Journal of Applied Geophysics 215 (2023) 105120

Contents lists available at ScienceDirect

Journal of Applied Geophysics

journal homepage: www.elsevier.com/locate/jappgeo

Spectral recomposition for optimizing starting points in Full-Waveform Inversion

Nelson Ricardo Coelho Flores Zuniga [a,*], Rafael dos Santos Gioria [a], Bruno Souza Carmo [b]

[a] Department of Mining and Petroleum Engineering, Polytechnic School, University of São Paulo, São Paulo 05508-010, Brazil
[b] Department of Mechanical Engineering, Polytechnic School, University of São Paulo, São Paulo 05508-010, Brazil

ARTICLE INFO

Keywords:
Finite difference method
Full-Waveform Inversion
Spectral recomposition

ABSTRACT

By optimizing the wavelet fitting of the arrivals in a reflection event, it is possible to perform significant enhancements for Full-Waveform Inversion (FWI) predictions and its processing time. For this reason, approaches capable of reducing the number of iterations for FWI — without decreasing the quality of the prediction — are of interest. Once the initial guess for FWI is better estimated, it is possible to predict the velocity model within a determined accuracy with fewer iterations. We thus propose an approach which can perform this estimation, based on spectral recomposition of seismic data. We design an inversion scheme to reconstruct the seismic spectrum of wavelets of a reflection event, which subsequently allows estimating the position in time of each wavelet in a seismogram. After finding the position in time of each wavelet, we can guide the calculated wavelet to fit the corresponding observed signal, starting from a closer initial point. Our approach leads to quite accurate predictions of velocity models with fewer iterations.

1. Introduction

The wavelet properties of a seismic reflection event at a subsurface boundary are an important piece of information for performing analyses of a seismic dataset. In comparison to other techniques, in Full-Waveform Inversion (FWI), increasing the knowledge of wavelet information can have a stronger impact on processing time (Brittan et al., 2013; Virieux et al., 2017). To obtain a better wavelet fitting, the position in time of a wavelet should be considered during the inversion procedure (Jones, 2010). More detailed information about wavelets is usually disregarded as a priori information, which demands better geological knowledge of the region for building more reliable initial guesses of velocity models for performing FWI (Asnaashari et al., 2013). Uncertainties in building an initial guess of velocity models would result in uncertainties in the prediction of the velocity model (Jones, 2010; Asnaashari et al., 2013). Therefore, it is desirable to have a method that considers more wavelet information in a seismogram for optimizing starting points for minimizing wavelets in the FWI technique.

The technique known as spectral recomposition can be used to extract useful information from a seismic spectrum (Tomasso et al., 2010; Li et al., 2011; Cai et al., 2013). Spectral recomposition is usually performed by estimating fundamental signal properties (i.e., frequency, amplitude and phase) to reconstruct the seismic spectrum (Castagna et al., 2003; Tomasso et al., 2010; Cai et al., 2013). Note that the method for estimating the fundamental properties of a signal can be based on the Prony decomposition, which makes it possible to estimate not only the frequency, amplitude, and phase of the seismic signal but also its attenuation coefficient (Fomel, 2013; Mitrofanov and Priimenko, 2013; Mitrofanov and Priimenko, 2015). Assuming a mathematical description of the analyzed wavelet, an inversion procedure can be used to reconstruct the seismic spectrum of the analyzed wavelet. This allows performing the inversion to recover the fundamental signal properties (i.e., frequency, amplitude and phase).

Even in the presence of random noise, spectral recomposition allows recovering reliable values of fundamental signal properties (Zuniga and Priimenko, 2022). For this reason, an approach which allows building more reliable initial models for FWI by using spectral recomposition for obtaining signal properties can be useful.

In this paper, we develop an inversion procedure to reconstruct the seismic spectrum of wavelets in a seismogram and to subsequently estimate better starting points for performing wavelet minimization in FWI. For this, we fit a calculated spectrum to an observed one, whereby the calculated curve is the mathematical representation of a Ricker wavelet (Ricker, 1953) and the observed curve is the wavelet of a

* Corresponding author.
E-mail addresses: nelson.zuniga@iag.usp.br (N.R.C.F. Zuniga), rafaelgioria@usp.br (R.S. Gioria), bruno.carmo@usp.br (B.S. Carmo).

https://doi.org/10.1016/j.jappgeo.2023.105120
Received 16 January 2023; Received in revised form 12 June 2023; Accepted 27 June 2023
Available online 28 June 2023

reflection event in a recorded trace. The curve fitting is performed for each wavelet in a common source gather. Each signal parameter of each wavelet in the gather is thus estimated and the spectrum can be reconstructed. This allows picking the position in time for each wavelet.

Since the reconstruction of a seismic spectrum deals with amplitude and phase spectrum in the frequency domain. Fourier transform must be performed for transforming the wavelet in time domain into an amplitude and phase in the frequency domain. Next, the inverse Fourier transform is performed to bring the inverted spectrum back to the time domain as a wavelet. Note that, contrary to other domain transformations (e.g., transformation into τ-p domain), neither transformation produce information loss (Wood, 1974; Yang et al., 2016).

Notice that our approach does not require previous geological information and estimates the signal parameters directly from data. Contrary to other approaches for building velocity models as initial guesses for FWI, our approach only uses information contained in the seismogram.

The proposed method works on a wavelet-by-wavelet basis, which allows establishing the number of wavelets we would apply our approach to. Since one can set up the number of wavelets to be inverted per trace and per time sample, other data could be used to perform fewer wavelet inversion procedures according to the a priori information available.

1.1. Spectral recomposition

As proposed by Tomasso et al. (2010), a sum of different Ricker-wavelet amplitude and phase spectra can represent the amplitude and phase spectrum of a seismic trace.

$$d(f) \approx \sum_{i=1}^{n} a_i \psi_i(m_i, f), \tag{1}$$

where $d(f)$ is the amplitude and phase spectrum of a seismic trace, f is frequency, and a_i and m_i are the amplitude and the peak frequency of the i-th Ricker-wavelet spectrum, respectively. A Ricker-wavelet amplitude and phase spectrum with a peak frequency at m is given by

$$R(f) = a\psi(m, f) = a\frac{f^2}{m^2} exp\left(-\frac{f^2}{m^2}\right). \tag{2}$$

The amplitude, peak frequency and phase of the spectra in a seismic trace can be obtained by using spectral recomposition to reconstruct the seismic spectrum of that trace. Contrary to other approaches, this one does not decompose the amplitude and phase spectrum (Huang et al., 1998; Castagna et al., 2003; Li et al., 2011; Cai et al., 2013).

Using the mathematical description given by Eqs. (1) and (2), which describe the amplitude and phase spectrum of a Ricker wavelet, it is possible to treat the problem in an inverse manner aiming to fit a calculated amplitude and phase spectrum to that obtained from an observed wavelet in each seismic trace. This allows recovering the parameters related to phase, peak frequency, and amplitude from the spectrum, and hence, identify wavelet properties and its position in time.

The model is a linear combination of Ricker wavelet spectra; each spectrum is a nonlinear function that depends on two parameters: the coefficients a (amplitude) and m (peak frequency). We can estimate the error using

$$r_j = d(f_j) - \sum_{i=1}^{n} a_i(m_i)\psi(m_i, f_j), \tag{3}$$

where f_j is the vector of frequencies related to each sampled point of the linear combination of Ricker wavelet spectra.

The objective function is then formulated as a least-squares estimation:

$$\min_{a,m} \| r(a, m) \|_2^2 \tag{4}$$

Note that this objective function is the one for performing the spectral recomposition of the spectra in a trace and not the same used for the inversion performed in FWI.

Concerning the numerical method, parameters of m are assumed to be nonlinear, while parameters of a are assumed as linear. They can be obtained by solving the nonlinear least-squares problem,

$$a = \psi(m)^* d, \tag{5}$$

where $\psi(m)$ is the matrix composed by $\psi_i = (m_i, f_i)$ and $\psi(m)^*$ is the Moore-Penrose generalized inverse of the $\psi(m)$ matrix (Moore, 1920; Penrose, 1955). When Eq. (5) is substituted in Eq. (4), the minimization problem takes the form

$$\min_{m} \| (I - \psi(m)\psi(m)^*)d \|_2^2, \tag{6}$$

where the linear parameters have been eliminated. In this case, the Gauss-Newton method can be used to linearize the problem, leading to

$$d(f_i) \approx \sum_i R_j(m_i, f_i) + \sum_i \frac{\partial R_j}{\partial m_i}, \tag{7}$$

where $R_j(m_i, f_i)$ represents one of the Ricker wavelet spectra of the linear combination for the same vector of frequencies related to each sampled point.

Substituting Eq. (2) in Eq. (7) results in

$$d(f_i) \approx \sum_i a_i \psi_i(m_i, f_i) + \sum_i \left[a_i' \psi_i(m_i, f_i) + a_i \psi_i'(m_i, f_i) \right] \Delta m_i. \tag{8}$$

With initial values of m_i, Eq. (5) can be used to solve for a_i and a_i'.

