Introduction

Full waveform inversion is a process of estimating the velocity of the earth from measured wave-field data. The method involves a forward modeling in time or frequency, adjusting the velocity field by some optimization algorithm in order to fit the measured data. Although the method has been investigated in the 80's (Tarantola, 1987) it has regained popularity in the last decade, when both high quality data, computing power and algorithms that can handle large volume of data allow to better tackle the problem (Pratt, 1999; Epanomeritakis et al., 2008; Krebs et al., 2009; Biondi and Almomin, 2014). Many different algorithms have been proposed for the solution of the problem, however, it remained difficult to solve in general.

Most algorithms start by solving the problem for low frequencies obtaining a spatially smooth model that contains mainly low spatial frequencies. A frequency continuation is then employed to obtain higher resolution. It is fairly well known that initializing the solution with low frequencies is crucial if we are to converge to the global minima (Pratt, 1999). However, in many realistic scenarios this is not possible, due to the lack of low frequency components in the data. To mathematically explain the difficulty, assume that the forward problem in frequency is $F(m,\omega)$ and the measured data is $d(\omega)$. Assume that we measure the distance between the computed data and the observed data by some misfit functional $S(F(m,\omega),d(\omega))$. For example, S can be a least square function or a more complicated one that is based on phase shift (Virieux and Operto, 2009) or a Wiener transform (Warner et al., 2013). FWI algorithms attempt to minimize S (usually, with additional regularization), that is, find a model m, such that the misfit function, S, is small. Any gradient-based method uses the gradient of the misfit with respect to the model in order to compute a step (either directly or scaled by an approximation of the Hessian). It is straight-forward to see that the gradient is simply, $g = J^T \partial S$ where J is the sensitivity matrix, $J = \partial_m F(m)$. For problems where no low frequency is recorded, the singular vectors of J do not contain low spatial frequencies. As a result, the gradient does not contain these low frequencies, and this is independent of the choice of S. The lack of low frequency measurements in the data can thus lead to non-geological models and to models that yield a local minima of the misfit function. In this paper we discuss a method to overcome this problem by obtaining low spatial frequencies from high frequency data. Our method is based on the extraction of travel time and placing it instead of the low frequency data, using the eikonal equation as a forward modeling equation. This yields two different inverse problems to be solved. The first one, known as Travel Time Tomography is based on the eikonal equation (Li et al., 2013; Benaichouche et al., 2015), and the second one is the waveform inversion which is based on the Helmholtz equation—the wave equation in Fourier domain. We then develop a joint inversion algorithm that incorporates both problems within a single computational framework. We show that our approach can overcome cases where the data does not contain low frequencies and yield models that are faithful to both eikonal and Helmholtz equations and their associated data samples.

Method and/or Theory

We consider the forward problem that is given by the Helmholtz equation for a constant density media

$$\nabla^2 u + m\omega^2 u = \delta(x - x_s). \tag{1}$$

Here, u is the wavefield, m is the model for the squared slowness, and ω is the frequency. The equation is given with some absorbing boundary conditions that mimics the propagation of a wave in an open domain. The source is assumed to be a delta function that is located at x_s. We consider the data, d,

$$d(\omega, x_r, x_s) = (p_r, u(\omega, x_s)) + \varepsilon \tag{2}$$

where p_r is a sampling operator that measures the field u at location x_r and $(\cdot,\cdot)$ is an inner product. The data is typically noisy and we assume that the noise, ε, is iid, Gaussian and with standard deviation σ^2. Given data that is collected in a number of receiver locations and a spectrum of frequencies we aim to estimate the model, m. Using a penalized least squares approach, this is done by solving the optimization problem

$$\min_{m_L \le m \le m_H} \mathscr{J}(m) \quad = \quad \frac{1}{2}\sum_{i,j,k} (d(\omega_i, x_{r_j}, x_{s_k}) - (p_{r_j}, u_{ik}))^2 + \alpha R(m) \tag{3}$$

$$\text{s.t} \qquad \nabla^2 u_{ik} + m\omega_i^2 u_{ik} = \delta(x - x_{s_k})$$

Here, $R(m)$ is a regularization term (we use either smoothness or total variation (Vogel, 2001)) and the bounds $m_L > 0$ and $m_H > 0$ are bounds that keep the model physical. To solve the optimization problem a variety of methods are typically used. First order methods such as nonlinear conjugate gradient and limited memory BFGS (Nocedal and Wright, 1999) has the advantage of low memory but converge slowly. Our method of choice is the Gauss Newton method (Pratt et al., 1998), that takes some curvature information and converge faster, especially if the noise level is low.

Solving the problem for all frequencies at once typically yields local minima. Therefore, it is common to solve the problem by frequency continuation, solving first the lowest frequency obtaining a model $m(\omega_1)$ and then, solving again the problem for 2 frequencies, starting from $m(\omega_1)$, and continuing forward by adding more frequencies, each time starting from the previous solution. This yields a stable process that converges to the global minima of the objective function.

