ACOUSTIC MULTI-SOURCE FULL WAVEFORM INVERSION WITH DEBLURRING

GE ZHAN, WEI DAI, CHAIWOOT BOONYASIRIWAT and GERARD T. SCHUSTER

Department of Geology and Geophysics, University of Utah, Salt Lake City, UT 84112, U.S.A.
ge.zhan@utah.edu

(Received November 2, 2011; revised version accepted September 8, 2013)

ABSTRACT

Zhan, G., Dai, W., Boonyasiriwat, C. and Schuster, G.T., 2013. Acoustic multi-source full waveform inversion with deblurring. *Journal of Seismic Exploration*, 22: 477-488.

The theory of preconditioned multi-source full waveform inversion (FWI) is presented where many shot gathers are simultaneously back-propagated to form the multi-source gradient of the misfit function. Synthetic tests on the 2D Marmousi data set show that multi-source full waveform inversion using an encoded multi-source deblurring filter as a preconditioner can provide an accurate velocity model at 1/100 the computational cost of conventional FWI.

KEY WORDS: full waveform inversion, multi-source, preconditioning, deblurring.

INTRODUCTION

Acoustic full waveform inversion (FWI) has the potential to provide estimates of velocity models with significantly higher resolution compared to traveltime tomography. However, FWI 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 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).

0963-0651/13/$5.00

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 the crosstalk noise. Recent work by Aoki and Schuster (2009) and Dai et al. (2011) 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 FWI to provide a more accurate misfit gradient with fewer artifacts and thus accelerate the inversion process. Synthetic tests on the 2D Marmousi model (Martin et al., 2006) shows that multi-source FWI with an encoded multisource deblurring filter can provide a good inversion result and reduce the computational time by two orders of magnitude.

The first part of this paper presents the theory of multi-source FWI, and the second part presents numerical results. The multi-source waveform results are obtained using synthetic data associated with the 2D Marmousi model. The encoded multi-source deblurring filter is applied during the inversion, and the theory of it is shown in Appendix. After that a summary is presented.

THEORY

Full waveform inversion updates the 2D velocity model V(x,z) by matching the calculated seismograms $P_{cal}(\mathbf{s},\mathbf{r},\omega)$ to the observed seismograms $P_{obs}(\mathbf{s},\mathbf{r},\omega)$, 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)

$$f = \tfrac{1}{2} \sum_{\omega} \sum_{\mathbf{s}} \sum_{\mathbf{r}} \| P_{obs}(\mathbf{s},\mathbf{r},\omega) - P_{cal}(\mathbf{s},\mathbf{r},\omega) \|^2 \quad , \tag{1}$$

where summations are over source and receiver locations and over the frequency variable ω. In this paper, all the wavefield propagations are implemented in the time domain, but for simplicity, we show the equations in the frequency domain.

The full waveform inversion is summarized in the following three main steps (Vigh et al., 2009). First, compute the waveform residual by calculating the residual vector, which is the difference between the observed data and the calculated data using the current velocity model. Second, cross-correlate the back-propagated 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.

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

$$g(x,z) = \sum_{\omega} S^{*}(x,z,\omega)R(x,z,\omega) \quad , \tag{2}$$

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

$$\tilde{S}(x,z,\omega) = \sum_{j=1}^{N} a_j(\omega)S_j(x,z,\omega) \quad , \tag{3}$$

and

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

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 eqs. (3) and (4) into eq. (2)

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

$$= \sum_{j=1}^{N} \sum_{\omega} |a_j(\omega)|^2 S_j^{*}(x,z,\omega)\tilde{R}_j(x,z,\omega)$$

$$+\sum_{j\neq k}^{N}\sum_{k=1}\sum_{\omega}\overbrace{a_j^*(\omega)a_k(\omega)S_j^*(x,z,\omega)\tilde{R}_k(x,z,\omega)}^{\text{cross-talk}} \quad . \tag{5}$$

If the phase-encoding factors are orthogonal, i.e.,

$$\sum_{j}\sum_{k} a_j^*(\omega)a_k(\omega) = \delta_{jk} \quad . \tag{5}$$

then the first summation in eq. (5) reduces to the correct misfit gradient [eq. (3)]. However, phase-encoding terms are typically not orthogonal so the unwanted $j \neq k$ cross-terms [the second term in eq. (5)] 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 multi-source misfit gradient. The multi-source supergather is composed of a sum of single shot gathers with random time delays. That is, the phase encoding for each trace is a simple time shift rather than white noise series.

