CORE
Metadata, citation and similar papers at core.ac.uk
Provided by DSpace@MIT

Extrapolated Full waveform inversion (EFWI) with synthesized low frequency data

Yunyue Elita Li and Laurent Demanet, Department of Mathematics and Earth Resources Laboratory, MIT

SUMMARY

The availability of low frequency data is an important factor in the success of full waveform inversion (FWI) in the acoustic regime. The low frequencies help determine the kinematically relevant, low-wavenumber components of the velocity model, which are in turn needed to avoid convergence of FWI to spurious local minima. However, acquiring data below 2 or 3 Hz from the field is a challenging and expensive task. In this paper we explore the possibility of synthesizing the low frequencies computationally from high-frequency data, and use the resulting prediction of the missing data to seed the frequency sweep of FWI. To demonstrate the reliability of bandwidth extension in the context of FWI, we first use the low frequencies in the extrapolated band as data substitute, in order to create the low-wavenumber background velocity model, and then switch to recorded data in the available band for the rest of the iterations. The resulting method, extrapolated FWI (EFWI), demonstrates surprising robustness to the inaccuracies in the extrapolated low frequency data. With a synthetic Marmousi model, we demonstrate that FWI based on an extrapolated $[1,5]$ Hz band, itself generated from data available in the $[5,15]$ Hz band, can produce reasonable estimations of the low wavenumber velocity models.

INTRODUCTION

Since proposed by Tarantola (1984), full waveform inversion (FWI) has established itself as the default wave-equation-based inversion method for subsurface model building. In contrast to ray-based traveltime tomography (e.g. Woodward et al. (2008)), FWI includes both phase and amplitude information in the seismograms for elastic parameter estimation. When working with reflection data, the model update of FWI is similar to a migration image of the data residual, given the background propagation velocity (Claerbout, 1985). The accuracy of the resolved model is controlled by the frequency band in the data and the accuracy of the initial (background, macro) model. When this initial macromodel is not sufficiently accurate, the iterative process of FWI gets trapped in undesirable local minima or valleys (Virieux and Operto (2009)). The lack of convexity is intrinsic and owes to the relatively high frequencies of the seismic waveforms.

In spite of an extensive literature on the subject, a convincing solution has yet to emerge for mitigating these convergence issues. So far, the community's efforts can be grouped into three categories. In the first category, misfit functions different from least-squares have been proposed to emulate traveltime shifts between the modeled and recorded waveforms (Luo and Schuster, 1991; Ma and Hale, 2013). However, these methods move away from the attractive simple form of the least-squares formulation and require additional data processing steps which are themselves not guaranteed to succeed in complex propagation geometries. In the second category, additional degrees of freedom are introduced to (attempt to) convexify the waveform inversion in higher dimensions (Symes and Carazzone, 1991; Shen, 2004; Biondi and Almomin, 2014). These methods rely on an iterative formulation to gradually restrict the extended nonphysical model space to the physical model space, which is often a delicate process. Moreover, they introduce significant computational cost and memory usage in addition to the already very expensive FWI. In the third category, the tomographic and migration components in the FWI gradient are separated and enhanced at different stages of the iterations (Mora, 1989; Tang et al., 2013; Alkhalifah, 2015). Although these methods do enhance the low wavenumber components of the FWI gradient, the essential difficulty of ensuring correctness of the tomographic component is still mostly untouched.

The most straightforward way to increase the basin of attraction of the least-squares FWI objective function is to seed it with low frequency data only, and slowly enlarge the data bandwidth as the descent iterations progress. However, until several years ago, the low frequency energy below 5 Hz was often missing due to instrument limitations. More recently, as the importance of the low frequencies became widely recognized by the industry, broadband seismic data with high signal-to-noise ratio between 1.5 Hz and 5 Hz started being acquired at a significantly higher cost than previously.

