Multiscale full-waveform inversion using flux corrected transport

Mahesh Kalita and Tariq Alkhalifah

King Abdullah University of Science and Technology

SUMMARY

Full-waveform inversion (FWI) iteratively recovers the unknown model parameters from seismic data. In practice, a successful FWI implementation often follows a multistage recovery approach: starting from the retrieval of the lower model wavenumbers (tomography) followed by the higher resolution ones (imaging). On that account, we propose a new method based on the flux-corrected transport (FCT) technique, used often in computational fluid dynamics owing to the removal of instabilities in a shock profile. FCT involves three finite-difference steps: a transport, a diffusion, and followed by an anti-diffusion. The third step, however, involves non-linear operators such as maximum and minimum, which are non-differentiable in a classic sense. However, since the seismic source wavelet and the corresponding wavefield are relatively smooth and continuous in nature, and does not yield any strong ripples like shock waves, we unsubscribe to the non-linear step from FCT, which allows us to evaluate the FWI gradient. As a result, it accentuates no trouble in achieving a converging FWI model by gradually reducing the diffusive flux-correction amount. Those features are demonstrated on a dataset from the Marmousi II model with no frequency content less than 5 Hz. We initiate the inversion process for the remaining full-bandwidth of the dataset with a linear $v(z)$ model. In addition, we show the versatility of the FCT based FWI on a marine field dataset from offshore Australia.

INTRODUCTION

Full-waveform inversion (FWI) aims to recover the unknown model parameters with high resolution from the recorded seismic data by iteratively updating the initial model using the gradient of a data misfit function (Tarantola, 1984). Its success often summons the gradient to be smooth, especially at its early iterations. All of the popular available strategies such as a hierarchical scaling of the available frequencies in the dataset (Bunks et al., 1995; Yuan and Simons, 2014); or preconditioning of the gradient (e.g. application of Gaussian kernel smoothing (Virieux and Operto, 2009) or scattering angle filter (Alkhalifah, 2015)); or modifications in the objective function (e.g., envelope (Bozdağ et al., 2011) or optimal transport (Engquist and Froese, 2013) or auxiliary bump functional (Bharadwaj et al., 2016)) and so forth attempts to maneuver a gradual retrieval of model wavenumbers in ascending order (from low to high). However, they are plagued with computational overburden, or implementation difficulties, or non-physical pre-

conditioning (in case of Gaussian filtering), which often incur instability in the inversion process. In this article, we precondition the gradient using a flux correction to the solution of the wave equation.

Thus, in this abstract, we develop an extension of Fei and Larner (1995) to FWI incorporating the required modifications, as mentioned in the theory section. We utilize the flux corrected transport algorithm (FCT), which substitutes the numerical failure with physics (Boris and Book, 1973) owing to remove the numerical dispersion. We exploit this technique in order to obtain a smooth wavefield, and hence a smooth gradient. Therefore, FCT renders an improved smooth gradient (because of the smooth wavefield), which is more physical and dynamic compared to Gaussian kernel smoothing. Over the iterations, we gradually decrease the degree of smoothness by reducing the flux correction and for no correction, the proposed method becomes conventional FWI for the entire bandwidth of the data-set. We demonstrate its versatility on a synthetic, noise-free, dataset obtained from the Marmousi II model and also on a field dataset from offshore Australia.

THEORY

Figure 1: Marmousi II model. Wavefield is recorded at the point marked in white ink.