NUMERICAL RESULTS

Synthetic shot gathers associated with the 2D Marmousi model are used in our numerical tests. For the purpose of efficiency, the original Marmousi model was subsampled from 2301 × 751 cells at 4 m cell size to 801 × 151 cells at 20 m cell size. As shown in Fig. 1a, the modified Marmousi model shows a very complicated geological structure, which is a challenge for FWI.

The modeled 2D acoustic data are generated for 8 multi-source supergathers. In each supergather, 100 sources with a shot position interval of 160 m were shot simultaneously with a random time delay using one FD simulation. There are 800 fixed-spread receivers on the surface and the receiver interval is 20 m. These data were generated with a 7.5 Hz Ricker wavelet. A free surface boundary condition was used at the top of the model and perfectly matched layer boundary conditions were used at other three sides during forward modeling. The starting velocity model shown in Fig. 1b was created by smoothing the slowness model using a moving average filter, where the window size is 20 × 20 samples. We only invert for the velocity distribution while relating the density to the velocity using Gardner's equation (Gardner et al., 1974).

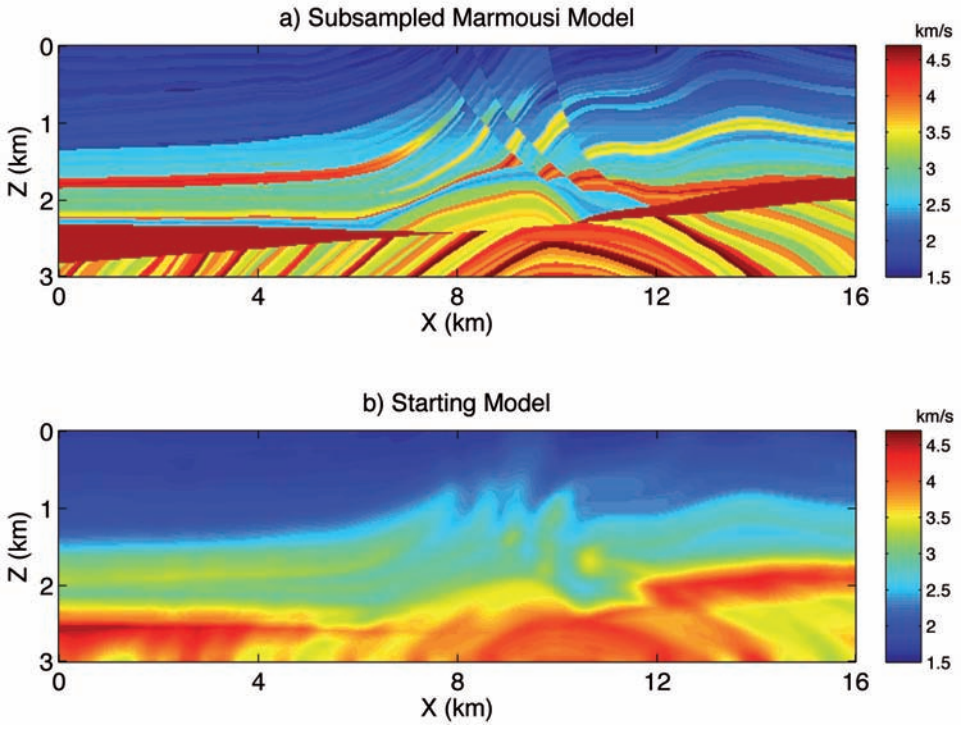

Fig. 1. The (a) 2D Marmousi velocity model and (b) the starting velocity model for the FWI.