In the next section, we explain how to perform the inversion for recovering phase information and for estimating the position in time for each wavelet; then, in the following section we show how the results of the inversion by means of spectral recomposition is incorporated to the FWI for guiding its gradient.

1.2. Finding position in time for each wavelet

Using the mathematical description presented in the previous section, i.e., Eqs. (1) and (2), one can calculate the amplitude and phase spectrum of the Ricker-wavelet. After time-windowing a reflection event in an observed trace (observed wavelet), and transforming it into frequency domain, it is possible to fit a spectrum calculated from Eq. (8) to the observed one.

We developed an inversion scheme which allows performing the optimization using curve fitting. Since the Ricker-wavelet spectrum can be obtained from the observed data, the parameters to be recovered by the inversion procedure are the peak frequency, amplitude, and phase.

Rather than using global-search optimization algorithms, we apply a multi-start procedure, which applies a local-search routine (Terlaky and Sotirov, 2010). From random starting points, our approach leads to a statistical distribution of the global and local minimum regions. Note that the quality of the statistical distribution is dependent on the quantity of iterations. We perform least-square minimization between the Ricker-wavelet spectrum of the observed trace and the theoretical spectrum obtained from Eqs. (1) and (2). We show the pseudocode of the multi-start procedure in the Appendix.

Each iteration results in a different minimum value (considering the process to be stochastic), which allows comparing the minima of each iteration. Finally, it is possible to select the lowest value of the minimum. The number of iterations can be set to adapt the complexity of the objective function of an analyzed spectrum. This increases the accuracy when recovering the amplitude, peak frequency, and phase corresponding to the global minimum region. It also reduces the number of

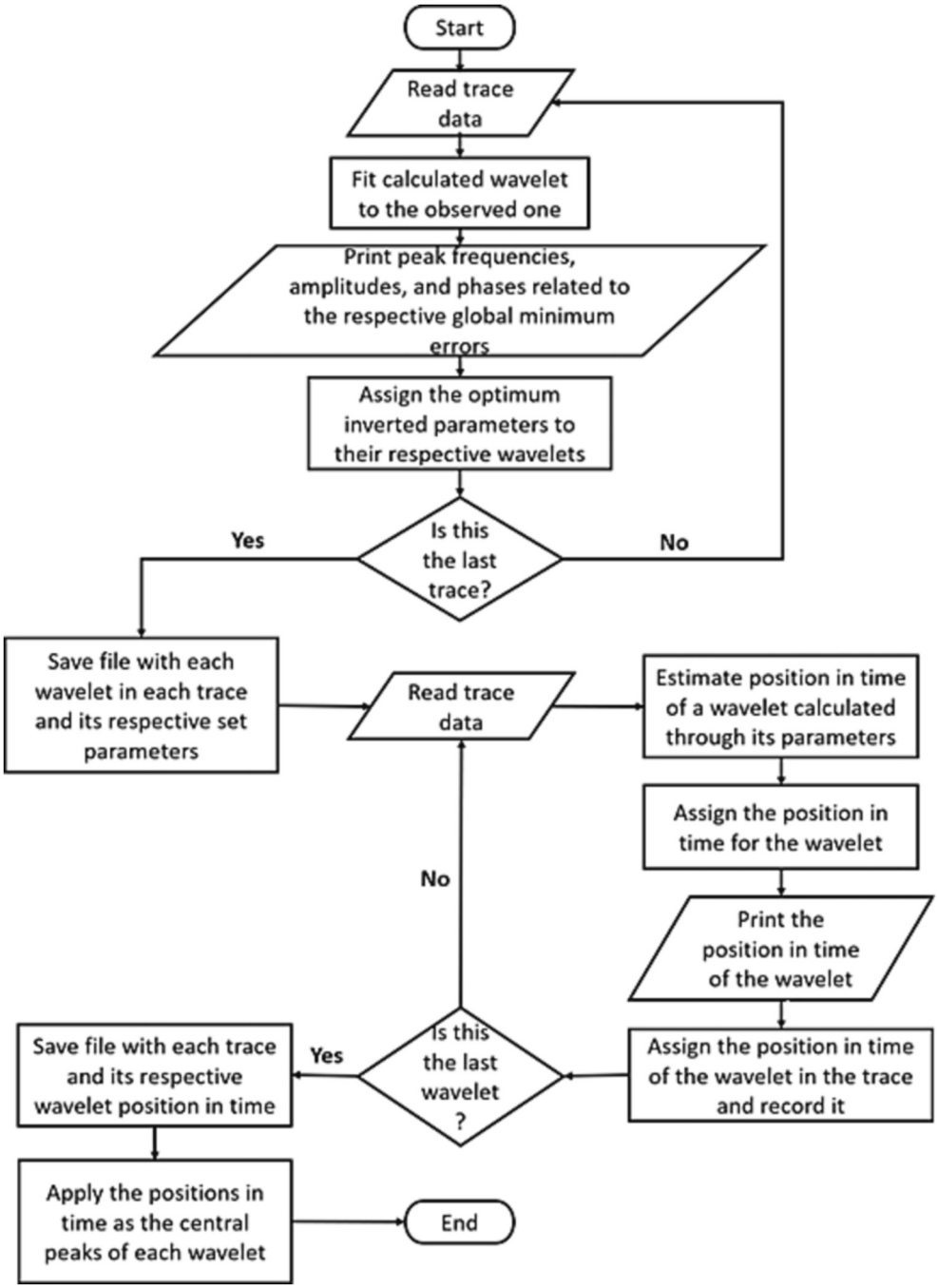

Fig. 1. Flowchart of the proposed algorithm for spectral recomposition.

iterations needed as shown in the results section.

This type of objective function mapping provides a deeper understanding of the complexity of the functional topology, which allows an improved inversion by the selection of a more appropriate optimization algorithm or minimization method.

This inversion is performed for each wavelet in a trace contained in a common source gather, resulting in a set of peak frequencies, amplitudes, and phases. Each optimum set of parameters recovered for each wavelet in a trace presents a residual error, which implies not marking the exact position in time of the central peak of a wavelet. Nevertheless, this marked position is a good reference for guiding the wavelet fitting during the minimization process in FWI.

Considering the description given by Eq. (2), the inversion enables us to estimate the peak frequency (m) and amplitude (a) by inverting both parameters for the wavelet function. It is not interesting to perform the inversion for the phase simultaneously, since the phase (φ) is function of both peak frequency and amplitude. For this reason, the inversion for recovering the phase, is performed after the inversion performed to recover m and a. So, after minimizing Eq. (8) to find m and a, it is necessary to find the phase.

The spectrum is fitted until it reaches its minimum with the objective function so that a_i and m_i can be estimated. With this, the estimation of the phase can be performed. To estimate the phase of this wavelet, one uses the argument of Eq. (2):

$$arg[R(f)] = arg\left[a\frac{f^2}{m^2}exp\left(-\frac{f^2}{m^2}\right)\right]. \tag{9}$$

From Eq. (9), it is found that

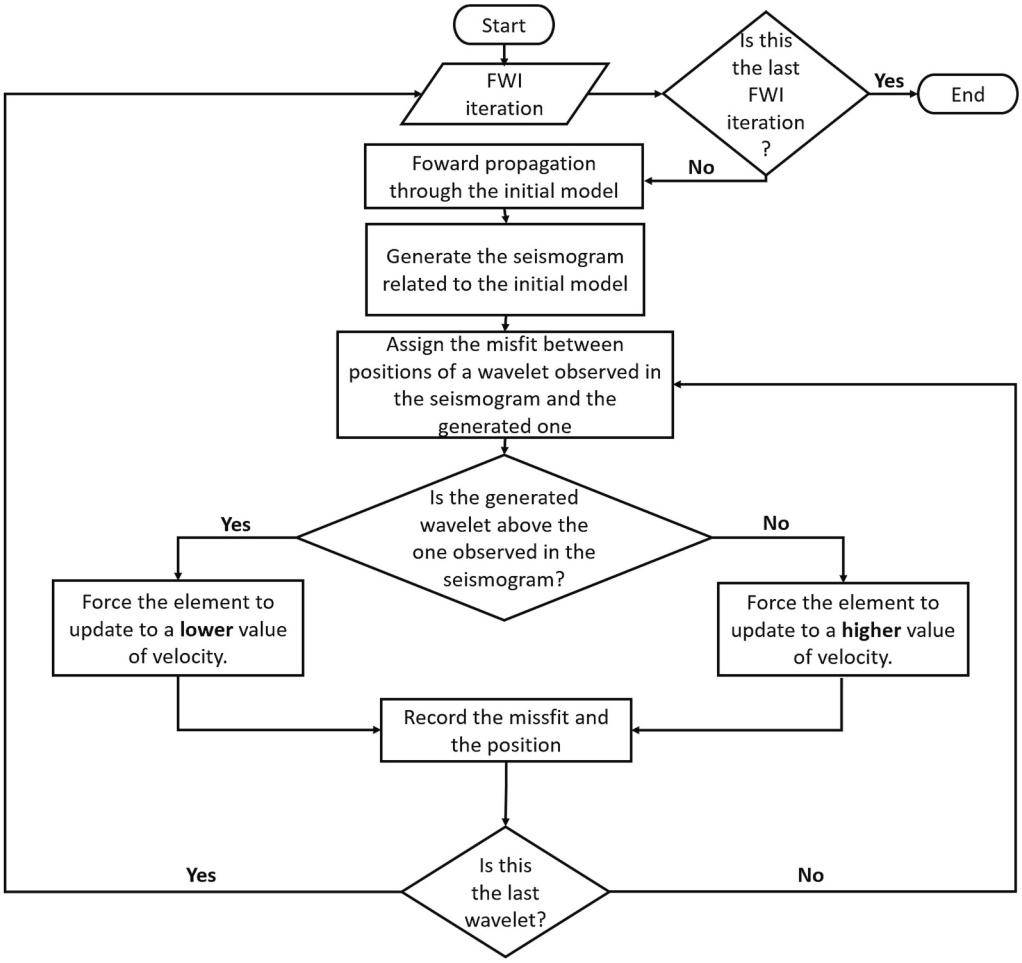

Fig. 2. Flowchart of the proposed algorithm for FWI using spectral recomposition.

