Preconditioning FWI with approximate receiver Green's functions

Tim Burgess and Mike Warner, Imperial College London*

SUMMARY

Preconditioning of full-waveform inversion (FWI) using approximations to the inverse Hessian is well established. Different approximations to the Hessian give different rates of convergence and therefore, given that the iteration count is typically limited by available compute resources, different levels of quality in the final results. In this paper, we demonstrate a low-cost, general-purpose, diagonal Hessian approximation which includes factors related to the receiver Green's functions and provides significant uplift in image quality compared to approximations not including these factors.

INTRODUCTION

Full-waveform inversion (FWI) is an iterative method for finding an earth model that directly minimises the difference, defined as a scalar objective function, between observed seismic data and those predicted from the model. Using the method described in Tarantola (1984), the derivative of the objective function with respect to the model parameters can be obtained at each iteration. Optimisation schemes such as conjugate gradient or L-BFGS may then be used to calculate improved models which minimise the objective. Due to the poor conditioning of the FWI problem, these iterations can converge extremely slowly if some kind of preconditioning is not applied.

Typically a preconditioner is constructed by calculating an easily-invertible approximation to the Hessian matrix. One of the simplest such approximations is a matrix containing only the diagonal entries of the Hessian, however even these may be expensive to compute and so approximations must be used. In this paper we present results for a new approximation, based on the reverse-time migration work of Plessix and Mulder (2004), which makes use of receiver position information and improves the convergence rate by approximately a factor of two on the demonstrated problems.

METHOD

The Hessian provides us with a measure of how sensitive the objective function is to 2nd-order updates in the model parameters. If a particular component of the gradient is large, this tells us that the objective function will change rapidly for a small change in the parameter. This does not give any understanding of whether the desired update of the parameter is large or small; it is often the case that those parameters with the smallest gradient components require the largest updates to fully minimise the objective.

For a conventional FWI objective function $f = \frac{1}{2}||(\mathbf{d}(\mathbf{m}) - \mathbf{o})||^2$, the diagonal elements of the Hessian may be written:

$$H_{ii} = \frac{\partial^2 f}{\partial {m_i}^2} = \mathrm{Re}\left[\left(\frac{\partial \mathbf{d}}{\partial m_i}\right)^\dagger \left(\frac{\partial \mathbf{d}}{\partial m_i}\right) + (\mathbf{d} - \mathbf{o})^\dagger \frac{\partial^2 \mathbf{d}}{\partial {m_i}^2}\right] \quad (1)$$

The second term in the above equation is often ignored using the justification that as the inversion proceeds the residuals $\mathbf{d} - \mathbf{o}$ become smaller (Virieux and Operto, 2009). Focussing on the first term, we see that the diagonal Hessian elements may be computed using the already-calculated forward wavefield $\overrightarrow{\mathbf{u}}$:

$$\left(\frac{\partial \mathbf{d}}{\partial m_i}\right)^\dagger \left(\frac{\partial \mathbf{d}}{\partial m_i}\right) = \overrightarrow{\mathbf{u}}^\dagger \left(\frac{\partial \mathbf{A}}{\partial m_i}\right)^\dagger \mathbf{A}^{-\dagger}\mathbf{D}^\dagger\mathbf{D}\mathbf{A}^{-1} \left(\frac{\partial \mathbf{A}}{\partial m_i}\right) \overrightarrow{\mathbf{u}} \quad (2)$$

The central factors in this expression are the expensive ones to calculate - for each i, we must:

- solve the forward wave equation $\mathbf{A}\mathbf{u}' = (\partial \mathbf{A} / \partial m_i)\overrightarrow{\mathbf{u}}$ for a source located at the corresponding model point,
- sample the resulting wavefield $\mathbf{u}'$ at the receiver locations with the receiver sampling operator $\mathbf{D}$,
- solve the adjoint (i.e. backpropagating) wave equation with the sampled wavefields reinjected as sources at the receiver locations, and
- sample the resulting backpropagated wavefield at the model location i.

Even allowing for the time-invariance of the wave equation and the reciprocity of the Green's functions, at a minimum this requires solving and storing the solution of the wave equation for each unique receiver location. Plessix and Mulder (2004), discussing the case of reverse-time migration, show that ignoring these factors (as in Shin et al., 2001) is equivalent to assuming infinite receiver coverage. They go on to demonstrate the impact of this assumption on migrated images, and discuss some superior approximations. In a similar way, we hope to find an approximation for these factors which takes the finite and irregular receiver coverage into account during FWI.