The premise of this paper is that the phase tracking method, proposed in Li and Demanet (2015), is a reasonably effective algorithm for extrapolating the low frequency data based on the phases and amplitudes in the observed frequency band. A tracking algorithm is able to separate each seismic record into atomic events, the amplitude and phase functions of which are smooth in both space and frequency. With this explicit parameterization, the user can now fit smooth non-oscillatory functions to represent and extrapolate the wave physics to the unrecorded frequency band. Although the resulting extrapolated data can only be expected to accurately reproduce the low frequency recordings in very controlled situations, they are nevertheless adequate substitutes that appear physically plausible in a broad range of scenarios. We are not aware that there is any other attempt at synthesizing low frequency data in the seismic literature. (The mathematical problem of providing tight guarantees concerning extrapolation of smooth functions from the knowledge of their noisy samples has however been solved in our companion paper (Demanet and Townsend, 2016).)

In this paper, we test the reliability of the extrapolated low frequency data on a synthetic Marmousi example in the constant-density acoustic regime. The low frequencies between 1 Hz and 5 Hz are extrapolated from the recordings at 5 Hz and above. We demonstrate that although the extrapolated low frequencies are sometimes far from exact, the low wavenumber models obtained from the extrapolated low frequencies are often suitable for initialization of FWI at higher frequencies.

METHOD

Review of waveform inversion with truncated Gauss-Newton iterations

Conventional FWI is formulated in data space via the minimization of the least-squares mismatch between the modeled seismic record u with the observed seismic record d,

$$J(m) = \frac{1}{2}\sum_{r,s,t}(u_s(x_r,t;m) - d_{r,s,t})^2, \tag{1}$$

where m is the slowness of the pressure wave, s indexes the shots, and x_r are the receiver locations. We also write sampling at the receiver locations with the sampling operator S as $u_s(x_r,t;m) = Su_s(x,t;m)$. The modeled wavefield $u_s(x,t;m)$ is the solution of a wave equation (discretized via finite differences in both space and time):

$$\left(m^2\frac{\partial^2}{\partial t^2} - \nabla^2\right)u_s = f_s, \tag{2}$$

with f_s a source wavelet at location x_s, and $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ the Laplacian operator.

Starting with an initial model $m^{(0)}$, we use a gradient-based iterative scheme to update the model

$$m^{(i+1)} = m^{(i)} - \alpha\left(\mathbf{J}_r^{\mathbf{T}}\mathbf{J}_r\right)^{-1}\mathbf{J}_r^{\mathbf{T}}r(m^{(i)}), \tag{3}$$

where $m^{(i)}$ is the model at the i^{th} iteration, $r = Su(m^{(i)}) - d$ is the data residual, $\mathbf{J}_r$ is the Jacobian matrix, and α is the step length for the update.

In practice, the normal matrix $\mathbf{J_r^T J_r}$ is too large to build explicitly and is often approximated by an identity matrix. In this paper we choose to precondition this matrix by approximately solving the following system using a few iterations of the conjugate gradient method (Claerbout, 1985; Metivier et al., 2013),

$$\min_{\delta m}||r(m^{(i)}) - \mathbf{J}_r(m^{(i)})\delta m||_2^2, \tag{4}$$

where δm is the unknown increment in model. The linear iteration in (4) is also known as Least-Squares Reverse Time Migration (LSRTM), which effectively removes the source signature and produces "true-amplitude" velocity perturbation at convergence. Due to the limited computational resource and the ill-conditioned matrix $\mathbf{J_r^T J_r}$, we truncate the Gauss-Newton inversion at 3 iterations for each nonlinear step. Then a careful line search is performed to make sure the objective function (Equation 1) decreases in the nonlinear iterations. (Another effective way of preconditioning the normal operator is to use randomized matrix probing, see Demanet et al. (2012).)

To help the iterative inversion avoid local minima, we perform frequency-continuation FWI starting from the lowest available frequency with a growing window.

