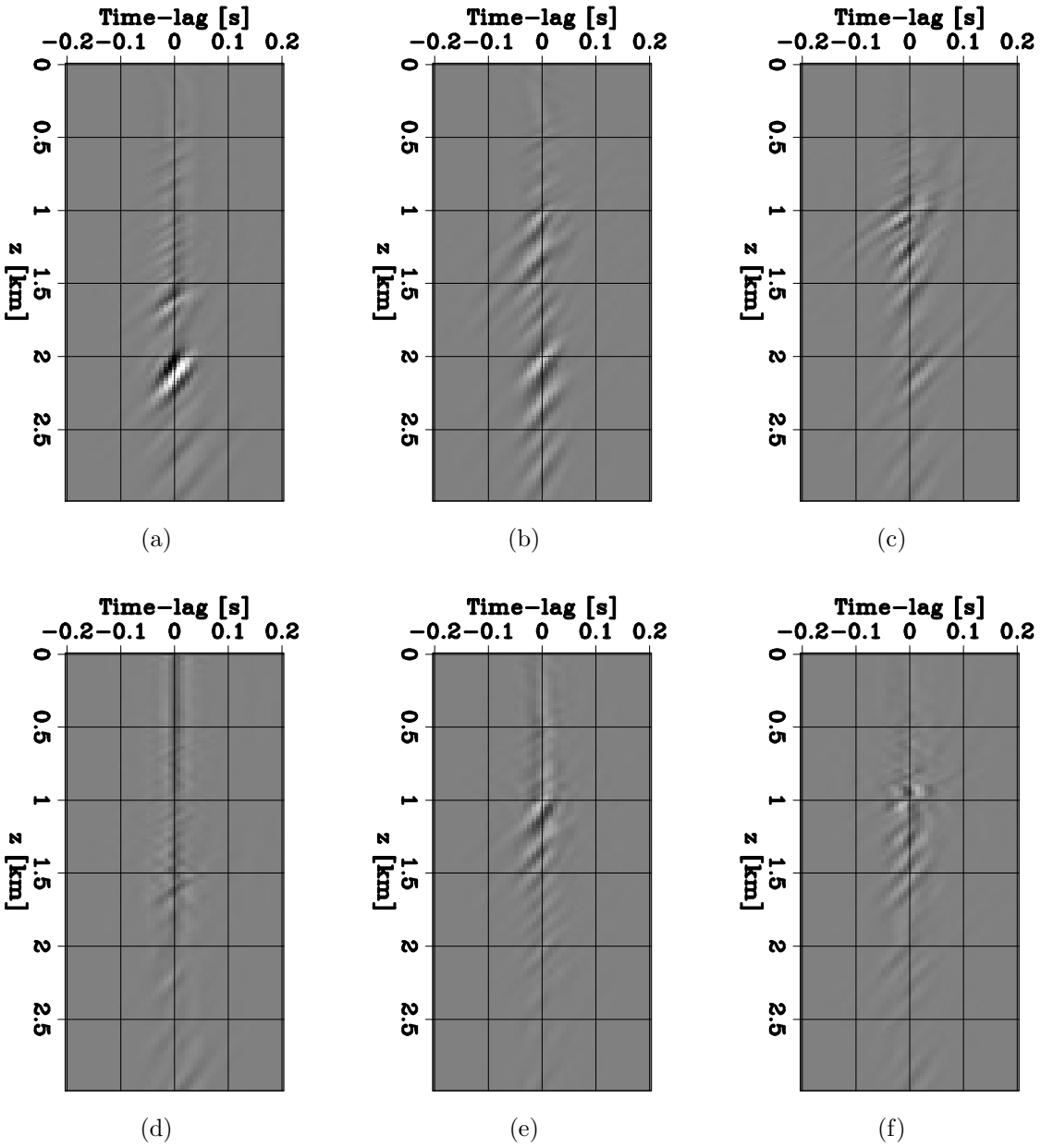


Figure 20: Time-lag CIGs extracted at $x = 3.0$ km (first column), $x = 5.5$ km (second column), and $x = 7.5$ km (third column) from $\tilde{\mathbf{p}}^{\epsilon}_{opt}$. CIGs from the first row were extracted from $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ computed at the first step of the last FWIME inversion (last spline grid parametrization). Figures from the second row were extracted from $\tilde{\mathbf{p}}^{\epsilon}_{opt}$ computed at the last step of our FWIME inversion. All CIGs are displayed on the same scale. [**CR**]

CONCLUSIONS, CHALLENGES AND OPPORTUNITIES

We develop a model-space multi-scale algorithm using spline interpolation and incorporate it into our FWIME workflow. We show on two numerical examples how it can be applied to gradually build the wavenumber components of the velocity model in order to mitigate the risk of converging to a local minimum. We then showed an application of FWIME on the Marmousi model, where we successfully invert pure reflection data with no available signal below 4 Hz while starting with an inaccurate velocity model. The output of FWIME was then used as an initial guess for conventional FWI and the final recovered model is very accurate.

There are clearly opportunities to improve our FWIME result, and more experiments need to be conducted. For instance, we are currently testing our workflow on the simultaneous inversion of refracted and reflected energy. Moreover, these tests should provide us a better insight on how to efficiently refine the spline grid and adjust the trade-off parameter between the FWIME data-fitting and the model-fitting components of the objective function. We believe there exists a consistent way to automatically choose these parameters.

Finally, the feasibility and application of FWIME on 3D field data highly depends on our ability to reduce the number of linear iterations during the variable projection step. Therefore, we are currently exploring methods to precondition the linear problem stemming from the variable projection step.

REFERENCES

Barnier, G., E. Biondi, and B. Biondi, 2018, Full waveform inversion by mode extension: Presented at the SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists.

Barnier, G., E. Biondi, and R. Clapp, 2019, Model-space multi-scale approach for waveform inversion using spline interpolation.

Biondi, B., and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity anaylisis, *in* SEG Technical Program Expanded Abstracts 2012: Society of Exploration Geophysicists, 1–5.

———, 2013, Tomographic full waveform inversion (TFWI) by extending the velocity model along the time-lag axis, *in* SEG Technical Program Expanded Abstracts 2013: Society of Exploration Geophysicists, 1031–1036.

———, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: Geophysics, **79**, WA129–WA140.

Biondi, E., R. G. Clapp, and G. Barnier, 2019, A flexible library for geophysical inverse problems – structure and usage.

Golub, G. H., and V. Pereyra, 1973, The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate: SIAM Journal on numerical analysis, **10**, 413–432.

Huang, Y., and W. W. Symes, 2015, Born waveform inversion via variable projection and shot record model extension, *in* SEG Technical Program Expanded Abstracts 2015: Society of Exploration Geophysicists, 1326–1331.

Martin, G. S., R. Wiley, and K. J. Marfurt, 2006, Marmousi2: An elastic upgrade for marmousi: The Leading Edge, **25**, 156–166.

Rickett, J., 2013, The variable projection method for waveform inversion with an unknown source function: Geophysical Prospecting, **61**, 874–881.

Sava, P., and B. Biondi, 2004, Wave-equation migration velocity analysis. i. theory: Geophysical Prospecting, **52**, 593–606.

Symes, W. W., and M. Kern, 1994, Inversion of reflection seismograms by differential semblance analysis: Algorithm structure and synthetic examples: Geophysical Prospecting, **42**, 565–614.

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