Instead of ignoring these factors, we therefore propose using an approximation to the Green's function based on a homogenous background velocity. For point receivers the kinematic effects are cancelled out by the forward/backward wave propagation pairing, leaving only amplitude variations to take into account. By introducing the constant-velocity approximation, the $\mathbf{A}^{-\dagger}\mathbf{D}^\dagger\mathbf{D}\mathbf{A}^{-1}$ in equation (2) can be replaced with a scale S_i, expressed as (for the 2D case):

$$S_i = \sum_{\mathbf{x}_r} ||\mathbf{x}_r - \mathbf{x}_i||^{-1} \quad (3)$$

Even with this simple approximation, the large number of model parameters i and receiver locations $\mathbf{x}_r$ make this an expensive function to calculate. However, using a k-d tree of

DOI http://dx.doi.org/10.1190/segam2015-5844253.1

receiver locations and the multipole expansion, as shown in Figure 3, we can reduce the calculation time to approximately $O(n_i \log n_r)$. In practise this takes significantly less computation than one solution of the wave equation.

RESULTS

The above preconditioner was tested by performing inversions on the Marmousi model, an OBC survey from the Tommeliten Alpha field, and the Chevron/SEG 2014 2D Marine Isotropic Elastic Synthetic Seismic Benchmark. The Chevron/SEG 2014 results are shown below. Display permission is pending for the field data.

The Chevron/SEG 2014 dataset is created from a highly realistic 50km x 6km seismic model, based upon field seismic data, well logs and rock physics. Using this model, Chevron generated synthetic data simulating a conventional 2D towed-streamer marine acquisition geometry with a maximum offset of 8000m. The model has a non-deterministic, realistic relationship between P-wave velocity, S-wave velocity and density. It is locally isotropic, has limited bandwidth, ambient noise, surface ghosts and surface multiples, and the provided data was modelled using the full visco-elastic two way wave equation. Chevron released this dataset as a benchmark to test FWI and related technology, and have not released the true model. It therefore presents a realistic inversion challenge against which algorithmic changes can be tested, and allows experiments to avoid the so-called "inversion crime" of inverting with the same modelling assumptions that were used to generate the original data.

In addition to the synthetic seismic data, Chevron also provided a source signature, a sonic profile in one well with a depth range of 1000-2450m, a 1-D starting p-wave velocity model, the velocity of the water column, and the geometry of the seabed. Since the provided starting model is badly cycle-skipped at the lowest usable frequencies in the data, and addressing cycle-skipping is not the topic of this abstract, for demonstration purposes we have used a starting model derived from the adaptive waveform inversion (AWI) results presented by Warner and Guasch at the 2014 FWI SEG workshop. By using a objective function based on matching filters, AWI is able to perform inversion even when the starting model predicts traces which are cycle-skipped compared to the observed data (Warner and Guasch, 2014). A gaussian smoother was applied to the final AWI output model before inversion.

The results of two inversions are shown. Both were performed using conjugate gradient, in the time domain, with data filtered to below 11Hz. The first inversion was performed using a preconditioner based only on the source wavefield (i.e. using the assumption of Shin et al., 2001). The second inversion incorporated the approximate receiver Green's functions described above. The convergence rate of the inversion with the approximate receiver Green's functions in the preconditioner is significantly higher than the inversion preconditioned by the source wavefield alone. Significantly larger updates are made in the deeper part of the model, without adversely affecting the near-surface section, and the resolution of the recovered model is significantly higher. The starting model and a section of the result for iterations 15 and 40 for each case is shown in Figure 1, and a zoomed-in section for the starting model and iteration 40 of each test is shown in Figure 2.

The observed increase in the quality of the result, particularly in the deeper part of the model, is consistent with the preconditioner amplifying updates in the deeper part of the model compared to that part nearer the surface. This amplification with depth, approximately by a factor of $\sqrt{z}$ in 2D, is one of two significant effects expected with the introduction of these extra receiver-location dependent factors into the preconditioner. The second effect is to suppress spuriously large update components very close to the receivers. In this case, in 2D and with extremely dense receiver spacing, this second effect is not visible. We expect that this second effect may help to suppress sail-line oriented artifacts observed when inverting 3D streamer data, but are yet to confirm this.