The artifacts in the misfit gradient are enemies to the FWI process. If the misfit gradient with such noise is directly used to update the velocity, the FWI 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 and Schuster, 2009; Zhan et al., 2010; Dai et al., 2011) as a preconditioner in the multi-source FWI process.

Before starting FWI, we first compute the phase-encoded multi-source deblurring filter, which is constructed in the following way.

1. Decompose the model space into a checkerboard of subsections with a point scatterer at the center of each subsection (Fig. 2a).

2. Take the starting velocity model shown in Fig. 1b as the background velocity model. Place sources along the surface of the model and compute the encoded multi-source supergathers generated by these point scatterers shown in Fig. 2a with one FD simulation.

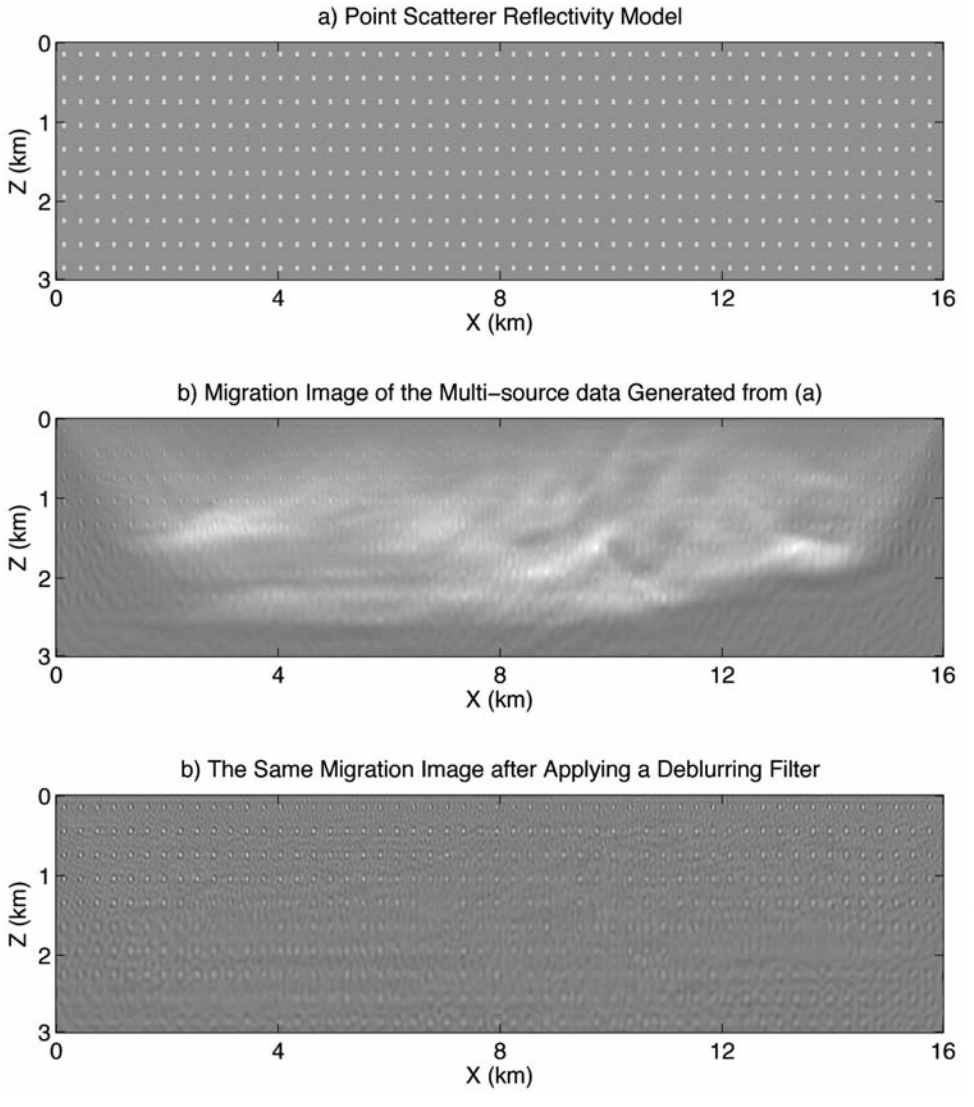

Fig. 2. Procedure for constructing the deblurring filter: a) Decompose the model into a checkerboard of subsections with a point scatterer at the center of each subsection. b) Take the starting model shown in Fig. 1b as the background model. Place sources along the surface of the model and compute the encoded multi-source gathers with one FD simulation to get the multi-source data. Migrate the multi-source data to get the migration image of these point scatterers. c) In every subsection, find the local series of deblurring filters that collapse the migration butterflies (Hu and Schuster, 2001) to the points in Fig. 2a. Then apply these deblurring filters to image Fig. 2b to get the deblurred migration image (Zhan et al., 2010).

3. Migrate the supergathers to get the migration image of these point scatterers (Fig. 2b).

4. Find the local deblurring filters by matching Fig. 2a and Fig. 2b in each subsection. Fig. 2c shows the migration image after applying such local filters to Fig. 2b subsection by subsection.

In the multi-source FWI, we simultaneously back-propagate the 100 shot gathers in a supergather to form the multi-source gradient of the misfit function. After getting the first gradient before iterations (Fig. 3a), we apply the deblurring filter to this gradient for both collapsing migration butterflies and attenuating unrelated crosstalk noise caused by multi-source inversion. Fig. 3 shows the comparison between the misfit gradients before and after applying this filter. The filtered misfit gradient shown in Fig. 3b is less contaminated by the artifacts compared to Fig. 3a. With the deblurring filter, we get a better gradient with less noise which results in a more accurate update of the velocity.