$$arg[R(f)] = -\frac{f^2}{m^2}. \tag{10}$$

Being n the number of traces, we can rewrite Eq. (10) as

$$arg[R_n(f)] = -\frac{f^2}{{m_n}^2}, \tag{11}$$

where R_n and m_n are the values of R and m corresponding to the trace with number n, (respectively)

Note that with m_n and a_n previously recovered, the inversion can be performed for recovering only the phase, as described in Eq. (11), since $arg[R(f)]$ is the instantaneous phase. In this case, the inversion is simpler than performing an inversion with three parameters in which one of them is a function of the other two.

Since the phases rotate as long as the offsets increase, a correction must be performed to determine the centre of the wavelet in a more accurate.

As each wavelet in a given reflection event presents a specific phase shift along the offset, one can compute the n-th phase shift for this event relative to the phase in the first ($n = 1$) trace as follows. For the first trace, the phase shift is zero:

$$arg[R_1(f)] = -\frac{f^2}{{m_1}^2} + \varphi_1 = -\frac{f^2}{{m_1}^2} + 0, \tag{12}$$

where φ is the instantaneous phase shift of the first trace.

For the following traces, $n > 1$, there is a different peak frequency (m_n) value. Therefore, the term $-f^2/{m_n}^2$ is unique of each trace. Then, φ_n represents how much the phase of the wavelet in the n-th trace differs in comparison to that of the wavelet in the first trace for the same reflection event (i.e., instantaneous phase shift):

$$arg[R_n(f)] = -\frac{f^2}{{m_n}^2} + \varphi_n. \tag{13}$$

In Eq. (13), the term $arg[R_n(f)]$ represents the observed instantaneous phase, while the right-hand of the equality represents the calculated instantaneous phase.

Finding the phase shift for each wavelet allows correcting the phase in each trace by rotating it until it gets close to the wavelet in the first trace. This enables using the reflection information at offsets beyond the critical angle, which is otherwise not possible. This allows estimating the set of three needed parameters m, a, and φ. This set of parameters is computed for both the reference seismic trace and the synthetic one. The observed difference is then used in seismic inversion procedures.

With parameters m, a, and φ, we perform the correction of the phase and estimate the position of the centre of the wavelet.

The maximum amplitude of that wavelet a_{max} is defined as the centre of that wavelet and it can be obtained after knowing the phase information. a_{max} must be minimized for each point in $R(f)$, and the minimum will occur when $a = a_{max}$ in the linear relation for vector $f : a = a_{max}$:

$$R(f) = a_k \psi(m, f_k) = a_{max} \psi(m, f_k) \tag{14}$$

The flowchart of the proposed algorithm is presented in Fig. 1.

1.3. Guiding the minimization process in FWI

With the position of the central peak of each Ricker-wavelet in a

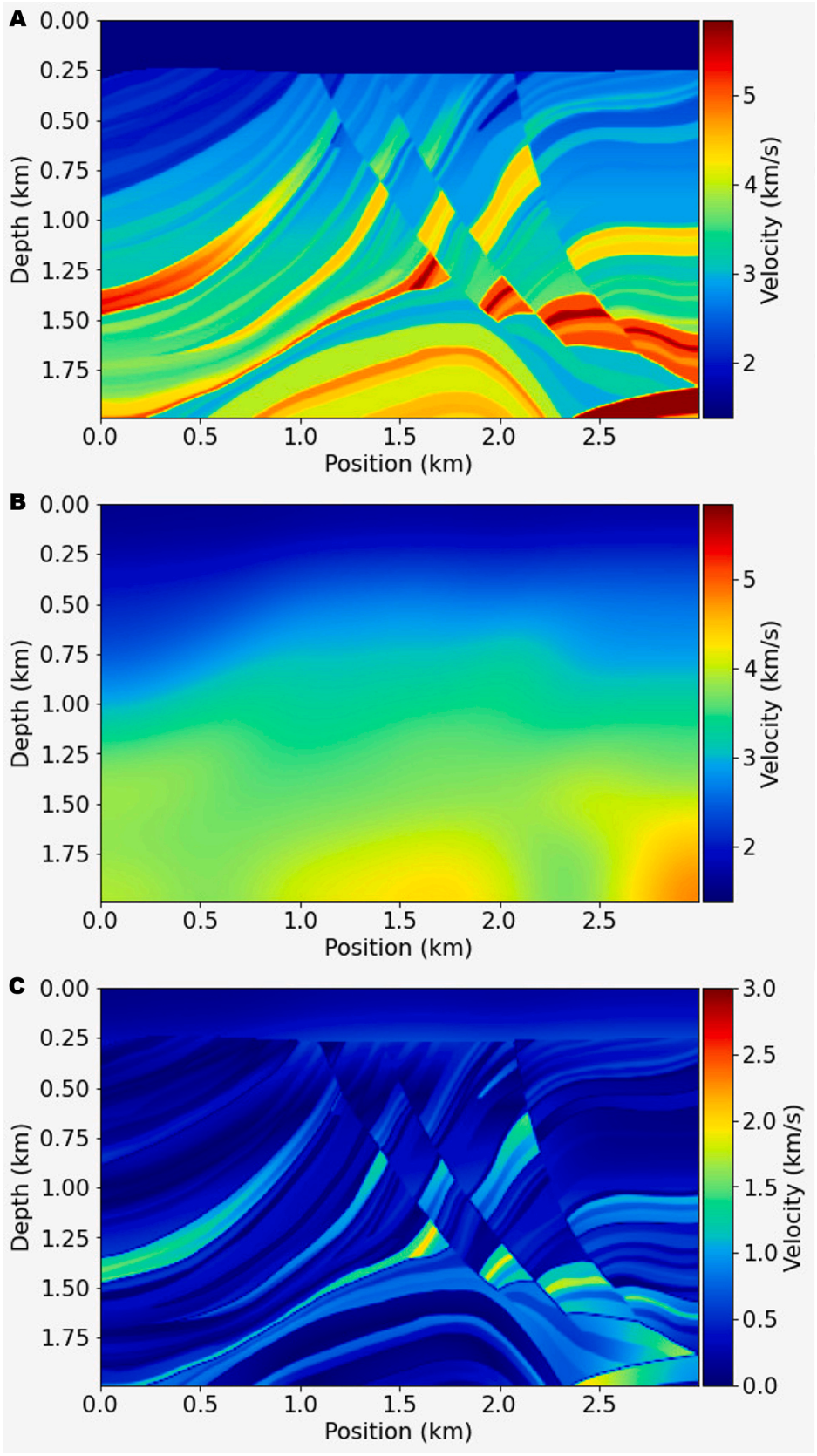

Fig. 3. (A) Resized crop from Marmousi II with a water layer, which is the ground truth of this experiment. (B) Smoothed model for being the initial guess. (C) Absolute difference between the ground truth and the initial guess.

