**Multi-source Waveform Inversion with Deblurring**

Ge Zhan and Gerard T. Schuster

**ABSTRACT**

The theory of preconditioned multi-source waveform inversion is presented where many shot gathers are simultaneously back-propagated to form the gradient of the misfit function. The implication is that, relative to standard waveform inversion, an order of magnitude increase in computational efficiency can be achieved by multi-source waveform inversion.

**INTRODUCTION**

Time-domain waveform inversion has the potential to provide estimates of velocity models with significantly higher resolution compared to traveltime tomography. However, waveform inversion is computer intensive due to the multiple iterations of forward modeling and residual wavefield back-propagation. As a partial remedy to the expense of reverse time migration (RTM), Morton (1998) proposed phase-encoding of shot records to simultaneously migrate a number of shot gathers within a single migration. This results in an increase in computational efficiency but the penalties are additional noise in the misfit gradient and inaccuracy in the inverted velocity model (Romero et al., 2000). In this procedure, each shot gather is encoded with a unique random time series and the result is summed together to form an encoded multi-source gather. Here, the unique time series assigned to a shot gather is approximately orthogonal to any of the other random time series. In theory, only a single phase-encoded back-propagation operation should be needed to generate the misfit gradient for velocity updating. The problem is that a phase-encoded finite-difference (FD) simulation with insufficient temporal duration yields noticeable artifacts in the misfit gradient and so it is not widely adopted in the industry.

To overcome this limitation, we develop an encoded multi-source deblurring filter to limit these cross-terms. Recent work by Aoki (2008), Aoki and Schuster (2008) and Dai and Schuster (2009) have shown that the use of deblurring filters as preconditioners in migration deconvolution (MD) and least squares migration (LSM) reduces migration artifacts and accelerates convergence. Here we successfully apply it to multi-source waveform inversion to provide a more accurate misfit gradient with fewer artifacts and thus accelerate the inversion process. Synthetic tests on 2D Marmousi model show that multi-source waveform inversion with an encoded multi-source deblurring filter can provide nearly the same result as single-source waveform inversion, but with an order of magnitude increase in computational efficiency.

This paper is organized into three sections. First, the theory of multi-source waveform inversion is introduced followed by an application of the encoded multi-source deblurring filter. Then the multi-source waveform results are compared with single-source inversion results using synthetic Marmousi data. Finally, a summary is presented.

**THEORY**

Waveform inversion updates the 2D velocity model \( V(x, z) \) by matching the calculated seismograms \( P_{\text{cal}}(s, r, \omega) \) to the observed seismograms \( P_{\text{obs}}(s, r, \omega) \), where \( s \) and \( r \) denote the source and receiver vectors, respectively. This can be accomplished by minimizing the waveform misfit function (Lailly, 1983; Tarantola, 1984):

\[
f = \frac{1}{2} \sum_{\omega} \sum_{s} \sum_{r} \| P_{\text{obs}}(s, r, \omega) - P_{\text{cal}}(s, r, \omega) \|^2,
\]

where summations are over source and receiver locations and over the frequency variable \( \omega \). Instead of minimizing the above misfit function, we incorporate smoothness into the inversion by adding 2nd-order spatial derivatives of the velocity to the misfit function

\[
f = \frac{1}{2} \sum_{\omega} \sum_{s} \sum_{r} \| P_{\text{obs}}(s, r, \omega) - P_{\text{cal}}(s, r, \omega) \|^2 + \alpha \| \frac{\partial^2 V(x, z)}{\partial x^2} \|^2 + \beta \| \frac{\partial^2 V(x, z)}{\partial z^2} \|^2 + \gamma \| V(x, z) - V_{\text{appr}} \|,
\]

where \( \alpha \), \( \beta \), and \( \gamma \) are regularization parameters.
where $\alpha$ and $\beta$ are regularization coefficients, the purpose of which is to control the trade-off between model smoothness and data fit (Scales et al., 1990). Here $\gamma \| V(x, z) - V_{\text{apriori}} \|$ is the regularization term that penalizes estimates of velocity models that deviate far from the apriori velocity model $V_{\text{apriori}}$.