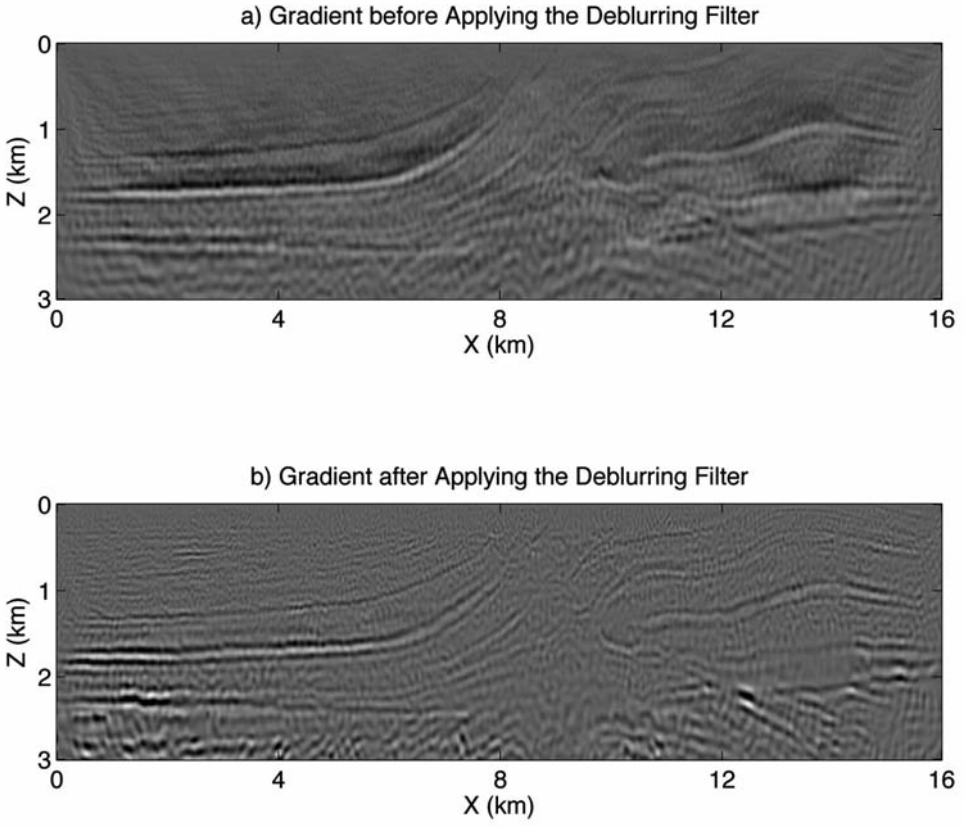

Fig. 3. a) The first calculated gradient a) before and b) after applying the deblurring filter.

The encoded multi-source deblurring filter is then updated and applied at every fifth iteration during the first 30 iterations of the multi-source FWI in order to suppress both migration artifacts and cross-terms in the misfit gradient. These cross-terms are caused by cross-correlations between the unrelated source and residual wavefields. A total of 7 deblurring filters is computed and applied for the inversion. There are three reasons why we stop using the deblurring filter after thirty iterations. The first and the most important thing is that the deblurring filter employed in our case is sensitive to the background velocity model. The smoother the background velocity model is, the better the filtered