FCT is a popular technique in computational fluid dynamics because of its effect on limiting the solution gradient near shocks. It removes the undesired non-physical ripples near discontinuities by performing two major steps: a conventional finite difference scheme, followed by a limited flux correction to the troublesome region. Since there is no a-priori knowledge about how much diffusion is required per grid point, the practical implementation attempts to supply sufficient diffusion everywhere, followed by limiting the flux wherever overcorrection happens. Without loss of generality, we consider the 2nd order seismic wave equation in a 2D acoustic medium to describe the methodology of the FCT technique. More details in the context of seismic modeling are reported in Fei and Larner (1995). We denote the

FWI with FCT

wavefield at any arbitrary grid point (i, j) with $u_{i,j}^n$ for any arbitrary n^{th} time step. Its numerical flux is defined as $f_{i+1/2,j}^n = u_{i+1,j}^n - u_{i,j}^n$ and $f_{i,j+1/2}^n = u_{i,j+1}^n - u_{i,j}^n$ along x and z directions, respectively. Next, we compute the associated diffusive flux defined as:

$$\begin{aligned} q_{i+1/2,j}^{n-1/2} &= \xi (f_{i+1/2,j}^n - f_{i+1/2,j}^{n-1}) \quad \& \\ q_{i,j+1/2}^{n-1/2} &= \xi (f_{i,j+1/2}^n - f_{i,j+1/2}^{n-1}); \quad 0 \leq \xi \leq 1. \end{aligned} \quad (1)$$

Therefore, the diffusion flux of u is evaluated indulging two finite-difference operations: spatial derivative followed by a temporal one. Now using equation (1), we modify the wavefield at time $(n+1)^{th}$ as follows:

$$\begin{aligned} \tilde{u}_{i,j}^{n+1} &= u_{i,j}^{n+1} + (q_{i+1/2,j}^{n-1/2} - q_{i-1/2,j}^{n-1/2}) \\ &\quad + (q_{i,j+1/2}^{n-1/2} - q_{i,j-1/2}^{n-1/2}) \end{aligned} \quad (2)$$

Figure 2: Recorded wavefield (normalized) and its frequency content (normalized) at the point of consideration in Figure 1 for different ξ values.

Figure 3: Overlay of seismic traces (normalized) with and without the non-linear step in FCT. They are kinematically almost similar.

Figure 4: Acquisition coverage of receivers per shot. The east part of the model has poor acquisition sampling.

$\tilde{u}$ is used for its propagation to the next time-step. Thus, the wavefield extrapolation enjoys an addition of a small corrective quantity (a function of both $\mathbf{x}$ and t) at every time step, yielding a smooth solution with a loss in amplitude. The corrective measure is carried out in a conservative way (except at the boundaries); that is, whenever a quantity of the diffusive flux is added at one point, that same amount is subtracted from somewhere else. Thus, it is shoved from point to point locally, but no net loss or gain as a whole to the system. However, the smoothness order depends on the parameter ξ , which varies with the scheme of finite difference under consideration. ξ should not be too big to violate the CFL criteria. To depict the role of ξ in seismic modeling, we show the wavefield recorded at a grid point of Marmousi II model (marked in white ink, Figure 1) from a source placed at $(x, z) = (8, 0)$ km in Figure 2a. It shows that the larger ξ yields a smoother wavefield. Moreover, larger ξ in FCT induces artificial low frequencies, not necessarily its accurate reconstruction, in the propagated wavefield from a high-frequency source. Thanks to the power of addition of diffusive fluxes at all model points and time steps, FCT operates as a high-cut filter and an artificial low frequency cultivator in the resulting wavefield. We plot frequency spectra of all the traces (Figure 2a) for various ξ values in Figure 2b.

Figure 5: A shot record from the source at $x = 6$ km and its frequency content.

We unsubscribe the essential limiting step from FCT (Fei and Lerner, 1995) in the context of obtaining the FWI gradient. The primary reasons behind this exclusion are as follows: (a) it involves non differentiable operators such as maximum and minimum; (b) it is highly non-linear; (c) seismic source wavelet being continuous and simple in nature is very unlikely to yield very strong ripples like shock-waves. This circumvention effectively makes the algorithm very easy to implement in FWI realm without aggravating its convergence issues if any. Figure 3 shows that the FCT yields almost identical solution even in the absence of non-linear step, particularly in the context of smoothing wavefield.

