WS05-D03

An Auxiliary Bump Functional to Overcome Cycle Skipping in Waveform Inversion

P. Bharadwaj* (Delft University of Technology), W.A. Mulder (Shell GSI BV & Delft University of Technology) & G.G. Drijkoningen (Delft University of Technology)

SUMMARY

To overcome the local minima problem in FWI, we propose to use an auxiliary data-domain objective function during inversion. It reduces the data to a simpler form by squaring, followed by blurring to ensure that events that are too far apart can still interact during the inversion. As it effectively replaces seismic arrivals by bumps, we call it the bump functional. This objective function is less sensitive to cycle skipping. Its rôle is to guide the inversion towards the global minimum by pulling the trapped solution out of the local minima associated with the least-squares functional whenever necessary.

Waveform inversion cannot be performed with only the auxiliary objective function because it is insensitive to the polarity of the arrivals and the source signature. Therefore, we alternate between minimization with this and the classic least-squares functional. We confirm the validity of the approach using a simple numerical example with reflection data.

Introduction

Full waveform inversion (FWI) is a non-linear optimization procedure that estimates the some of the Earth's parameters by least-squares fitting of the recorded arrivals in the seismic data (Tarantola, 1984; Virieux and Operto, 2009; Fichtner, 2010). Due to the computational cost of the wave-equation modelling, the optimization is usually performed with gradient-based techniques. While fitting the observed and the modelled seismic arrivals with the conventional least-squares objective function, the gradient-based optimization gets trapped in the nearest local minimum when the error in the arrival time exceeds about half a period of the signal. Here, period refers to the dominant frequency of the data. In other words, the least-squares inversion cannot reconstruct velocity anomalies that cause shifts in the arrival times larger than half a period. Usually, the velocity anomalies accounting for the arrival times have relatively low wave-numbers. Hence, inversion of the low-frequency seismic signals with larger periods is easier and they help in the reconstruction of a kinematically correct velocity model.

Many authors have formulated alternative data-domain functionals to achieve global convergence in the absence of sufficiently low frequencies. These can be classified into the following two major classes:

1. Functionals that give more weight to the kinematic error than to the amplitude error between the seismic waveforms. They often involve the cross-correlation between the observed and modelled waveforms (Luo and Schuster, 1991; van Leeuwen and Mulder, 2010; van Leeuwen, 2010). After cross-correlation the arrival-time error can be picked by hand, which is tedious. More automatic methods involve proper normalization of the waveforms prior to cross-correlation in order to ensure that the energy at non-zero lags quantifies the arrival-time error. These automatic functionals, however, suffer from cross-talk between multiple arrivals. Hence, the waveforms are assumed to have only single arrivals or strong first arrivals and their primary application lies in crosswell data and tomographic inversion.
2. Functionals that aim to match the data in some reduced form that avoids strong non-linearity. These functionals can handle multiple arrivals in the data and do suffer less or not at all from cycle skipping. They have a basin of attraction with a larger size than that of the least-squares functional. An example is the functional that measures the misfit between the envelopes of the observed and the modelled waveforms (Bozdağ et al., 2011). Unfortunately, the reduction of the data may lead to non-uniqueness in the reconstruction of the Earth model. Also, the envelope misfit may perform worse than a correlation-based functional, even in the case of a single arrival, because convergence with a gradient-based method can only be obtained if the envelopes partially overlap with each other. In other words, events should not be separated in time by roughly more than the dominant period.

In this abstract, we propose a new data-domain objective function that uses the data in a reduced form. It allows for a time separation between events larger than just a period by applying additional smoothing or blurring to the reduced data. In our case, the reduction consists in simple squaring of the data, which removes polarity information from the data but does not simplify them as much as replacement by the envelope. In the extreme case, an indicator function would only keep information about the support of seismic events. Squaring maintains amplitude information and keeps the gradient computation simple. This nonlinear operation creates artificial low frequencies in the data. In that respect, it bears some similarity to other methods such as spiking deconvolution, travel-time picking, and a differential-semblance variant that recovers missing low frequencies (Sun and Symes, 2012). The rôle of the blurring is to increase the size of the basin of attraction of the objective function. It enables isolated pulses in the observed and modelled data to *sense* each other, even if they have no overlap before blurring. As the evaluation of the functional does not involve cross-correlation, it should be applicable to data containing multiple arrivals. Therefore, this new functional does not suffer from the limitations of the two classes discussed above. The convergence robustness of the functional should be comparable to that

of the envelope-based misfit. Note that inversion with only the proposed objective function based on squared data suffers from non-uniqueness and lack of resolution. Therefore, we also have to use it in combination with the classic least-squares functional to achieve global convergence in the absence of low frequencies.