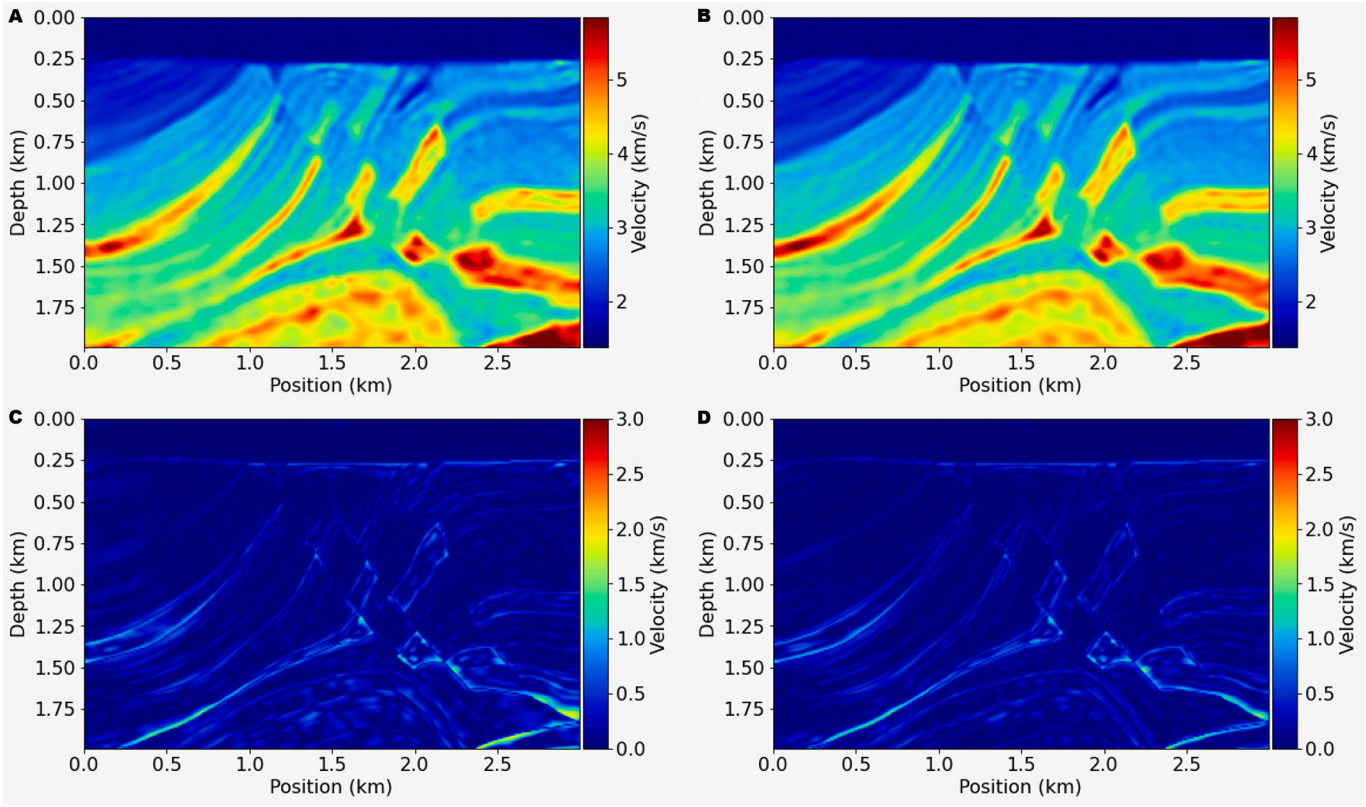

Fig. 4. Results from using the model in Fig. 3B as an initial guess, for (A) the conventional multiscale FWI and (B) FWI with our approach. (C) Absolute difference between the conventional multiscale FWI and the ground truth. (D) Absolute difference between FWI with our approach and the ground truth.

common source gather, we can now introduce boundary conditions as guides to performing the minimization process in each FWI iteration. Note that our algorithm for finding the position in time for each wavelet is applied to each FWI iteration. For this, we need to define conditions so that the model properties updates make the wavelet central peak of the forward propagation closer to the wavelet central peak marked in the observed common source gather.

The FWI inversion problem using Devito (Louboutin et al., 2019) can be defined in time domain, for an acoustic isotropic case and a spatially varying model, as

$$\begin{cases} m\dfrac{\partial^2 u(x,t)}{\partial t^2} - \nabla^2 u(x,t) = q \\ u(,0) = 0 \\ \dfrac{\partial u(x,t)}{\partial t}\Big|_{t=0} = 0 \end{cases} \tag{15}$$

where u is the wavefield, m is the square slowness, q is the source, and ∇^2 is the Laplacian. This system is solved with a time marching scheme, which leads to the seismic inversion problem

$$min_{\vec{m}}\ \Phi_s(\vec{m}) = \frac{1}{2}\left\| \left(P_r A^{-1}(\vec{m})\vec{q} - \vec{d}\right\|_2^2, \right. \tag{16}$$

where A is the discrete wave equation operator, $\vec{d}$ is the field measured data, $\vec{m}$ is the vector of square slowness, $\vec{q}$ is the vector of the source term, lower index s represents the size of a discrete time step, and P_r is the sampling operator to the measurement locations. This problem can be solved by performing the optimization with the gradient-descent method. The adjoint-state method is used for evaluating the gradient $\nabla\Phi_s(\vec{m})$:

$$\nabla\Phi_s(\vec{m}) = \sum_{t=1}^{n_t} \vec{u}[\vec{t}]\,\vec{v}_{tt}[\vec{t}] = J^T\delta\vec{d}, \tag{17}$$

where $\delta\vec{d} = \left(P_r\vec{u} - \vec{d}\right)$ is the residual data, $\vec{u}$ is the discrete wavefield, $\vec{v}$ is the adjoint wave equation computed backwards in time which is the solution of the discrete adjoint wave equation for an adjoint source, and $\vec{v}_{tt}$ is the second-order time derivative of the adjoint wave equation computed backwards in time:

$$A^*(\vec{m})\vec{v} = P_r^*\delta\vec{d}_s \tag{18}$$

The adjoint-state method demands the wave equation to be solved for both the forward and adjoint wavefield and the full storage of the forward wavefield $\vec{u}$ in order to compute the gradients.

This objective function described in Eq. (16) is used as a metric for optimizing FWI with the least-squares as the minimization method.

We propose an approach that uses the marked positions in time of the central peak of each Ricker-wavelet as a guide to each minimization process in FWI by restricting the variation of the properties (e.g., P-wave velocity) in an element in the matrix during the model update. This prevents the model from updating the elements properties to values that make the next forward propagation seismogram less befitted to the observed one. With an event in time defined for the central peak of a wavelet, we can then tune next properties update to lead the new wavelet generated by the forward propagation upwards or downwards in a trace, i.e., leading to a lower or higher value in time.

Considering two scenarios to set the conditions that fit as a guide for the model update:

1-The wavelet generated by the previous forward propagation presents a misfit being above the observed wavelet in a trace from the common shot gather. In this case, the updated model presents a higher velocity value in that element than it should;

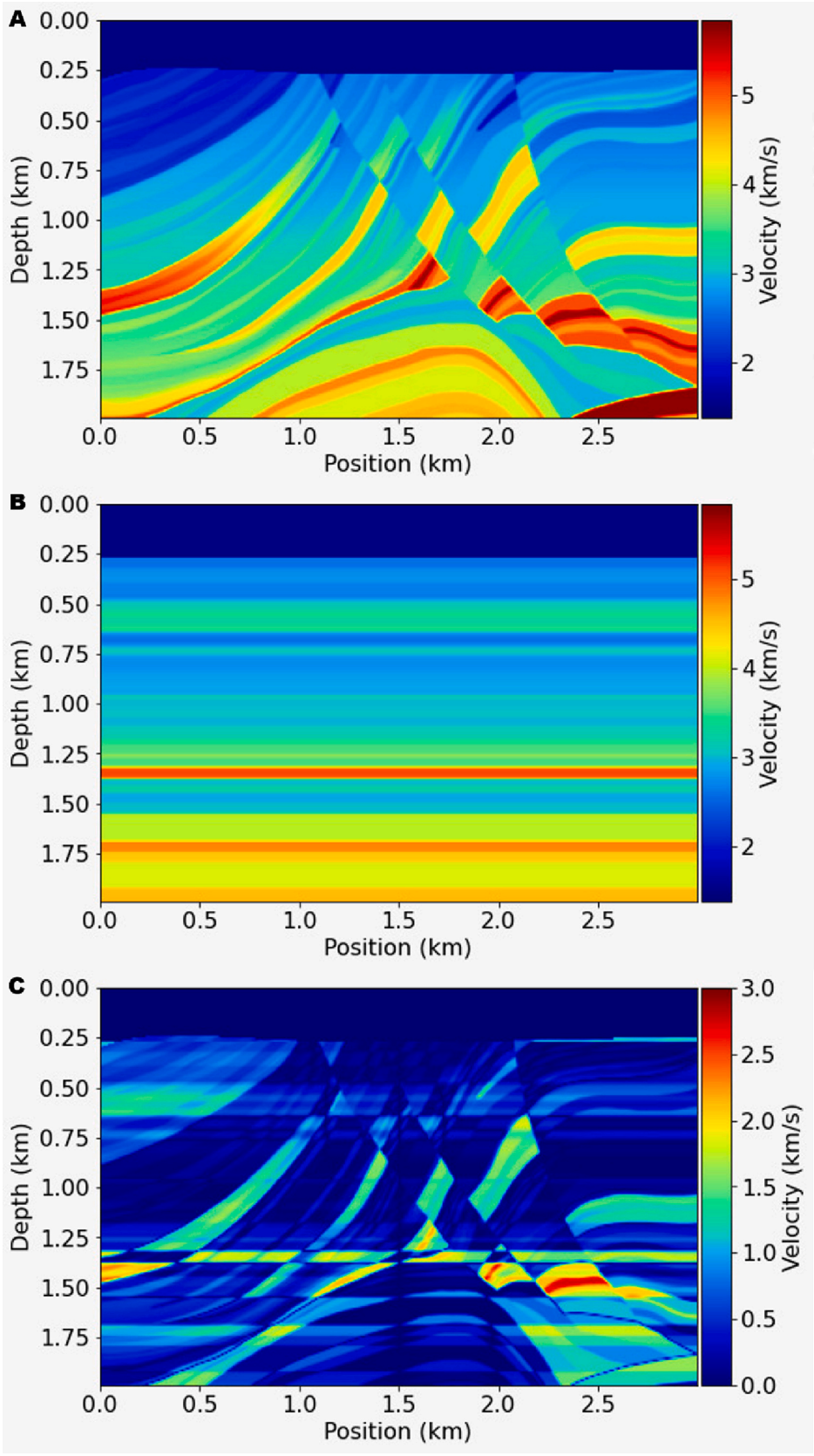

Fig. 5. (A) Resized crop from Marmousi II with a water layer, which is the ground truth of this experiment. (B) Model made by horizontally extended velocity profile for being the initial guess. (C) Absolute difference between the ground truth and the initial guess.