Waveform inversion is summarized in the following three main steps (Vigh et al., 2009). First, compute the waveform residual by calculating the difference between the observed data and the calculated data using the current velocity model. The residual data exhibits the accuracy of the current model. Second, cross-correlating the backpropagated residual wavefield with the corresponding forward modeled source wavefield at each time step and summing over all time steps yields the misfit gradient. Finally, update the velocity model by using the misfit gradient in a non-linear iterative method. The amplitude of the misfit gradient at each spatial point is proportional to the velocity change. To avoid the local minima, a multiscale approach is implemented (Bunks et al., 1995; Sirgue and Pratt, 2004; Boonyasiriwat et al., 2008).

The misfit gradient calculation for waveform inversion is similar to RTM, where we can represent the misfit gradient as

$$ g(x, z) = \sum_{\omega} S^*(x, z, \omega) R(x, z, \omega), \quad (3) $$

where $S(x, z, \omega)$ and $R(x, z, \omega)$ represent the source and residual wavefields, respectively; $g(x, z)$ is the misfit gradient at $(x, z)$, and $*$ represents the complex conjugate.

There is tremendous potential for computational speedup if we can perform the summation over all (or partial) shot gathers before applying this imaging condition. That is, the composite (or multi-source) wavefields are defined as

$$ \tilde{S}(x, z, \omega) = \sum_{j=1}^{N} a_j(\omega) S_j(x, z, \omega), \quad (4) $$

and

$$ \tilde{R}(x, z, \omega) = \sum_{j=1}^{N} a_j(\omega) R_j(x, z, \omega), \quad (5) $$

where $N$ is the number of shot combined together, and $a_j$ is the phase-encoding factor. However, this approach breaks down when we insert equations 4 and 5 into equation 3:

$$ \tilde{g}(x, z) = \sum_{\omega} \tilde{S}^*(x, z, \omega) \tilde{R}(x, z, \omega) $$

$$ = \sum_{j=1}^{N} \sum_{\omega} |a_j(\omega)|^2 S_j^*(x, z, \omega) R_j(x, z, \omega) + \sum_{j \neq k}^{N} \sum_{\omega} a_j^*(\omega) a_k(\omega) S_j^*(x, z, \omega) R_k(x, z, \omega). \quad (6) $$

If the phase-encoding factors are orthogonal (i.e., $a_j^* a_k = \delta_{jk}$), then the first summation in equation 6 reduces to the correct misfit gradient (equation 3). However, phase-encoding terms are typically not orthogonal so the unwanted $j \neq k$ cross-terms (the second term in equation 6) are unphysical cross-correlations between unrelated source and residual wavefields. If these cross-talk terms are strong enough then they make the migration result (i.e., the misfit gradient) unacceptable.

In this paper, we partly overcome the cross-talk problem by applying a multi-source preconditioner to the multisource misfit gradient. The multi-source shot gather is composed of a sum of single shot gathers with random time delays. For our tests, the preconditioned cross-terms become mostly uncorrelated with one another for a multi-source shot gather composed of about 12 single shot gatherers. A large physical separation between shot locations can further reduce the cross-correlations between unrelated wavefields in the multi-source shot gather (Morton, 1998).

**NUMERICAL RESULTS**

Synthetic shot gathers associated with the 2D Marmousi model are used in our synthetic test. As shown in Figure 1a, the Marmousi model shows a very complicated geological structure, which is a challenge for waveform inversion.

The modeled 2D acoustic data are generated for 192 shot gathers and 192 receivers with the same shot interval and receiver interval of 5 m. These data were generated with a 30-Hz Ricker wavelet and the starting velocity model shown in Figure 1b is a smoothed version of the true velocity model. We only invert for the velocity distribution while relating the density to the velocity using Gardner’s equation (Gardner et al., 1974). First, we apply the multi-scale single-source waveform inversion which uses the single-source data compared to the multi-source data. The inverted velocity image after 50 iterations is shown in Figure 1c, and the misfit gradient calculated with the starting model is shown in Figure 2. We will consider it as a template for comparison with multi-source waveform inversion results.

