Time-domain incomplete Gauss-Newton full-waveform inversion of Gulf of Mexico data

Abdullah AlTheyab, Xin Wang, Gerard T. Schuster, King Abdullah University of Science and Technology*

SUMMARY

We apply the incomplete Gauss-Newton full-waveform inversion (TDIGN-FWI) to Gulf of Mexico (GOM) data in the space-time domain. In our application, iterative least-squares reverse-time migration (LSRTM) is used to estimate the model update at each non-linear iteration, and the number of LSRTM iterations is progressively increased after each non-linear iteration. With this method, model updating along deep reflection wavepaths are automatically enhanced, which in turn improves imaging below the reach of diving-waves. The forward and adjoint operators are implemented in the space-time domain to simultaneously invert the data over a range of frequencies. A multiscale approach is used where higher frequencies are down-weighted significantly at early iterations, and gradually included in the inversion.

Synthetic data results demonstrate the effectiveness of reconstructing both the high- and low-wavenumber features in the model without relying on diving waves in the inversion. Results with Gulf of Mexico field data show a significantly improved migration image in both the shallow and deep sections.

INTRODUCTION

For deep subsurface imaging, waveform inversion (Tarantola, 1984a) should invert deeper reflections and later-arrival refractions. Unfortunately, standard FWI has low sensitivity to waveform residuals related to relatively weak deeper reflections compared to the stronger amplitude diving waves. The consequence is slow and often inadequate FWI convergence for reconstructing deep portions of the slowness model. To enhance the effectiveness of FWI for deeper reflections, we use a linear-inversion scheme instead of reverse-time-migration (RTM) (Baysal et al., 1983) for calculating slowness updates. Using this linear inversion, sharp boundaries are incorporated into the slowness model so that they implicitly enhance the model updating along the reflection wavepaths at subsequent iterations. This linear inversion is least-squares reverse-time migration (LSRTM) (Tarantola, 1984b; Plessix and Mulder, 2004; Dai et al., 2012).

Using LSRTM, a slowness-perturbation model is computed based on the Born approximation, where the background slowness is fixed during the linear inversion. The slowness model is then updated with the inverted slowness-perturbation model. After that, the linear inversion is repeated with the updated slowness model as a background slowness. Each linear-inversion and updating of the slowness model constitutes a non-linear iteration. This combined linear and non-linear inversion procedure is cyclically repeated until acceptable convergence.

The above procedure is a variation of Gauss-Newton optimization for FWI (Akcelik, 2002; Akcelik et al., 2002; Erlangga and Herrmann, 2009; Virieux and Operto, 2009). To avoid high computational and memory costs, the linear inversion is computed by an iterative conjugate gradient (CG) solver. The number of CG iterations is increased after each non-linear iteration, and is essential for an accurate model reconstruction. The algorithm is implemented in the time-space domain, and a multiscale approach is used to invert the data for a band of frequencies (Bunks et al., 1995; Boonyasiriwat et al., 2010); starting from a narrow band of low frequencies, and progressively including higher frequencies into the inversion.

In this paper, we review the algorithm, and illustrate its effectiveness on synthetic data that do not contain diving waves. Then, the inversion is applied to Gulf of Mexico (GOM) data. The resulting tomograms show significant improvements in the deeper section compared to the starting model.

THEORY

Newton's method (Pratt et al., 1998; Nocedal and Wright, 2006) for minimizing the residual difference $\mathbf{r}$ between the calculated and observed data can be written algebraically as

$$\mathbf{s}_{k+1} = \mathbf{s}_k - \mathbf{H}_f^{-1}(\mathbf{s}_k)\nabla f(\mathbf{s}_k), \tag{1}$$

where $\mathbf{s}_k$ is the slowness model, $\mathbf{H}_f$ is the Hessian matrix and $\nabla f(\mathbf{s}_k)$ is the gradient of the objective function $f(\mathbf{s}_k) = \frac{1}{2}\|\mathbf{r}(\mathbf{s}_k)\|_2^2$ at the $k-$th iteration. By approximating the Hessian as

$$\mathbf{H} \approx \left(\mathbf{J}^\dagger \mathbf{J}\right),$$

where $\mathbf{J}$ is the Jacobian matrix, we get the Gauss-Newton optimization formula

$$\mathbf{s}_{k+1} = \mathbf{s}_k - \alpha_k \left(\mathbf{J}_k^\dagger \mathbf{J}_k\right)^{-1} \mathbf{J}_k^\dagger \mathbf{r}_k. \tag{2}$$