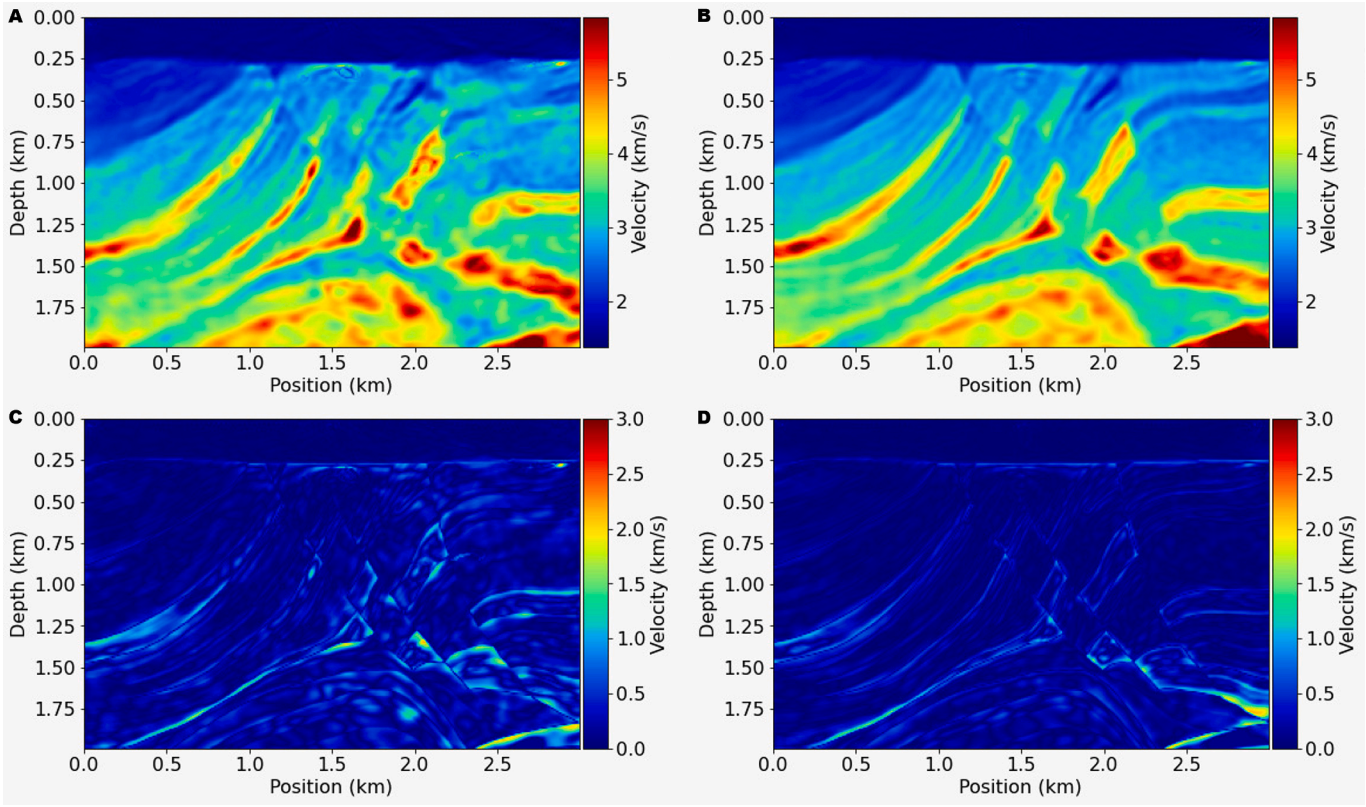

Fig. 6. Results from using the model in Fig. 5B as an initial guess, for (A) the conventional multiscale FWI and (B) FWI with our approach. (C) Absolute difference between the conventional multiscale FWI and the ground truth. (D) Absolute difference between FWI with our approach and the ground truth.

2-The wavelet generated by the previous forward propagation presents a misfit being under the observed wavelet in a trace from the common shot gather. In this case, the updated model presents a lower velocity value in that element than it should.

For example, in the first scenario, the next model update must have a lower velocity in that element. Therefore, we prevent the velocity from increasing for that element, forcing the element to update to a lower velocity.

The conventional FWI is performed and, then, each gradient value is computed and compared to the information provided by spectral recomposition. In the case of the gradient computed by the FWI matches the gradient calculated by the spectral recomposition, the algorithm proceeds; however, in the case the gradients do not match, the gradient, the FWI computed gradient is disregarded and considered zero to avoid the increase of the error. Note that the spectral recomposition does not modify the conventional multiscale FWI technique itself, it provides a priori information for restricting the potential global minimum region of the mapped objective function. It limits its gradients by restricting the objective function's region to a region closer to the global minima. The position in time in a trace found by the spectral recomposition provides a narrower potential global minimum region, enabling us to perform an inversion procedure with a lower number of non-global solutions — without modifying FWI objective function itself.

There is certainly a correlation in each wavelet above and under the analyzed wavelet, which makes it much more complex to estimate the contribution of each property in each element related to a trace. However, this approach by bounding reduces the number of possibilities that are more likely to provide worse fittings. The proposed approach provides an efficient manner to reduce iterations by preventing the Ricker-wavelet from fitting local minima (lateral peaks of Ricker-wavelet) instead of the global one (central peak of the Ricker-wavelet). The flowchart of the proposed algorithm is presented in Fig. 2.

2. Results

Our approach can enhance the efficiency of FWI predictions by guiding its minimization process, which allows decreasing the number of iterations executed during the inversion procedure. This approach shows to be able to reach more befitted values of P-wave velocity with fewer iterations in a multiscale technique. The multiscale technique shows to be effective for preventing cycle skipping.

We tested our approach by analyzing P-wave reflection events in a model that is a crop from Marmousi II. This model represents a window cropped from position 1.44 Km up to position 5.25 Km and from depth 0.3 Km up to depth 1.42 Km. This crop is resized to preserve the complexity regarding the depth, and we also added a water layer of average depth of 0.25 Km with depth variation of $\pm$0.03 Km. We employed a grid with 300 $\times$ 200 elements, and the mesh spacing was 0.01 km. We used 100 equispaced receivers and 10 equispaced sources located at the top of the domain. The time sampling was performed at every 0.001 s during 3 s. The frequency of the source was 15 Hz with smoothed spectrum varying from 7 Hz to 21 Hz. We used a multiscale technique in which the cutoff frequency was 15 Hz, the initial frequency was 7.5 Hz, and the frequency interval was 2.5 Hz. It is worth mentioning that multiscale technique helps to improve the accuracy of predicting fractures, faults, and other high-angle features (Pan et al., 2020), which are commonly found in Marmousi II and other hydrocarbon reservoir scenarios.

For the forward seismograms generation during each FWI iteration, we simulated a P-wave survey using Devito (Louboutin et al., 2019). Surface waves and free-surface multiples were not modelled. In a case of field data, this means that surface-wave suppression would have been applied, for example, using data-driven interferometric surface-wave suppression (Balestrini et al., 2020), followed by surface-related multiple suppression in the raw field data (Verschuur et al., 1992; Ghose and Goudswaard, 2004). In many approaches, amplitude information is