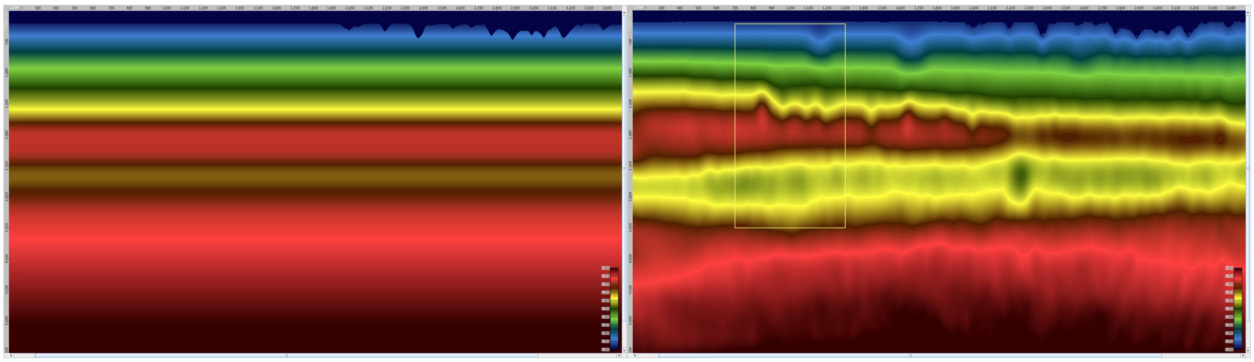
CONCLUSIONS

We have demonstrated a low-cost, general-purpose preconditioner which improves the convergence of full-waveform inversion by incorporating an approximation to the receiver Green's functions. This improved performance can be viewed either as a way to achieve similar results at lower computational cost, or to obtain superior results for equivalent cost.

ACKNOWLEDGEMENTS

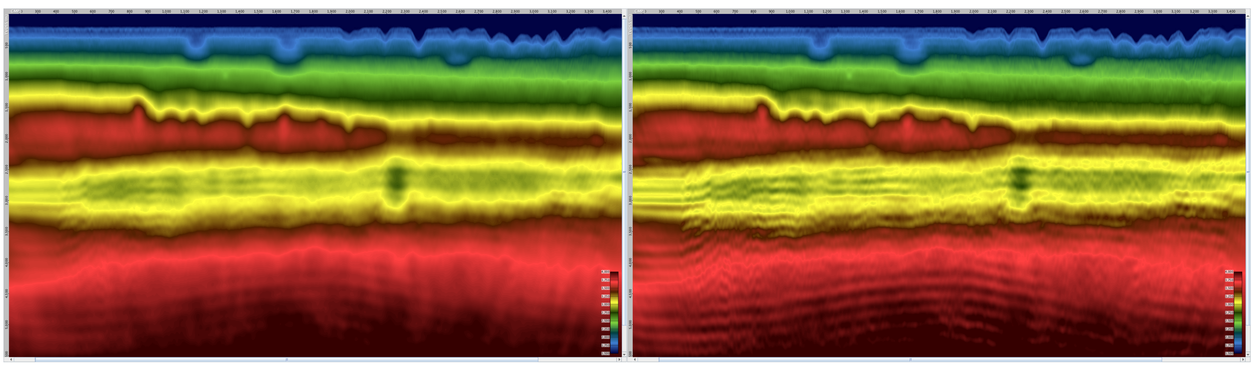
We thank Chevron for making these data available and for designing a challenging and useful dataset, and Lluis Guasch for his work in performing the adaptive waveform inversion from which we derived the starting model. This work was undertaken as part of the Fullwave Game Changer consortium; we thank the sponsors of that project for their support.