A line search is used to estimate the step length α_k because the approximation of the Hessian might not be an accurate estimate of the curvature for the non-linear misfit function. Instead of inverting the Hessian matrix, we iteratively solve the overdeterimined system of equations

$$\mathbf{J}_k \mathbf{g}_k = \mathbf{r}_k, \tag{3}$$

using the same slowness model to get the search direction $\mathbf{g}$. In other words, LSRTM is used to compute the search direction $\mathbf{g}$ instead of RTM. Once the search direct direction $\mathbf{g}$ and the line-search parameter α are computed, the slowness model is updated using

$$\mathbf{s}_{k+1} = \mathbf{s}_k - \alpha_k \mathbf{g}_k, \tag{4}$$

and the Jacobian operator and the Hessian matrix are also updated according to the new slowness model. In the following section, we review the implementation of the Jacobian operator and its adjoint.

DOI http://dx.doi.org/10.1190/segam2013-1478.1

Time-domain implementation of the Jacobian and its adjoint

We follow a similar procedure to that of Dai et al. (2012) in deriving a time-domain implementation of applying the Jacobian matrix $\mathbf{J}$ to the slowness perturbation vector $\triangle\mathbf{s}$. We will describe our procedure for a single frequency, with the understanding that the method is fully implemented in the space-time domain. Each row of the matrix operation $\triangle\mathbf{p} = \mathbf{J}\triangle\mathbf{s}$ for calculating the wavefield perturbation $\triangle\mathbf{p}$ (indexed by receiver position $\mathbf{x}_r$ and frequency ω) from the slowness perturbation $\triangle\mathbf{s}$ (indexed by the spatial position $\mathbf{x}$) for a given source at position $\mathbf{x}_s$ with a source wavelet $q(\omega)$ is a discretization of the integral equation

$$\delta p(\mathbf{x}_r,\omega) = 2\int \omega^2 s_0(\mathbf{x})\,\delta s(\mathbf{x})\,p_0(\mathbf{x},\omega)\,G_0(\mathbf{x}|\mathbf{x}_r,\omega)\,d\mathbf{x}, \tag{5}$$

where ω is the frequency, s_0 is the background slowness, δs is the slowness perturbation, p_0 is the background incident wavefield from the source, and G_0 is the Green's function. This equation is the solution to the following system of partial differential equations

$$\left(\nabla^2+\omega^2 s_0^2\right)p_0(\mathbf{x},\omega) = -\delta(\mathbf{x}-\mathbf{x}_s)\,q(\omega), \tag{6}$$
$$\left(\nabla^2+\omega^2 s_0^2\right)\delta p(\mathbf{x},\omega) = -2\omega^2 s_0\delta s(\mathbf{x})\,p_0(\mathbf{x},\omega), \tag{7}$$

which indicate that we can evaluate the integral in equation 5 by having two wave-propagation simulations.

Similarly, for the adjoint operation $\triangle\mathbf{s} = \mathbf{J}^\dagger\triangle\mathbf{p}$ each row-vector multiplication is the discrete approximation to the integral

$$\delta s(\mathbf{x}) = 2\iint \omega^2 s_0(\mathbf{x})\,p_0(\mathbf{x},\omega)\times G_0(\mathbf{x}|\mathbf{x}_r,\omega)\,\delta p^*(\mathbf{x}_r,\omega)\,d\mathbf{x}_r d\omega. \tag{8}$$