gradient will be. After 30 iterations, the updated velocity model becomes much sharper compared to the smoothed starting model. Second, the deblurring filter not only removes migration artifacts, but also introduces some noise as well. If we use these deblurring filters many times, this kind of noise will accumulate and finally harm the velocity updating in later iterations. Finally, the computational cost for constructing the deblurring filter is not negligible. The cost of constructing a deblurring filter equals that of a multi-source forward modeling plus a RTM of the model with the same size. If we keep using it for all iterations, the computational cost of the deblurring filter can not be ignored.

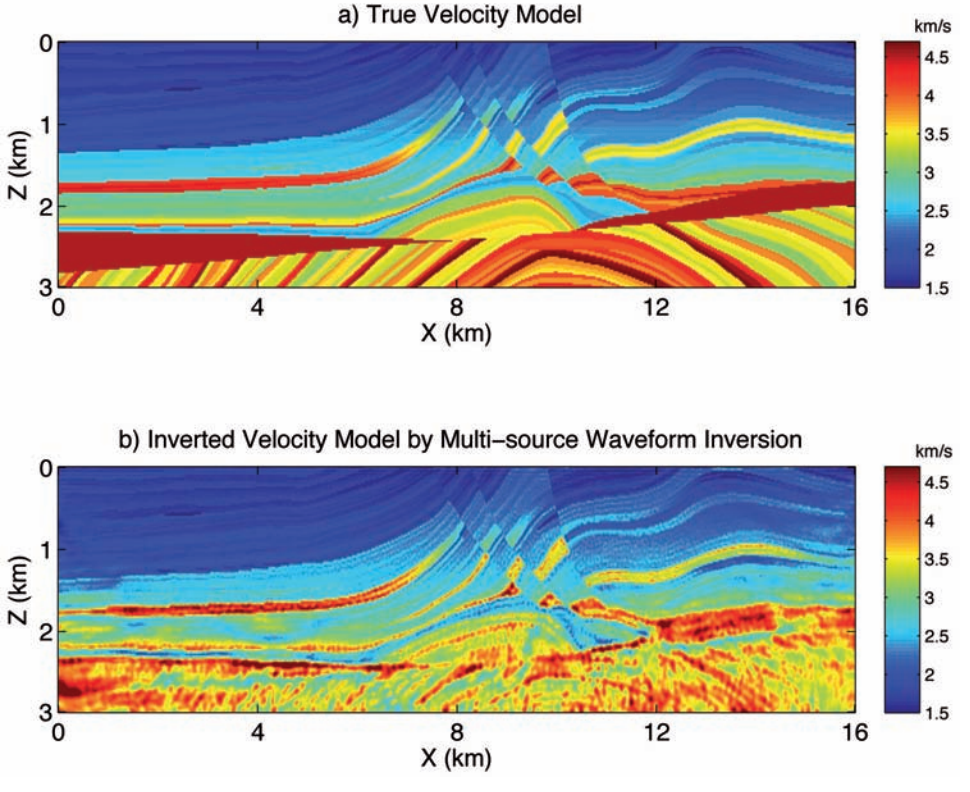

Fig. 4. a) The same velocity model as Fig. 1a. b) The multi-source FWI result after 300 iterations.