Review of phase tracking and frequency extrapolation

In a previous paper (Li and Demanet, 2015), we demonstrated that there exist interesting physical scenarios in which low frequency data can be synthesized from the band-limited field recordings using nonlinear signal processing. This processing step is performed before full waveform inversion in the frequency domain. To extrapolate the data from the recorded frequency band to lower (and higher) frequencies, the phase tracking method consists in solving the following minimization problem to separate the measured data to its atomic event-components:

$$\begin{aligned} J_{tracking}(\{a_j,b_j\}) &= \tfrac{1}{2}||\widehat{u}(\omega,x) - \widehat{d}(\omega,x)||_2^2 \\ +\lambda\textstyle\sum_j||\nabla_\omega^2 b_j(\omega,x)||_2^2 &+ \mu\textstyle\sum_j||\nabla_x b_j(\omega,x)||_2^2 \\ +\gamma\textstyle\sum_j||\nabla_{\omega,x}a_j(\omega,x)||_2^2, & \end{aligned} \tag{5}$$

where $\widehat{d}$ are the measured data in the frequency domain; ∇_k and ∇_k^2, with $k = \omega, x$, respectively denote first-order and second-order partial derivatives; $\nabla_{\omega,x}$ denotes the full gradient; and the predicted data record $\widehat{u}$ is modeled by the summation of r individual events:

$$\widehat{u}(\omega,x) = \sum_{j=1}^{r}\widehat{v}_j = \sum_{j=1}^{r}\widehat{w}(\omega)a_j(\omega,x)e^{ib_j(\omega,x)}, \tag{6}$$

where the wavelet $\widehat{w}(\omega)$ is assumed known to a certain level of accuracy. The constants λ, μ, and γ are chosen empirically.

The optimization problem for event tracking here is reminiscent of full waveform inversion with high-frequency data, hence shares a similar level of nonconvexity. Yet, by posing it as a data processing problem, the nonconvexity can be empirically overcome with an explicit initialization scheme using Multiple Signal Classification (MUSIC), coupled with a careful trust-region "expansion and refinement" scheme to track the smooth phase and amplitude. We refer the reader to the detailed algorithm in the previous paper (Li and Demanet, 2015).

Having obtained the individual events, we make explicit assumptions about their phase and amplitude functions in order to extrapolate outside of the recorded frequency band. Namely, we assume that the Earth is nondispersive, i.e., the phase is affine (constant + linear) in frequency, and the amplitude is to a good approximation constant in frequency – though both are variable in x, of course. A least-squares fit is then performed to find the best constant approximations $a_j(\omega,x) \simeq \alpha_j(x)$, and the best affine approximations $b_j(\omega,x) \simeq \omega\beta_j(x) + \phi_j(x)$, from values of ω within the useful frequency band. These phase and amplitude approximations can be evaluated at values of ω outside this band, to yield synthetic flat-spectrum atomic events of the form

$$\widehat{v}_j^e(\omega,x) = \alpha_j(x)e^{i(\omega\beta_j(x)+\phi_j(x))}. \tag{7}$$

These synthetic events are multiplied by a band-limited wavelet, and summed up, to create a synthetic dataset.

The effectiveness of this method for event identification is limited by many factors, chiefly the resolution of the MUSIC algorithm and the signal-to-noise ratio of the data. The algorithm often tracks the strong events and treats the weak events as noise. Moreover, the amplitudes of the events are less predictable than the phases, due to propagation and interfering effects. Therefore, the extrapolated data record is inexact, typically with higher fidelity in phase than in amplitude. In the

following section, we test the reliability of the extrapolated low frequencies by initializing the frequency continuation of FWI. Our goal is to bring reliable low wavenumber information in the model by fitting the phase of the extrapolated data and to help enlarge the basin of attraction for FWI when the low frequencies are missing from data.

NUMERICAL EXAMPLES

In this section, we demonstrate the reliability of the extrapolated low frequencies on the Marmousi model. In this example, we compare three cases while keeping the initial models fixed:

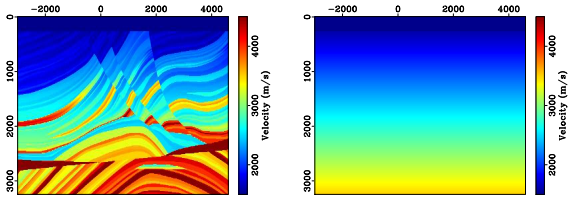
- In the *control* case, frequencies from 1 Hz to 15 Hz are used in the frequency continuation of FWI.
- In the *extrapolated* FWI case, we first extrapolate the data between 5 Hz and 15 Hz to the frequency band between 1 Hz and 5 Hz. The extrapolated data are used to build the low wavenumber model to further initialize the frequency continuation starting at 5 Hz.
- In the *missing low-frequency* case, frequency continuation of FWI starts from the lowest frequency at 5 Hz.