To evaluate this integral, two wavefields are simultaneously simulated by solving the two wave equations:

$$\left(\nabla^2+\omega^2 s_0(\mathbf{x})^2\right)p_0(\mathbf{x},\omega) = -\delta(\mathbf{x}-\mathbf{x}_s)\,q(\omega), \tag{9}$$
$$\left(\nabla^2+\omega^2 s_0(\mathbf{x})^2\right)R^*(\mathbf{x},\omega) = -\delta(\mathbf{x}-\mathbf{x}_r)\,2\omega^2\delta p^*(\mathbf{x}_r,\omega). \tag{10}$$

The solution to equation 10 is

$$R^*(\mathbf{x},\omega) = 2\int \omega^2 G_0(\mathbf{x}|\mathbf{x}_r,\omega)\,\delta p^*(\mathbf{x}_r,\omega)\,d\mathbf{x}_r. \tag{11}$$

By taking the zero-lag correlation and scaling by the background slowness, we get the integral in equation 8, i.e

$$\delta s(\mathbf{x}) = \int s_0(\mathbf{x})\,R^*(\mathbf{x},\omega)\,p_0(\mathbf{x},\omega)\,d\omega =$$
$$2\iint \omega^2 s_0(\mathbf{x})\,p_0(\mathbf{x},\omega)\,G_0(\mathbf{x}|\mathbf{x}_r,\omega)\,\delta p^*(\mathbf{x}_r,\omega)\,d\mathbf{x}_r d\omega. \tag{12}$$

Similar to the forward modeling operator, the integral in equation 8 can be computed by conducting two wave-propagation simulations and applying the zero-lag cross-correlation to the two wavefields. The above equations are implemented in the space-time domain for the applications in the following sections.

Physical interpretation of TDIGN-FWI

In this section, we develop by example an intuitive reasoning that explains why TDIGN-FWI provides more accurate tomograms than standard FWI. Figure 1a shows a block-velocity model which contains one deeper reflector to constrain the shallow velocity anomalies. Because the deep layer has a slower velocity than the shallow layer, it will not generate refractions. Frequencies below 5 Hz are absent from the data as shown in Figure 1b, and the maximum source-receiver offset of 3 km is used for the synthetic data. The initial velocity model is homogeneous with a constant velocity of 2000 m/s. The shallow rectangular anomalies are larger in size than the minimum effective wavelength.

The block model does not generate refractions from the deep interface so that only the reflections from the deeper interface will be employed to reconstruct the square-shaped anomalies. FWI relies on reflections for reconstructing the shallow anomalies, and the velocity updates are attributed to the reflection wavepaths associated with the deep reflector and the boundaries of the anomalies. The construction of reflection wavepaths is dependent on the presence of sharp reflectors in the velocity model. Without the sharp boundaries, standard FWI fails to reconstruct the shallow anomalies as shown in Figure 1c.

LSRTM is known to focus reflections and diffractions into a sharp interfaces in the subsurface model. If such sharp interfaces are incorporated into the velocity model as is the case for TDIGN-FWI, these highly resolved reflectors and diffractors generate the wavepaths needed for reconstructing the shallow anomalies. As shown in Figure 1d, TDIGN-FWI make use of reflections, diffractions, multiples and prism waves to construct the anomalies and delineate the boundaries with high resolution. The mispositioning of the deeper reflector due to the shallow velocity error is reduced and the reconstructed reflector is nearly flat. In contrast, the reflectors in the standard FWI tomogram in Figure 1c are more distorted because the deeper reflections are not fully utilized for the same number of iterations. A prohibitively large number of iterations would be needed to accomplish the same results using a non-linear steepest descent optimization method.

APPLICATION TO GOM STREAMER DATASET

The TDIGN-FWI is applied streamer data from the Gulf of Mexico. There are 515 shots with a 37.5 meter shooting interval, and the source-receiver offsets are from 198 meters to 6 kilometers, with a 12.5 meter receiver spacing. The trace length is 10 seconds with a 2 ms sampling interval. Prior to inversion, the data spectra are filtered by $\sqrt{i/\omega}$ and gained by $\sqrt{t}$ in the time domain to transform 3D to 2D geometric spreading. The source wavelet is estimated by stacking early arrivals from the near-offset traces.