Fig. 4b shows the final inverted velocity after 300 iterations of inverting 8 supergathers. Each supergather is a 100-fold gather formed by stacking 100 shot gathers with random source statics. Compared to the true velocity model (Fig. 4a), the inverted one matches the true one quite well especially above 2 km. Importantly, we gain a significant speedup in computation time, which is almost equal to the number of shot gathers combined (i.e., 100). However, the lower part of the tomogram is not as accurate as the upper part. Thus the deblurring filter is more powerful in the upper part and less effective in the lower part. Besides, the lower part of the starting model (Fig. 1b) is too far away from the true one, we believe that the lower part of the inverted velocity should be improved a lot if the initial model is closer to the true model. Another reason might be that Schuster et al.'s (2011) analytic formula predicts stronger crosstalk with deeper parts of the image.

The plot of the multi-source waveform residual versus iteration number is shown in Fig. 5. From the comparison of the residual curves with and without applying the deblurring filter, we conclude that the deblurring filter accelerates convergence in the early iterations. This residual curve is also a good indicator of convergence rate of the multi-source FWI and shows that data residual decreased substantially with iteration number.

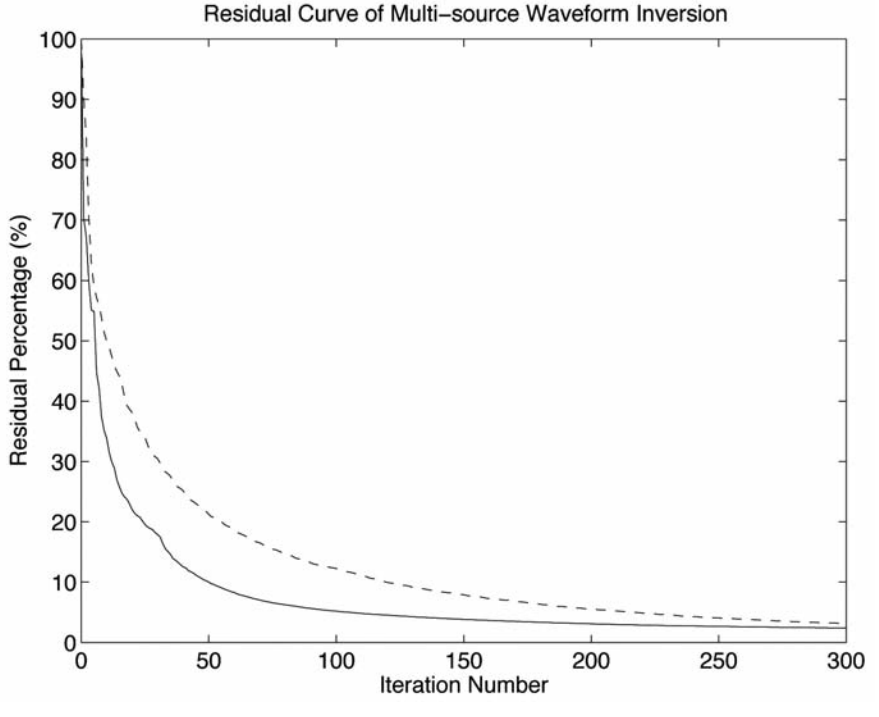

Fig. 5. The normalized waveform residuals versus iterations. The dashed one is the residual curve without the deblurring filter. The solid one is the residual curve with the deblurring filter.