To generate the multi-source data, we first apply a random time delay to all 192 shot gatherers, then combine each of two unique shot gatherers with sources that are widely spaced as possible. For example, shot gatherers 1 and 97 (out of 192) are randomly delayed and stacked together to produce one multi-source gather. This is followed with shot gatherers 2 and 98, and so on, up to shot gatherers 96 and 192, producing a total of 96 multi-source gatherers from the original 192 single-source shot gatherers. When three shot gatherers are combined together, shot gatherers 1, 65 and 129 are stacked together and this is repeated until there are a total of 64 multi-source gatherers.

Combining two separate shot gatherers for waveform inversion, we could theoretically reduce the computation cost by half. A further savings in run time can be ob-
Multi-source Waveform Inversion

Figure 1: The (a) 2D Marmousi model and (b) the starting model for the waveform inversion. (c) The inverted velocity using the starting model after 50 iterations.

Figure 2: The first calculated misfit gradient using the starting model in single-source waveform inversion.

Figure 3: The $L_2$ norm difference of gradients for different numbers of shot gathers.

Figure 4: Relative Difference vs. Number of shots

Unpreconditioned Multi-source WFI

Figure 3 shows the plot of the $L_2$ norm of the difference between the misfit gradient calculated from single-source waveform inversion and the misfit gradient derived from several randomly delayed shot gathers combined. The number of shot gathers combined is also the factor by which the computational cost is reduced.

Figure 3 demonstrates that the greater the number of combined shot gathers, the larger the artifacts or cross-terms generated in the misfit gradient. This is also demonstrated in Figure 4, which compares misfit gradients using different kinds of multi-source gathers. On the left are shown misfit gradients calculated with a different number of shot combined together using the same starting model (Figure 1b) from the first iteration of waveform inversion. On the right are the differences between these misfit gradients and the misfit gradient calculated from the single-source waveform inversion. All plots are displayed at the same scale. Artifacts are not obvious when only two shot gathers separated by a large distance are combined. Unfortunately, as the number of shot increases, the artifacts or cross-terms are as large as the misfit gradient itself, as shown in Figure 4d when we combine all 192 shot gathers together.
Figure 4: The misfit gradients calculated from the starting model combining (a) 2, (b) 12, (c) 48, and (d) 192 shot gathers. The left column is the misfit gradient, the right column is the difference between the multi-source and single-source misfit gradients.
Preconditioned Multi-source WFI

The artifacts in the misfit gradient are an enemy to the waveform inversion process. If the misfit gradient with such noise is directly used to update the velocity, the waveform inversion process will attempt to incorrectly change the inverted attribute so that the modeled data matches the observed data in a least squares sense (Vigh et al., 2009). So it is necessary to remove or attenuate these artifacts from the misfit gradient before updating the velocity. In this regard, we use an encoded multi-source deblurring filter (Aoki, 2008; Aoki and Schuster, 2008; Dai and Schuster, 2009; Schuster, 2009) as a preconditioner in the multi-source waveform inversion process. The theory for this filter is presented in the Appendix. We also use regularization to reduce artifacts which incorporates smoothness into the inversion (equation 2).

Here, we only investigate the multi-source waveform inversion case which combines 12 multiple shot gathers. The encoded multi-source deblurring filter is used and updated at every fifth iteration during waveform inversion in order to suppress both migration artifacts and cross-terms in the misfit gradient caused by cross-correlations between the unrelated source and residual wavefields. Figure 5 shows the comparison between misfit gradients before and after applying this filter. The misfit gradient after filtering is less contaminated by the artifacts in Figure 5b compared to Figure 5a. Here the $L_2$ norm difference on the right is reduced from 19.5% to 7.1% compared to the single-source misfit gradient.