DOI http://dx.doi.org/10.1190/segam2013-1478.1

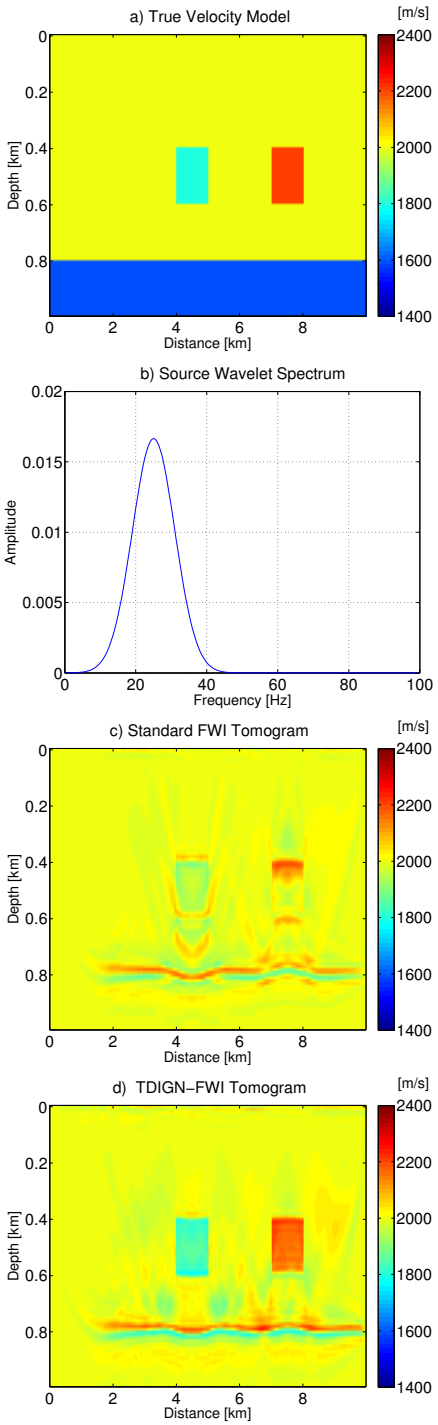

Figure 1: a) Test velocity model, b) the source wavelet spectrum for forward-modeling the synthetic data and inversion, c) steepest descent FWI and d) TDIGN-FWI tomograms. The starting velocity model is a constant velocity model with the velocity 2000 m/s, and both of the inversion results have the same computational cost.

We start the inversion with the data bandpass-filtered from 0-4 Hz, because there is reliable signal at 4 Hz. At later iterations, we widened the band of data frequencies to 10 Hz, and Figure 3 shows the initial velocity model and the final tomogram. The grid size for the tomograms is 301 by 1600 grid points in the vertical and horizontal directions, respectively, and the grid spacing is 12.5 meters. Figure 2 shows the convergence curve after the last reset, where the data residual decreases by more than 60 percent.

Figure 4 shows the migration images using the initial velocity model and the TDIGN-FWI tomogram. The TDIGN-FWI image is more focused. The spliced common-image-gathers (CIG) are flat in the final image, while the CIG's are not flat for the initial image. This highlights the significant improvement to the velocity model. Figure 5 shows two shot gathers from the observed and calculated data and the match is generally good for early arrivals and most of the deeper reflections.

LIMITATIONS AND FUTURE IMPROVEMENTS

A problem with our approach is that the density is assumed to be a constant so that the TDIGN-FWI will introduce sharp velocity boundaries with the wrong velocity values. Those boundaries still help in updating the background velocity, which will improve the migration image. Such sharp boundaries can be removed before applying an FWI algorithm which inverts for more subsurface parameters than the acoustic velocity.

We chose to start the inversion with two LSRTM iterations and increase the number of LSRTM iterations by one for every non-linear iteration. Our choice is heuristic based on tests with synthetic data. More work is needed for choosing optimal TDIGN-FWI inversion parameters.