CONCLUSIONS

We introduce the theory of multi-source full waveform inversion with multi-source preconditioners. Results with the 2D Marmousi model show that preconditioned misfit gradients with inputs of 100 delayed shot gathers can estimate the velocity model with good accuracy at 1/100 the computational cost of conventional FWI. 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. To decrease noise, a multi-scale approach is suggested by decreasing a multi-source gather's fold with increasing iteration number.

ACKNOWLEDGMENTS

We thank the 2009 sponsors of the University of Utah Tomography and Model/Migration (UTAM) Consortium for their support. We appreciate the reviews by André Kurzmann, Stefan Jetschny, and Jürgen Mann, who made a number of helpful suggestions that improved the quality of our manuscript.

REFERENCES

Aoki, N. and Schuster, G., 2009. Fast least-squares migration with a deblurring filter. Geophysics, 74: WCA83-WCA93.

Dai, W., Wang, X. and Schuster, G., 2011. Least-squares migration of multisource data with a deblurring filter. Geophysics, 76: R135-R146.

Gardner, G.H.F., Gardner, L.W. and Gregory, A.R., 1974. Formation velocity and density - the diagnostic basis of stratigraphic traps. Geophysics, 39: 770-780.

Hu, J. and Schuster, G.T., 2001. Poststack migration deconvolution. Geophysics, 66: 939-952.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations. SIAM Philadelphia Conf. Inverse Scattering: Theory and Application: 206-220.

Martin, G., Wiley, R. and Marfurt, K., 2006. Marmousi2: An elastic upgrade for marmousi. The Leading Edge, 25: 156-166.

Morton, S.A., 1998. Faster shot-record depth migrations using phase encoding. Expanded Abstr., 68th Ann. Internat. SEG Mtg., New Orleans: 1131-1134.

Romero, L.A., Ghiglia, D.C., Ober, C.C. and Morton, S.A., 2000. Phase encoding of shot records in prestack migration. Geophysics, 65: 426-436.

Schuster, G.T., Wang, X., Huang, Y., Dai, W. and Boonyasiriwat, C., 2011. Theory of multisource crosstalk reduction by phase-encoded statics. Geophys. J. Internat., 184: 1289-1303.

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

Vigh, D., Starr, E.W. and Kapoor, J., 2009. Developing earth models with full waveform inversion. The Leading Edge, 28: 432-435.

Zhan, G., Dai, W., Boonyasiriwat, C. and Schuster, G.T., 2010. Acoustic multi-source waveform inversion with deblurring. Extended Abstr., 72nd EAGE Conf., Barcelona: G002.

APPENDIX

ENCODED MULTI-SOURCE DEBLURRING FILTER

The phase-encoded multi-source deblurring filter (Aoki and Schuster, 2009; Zhan et al., 2010; Dai et al., 2011) is constructed in the following way.

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

2. Place sources along the surface of the background model and compute the encoded multi-source gather of traces with one FD simulation to get d_1 in Fig. A-1c. Also do the same FD simulation in the model with point scatterers to get another encoded multi-source gather of traces d_2 in Fig. A-1d.

3. Subtract d_1 from d_2 to get a resulting data related to point scatterers only and set this equal to d.

4. Apply encoded multi-source migration operator L^T to d to get the migration image $\mathbf{m}_{mig} = L^T\mathbf{d}$, as shown in Fig. A-1d.

5. Find the local series of deblurring filters $\mathbf{f}_i$ that collapse the migration butterflies (Hu and Schuster, 2001) shown in Fig. A-1d to the points in Fig. A-1e. In this case there will be six different filters $\mathbf{f}_i; i\varepsilon[1,2,3,4,5,6]$, each with dimension K × L. For example, Aoki and Schuster (2009) uses K = 5 and L to be equal to the number of points that represent two wavelengths in the 2D case. Each $\mathbf{f}_i$ approximates the inverse to one of the migration butterflies shown in Fig. A-1d, and are assembled into the deblurring matrix **F** shown in Fig. A-1f. The assumption is that the $\mathbf{f}_i$ 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.