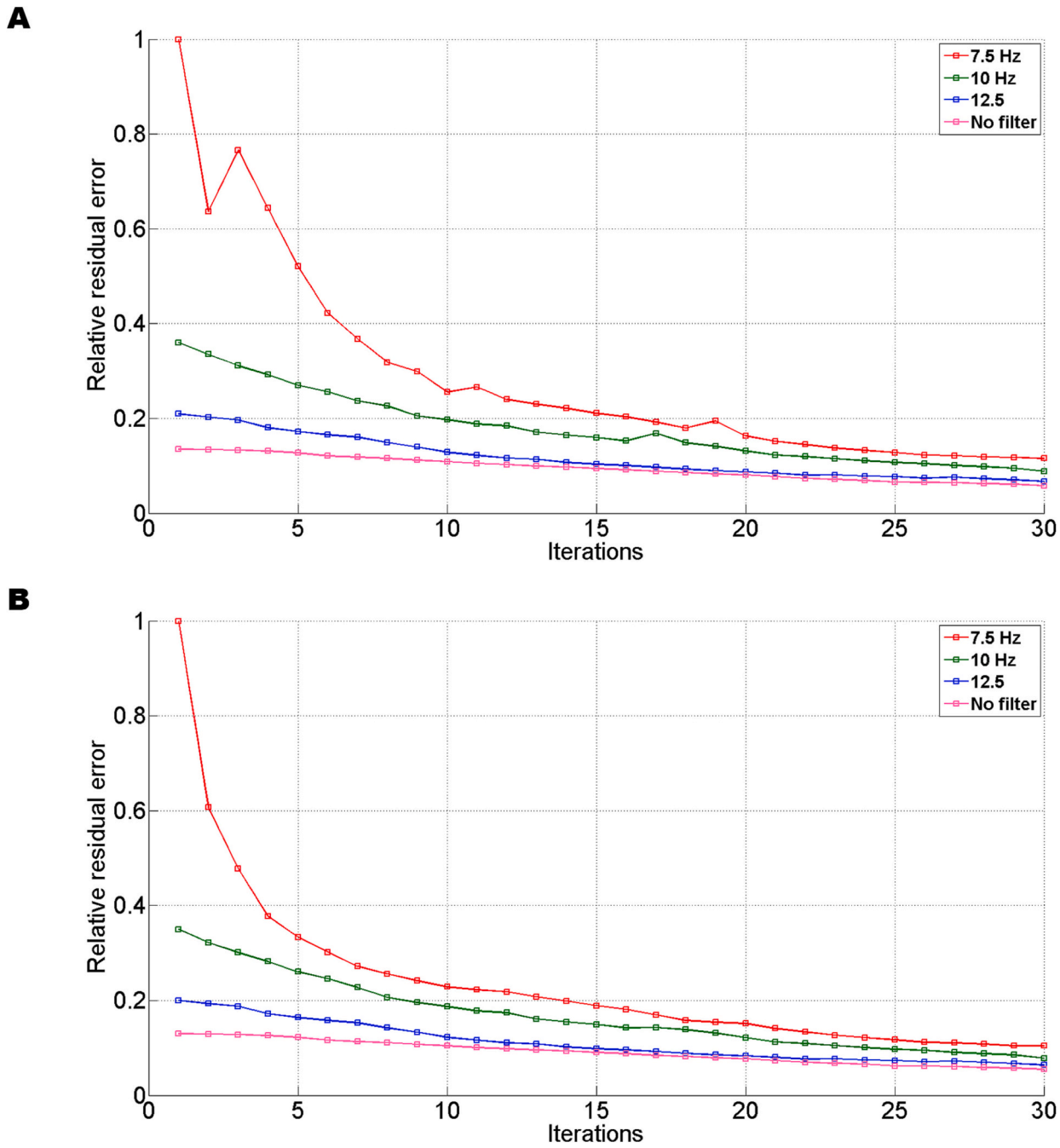

Fig. 7. Variation of the normalized objective function for each iteration and each frequency scale of the experiment that uses the model represented by Fig. 3B (A) using the conventional multiscale FWI and (B) using FWI with our approach.

partially neglected. However, it plays an important role in helping to improve the inversion for other signal attributes (Luo et al., 2020) and for noise suppression and elimination (Zhong et al., 2022).

For the first experiment, we used a Gaussian filter for smoothing 20 points of the model vertically and horizontally (Fig. 3). The smoothed model is the initial guess for performing the first iteration of the conventional multiscale FWI. We then used the same initial condition for performing the FWI with spectral recomposition.

Fig. 4A shows the result for the conventional multiscale FWI, using the smoothed ground truth (Fig. 3B) as an initial guess. Fig. 4C shows the difference between this result in Fig. 4A and the ground truth (Fig. 3A). Fig. 4B presents the result from the FWI with our approach. Fig. 4D shows the difference between the result with our approach (Fig. 4B) and the ground truth (Fig. 3A). We see that spectral recomposition was able to provide a slightly lower error throughout the model in general, but concerning the interfaces, it provided a more significant lower error when compared to the conventional multiscale FWI.

For the second experiment, we simulated a well log performed at position 1.5 Km in the model represented by Fig. 3A. With the values of P-wave velocity extracted from each depth at position 1.5 Km, we made an initial guess built as a horizontally extended velocity profile (Fig. 5).

Fig. 6A presents the result for the conventional multiscale FWI, using the velocity profile at position 1.5 Km extended horizontally (Fig. 5B) as an initial guess. Fig. 4C presents the difference between this result in Fig. 6A and the ground truth (Fig. 3A). Fig. 6C shows the result for FWI with our approach. Fig. 6D shows the difference between the result with our approach (Fig. 6B) and the ground truth (Fig. 3A). Our approach provides a lower error throughout the model in general. However,

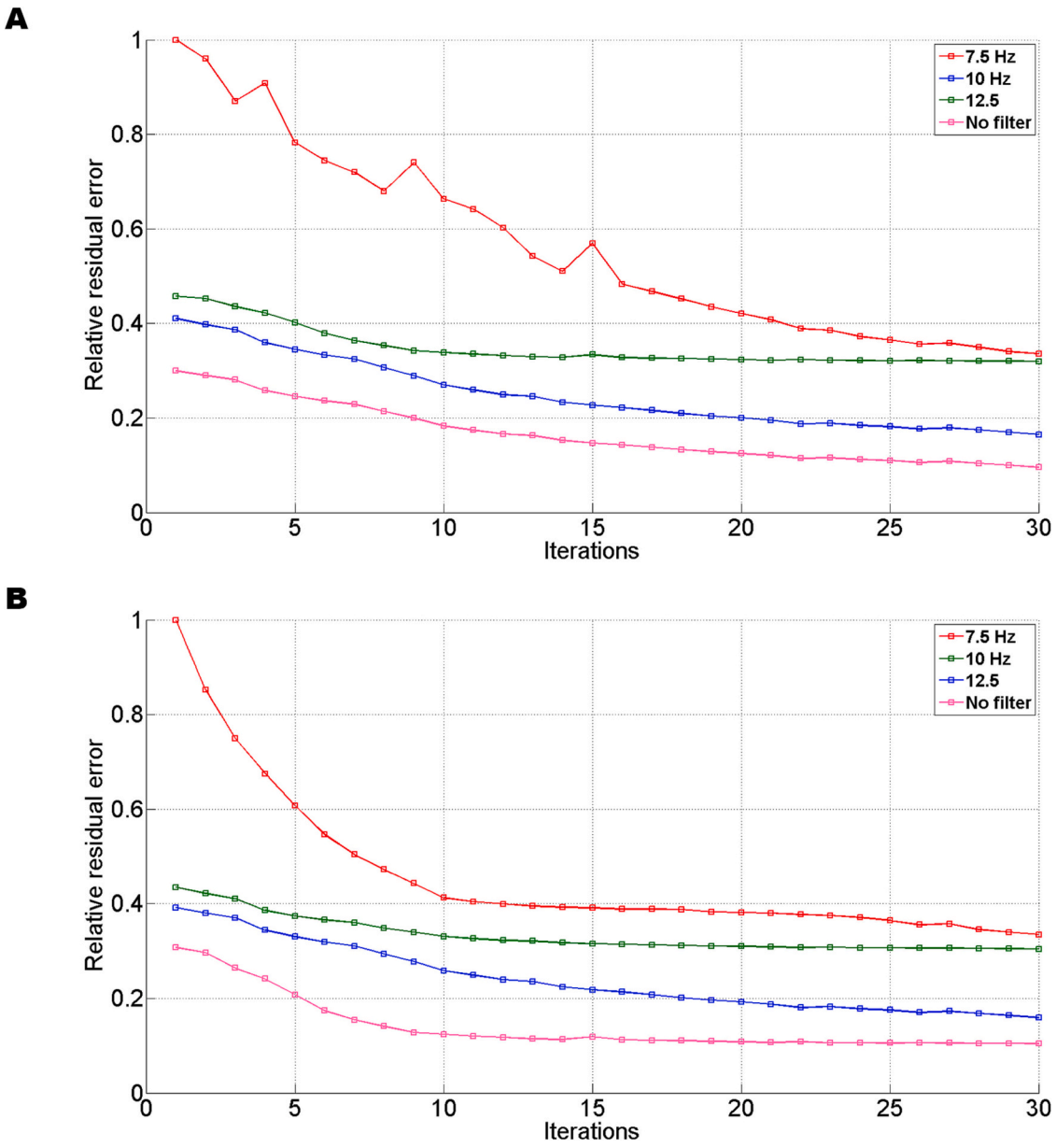

Fig. 8. Variation of the normalized objective function for each iteration and each frequency scale of the experiment that uses the model represented by Fig. 5B (A) using the conventional multiscale FWI and (B) using FWI with our approach.

similarly to our previous results, concerning the interfaces, our approach provided a lower error when compared to the conventional multiscale FWI. For this case, the difference in error between the conventional multiscale FWI and the FWI with our approach is substantial, which shows our approach is able to work with initial guesses that are more befitted to more realistic initial models.

With the results for each initial model with the conventional multiscale FWI and with FWI with our approach, we can now analyze the results for each scale independently, in order to identify the main contributions of our approach.