CONCLUSION

We implemented and applied the TDIGN-FWI to a GOM dataset. The algorithm uses the LSRTM images as the slowness updates instead of the RTM images. The TDIGN-FWI uses the deep reflection data to define sharp boundaries in the velocity model. Those sharp boundaries generate wavepaths that are used by the inversion to build velocity updates for the deeper section. The definition of sharp boundaries and using them in calculating slowness updates are implicit within the algorithm. As a result, the quality of the migration images computed with the TDIGN-FWI tomogram appears to be highly resolved at both the shallow and deeper parts.

DOI http://dx.doi.org/10.1190/segam2013-1478.1

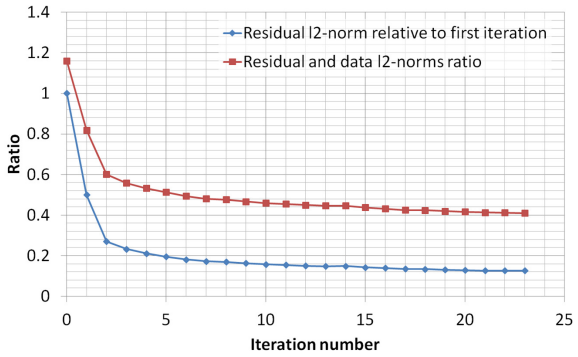

Figure 2: The convergence curves for TDIGN-FWI applied to 0-10 Hz GOM data.

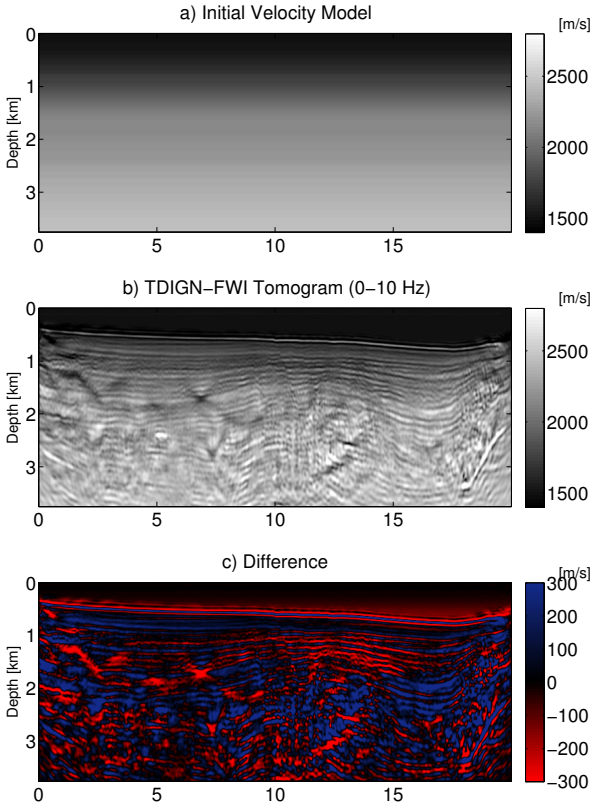

Figure 3: a) the initial velocity model, and b) final TDIGN-FWI tomogram of the band 0-10 Hz.