The bump functional

The classic least-squares inversion is performed by minimizing the functional (Tarantola, 1984),

$$J^{LS} = \frac{1}{2}\sum_s\sum_r\sum_t \left[p(\mathbf{x}_r,\mathbf{x}_s,t) - q(\mathbf{x}_r,\mathbf{x}_s,t)\right]^2. \tag{1}$$

Here, $q(\mathbf{x}_r,\mathbf{x}_s,t)$ is the observed pressure for a source at position $\mathbf{x}_s$ and receiver at $\mathbf{x}_r$ as a function of time t. The modelled data are denoted by $p(\mathbf{x}_r,\mathbf{x}_s,t)$. For the examples further on, we will consider the 2-D constant-density acoustic wave equation

$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t} - \Delta p = w(t)\delta(\mathbf{x}-\mathbf{x}_s),$$

with pressure $p(\mathbf{x},t)$, sound speed $c(\mathbf{x})$, source signature $w(t)$ and Laplace operator Δ. In 2D, $\mathbf{x} = (x,z)$.

Computation of the bump functional starts with squaring the modelled and observed data. This removes some phase information but artificially creates low frequencies, similar to replacement with the envelope or a spiking deconvolution. After squaring, a single pulse in the modelled data will still not be able to sense a corresponding pulse in the observed data if they are separated. We therefore smooth or blur both until some overlap is created, which then allows for convergence. Squaring and blurring results in *bumpy* data. The functional is given by

$$J^B = \frac{1}{2}\sum_s\sum_r\sum_t \left\{b(t,\sigma) *_t \left[p^2(\mathbf{x}_r,\mathbf{x}_s,t) - q^2(\mathbf{x}_r,\mathbf{x}_s,t)\right]\right\}^2. \tag{2}$$

Here, $b(t,\sigma)$ is a blurring function, for which we choose a Gaussian with standard deviation σ, and $*_t$ denotes convolution in time. The parameter σ controls the amount of blurring applied to the squared data. For example, choosing $\sigma = T$, where $T = 1/f_{\text{peak}}$ is the period of the data corresponding to the dominant frequency f_{peak}, would blur an impulsive arrival in the squared data such that it is spread roughly over $2T$. Blurring reduces the detail in the squared data and increases the convergence robustness of J^B in comparison to the envelope-based misfit (Bozdağ et al., 2011). The size of the basin of attraction depends on the parameter σ. A larger σ, however, decreases the resolution.

Numerical example

In this section, we show an example demonstrating the use of the bump functional for full waveform inversion. The example aims to reconstruct the subsurface velocity distribution using reflected arrivals. For simplicity, we assume the source wavelet to be known. Also, the surface located at $z = 0$ is assumed to be non-reflecting. We use a time-domain staggered-grid finite-difference code for the constant-density acoustic wave equation to model the data and to perform the adjoint wavefield computations required for calculating the gradients (Tarantola, 1984; Fichtner, 2010). The optimization was performed by the conjugate-gradient method. Both the background velocity and the position of the reflectors in the model have to be estimated. The offset-based moveout in the data contains the information needed to determine the background velocity. We show that the J^B functional cannot update the background velocity of the model by itself. It only changes the reflectivity so that the observed and the modelled data match in their bumpy form. This is due to non-uniqueness: different background velocity models can produce data that match the observed data after squaring and blurring. However, matching the bumpy data is enough to pull the solution out of the local minimum where it was trapped with the least-squares functional. This is the rôle of the bump objective function in the inversion.