Further simplification of equation (2) enables us to formulate the FCT on matrix-vector form as follows:

$$\tilde{\mathbf{u}}^{n+1} = \mathbf{u}^{n+1} + D(\tilde{\mathbf{u}}^n + \tilde{\mathbf{u}}^{n-1}) \text{ with } D = \xi \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & 1 \\ 0 & 1 & 0 \end{pmatrix}; \quad (3)$$

where the superscript $n > 1$ is a time step; $\tilde{\mathbf{u}}$ is the flux-corrected wavefield that yields the transport solution $\mathbf{u}$ for the next time-step. The 3×3 kernel D is a self-adjoint linear operation due to ignoring the limiting of the flux step. In a nutshell, the FCT operator $\tilde{L}$ is just the wave equation operator L followed by D . Therefore, the gradient direction of FWI ($\tilde{\mathbf{g}}$), based on the L_2 norm misfit function is given by:

$$\tilde{\mathbf{g}}(\mathbf{x}) = \langle \tilde{\mathbf{u}}, \tilde{\lambda} \rangle \text{ with } \tilde{L}^* \tilde{\lambda} = \Delta \mathbf{d}, \Delta \mathbf{d} = \mathbf{d}_o - \mathbf{d}_m; \quad (4)$$

where $\mathbf{d}_o$ and $\mathbf{d}_m$ represent full-bandwidth observed and modeled data, respectively. Therefore, the gradient based on FCT is just a preconditioner due to the modification in the wave-equation operator, controlled by the parameter ξ . We start the inversion with a large ξ and relax it progressively to zero so that the final inverted model is principally achieved from conventional FWI. In the following examples of our consideration, we set $\xi = \{0.15, 0.1, 0.05, 0.01, 0.005, 0.001, 0\}$ over the iterations.

EXAMPLE

Marmousi II: We display the true model in Figure 1. Data are recorded for 8 seconds in a 10 km long streamer starting from 0.2 km at the near offset. Figure 4 confirms the poor acquisition coverage towards the east part of the model. A Ricker wavelet of 10 Hz dominant frequency and a low cut filter at 5 Hz is used in order to prepare the observed data. Figure 5 shows a shot gather from the source at $x = 6.0$ km and its frequency content. We start the inversion process with a linearly increasing velocity model (Figure 6a). FCT produces a smoother gradient (Figure 6b) compared to the conventional method (Figure 6c) and hence, enjoys a preferred reconstruction of the velocity model (Figure 7a(e)). On the other hand, conventional FWI with a multi-scale approach in frequency (Bunks et al., 1995) fails miserably as shown in Figure 7b. Taking the acquired seis-

Figure 6: (a) Initial model, and gradient at the first iteration using (b) FCT with $\xi = 0.15$ and (c) conventional multiscale strategy (Bunks et al., 1995).

Figure 7: Inverted model using (a) FCT, (b) conventional multiscale strategy (Bunks et al., 1995).

FWI with FCT

mic dataset's innate limitations such as its finite frequency bandwidth, incomplete acquisition aperture and so forth into account, Figure 7a suggests that the proposed method successfully retrieves the most, if not all, the true velocity values.

Field-data: We demonstrate the versatility of the proposed method in a 2D marine dataset from North-Western Australia Continental shelf, acquired by CGG using a Broadseis acquisition system with a variable depth streamer (Soubaras and Dowle, 2010). CGG filtered out frequencies lower than 2.7 Hz because of poor signal-to-noise ratio. We perform FWI to minimize the normalized zero-lag cross-correlation objective function (Choi and Alkhalifah, 2012) for a region of 11.6 km long and 3.75 km deep. The streamer is 8.3 km long containing 648 receivers with the near offset at 0.169 km. Borrowing the inverted wavelet from Kalita and Alkhalifah (2017), we start the inversion with a model (Figure 8a), derived from the given brute-stack time-domain velocity information. We obtain the inverted model as shown in Figure 8b, featuring many sharp contrasting provinces. We validate the inversion result by matching the synthetic data with observed counterparts in Figure 9. It confirms that the inverted velocity compared to the starting model produces a dataset which enjoys better resemblance to its observed counterpart. Moreover, the inverted model delineates the reflectors better compared to the initial model as indicated in the RTM images in Figure 10.