DOI http://dx.doi.org/10.1190/segam2015-5844253.1



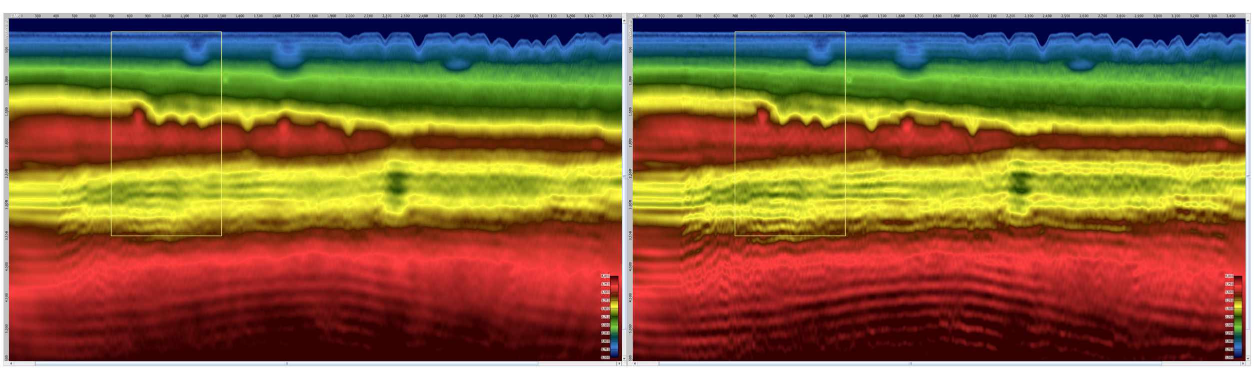
(a) Starting Model

(b) Smoothed AWI result



(c) Source Wavefield Preconditioning only, iteration 15

(d) ARGF Preconditioning, iteration 15

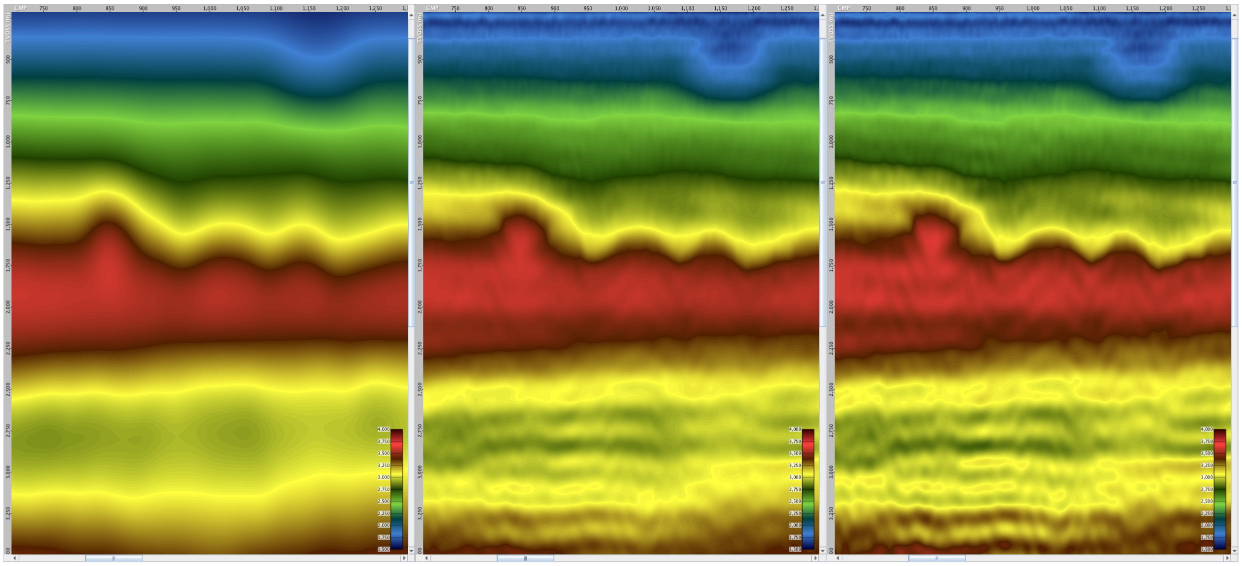


(e) Source Wavefield Preconditioning only, iteration 40

(f) ARGF Preconditioning, iteration 40

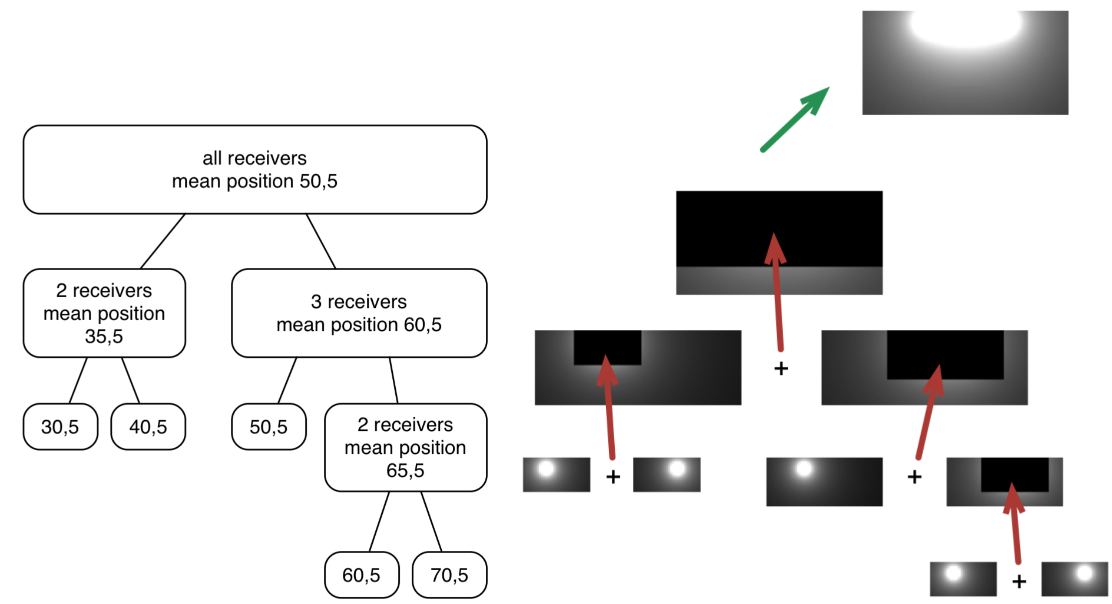
Figure 1: Comparison between the starting model (top left), the smoothed result from adaptive waveform inversion (top right), and subsequent FWI inversion results after iterations 15 (centre) and 40 (bottom), with (right) and without (left) the new preconditioner. When the inversion is preconditioned with approximate receiver Green's functions, the resulting models are significantly sharper. Note the similarity betweeen (1d) and (1e): the preconditioning can be viewed as either a performance increase for results of the same quality, or a quality increase for equivalent computational expense.

DOI http://dx.doi.org/10.1190/segam2015-5844253.1



(a) Starting Model (b) Source Wavefield Preconditioning only, iteration 40 (c) ARGF Preconditioning, iteration 40

Figure 2: Zoomed-in view of the highlighted area from Figure 1. For the same computational expense, inversion performed with the new preconditioner gives significantly sharper results.



(a) (b)

Figure 3: Example illustrating the efficient construction of the preconditioner on a 101x51 grid containing 5 receivers. (3a) First, the receiver locations are arranged into a tree data structure which groups nearby receivers together. Individual receiver locations become leaf nodes of the tree, while nodes at higher levels of the tree store the mean position and quadrupole moments of all the receiver locations under them. These give a cheap approximation of the preconditioner contribution from