We restrict FWI to reflection events only by limiting the maximum offset to 500 m in the Marmousi model (Figure 1a). The starting model for FWI is a 1.5 D linearly increasing velocity profile (Figure 1b). It is an extremely challenging task for FWI to recover large low-wavenumber discrepancies between the initial model and the true model, especially in the deeper section (below 2 km).

Reliability of the extrapolated low frequencies ($1-5$ Hz) are tested with FWI at these low frequencies. Figure 2 shows the inverted velocity model using modeled data (a) and using extrapolated data (b). Although not as detailed as Figure 2a, the velocity model inverted using the extrapolated data correctly captures the very low wavenumber component of the true model. These models are used to initialize FWI with data at higher frequencies.

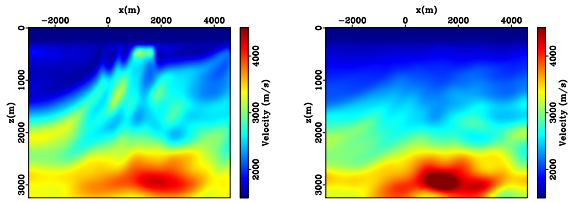
Figure 3 compares the final inverted results after a full bandwidth FWI continuation. In the shallow region (above 2 km), velocity models resolved in the control case and the extrapolated case are very similar with accurately imaged fine layers and normal faults. Both models have trouble resolving a high resolution and accurate velocity model in the deep region, because reflections from the dipping reflectors and the anticline structure have not been sufficiently recorded due to the limited offset. In comparison, FWI starting at 5 Hz yields little meaningful information about the subsurface. The inversion failed to update the low wavenumber structure of the velocity model and placed reflectors at wrong positions.

Figure 4 compares pseudo velocity logs at three surface locations. Velocity models in both control and extrapolated cases recover the true velocity model very well above 1 km. Quality of the inverted model degrades with depth. However, both velocity models capture the low wavenumber components of the velocity model. The maximum updates in the deeper section are as high as 1000 m/s. The huge velocity error prevents FWI starting at 5 Hz from converging to the true model.



(a) (b)

Figure 1: Marmousi model (a) and the starting model for FWI (b). The starting model is a 1.5 D linearly increasing velocity profile from 1500 m/s at the water bottom to 3500 m/s at 3.2 km. The initial model is far from the true especially in the deeper section.



(a) (b)

Figure 2: Comparison between the inverted model after FWI using modeled low frequency ($1-5$ Hz) data (a) and using extrapolated low frequency data ($1-5$ Hz) (b). Both models capture the low wavenumber structure of the Marmousi model, although the inverted model using modeled data contains more details in the shallow part.

DISCUSSION

There are two main reasons that both the amplitude and the phase of the extrapolated data are inexact. First, the tracking algorithm determines the number of individual events as an initialization step. This number may decrease (for event truncation), but may not increase (for even bifurcation) as the tracking expands. These untracked events are the main contributors to the errors in the extrapolated phase function. An aggregation method, with event fragments tracked in subsets of traces and merged into actual composite events, may help improve the accuracy of the event tracking and phase extrapolation. Second, compared with the phase extrapolation based on non-attenuative physics, the amplitude extrapolation is less constrained by physical principles. To improve the accuracy of the extrapolated amplitudes, a higher order polynomial or rational function can be used to approximate the amplitude variation with respect to frequency. However, we do not expect the extrapolation to perfectly reconstruct the amplitudes. Consequently, although the extrapolated data are adequate for initializing FWI, they are not suitable for absolute impedance inversion and amplitude-based rock property interpretation.