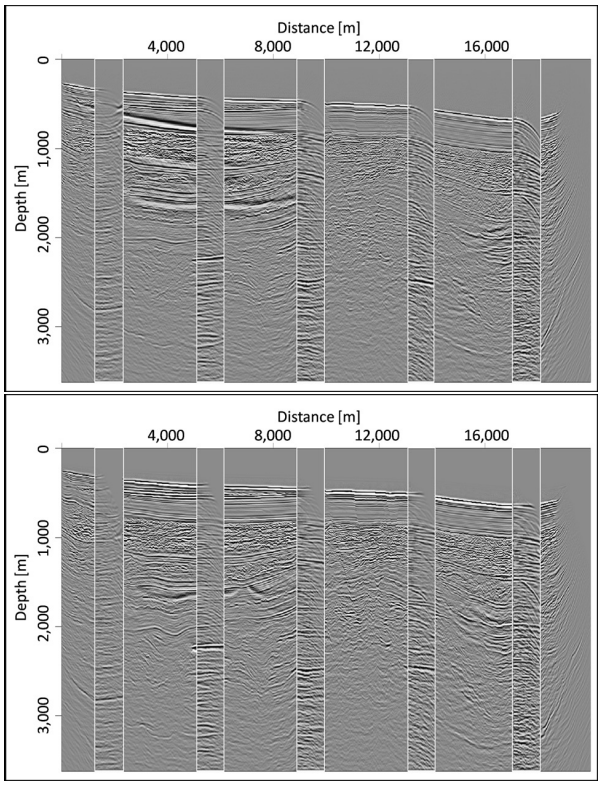

Figure 4: Kirchhoff migration images using the initial velocity model (upper) and the final velocity model (lower). The spliced narrow panels are common image gathers.

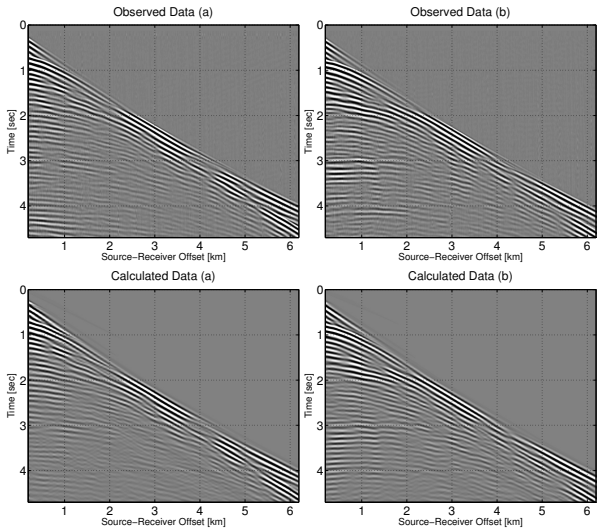

Figure 5: 0-10 Hz shot gathers from different parts of the survey. The observed data are shown in the top panels and the corresponding calculated data at the bottom.

DOI http://dx.doi.org/10.1190/segam2013-1478.1

http://dx.doi.org/10.1190/segam2013-1478.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2013 SEG Technical Program Expanded Abstracts have been copy edited 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

Akcelik, V., 2002, Multiscale Newton–Krylov methods for inverse acoustic wave propagation: Ph.D. dissertation, Carnegie Mellon University.

Akcelik, V., G. Biros, and O. Ghattas, 2002, Parallel multiscale Gauss–Newton–Krylov methods for inverse wave propagation: Proceedings of the ACM/IEEE Conference on Supercomputing.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524, http://dx.doi.org/10.1190/1.1441434.

Boonyasiriwat, C., G. Schuster, P. Valasek, and W. Cao, 2010, Applications of multiscale waveform inversion to marine data using a flooding technique and dynamic early-arrival windows: Geophysics, **75**, no. 6, R129–R136, http://dx.doi.org/10.1190/1.3507237.

Bunks, C., F. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473, http://dx.doi.org/10.1190/1.1443880.

Dai, W., P. Fowler, and G. T. Schuster, 2012, Multisource least-squares reverse time migration: Geophysical Prospecting, **60**, 681–695, http://dx.doi.org/10.1111/j.1365-2478.2012.01092.x.

Erlangga, Y. A., and F. J. Herrmann, 2009, Seismic waveform inversion with Gauss-Newton-Krylov method: 79th Annual International Meeting, SEG, Expanded Abstracts, 2357–2361.

Nocedal, J., and S. J. Wright, 2006, Numerical optimization, 2nd ed.: Springer.

Plessix, R.-E., and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, **157**, 975–987, http://dx.doi.org/10.1111/j.1365-246X.2004.02282.x.

Pratt, G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency–space seismic waveform inversion: Geophysical Journal International, **133**, 341–362, http://dx.doi.org/10.1046/j.1365-246X.1998.00498.x.

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

Tarantola, A., 1984b, Linearized inversion of seismic reflection data: Geophysical Prospecting, **32**, 998–1015, http://dx.doi.org/10.1111/j.1365-2478.1984.tb00751.x.

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.