those receivers which is accurate for sufficiently distant points in the model. (3b) Recursive calculation of the preconditioner: the cheap approximation is used for points sufficiently distant from any of the receivers, and successively better approximations are used to fill in the remaining area closer to the receiver array. The total effort to compute the preconditioner scales logarithmically with the receiver count.

DOI http://dx.doi.org/10.1190/segam2015-5844253.1

REFERENCES

Plessix, R.-E., and W. a. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, **157**, 975–987.

Shin, C., S. Jang, and D. J. Min, 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical Prospecting, **49**, 592–606.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

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

Warner, M., and L. Guasch, 2014, Adaptive Waveform Inversion - FWI Without Cycle Skipping - Theory: Presented at the 76th EAGE Conference and Exhibition.

DOI http://dx.doi.org/10.1190/segam2015-5844253.1

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2015 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Plessix, R.-E., and W. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, **157**, no. 3, 975–987. http://dx.doi.org/10.1111/j.1365-246X.2004.02282.x.

Shin, C., S. Jang, and D. J. Min, 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical Prospecting, **49**, no. 5, 592–606. http://dx.doi.org/10.1046/j.1365-2478.2001.00279.x.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266. http://dx.doi.org/10.1190/1.1441754.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26. http://dx.doi.org/10.1190/1.3238367.

Warner, M., and L. Guasch, 2014, Adaptive waveform inversion - FWI without cycle skipping — Theory: Presented at the 76th Annual International Conference and Exhibition, EAGE, http://dx.doi.org/10.3997/2214-4609.20141092.