Due to the inaccuracy in their phase and amplitude, we do not allow FWI to fully fit the data at the extrapolated low frequen-

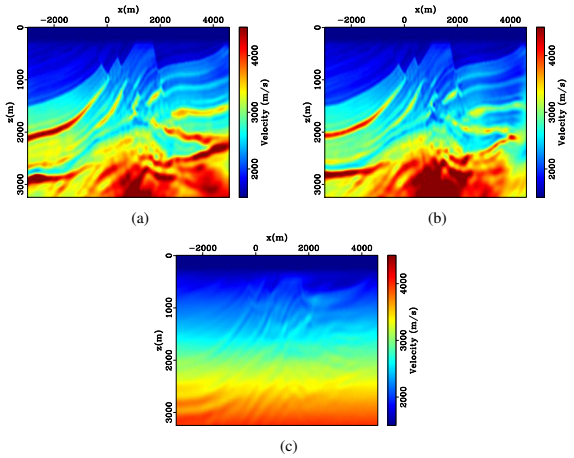


Figure 3: Comparison between the inverted model from FWI after a full bandwidth continuation. In (a), resulting model from the control case (frequency continuation from 1 to 15 Hz). In (b), resulting model from the extrapolated case (initialization using extrapolated low frequencies $(1-5)$ Hz and frequency continuation with recorded data from 5 to 15 Hz). In (c), resulting model from the missing low-frequency case (frequency continuation from 5 to 15 Hz). A better inverted model can be obtained in the control case if we iterate to convergence at the lowest frequencies. However, we limit the number of iterations in the control case to ensure a fair comparison with the extrapolated case.

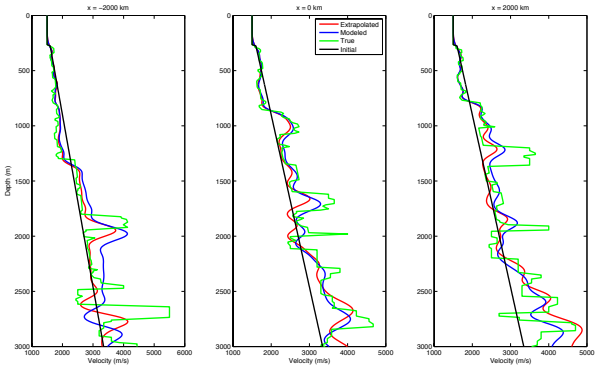


Figure 4: Pseudo velocity logs at three surface locations. The black line denotes the initial model. The green line denotes the true model. The black, green, red, and blue lines denote the pseudo log from the initial, true, control and extrapolated model.

cies. This limits the resolution of the inverted FWI model. With hundreds of more iterations in the control case, FWI starting at 1 Hz can reduce the data residual to 1% at each frequency band. Hence, the velocity model can be fully resolved (with all wavenumber components) because of the availability of both low frequency and long offsets. To the contrary, overfitting the extrapolated data at low frequencies would lead the inversion to undesired local minima and spurious models. FWI with limited number of iterations resolves a good estimate of the velocity model within the increased available frequency band, but it cannot perfectly resolve the model at all wavenumber due to both the slow convergence and the potential local minima introduced by the inaccuracies in the extrapolated data.

In our numerical examples, the extrapolated low frequencies are used only to initialize FWI in order to obtain a low wavenumber model. As soon as the frequency continuation moves to the recorded frequency band, the extrapolated low frequencies are abandoned. This leaves the low wavenumber components of the model space unconstrained in later FWI iterations. A proper combination of the extrapolated data and the recorded data needs to be studied to ensure a fully constrained inversion for velocity in the whole wavenumber band.

CONCLUSION

To mitigate the nonconvexity of FWI, we propose to start the frequency continuation using the extrapolated low frequency data. The extrapolation is only feasible after decomposing the seismic records into individual atomic events via phase tracking for each isolated arrival. Numerical examples demonstrate that full waveform inversion is surprisingly tolerant to inaccuracies in the amplitude and phase of the extrapolated events. Initializing with the extrapolated low frequencies mitigates the severe nonconvexity that FWI suffers from when only high frequency data are available. By explicitly obtaining the phase and amplitude of each individual event, our method shares an important feature with travel time tomography: its ability to extract kinematic information from high frequencies only.

ACKNOWLEDGEMENTS

This project was funded by Total S.A. Laurent Demanet is also grateful to AFOSR, ONR, and NSF for funding. Yunyue Elita Li acknowledges the sponsors of the Energy Resource Laboratory at MIT for funding.