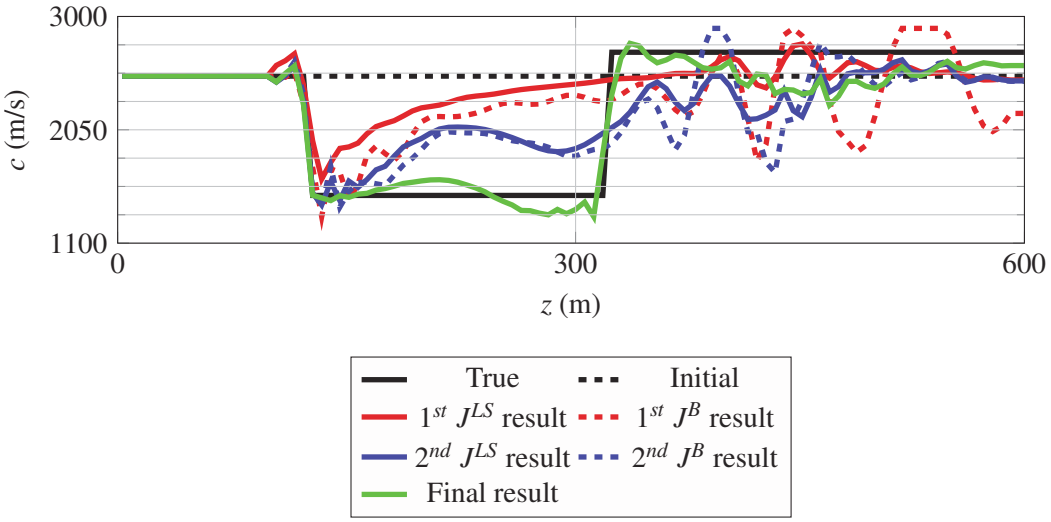

Figure 1: A reflection experiment using an assumed model with a negative velocity anomaly. The initial model (dashed black) for inversion is also plotted. The models after the first J^{LS} inversion (solid) and the first J^B inversion (dashed) are plotted in red. The models after the second alternating loop are plotted in blue. The final result is plotted in green.

We consider an assumed Earth models of 1500 m width and 600 m depth with a negative velocity anomaly. The velocity is assumed to be only varying with depth. We evenly placed a horizontal array of 20 sources at a depth of 20 m and 100 receivers at a depth of 30 m. We generated the data using a 17-Hz Ricker source wavelet, after applying a 10-Hz low-cut filter. The actual model, plotted in Figure 1, contains two reflectors at $z = 120$ m and $z = 320$ m. The velocity anomaly of the second layer is approximately 40%. The first reflector has a negative velocity contrast, causing head-wave arrivals to be absent from the data (plotted in Figure 2a). The initial model, plotted in Figure 1, for inversion is a homogeneous model with $c = 2500$ m/s. We now perform the LS inversion. It fails to reconstruct the background velocity of the model and the convergence stagnates. As seen in the Figure 1, after updating the first-layer reflectivity, the LS inversion updates the reflectivity of the model at depths between 400 and 450 m. This causes the modelled data, plotted in Figure 2c, to partially match the 'observed' data, primarily at the short offsets. We also see that these reflectors, positioned at the wrong depths, generate arrivals in the data that are cycle-skipped at larger offsets. The cycle skipping indicates that the solution is caught in a local minimum. In order to see if the bump functional can update the background velocity, we use the output from 30 iterations of LS inversion as an initial velocity model for J^B inversion with $\sigma = 0.8T$. Both the initial velocity model and the output of the J^B inversion are plotted in red in Figure 1. It can be seen that the J^B inversion has not updated the background velocity but only boosted the reflectivity of the model at depths between 400 and 550 m. Note that the velocity model that explains the bumpy data, plotted in Figure 2b, does not have to be unique. Next, we take the velocity model output from the J^B inversion as input to the J^{LS} inversion. Now, the J^{LS} inversion updates the background velocity. The output velocity model after the second J^{LS} inversion is plotted in Figure 1. We performed a second J^B inversion and noticed that the final modelled shot gather better explains the observed data.