Nonetheless, in the absent of low frequencies this process cannot be used, as convergence to local minima is often observed. To alleviate this problem we propose a different process. It is fairly well known, that assuming that the wave field has a solution of the form $u = a(x,\omega)\exp(i\omega T(x))$, substituting it into the the Helmholtz equation (1) we obtain the eikonal equation for high frequencies

$$|\nabla T|^2 - m = 0, \qquad T(x_s) = 0. \tag{4}$$

This equation models the first arrivals of the waves. Since the travel time is an integral of the model over the ray path, its Jacobian with respect to the model contains mainly low frequencies. Thus, a way to overcome the lack of low frequencies in the data d, is to extract it from the travel time T. To solve (4) we use the Fast Marching method in (Treister and Haber, pers. comm.) which is also useful for the sensitivity calculation.

To this end, consider the extraction of travel time, T from the *high frequency data* $u(\omega)$. Although this can be done manually, it can be automated (Saragiotis et al., 2013). Consider also the new data,

$$d_{\mathsf{T}}(x_{r_j}, x_{s_k}) = (p_j, T_k),$$

where as before, p_j is the j-th receiver and T_k is the field from the k-th transmitter. Using the new data we can now solve the optimization problem,

$$\begin{aligned}
\min_{m_L \le m \le m_H} \mathscr{J}(m) \quad &= \quad \frac{w_{\mathsf{fwi}}}{2}\sum_{i,j,k}(d(\omega_i, x_{r_j}, x_{s_k}) - (p_{r_j}, u_{ik}))^2 + \frac{w_{\mathsf{eik}}}{2}\sum_{j,k}(d_{\mathsf{T}}(x_{r_j}, x_{s_k}) - (p_{r_j}, T_k))^2 \qquad (5)\\
&+ \quad \alpha R(m)\\
\text{s.t} \qquad & \nabla^2 u_{ik} + m\omega_i^2 u_{ik} = \delta(x - x_{s_k})\\
& |\nabla T_k|^2 - m = 0, \qquad T_k(x_{s_k}) = 0.
\end{aligned} \tag{6}$$

Here, w_{fwi} and w_{eik} are the inverse standard deviation squared of the fields and travel time data.

There are two novel points in the optimization problem (5). First, assuming that the data does not contain low frequencies, the eikonal equation substitutes the equation for low frequencies. Second, the resulting model satisfies *both* the travel time equations and the full waveform equations, thus, it is consistent to different physical interpretation of the wavefield.

To solve the optimization problem we now use a process that is a kin to the frequency continuation previously discussed. We first solve the optimization problem with the eikonal equation alone obtaining a model, m_{eik}. This model is used to initialized an inversion with *both* the eikonal and lowest frequency data, obtaining a new model. We then proceed by adding frequencies to the process, obtaining a final model that fits both eikonal and full waveform data *and* contains low spatial frequencies. We note that computationally, solving the eikonal equation (4) and obtaining its sensitivities is trivial compared to the corresponding operations for the Helmholtz equation (1). Therefore, the additional computational cost that is involved in solving (5) instead of (3) is negligible.

Examples

To demonstrate the effectiveness of our method we use the SEG salt model plotted in figure 1. The model is discretized to 512×140 cells of $15.6m^2$ each. 30 equally spaced sources (located every 60.2m) and 140 receivers are placed on the top of the model, yielding 4200 data measurements for each frequency and 4200 travel times.

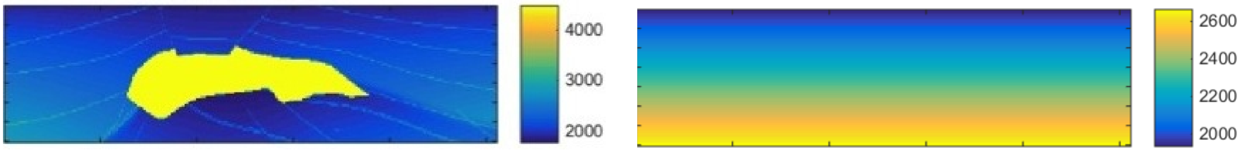


Figure 1 *The salt model used to test the inversion and the starting model*

We conducted three different experiments. In the first, we use frequencies that range from $0.75 - 14Hz$. In the second, we do not use the first two frequencies and use the range from $2.5 - 14Hz$, thus assuming that no low frequencies are recorded. In this third experiment we use the high frequency range $2.5 - 14Hz$ but add the travel time data to replace the low frequency component, jointly inverting the travel time and the full waveform data. All inversions start from a gradient velocity model in Figure 1. We use the inexact projected Gauss-Newton method for the solution of each optimization problem. At each iteration we solve the linear system using the Conjugate Gradient algorithm with a tolerance of 10^{-1}. Convergence was declared when the misfit was less than 1% (which is the noise level).

Results

The results of the 3 different experiments are presented in Figure 2. The left column of the figure shows the inversion result after the first iteration (using the lowest frequency or the eikonal). The right column shows the final result of the inversion. When low frequency is available, (first row) a blurred representation of the model is obtained in the first iteration. This blurred version is then sharpened when more frequencies are added. When the low frequencies are missing, the lowest frequency yields high frequency features in the initial model. Since high frequencies do not contain low frequency content, the inversion does not recover and the results are far from the true model. Finally, when the eikonal is used to replace the low frequency data, a low frequency representation of the model is also obtained. Even though this representation is incorrect at depth, the additional frequencies manage to overcome this and the final result is equivalent to the result obtained by using the low frequency data.