Figure 6: Waveform inversion result using (a) single-source shot gathers and (b) multi-source shot gathers. Both of them are after 50 iterations.

Figure 6b shows the final inverted velocity after 50 iterations from the multi-source waveform inversion. Compared to the single-source inversion result (Figure 6a) after the same iteration, they appear to be quite similar and both are very close to the true model. This is verified by Figure 7 which compares vertical slices at selected locations. From this comparison, we can see that with the encoded multi-source deblurring filter, the multi-source waveform inversion provides nearly the same result as single-source waveform inversion. Both of them match well with the true velocity. However, we gain a significant speedup in computation time, which is almost equal to the number of shot gathers combined (i.e., 12).

The plot of the multi-source waveform residual versus iteration number is shown in Figure 8. This residual curve is a good indicator of convergence rate and shows that data residual decreased substantially with iteration number.

Figure 8: The normalized waveform residuals versus iterations.

CONCLUSIONS

We introduced the theory of multi-source preconditioner for accelerating the convergence rate of iterative multi-source waveform inversion. Results with the 2D Marmousi model showed that preconditioned misfit gradients with inputs of 12 delayed shot gathers can estimate the velocity model with nearly the same accuracy as standard waveform inversion, but with 1/12 the computational cost. The Marmousi model has an acquisition geometry and a complexity that is similar to that seen in real earth models, so this suggests a similar computational efficiency for application to real data sets. For 3D models, there is an extra spatial dimension where we can combine delayed shot gathers so this suggests the possibility of even a greater speedup in convergence rate.

ACKNOWLEDGMENTS

We thank the 2009 sponsors of the University of Utah Tomography and Model/Migration (UTAM) Consortium for their support.
Figure 5: (a) The misfit gradient calculated from the starting model combining 12 shot gathers (the same as Figure 5b). (b) The misfit gradient after applying the encoded multi-source deblurring filter. On the right are the differences between the left and Figure 2. They are 19.5% and 7.1% respectively.

Figure 7: Comparison of vertical slices between the true velocity and inverted velocities at different locations.
APPENDIX: ENCODED MULTI-SOURCE DEBLURRING FILTER

The phase-encoded multi-source deblurring filter (Aoki, 2008; Aoki and Schuster, 2008; Dai and Schuster, 2009; Schuster, 2009) is constructed in the following way.

1. Take the migration velocity model as the background model without scatterers, and decompose the model into a checkerboard of subsections with a point scatterer at the center of each subsection (see Figure 9b).

2. Place sources along the surface of the model and compute the encoded multi-source gather of traces with one FD simulation to get $d$ in Figure 9c.

3. Subtract $d$ from the encoded multi-source FD simulation in the background model without point scatterers. Subtract the resulting data from $d$ and set this equal to the new $d$.

4. Apply encoded multi-source migration operator $L^T$ to $d$ to get $L^T d = m_{mig}$, as shown in Figure 9d.

5. As shown in Figure 9d-e, find the local series of deblurring filters $f_i$ that collapse the migration butterflies (Hu and Schuster, 2001) to the points in Figure 9d. In this case there will be six different filters $f_i$; $i\in\{1, 2, 3, 4, 5, 6\}$, each with dimension $K \times L$. For example, Aoki (2008) uses $K=5$ and $L$ to be equal to the number of points that represent two wavelengths in the 2D case. Each $f_i$ approximates the inverse to one of the migration butterflies shown in Figure 9d, and are assembled into the deblurring matrix $F$ shown in Figure 9f. The assumption is that the $f_i$ computed for one point scatterer in a subsection is valid for any location of a point scatterer in that subsection. This means that the computation of the multi-source deblurring filter is affordable, but is achieved at a loss of accuracy. The accuracy loss can be regained by the iterations in the MD or LSM algorithm.

REFERENCES

Figure 9: Steps for computing the encoded deblurring filter $F$ (Schuster, 2009).