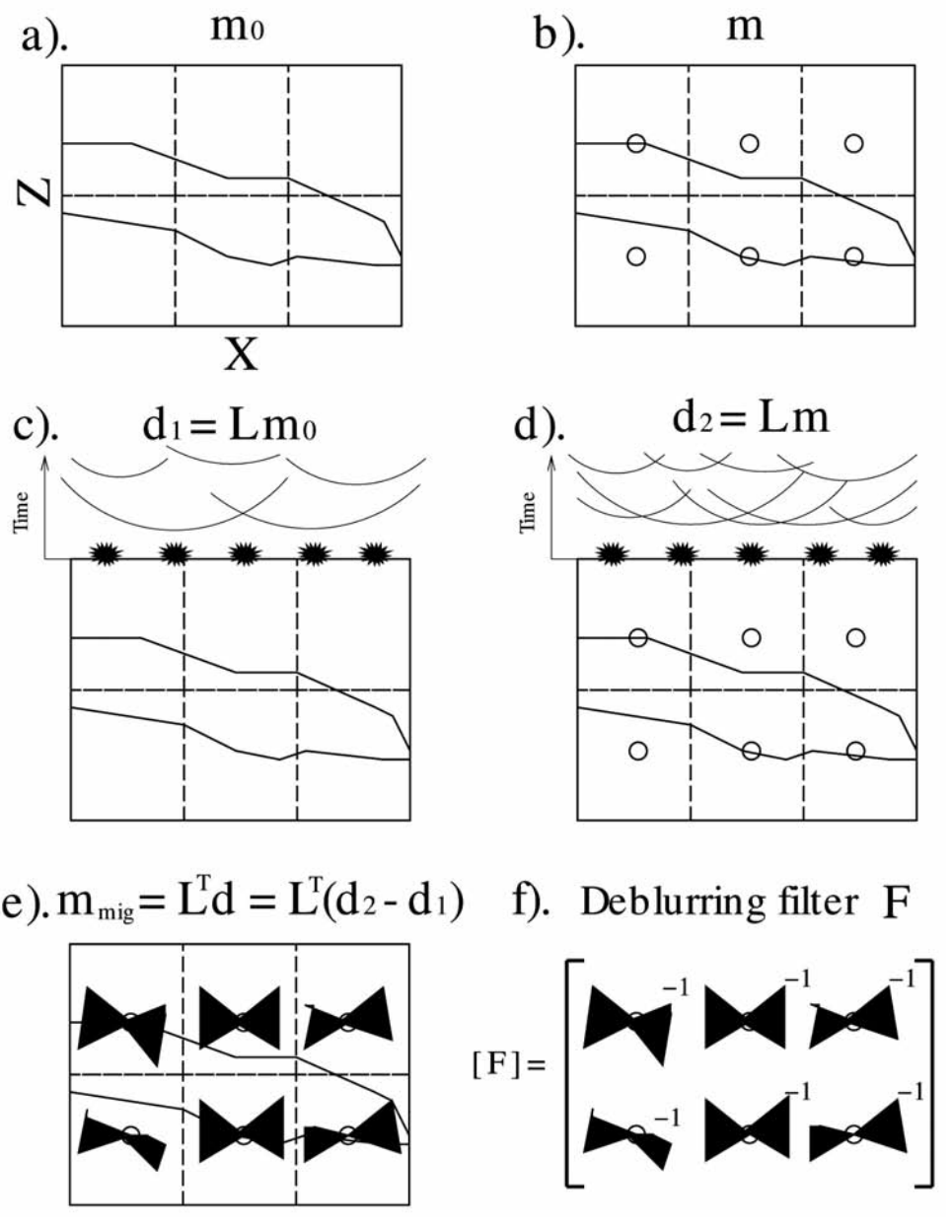

Fig. A-1. Steps for computing the encoded deblurring filter **F**.