Fig. 7 shows the variation of the normalized objective function for each of the 30 iterations in each frequency scale (7.5, 10, 12.5, and 15 Hz) of the experiment using the smoothed model shown in Fig. 3B. Fig. 7A presents the results for the conventional multiscale FWI, in which the objective function value is observed to increase for some iterations, especially for the first scale with a cutoff frequency of 7.5 Hz (e.g., for iterations number 3, 9, 11, and 19). Fig. 7B shows the results for the FWI with our approach, which provided a continuously decreasing value of objective function, especially for the cutoff frequency of 7.5 Hz. The results for the other frequency cutoffs (10, 12.5, and 15 Hz) do not show significant variation when our approach is compared to the conventional multiscale FWI. However, when comparing both approaches regarding the first cutoff frequency (7.5 Hz), our approach provides a manner to reach an optimum objective function for the first cutoff with a lower number of iterations.

Fig. 8, similarly to Fig. 7, shows the variation of the normalized objective function for each iteration in each cutoff frequency scale of the experiment using the horizontally extended velocity profile as the initial model shown in Fig. 5B. Fig. 8A shows the results for the conventional multiscale FWI; from which the objective function is observed to increase in iterations 4, 9, 15, and 27 for the first cutoff frequency (7.5 Hz). Fig. 8B shows the results for the FWI with our approach, which, similarly to the results shown in Fig. 7, showed a continuous decrease concerning the value of objective function. This is more visible in 7.5 Hz, the first cutoff frequency. Regarding the other frequency cutoffs, again, they do not show significant variation when the conventional approach is

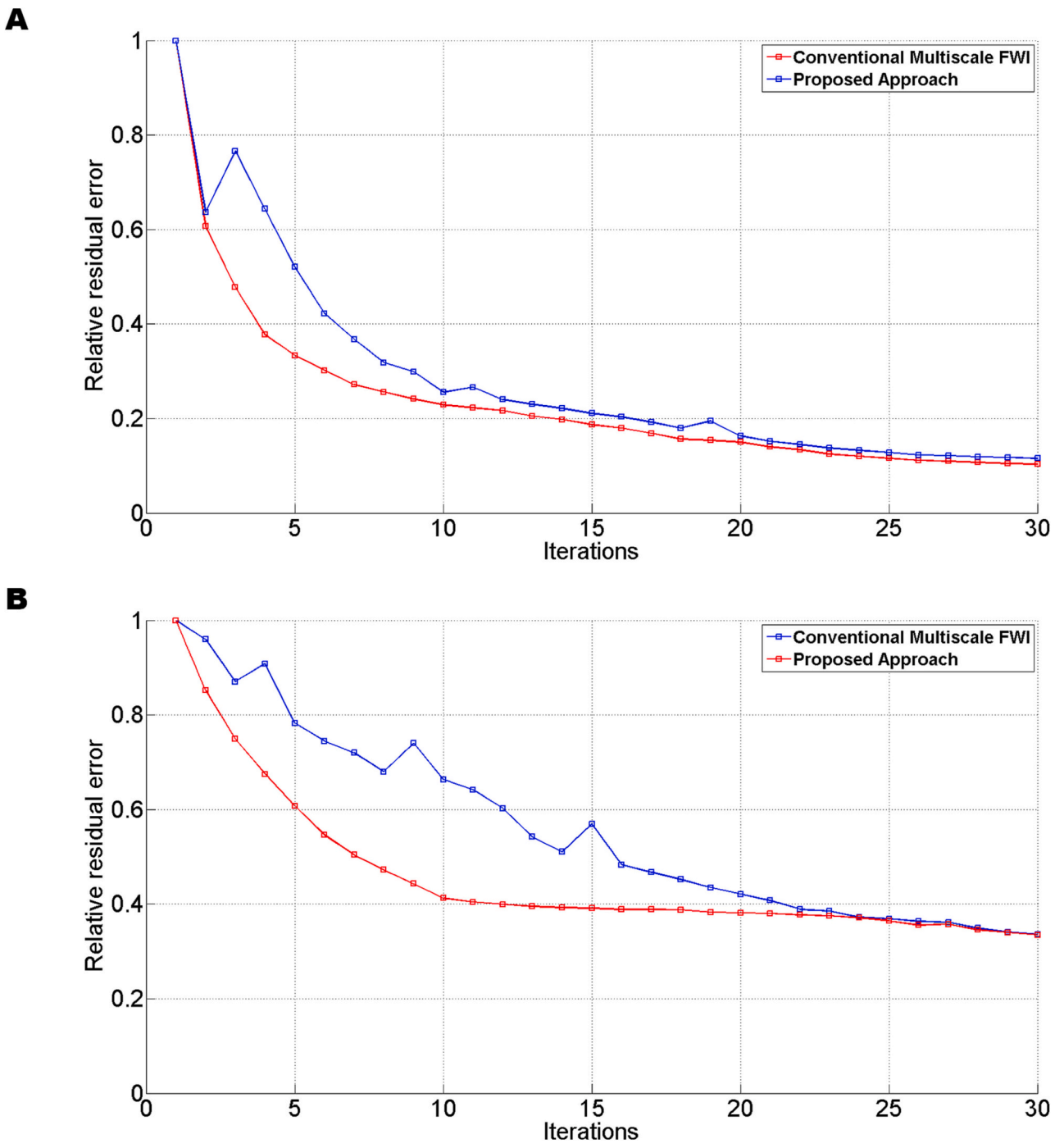

Fig. 9. Variation of the normalized objective function for each iteration of the first frequency scale (i.e., 7.5 Hz) using the conventional multiscale FWI and using FWI with our approach (A) for the experiment that uses the model represented by Fig. 3B and (B) for the experiment that uses the model represented by Fig. 5B.

compared to our approach. Nevertheless, when we compare the conventional multiscale FWI to our approach only for the first cutoff frequency, we notice that our approach reached an optimum value of objective function with a lower number of iterations.

Fig. 9 presents the comparison of the FWI with our approach and the conventional multiscale FWI regarding the variation of the normalized objective function for each iteration of the first frequency scale of 7.5 Hz. Fig. 9A shows the comparison for the experiment that uses the model represented by Fig. 3B, which is the smoothed ground truth. For this case, FWI with spectral recomposition shows to reach the results using a lower number of iterations (around 3 iterations). Fig. 9B compares the experiment with the model represented by Fig. 5B, which is the model built by horizontally extending the velocity profile. From this result, our approach notedly reaches an optimum in the objective function with a significantly lower number of iterations when compared to the conventional approach. For this case, we can consider using 10 fewer iterations by using our approach.

3. Conclusions

We proposed an alternative approach for reducing the number of iterations during an FWI procedure based on guiding the optimization of the iterations in a multiscale technique. This is performed by estimating the position in time of each wavelet by reconstructing the seismic

spectrum using an inversion based on spectral recomposition. We calculated a spectrum that best-fits the observed spectrum. The approach does not require any prior velocity information. Having estimated the position in time for each wavelet, the adapted velocity in each iteration of a multiscale FWI can be better estimated — which allows reducing the number of iterations.

For evaluating the capability of this approach, we tested it on a resized crop from Marmousi II with a water layer added. For seismograms in each iteration, we numerically modelled common-source gather data of P-wave reflections with Devito (Louboutin et al., 2019). We performed two experiments: the first one with the initial guess being a smoothed version of the model; and, the second one being a horizontally extended velocity profile of the model. For our first experiment, we estimated the velocities in the model with our proposed approach being slightly better than the conventional approach. Furthermore, our approach allows using fewer iterations during the first scale in the multiscale procedure for reaching the same result. For the second experiment, the results show that, using our approach, it was possible to estimate the velocity model significantly better at interfaces. Our approach was also able to reach the same results as the conventional approach, but with up to one-third fewer iterations for the first frequency cutoff. This offers a new means to quite accurately estimate the velocity model with significantly fewer iterations. The method requires prior pre-processing to reduce multiple and surface waves. Future improvements to this method should enable handling more complex wavefields and many subsurface layers.

CRediT authorship contribution statement

Nelson Ricardo Coelho Flores Zuniga: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Rafael dos Santos Gioria:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Data curation. **Bruno Souza Carmo:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Data curation.

Declaration of Competing Interest

The authors have no conflicts of interest to declare.

Data availability

The data used is public.

Acknowledgements

The authors gratefully acknowledge support from Shell Brasil through the ANP 20714-2, “Desenvolvimento de técnicas numéricas e software para problemas de inversão com aplicações em processamento sísmico” project at University of São Paulo and the strategic importance of the support given by ANP through the R&D levy regulation. The authors would like to also acknowledge the support of the Research Centre for Greenhouse Gas Innovation (RCGI), hosted by the University of São Paulo (USP) and sponsored by the FAPESP, the São Paulo state Research Foundation (2014/50279-4 and 2020/15230-5). B.S.C. thanks the Brazilian National Council for Scientific and Technological Development (CNPq) for financial support in the form of a productivity grant, number 314221/2021-2.

Appendix A. Appendix

Here, the sequential trust region algorithm is described (Table 1). The first algorithm proposed by Terlaky and Sotirov (2010) forces the convergence to a local minimum region from a remote starting point of a matrix of variables X. The computed optima do not have to satisfy integrality conditions, which allows improving the local optima.