Conclusions

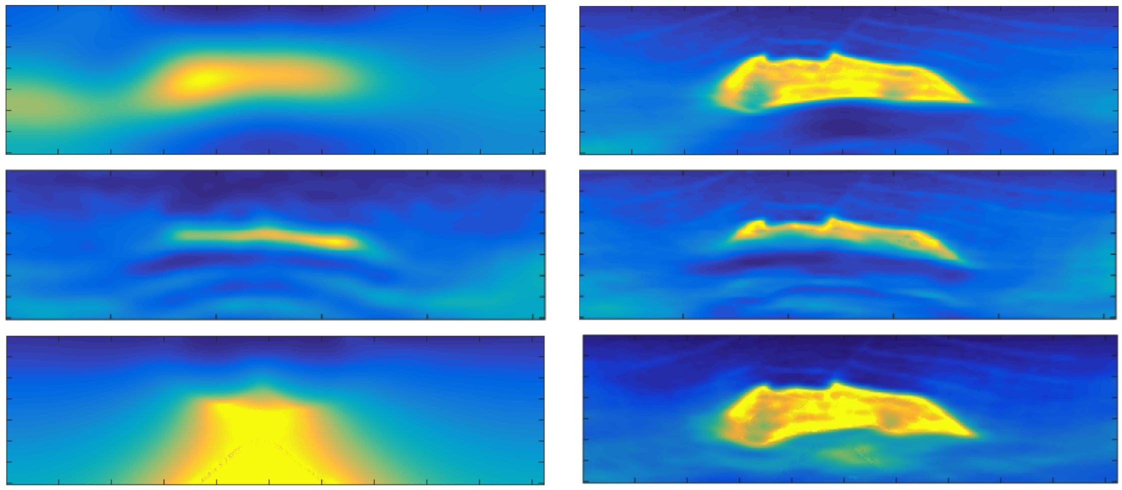
In this work we have explored a methodology that aids full waveform inversion to converge to the global minimum in the absence of low frequency data. The method is based on the extraction of travel time from high frequency data and using the eikonal equation in order to model the travel time. Next, the travel time data is used instead of the low frequency data and jointly inverted with the rest of the data using a frequency continuation process. Since the travel time inversion is sensitive to low spacial modes in the model, it yields an initial model that enable the recovery of a good approximation to the true model. Furthermore, since we jointly invert the full waveform and travel time data, our final model is consistant for both physical models.

While our method seems to be robust in the presence of noise, it has two main limitations. First, long offset data must be recorded in order to have a meaningful first arrival inversion. Second, our approach requires travel time picking. While doing so with marine data is relatively simple, it can be a much more involved process for land base data.

a

b

c



First Iteration Final Inversion

Figure 2 *First iteration and inversion of (a) All frequencies (low and high), (b) with 2 first frequencies missing and (c) with the eikonal equation replacing the low frequencies.*

Acknowledgements

This work was supported by Kyma Geoscience and MITACS.
The research leading to these results has received funding from the European Union's - Seventh Framework Programme (FP7/2007-2013) under grant agreement no 623212 - MC Multiscale Inversion.

References

Benaichouche, A., Noble, M. and Gesret, A. [2015] First Arrival Traveltime Tomography Using the Fast Marching Method and the Adjoint State Technique. In: *77th EAGE Conference Proceedings*.

Biondi, B. and Almomin, A. [2014] Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion. *Geophysics*, **79**(3), WA129–WA140.

Epanomeritakis, I., Akcelik, V., Ghattas, O. and Bielak, J. [2008] A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion. *Inverse Problems*.

Krebs, J.R., Anderson, J.E., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A. and Lacasse, M.D. [2009] Fast full-wavefield seismic inversion using encoded sources. *Geophysics*, **74**(6), WCC177–WCC188.

Li, S., Vladimirsky, A. and Fomel, S. [2013] First-break traveltime tomography with the double-square-root eikonal equation. *Geophysics*, **78**(6), U89–U101.

Nocedal, J. and Wright, S. [1999] *Numerical Optimization*. Springer, New York.

Pratt, R. [1999] Seismic waveform inversion in the frequency domain, Part 1: Theory, and verification in a physical scale model. *Geophysics*, **64**, 888–901.

Pratt, R.G., Shin, C. and Hick, G. [1998] Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion. *Geophysical Journal International*, **133**(2), 341–362.

Saragiotis, C., Alkhalifah, T. and Fomel, S. [2013] Automatic traveltime picking using instantaneous traveltime. *Geophysics*, **78**(2), T53–T58.

Tarantola, A. [1987] *Inverse problem theory*. Elsevier, Amsterdam.

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

Vogel, C. [2001] *Computational methods for inverse problem*. SIAM, Philadelphia.

Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Štekl, I., Guasch, L., Win, C. et al. [2013] Anisotropic 3D full-waveform inversion. *Geophysics*, **78**(2), R59–R80.