Figure 8: (a) Initial model for the field data of consideration, (b) inverted model using FCT.

CONCLUSIONS

We developed FWI multi-scale algorithm with flux correction, which serves as an improved alternative to the conventional hierarchical frequency scaling approach. It

Figure 9: Shot gather displaying interleaved synthetic and observed data using the (a) initial velocity and (b) final velocity. We intersperse 40 traces starting from the synthetic followed by observed data.

Figure 10: The RTM image overlaid over the migration velocity: (a) initial, (b) final.

involves a parameter ξ , which controls the smoothness of the wavefield, hence, the gradient. Moreover, FCT with large ξ attempts to synthesize artificial low frequencies in its absence from the high-frequency source. A gradual decrease of ξ to zero results in conventional FWI for the full bandwidth of the dataset. We validated the proposed method on the Marmousi II model and a field dataset.

ACKNOWLEDGMENTS

We would like to thank KAUST for its support, all members of Seismic Wave Analysis Group in KAUST and Dr. David Ketcheson for the fruitful discussions. For computer time, this research used the resources of the Supercomputing Laboratory and IT Research Computing at KAUST. We also thank CGG for providing the marine dataset example.

REFERENCES

- Alkhalifah, T., 2015, Scattering-angle based filtering of the waveform inversion gradients: *Geophysical Journal International*, **200**, 363–373, <https://doi.org/10.1093/gji/ggu379>.
- Bharadwaj, P., W. Mulder, and G. Drijkoningen, 2016, Full waveform inversion with an auxiliary bump functional: *Geophysical Journal International*, **206**, 1076–1092, <https://doi.org/10.1093/gji/ggw129>.
- Boris, J. P., and D. L. Book, 1973, Flux-corrected transport. I. SHASTA, a fluid transport algorithm that works: *Journal of Computational Physics*, **11**, 38–69, [https://doi.org/10.1016/0021-9991\(73\)90147-2](https://doi.org/10.1016/0021-9991(73)90147-2).
- Bozdağ, E., J. Trampert, and J. Tromp, 2011, Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements: *Geophysical Journal International*, **185**, 845–870, <https://doi.org/10.1111/j.1365-246x.2011.04970.x>.
- Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: *Geophysics*, **60**, 1457–1473, <https://doi.org/10.1190/1.1443880>.
- Choi, Y., and T. Alkhalifah, 2012, Application of multi-source waveform inversion to marine streamer data using the global correlation norm: *Geophysical Prospecting*, **60**, 748–758, <https://doi.org/10.1111/j.1365-2478.2012.01079.x>.
- Engquist, B., and B. D. Froese, 2013, Application of the Wasserstein metric to seismic signals: *Communications in Mathematical Sciences*, **12**, 979–988, <https://doi.org/10.4310/cms.2014.v12.n5.a7>.
- Fei, T., and K. Lerner, 1995, Elimination of numerical dispersion in finite-difference modeling and migration by flux-corrected transport: *Geophysics*, **60**, 1830–1842, <https://doi.org/10.1190/1.1443915>.
- Kalita, M., and T. Alkhalifah, 2017, Efficient full waveform inversion using the excitation representation of the source wavefield: *Geophysical Journal International*, **210**, 1581–1594, <https://doi.org/10.1093/gji/ggx214>.
- Soubaras, R., and R. Dowle, 2010, Variable-depth streamer – a broadband marine solution: *First Break*, 28.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266, <https://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, <https://doi.org/10.1190/1.3238367>.
- Yuan, Y. O., and F. J. Simons, 2014, Multiscale adjoint waveform-difference tomography using wavelets: *Geophysics*, **79**, no. 3, WA79–WA95, <https://doi.org/10.1190/geo2013-0383.1>.