Table 1
The sequential trust-region algorithm pseudocode used to force the convergence to a local minimum from a remote starting point with the computed optima without the need of satisfying integrality conditions.

Sequential trust-region algorithm

Input:
a starting point: X_0;
initial trust region radius: $\Delta_0 = 1$;
input parameters: $r_1 = 0.2$, $r_2 = 1.5$, $c_1 = 0.2$, $c_2 = 0.95$;
begin
$\Delta \leftarrow \Delta_0$;
$\bar{X} \leftarrow X_0$;
while stopping criteria is satisfied
solve (AR_{trust}): return $trust_X$ & $(h_{ij}, \forall_i, j, i \neq j)$;
new candidate point: $x_{ij} = \bar{x}_{ij} + h_{ij}, \forall_i, j, i \neq j$;
compute $det(x_{ij})$;
if $det(x_{ij}) < r_1$
$\Delta \leftarrow \Delta \cdot c_1$;
else
$\bar{X} \leftarrow X$;
end
end
return $\bar{X}$, $trust_X$, $det(x_{ij})$, $det(\bar{X})$;
end

The *input parameters* (r_1, r_2, c_1, and c_2) are used accordingly to the values tested by Terlaky and Sotirov (2010), and it is assumed that the given

matrix B does not need to be positive.

With this, the gradient ascent approach assumes that for a given initial feasible point, the *Sequential trust-region algorithm* returns a matrix $\overline{X}$, which contains elements whose absolute values are different than one. The simplified attractor-repeller semidefinite programming relaxation based on the algorithm proposed by Terlaky and Sotirov (2010) is defined as.

$$(AR)\begin{cases} min\left[-log\left[det(B\circ X) - \sum_{i\neq j} log\left(x_{ij}^2\right)\right]\right] \\ -1 \leq x_{ij} \leq 1, \forall i,j, i \neq j \end{cases} \tag{19}$$

where the relaxing constraints for the element x_{ij} are defined. So, in a given point of the matrix $\overline{X} = \left(\overline{x}_{ij}\right)$, and h_{ij} is the displacement variable in $\overline{x}_{ij}$.

The stop criterion used was when the trust region becomes $\leq 10^{-6}$ or $> 4\Delta_0$.

Applying the steepest ascent algorithm proposed by Terlaky and Sotirov (2010) allows making the absolute value of each off-diagonal element not exceed one. This enables the current point to be feasible. Note that with the application of the element-wise steepest ascent method, the current point from the local minimum that was computed by the *Sequential trust-region algorithm* is moved away.

The adaptative multi-start approach (Table 2) proposed by Terlaky and Sotirov (2010) can find the global minimum by restarting the trust-region algorithm from multiple starting points. The starting point is a matrix, which is random and feasible, or a matrix obtained from a local minimum, such as a matrix from which the elementwise steepest ascent algorithm is applied.

Table 2
The adaptative multi-start procedure algorithm pseudocode used for restarting *Sequential trust-region algorithm* from a new randomly automated picked starting point, in case of the local minimum is an integer or there are no improvements in the neighborhood of the nonintegral local minimum.

Adaptative multi-start procedure algorithm

Input:
data matrix: B;
$\beta = 0.9$;
$k \leftarrow 1$;
repeat for a given number of starting points
a random starting point X_0
start **Sequential trust-region algorithm** from X_0: $return\ \overline{X}$, $det\overline{X}$;
$br \leftarrow 1$;
if $\overline{X}$ is not integral
while $br \leq$ the prescribed number of local searches;
do steepest ascent for the non-integer elements till stopping criteria is satisfied
$\tilde{x}_{ij} \leftarrow \tilde{x}_{ij} + \beta \cdot inv\left(B \circ \tilde{X}\right)_{ij} \cdot det(B)B_{ij}$;
end
start **Sequential trust-region algorithm** from $\tilde{X}$: $return\overline{X}$, $det(\overline{X})$
$br \leftarrow br + 1$;
end
end
if $\overline{X} \in \{-1,1\}^n$ & $det(B \circ X) > 0$, **then** $X_k \leftarrow \overline{X}$; $k \leftarrow k + 1$;
end
end
$X^* = \min_k\{-log[det(B \circ X_k)]\}$;

Since $\left(\boldsymbol{B}\circ\widetilde{\boldsymbol{X}}\right)$ has an eigenvalue that is significantly close to zero, it is necessary to backtrack from the *Sequential trust-region algorithm* to a point that is in the interior of the semidefinite region. It is necessary to apply the steep ascent method to non-integer off-diagonal matrix elements of $\widetilde{\boldsymbol{X}}$ while holding the integer elements fixed. With this, the *Adaptative multi-start procedure algorithm* can provide new randomly starting points when the local minimum is an integer or when there are no improvements in the neighborhood of the nonintegral local minimum.

References

Asnaashari, A., Brossier, S., Garambois, F., Thorne, P., Virieux, J., 2013. Regularized seismic full waveform inversion with prior model information. Geophysics 72 (2), R25–R36.

Balestrini, F., Draganov, D., Malehmir, A., Marsden, P., Ghose, R., 2020. Improved target illumination at Ludvika mines of Sweden through seismic-interferometric noise reduction. Geophys. Prospect. 68, 200–213.

Brittan, J., Bai, J., Delome, H., Wang, C., Yingst, D., 2013. Full waveform inversion – the state of the art. First Break 31 (10), 75–81.

Cai, Y., Fomel, S., Zeng, H., 2013. Automated spectral recomposition with application in stratigraphic interpretation. Interpretation 1 (1), 109–116.

Castagna, J.P., Sun, S., Siegfried, R.W., 2003. Instantaneous spectral analysis: Detection of low-frequency shadows associated with hydrocarbons. Lead. Edge 22, 120–127.

Fomel, S., 2013. Seismic data decomposition into spectral components using regularized nonstationary autoregression. Geophysics 78 (6), O69–O76.

Ghose, R., Goudswaard, J., 2004. Integrating S-wave seismic-reflection data and cone penetration test data using a multiangle multiscale approach. Geophysics 69, 440–459.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In: Proceedings of the Royal Society A, 454 (1971).

Jones, I.F., 2010. An introduction to: Velocity model building. EAGE 296.

Li, Y., Zheng, X., Zhang, Y., 2011. High-frequency anomalies in carbonate reservoir characterization using spectral decomposition. Geophysics 76, V47–V57.
Louboutin, M., Lange, M., Luporini, F., Kukreja, N., Witte, P.A., Herrmann, F.J., Velesko, P., Gorman, G.J., 2019. Devito (v3.1.0): an embedded domain-specific language for finite differences and geophysical exploration. Geosci. Model Dev. 12 (3), 1165–1187.
Luo, C., Ba, J., Carcione, J.M., Huang, G., Guo, Q., 2020. Joint PP and PS pre-stack AVA inversion for VTI medium based on the exact Graebner equation. J. Pet. Sci. Eng. 194, 107416.
Mitrofanov, G.M., Priimenko, V.I., 2013. Prony filtering of sesmic data: Mathematical and physical modelling. Rev. Brasil. Geofís. 31 (1), 151–168.
Mitrofanov, G.M., Priimenko, V.I., 2015. Prony filtering of sesmic data. Acta Geophys. 63 (3), 652–678.
Moore, E.H., 1920. On the reciprocal of the general algebraic matrix. Bull. Am. Math. Soc. 26 (9), 385–396.
Pan, X., Li, L., Zhang, G., 2020. Multiscale frequency-domain seismic inversion for fracture weakness. J. Pet. Sci. Eng. 195, 107845.
Penrose, R., 1955. A generalized inverse for matrices. Math. Proc. Camb. Philos. Soc. 51 (3), 406–413.
Ricker, N., 1953. The form and laws off propagation of seismic wavelets. Geophysics 18, 10–40.
Terlaky, T., Sotirov, R., 2010. Multi-start approach for an integer determinant maximization problem. Optimization 61, 101–114.
Tomasso, M., Bouroullec, R., Pyles, D.R., 2010. The use of spectral recomposition in tailored forward seismic modeling of outcrop analogs. AAPG Bull. 94 (4), 457–474.
Verschuur, D.J., Berkhout, A.J., Wapenaar, C.P.A., 1992. Adaptive surface-related multiple elimination. Geophysics 57 (9), 1166–1177.
Virieux, J., Asnaashari, A., Brossier, R., Métivier, L., Ribodetti, A., Zhou, W., 2017. An introduction to full waveform inversion. Society of. Explor. Geophys. 40.
Wood, L.C., 1974. Seismic data compression methods. Geophysics 39 (4), 389–581.
Yang, P., Brossier, R., Virieux, J., 2016. Wavefield reconstruction by interpolating significantly decimated boundaries. Geophysics 81 (5), T197–T209.
Zhong, Y., Gu, H., Liu, Y., Mao, Q., 2022. A new joint reverse time migration method to improve vertical seismic profile image quality. J. Pet. Sci. Eng. 214, 110546.
Zuniga, N.R.C.F., Priimenko, V.I., 2022. Automated Travel-Time Picking using Spectral Recomposition. Rev. Brasil. Geofís. 39 (3), 1–18.