In order to reach the global minimum, we propose the following inversion strategy. The velocity model output of the J^{LS} inversion always acts as an input to the J^B inversion and vice-versa. The iterations are stopped when convergence stagnates. In the case of J^{LS} inversion, the stagnation indicates that a local minimum has been reached. The J^B iterations stop when the inversion finds one of the many non-unique solutions that explain the observed data in their bumpy form. The final output velocity model after alternating between the J^{LS} and J^B functionals 9 times is plotted in also Figure 1. The modelled data corresponding to the final output, plotted in Figure 2d, explain the observed data.

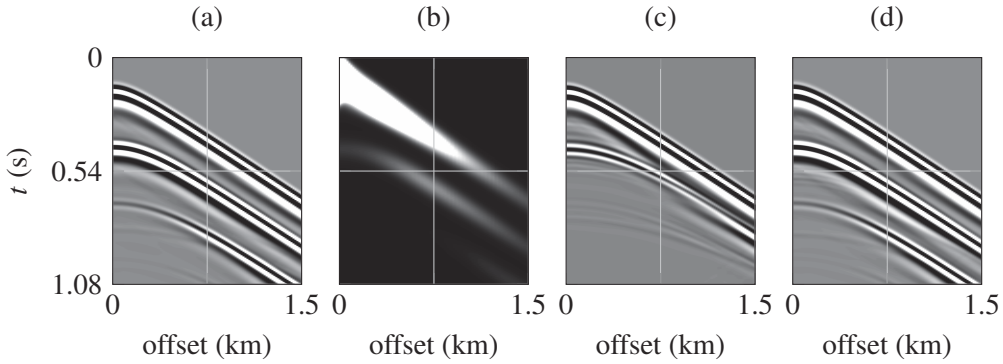

Figure 2: Data panels corresponding to the reflection experiment in Figure 1 for a source at $(0, 20)$. (a) Observed shot gather. (b) Bumpy data for J^B inversion. (c) Modelled shot gather after J^{LS} iterations, which is cycle skipped. (d) Modelled shot gather after inversion using both J^{LS} and J^B functionals.

Conclusions

We formulated a data-domain functional that is less sensitive to cycle skipping and takes multiple arrivals in the data into account. This functional matches the observed and the modelled data in a reduced form to avoid cycle skipping. The reduction results in bumpy data, which are data after squaring and blurring. Blurring or smoothing with a Gaussian is applied to the squared data to increase the size of the basin of attraction that corresponds to the functional. However, we showed that the proposed functional cannot be used to update the background velocity using only the reflected arrivals, as the models that generate given bumpy data can be non-unique. Hence we proposed an inversion scheme, where the proposed functional is used as an auxiliary functional. It pulls the trapped solution out of the local minimum whenever necessary. We have numerically validated this inversion scheme using a simple reflection experiment.

Acknowledgements

This work is a part of the NeTTUN project, funded from the European Commissions Seventh Framework Programme for Research, Technological Development and Demonstration (FP7 2007-2013) under Grant Agreement 280712. The computations were carried out on the Dutch national e-infrastructure with the support of the SURF Foundation (www.surfsara.nl).

References

Bozdağ, E., Trampert, J. and Tromp, J. [2011] Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, **185**(2), 845–870.

Fichtner, A. [2010] *Full seismic waveform modelling and inversion*. Springer.

Luo, Y. and Schuster, G. [1991] Wave-equation traveltime inversion. *Geophysics*, **56**(3), 645–653, doi: 10.1190/1.1443081.

Sun, D. and Symes, W.W. [2012] *Waveform Inversion via Nonlinear Differential Semblance Optimization*, chap. 497. 1–7, doi:10.1190/segam2012-1190.1.

Tarantola, A. [1984] Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259–1266, doi:10.1190/1.1441754.

van Leeuwen, T. [2010] *PhD thesis, Delft University of Technology.*

van Leeuwen, T. and Mulder, W.A. [2010] A correlation-based misfit criterion for wave-equation traveltime tomography. *Geophysical Journal International*, **182**(3), 1383–1394, doi:10.1111/j.1365-246X.2010.04681.x.

Virieux, J. and Operto, S. [2009] An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74**(6), WCC1–WCC26, doi:10.1190/1.3238367.